IJCNN 2019

Program with Abstracts

DETAILED PROGRAM

Sunday, July 14, 8:00AM-10:00AM

Tutorial: Physics of the Mind
Sunday, July 14, 8:00AM-10:00AM, Room: Sofitel Bellevue 1, Instructor: Harvard University Leonid I. Perlovsky

Tutorial: Modern Gaussian Processes: Scalable Inference and Novel Applications
Sunday, July 14, 8:00AM-10:00AM, Room: Sofitel Bellevue 2, Instructor: Data61, Australia Edwin V. Bonilla and EURECOM, France Maurizio Filippone

Tutorial: Task-Independent and Modality-Independent Developmental Learning Engines: From Theory to Programming (*)
Sunday, July 14, 8:00AM-10:00AM, Room: Sofitel Bellevue 3, Instructor: Juyang Weng and Michigan State University, Juan L. Castro-Garcia

Sunday, July 14, 10:00AM-10:20AM

Special Lecture: Coffee Break
Sunday, July 14, 10:00AM-10:20AM, Room: Sofitel

Sunday, July 14, 10:20AM-12:20PM

Tutorial: Beyond Deep Learning: How to get Fast, Interpretable and Highly Accurate Classifiers
Sunday, July 14, 10:20AM-12:20PM, Room: Sofitel Bellevue 1, Instructor: Lancaster University, UK Plamen Angelov

Tutorial: Deep Learning for Graphs
Sunday, July 14, 10:20AM-12:20PM, Room: Sofitel Bellevue 2, Instructor: Davide Bacciu (Universit{\'a} di Pisa)

Tutorial: Theory and Methodology of Transfer Learning
Sunday, July 14, 10:20AM-12:20PM, Room: Sofitel Bellevue 3, Instructor: AgroParisTech And France Pierre-Alexandre Murena, T{\'e}l{\'e}com ParisTech Antoine Cornuejols and AgroParisTech

Sunday, July 14, 12:20PM-1:30PM

Special Lecture: Lunch Break
Sunday, July 14, 12:20PM-1:30PM, Room: Various locations in the area

Sunday, July 14, 1:30PM-3:30PM

Tutorial: Deep Learning: Artificial Neural Networks and Kernel based Models
Sunday, July 14, 1:30PM-3:30PM, Room: Sofitel Bellevue 1, Instructor: DKE, Maastricht University, Johan A. K. Suykens, ESAT-STADIUS, KU Leuven, Belgium Siamak Mehrkanoon

Tutorial: Machine Learning methods in Spiking Neural Networks for classification problems
Sunday, July 14, 1:30PM-3:30PM, Room: Sofitel Bellevue 2, Instructor: Singapore), Savitha Ramasamy (Institute for Infocomm Research, A*STAR), Suresh Sundaram (Nanyang Technological University, Singapore) Abeegithan Jeyasothy (Nanyang Technological University

Tutorial: Universal Turing Machines and How They Emerge from DN Network
Sunday, July 14, 1:30PM-3:30PM, Room: Sofitel Bellevue 3, Instructor: Michigan State University Juyang Weng

Sunday, July 14, 3:30PM-3:50PM

Special Lecture: Coffee Break
Sunday, July 14, 3:30PM-3:50PM, Room: Sofitel

Sunday, July 14, 3:50PM-5:50PM

Tutorial: Tensor Decompositions for Big Data Analytics: Trends and Applications
Sunday, July 14, 3:50PM-5:50PM, Room: Sofitel Bellevue 1, Instructor: Ilia Kisil Danilo P. Mandic and , Imperial College London Giuseppe G. Calvi

Tutorial: Non-Iterative Learning Methods for Classification and Forecasting
Sunday, July 14, 3:50PM-5:50PM, Room: Sofitel Bellevue 3, Instructor: Technological University, Singapore. P. N. Suganthan

Sunday, July 14, 6:30PM-8:00PM

: Opening Reception

Sunday, July 14, 6:30PM-8:00PM, Room: Pre-function area Intercontinental, Chair: Irwing King

Monday, July 15, 8:10AM-9:30AM

1l: Deep neural networks, Cellular Computational Networks

Monday, July 15, 8:10AM-9:30AM, Room: Ballroom I, Chair: Vanika Singhal

8:10AM *Age and Gender Estimation via Deep Dictionary Learning Regression [#19486]*

Vanika Singhal and Angshul Majumdar, IIITD, India

The paper addresses the problem of estimating age and gender from frontal photos. Most prior studies on deep learning based estimation formulate it as a convolutional neural network based classification problem. For gender it is a two class problem; for age, several age brackets are created to form classes. In this work we formulate it as a regression problem. This is a natural way to handle both gender and age. Gender can be represented as a single variable possible of taking binary values (male or female) whereas age can be represented by a single variable taking non-negative real values. We formulate regression on the newly proposed deep dictionary learning framework. Prior work on this topic, is on unsupervised representation learning; in this work we in-built regression into the deep dictionary learning framework making the formulation supervised. Testing has been done on several state-of-the-art datasets - Adience, MORPH, ChaLearn LAP, LFWA and CelebA. Our method yields age and gender estimation results better than the state-of-the-art.

8:30AM *The Impact of Image Resolution on Facial Expression Analysis with CNNs [#19635]*

Asad Abbas and Stephan Chalup, The University of Newcastle, Australia

While deep learning has achieved state-of-the-art results on many computer vision tasks it is still challenged when interpreting human facial expressions, namely: poor generalisation ability of models across datasets, failure to account for individual differences in similar emotional states, and inability to recognise compound facial expressions and low-intensity or subtle emotional states. This study analyses how the resolution of face images that are input to various Convolutional Neural Network (CNN) models impacts on their ability to recognise compound and low-intensity emotions. Several high-resolution facial expression databases were combined to compile a simple dataset containing high-intensity emotions and a complex data set consisting of compound and low-intensity emotions. In the experiments, standard pre-trained CNN models that were further fine-tuned achieved higher validation accuracies than CNN models that were trained from scratch on the simple data set. However, when tested on the complex data the models trained from scratch generalised better than the refined pre-trained models. Using a technique of output visualisation we could show how our high-resolution CNN models were able to generalise to the complex data where they utilised small facial features that previously were not detectable.

8:50AM *Fast and Efficient Text Classification with Class-based Embeddings [#19584]*

Jonatas Wehrmann, Camila Kolling and Rodrigo Barros, PUCRS, Brazil

Current state-of-the-art approaches for Natural Language Processing tasks such as text classification are either based on Recurrent or Convolutional Neural Networks. Notwithstanding, those approaches often require a long time to train, or large amounts of memory to store the entire trained models. In this paper, we introduce a novel neural network architecture for ultra-fast, memory-efficient text classification. The proposed architecture is based on word embeddings trained directly over the class space, which allows for fast, efficient, and effective text classification. We divide the proposed architecture into four main variations that present distinct capabilities for learning temporal relations. We perform several experiments across four widely-used datasets, in which we achieve results comparable to the state-of-the-art while being much faster and lighter in terms of memory usage. We also present a thorough ablation study to demonstrate the importance of each component within each proposed model. Finally, we show that our model predictions can be visualized and thus easily explained.

9:10AM *Hardening Deep Neural Networks via Adversarial Model Cascades [#19213]*

Deepak Vijaykeerthy, Anshuman Suri, Sameep Mehta and Ponnurangam Kumaraguru, IBM Research, India; IIIT Delhi, India

Deep neural networks (DNNs) are vulnerable to malicious inputs crafted by an adversary to produce erroneous outputs. Works on securing neural networks against adversarial examples achieve high empirical robustness on simple datasets such as MNIST. However, these techniques are inadequate when empirically tested on complex data sets such as CIFAR-10 and SVHN. Further, existing techniques are designed to target specific attacks and fail to generalize across attacks. We propose Adversarial Model Cascades (AMC) as a way to tackle the above inadequacies. Our approach trains a cascade of models sequentially where each model is optimized to be robust towards a mixture of multiple attacks. Ultimately, it yields a single model which is secure against a wide range of attacks; namely FGSM, Elastic, Virtual Adversarial Perturbations and Madry. On an average, AMC increases the model's empirical robustness against various attacks simultaneously, by a significant margin (of 6.225% for MNIST, 5.075% for SVHN and 2.65% for CIFAR-10). At the same time, the model's performance on non-adversarial inputs is comparable to the state-of-the-art models.

2e: Deep learning

Monday, July 15, 8:10AM-9:30AM, Room: Ballroom II, Chair: Martin Pilat

8:10AM *Road Detection via Deep Residual Dense U-Net [#19735]*

Xiaofei Yang, Xutao Li, Yunming Ye, Xiaofeng Zhang, Haijun Zhang, Xiaohui Huang and Bowen Zhang, Harbin Institute of Technology, Shenzhen, China; School of Information Engineering East China Jiaotong University, China

Road extraction from aerial images is a hot research topic. With the advancement of convolutional neural network (CNN), several CNN-based road detection methods have been developed. However, most of them do not make full use of the hierarchical features from the original aerial images. In this paper, we propose a novel residual dense U- Net (RDUN), a semantic segmentation network which combines the strengths of residual learning, DenseNet, and U-Net, to overcome the drawback. Our proposed RDUN can fully exploit the hierarchical features from all the convolutional layers, which utilizes the residual dense blocks (RDB) to build up a U-Net architecture. The benefits of our model are two-fold. First, by using the RDB abundant local features can be extracted and fused effectively. Second, based the local features, hierarchical features are constructed by shortcut connections between layers in RDB. Extensive experiments are carried out on a real-world road detection dataset and the results demonstrate the proposed RDUN outperforms state-of-the-art competitors.

8:30AM *Using Local Convolutional Units to Defend Against Adversarial Examples [#20328]*

Matej Kocian and Martin Pilat, Charles University, Faculty of Mathematics and Physics, Czech Republic

Deep neural networks are known to be sensitive to adversarial examples -- inputs that are created in such a way that they are similar (if viewed by people) to clean inputs, but the neural network has high confidence that they belong to another class. In this paper, we study a new type of neural network unit similar to the convolutional units, but with a more local behavior. The unit is based on the Gaussian radial basis function. We show that if we replace the first convolutional layer in a convolutional network by the new layer (called RBFolutional), we obtain better robustness towards adversarial samples on the MNIST and CIFAR10 datasets, without sacrificing the performance on the clean examples.

8:50AM *Sparsity as the Implicit Gating Mechanism for Residual Blocks [#20428]*

Shaeke Salman and Xiuwen Liu, Florida State University, United States

Neural networks are the core component in the recent empirical successes of deep learning techniques in challenging tasks. Residual network (ResNet) architectures have been instrumental in improving performance in object recognition and other tasks by enabling training much deeper neural networks. Studies of residual networks reveal that they are robust to removing layers. However, it is still an open question of why residual networks behave well and how they make it feasible to train networks with many layers. In this paper, we show that sparsity of the residual blocks acts as the implicit gating mechanism. When a neuron is inactive, it behaves as a node in an information highway, allowing the information from the previous layer to pass to the next layer unchanged. As the identity function has a derivative of 1, it avoids the exploding or vanishing gradient problem that is known to contribute to the difficulty of training deep neural networks. When a neuron is active, it captures input-output relationships that are necessary to achieve good performance. By using the ReLu activation functions, residual blocks produce sparse outputs for typical inputs. We perform systematic experimental analysis on the residual blocks of trained ResNet models and show that sparsity acts as the implicit gate for deep residual networks.

9:10AM *Agile Domain Adaptation [#19077]*

Jingjing Li, Mengmeng Jing, Yue Xie, Ke Lu and Zi Huang, University of Electronic Science and Technology of China, China; The University of Queensland, Australia

Domain adaptation investigates the problem of leveraging knowledge from a well-labeled source domain to an unlabeled target domain, where the two domains are drawn from different data distributions. Because of the distribution shifts, different target samples have distinct degrees of difficulty in adaptation. However, existing domain adaptation approaches overwhelmingly neglect the degrees of difficulty and deploy exactly the same framework for all of the target samples. Generally, a simple or shadow framework is fast but rough. A sophisticated or deep framework, on the contrary, is accurate but slow. In this paper, we aim to challenge the fundamental contradiction between the accuracy and speed in domain adaptation tasks. We propose a novel approach, named agile domain adaptation, which agilely applies optimal frameworks to different target samples and classifies the target samples according to their adaptation difficulties. Specifically, we propose a paradigm which performs several early detections before the final classification. If a sample can be classified at one of the early stage with enough confidence, the sample would exit without the subsequent processes. Notably, the proposed method can significantly reduce the running cost of domain adaptation approaches, which can extend the application scenarios of domain adaptation to even mobile devices and real-time systems. Extensive experiments on two open benchmarks verify the effectiveness and efficiency of the proposed method.

8a: Applications of deep networks

Monday, July 15, 8:10AM-9:30AM, Room: Ballroom III, Chair: Plamen Angelov

8:10AM *Syntax Tree Aware Adversarial Question Rewriting for Answer Selection [#19990]*

Shuang Qin, Wenge Rong, Libin Shi, Jianxin Yang, Haodong Yang and Zhang Xiong, Beihang University, China; Microsoft, China

Answer selection is an important method to achieve better user experience in question answering (QA) systems and it is essential in ensuring better QA matching performance. Improving mutual information between QA pairs is a useful way to obtain the matching degree improvement and question rewriting has been proven a helpful way in utilizing mutual interaction between questions and answers. In this research, we focus on syntax tree aware question rewriting inspired by the thought of integrating syntactic information into question answering. Besides, to improve the quality of rewriting, we employ the generative adversarial network for rewriting optimization, which consists of a syntax tree aware rewriting model and a discriminator. The quality information given by the discriminator guides the optimizing of the rewriting model in the training phase. The experimental study has shown the effectiveness of syntax tree aware question rewriting and utilizing the generative adversarial network for rewriting.

8:30AM *Paraphrase Generation with Collaboration between the Forward and the Backward Decoder [#19669]*

Wang Qianlong and Ren Jiangtao, Sun Yat-sen University, China

Automatic paraphrase generation plays a key role in many natural language applications. The dominant paraphrase generation models are the encoder-decoder neural networks with attention, where the decoder uses the information of the source text while predicting target text. However, the outputs of these paraphrase models often suffer the semantic error problem. This problem is caused by the inadequate information of the decoder. In this work, we introduce a novel neural model to solve this problem, called Collaboration between the Forward and the Backward Decoder. Specifically,

the hidden states of the backward decoder are used as supplementary information of the forward decoder. Therefore, the forward decoder can generate more reasonable paraphrase text using the target-side future contextual. Conversely, the backward decoder employs the hidden states of the forward decoder to prevent the semantic error problem. As two experimental examples show, the proposed model can generate the high-quality paraphrase through this collaboration mechanism. The empirical study on two benchmark datasets demonstrates that our model outperforms some baselines and achieves the state-of-the-art performance.

8:50AM *Seq-DNC-seq: Context aware dialog generation system through external memory [#20383]*

Donghyun Kang and Minho Lee, School of Electronics Engineering, Kyungpook National Universty, Korea (South)

Most of the conventional Seq2seq based chit-chat models analyze and process at most one or two sentences at a time and are trained to answer to specific conversation patterns. However, in real chit-chat conversations, a single sentence can be interpreted in various ways according to the preceding context and it is hard to make a proper response for the same input sentence without knowing the previous long term context. Therefore, existing Seq2seq based chit-chat models have a limitation in understanding the previous context and properly responding to it in a multi-turn chit-chat conversation. To overcome this problem, dialogue generation models should be able to memorize the long term previous contextual information related to a current conversation. In this paper, we propose a new dialogue generation model, which uses Differentiable Neural Computer (DNC) in the conventional Seq2seq model named as Seq-DNC-seq. This model incorporates external memory into the conventional Seq2seq structure to generate an appropriate dialogue based on the memory of previous conversations. Experiments show that the proposed Seq-DNC-seq model, used in multi-turn chit-chat scenario, successfully generates the proper output sentences when given the same inputs with different context. Our model not only helps the agent overcome the existing limitations of chit-chat conversation models but also enables context understanding in longer conversations.

9:10AM *Robust and Accurate Short-Term Load Forecasting: A Cluster Oriented Ensemble Learning Approach [#20052]*

Fateme Fahiman, Sarah M. Erfani and Christopher Leckie, The University of Melbourne, Australia

One of the most critical tasks for operating a power system is load forecasting in order to keep balance between demand and supply and for planning infrastructure. Errors in load forecasting can result in significant cost increases for electricity suppliers and increase the chance of unexpected blackouts or brownouts. Improving the accuracy of short term load forecasting is a challenging open problem. This paper proposes a novel framework for short-term load forecasting using an effective new combination of c-Shape clustering, LSTM networks and Xgboost methods. In particular, our proposed approach introduces an ensemble process together with novel features that lead to improved accuracy of the load forecasting model. The performance of the proposed framework is validated with publicly available real-life data from Australian Energy Market Operator, as well as through on-site deployment, which has led to substantially higher accuracy over existing methods.

1h: Spiking neural networks

Monday, July 15, 8:10AM-9:30AM, Room: Duna Salon I, Chair: Kaushik Roy

8:10AM *A Comprehensive Analysis on Adversarial Robustness of Spiking Neural Networks [#20338]*

Saima Sharmin, Priyadarshini Panda, Syed Shakib Sarwar, Chankyu Lee, Wachirawit Ponghiran and Kaushik Roy, Purdue University, United States

In this era of machine learning models, their func- tionality is being threatened by adversarial attacks. In the face of this struggle for making artificial neural networks robust, finding a model, resilient to these attacks, is very important. In this work, we present, for the first time, a comprehensive analysis of the behavior of more bio-plausible networks, namely Spiking Neural Network (SNN) under state-of-the-art adversarial tests. We perform a comparative study of the accuracy degradation between conventional VGG-9 Artificial Neural Network (ANN) and equivalent spiking network with CIFAR-10 dataset in both whitebox and blackbox setting for different types of single-step and multi-step FGSM (Fast Gradient Sign Method) attacks. We demonstrate that SNNs tend to show more resiliency compared to ANN under blackbox attack scenario. Additionally, we find that SNN robustness is largely dependent on the corresponding train- ing mechanism. We observe that SNNs trained by spike-based backpropagation are more adversarially robust than the ones obtained by ANN-to-SNN conversion rules in several whitebox and blackbox scenarios. Finally, we also propose a simple, yet, effective framework for crafting adversarial attacks from SNNs. Our results suggest that attacks crafted from SNNs following our proposed method are much stronger than those crafted from ANNs.

8:30AM *Multi-layered Spiking Neural Network with Target Timestamp Threshold Adaptation and STDP [#20266]*

Pierre Falez, Pierre Tirilly, Ioan Marius Bilasco, Philippe Devienne and Pierre Boulet, Univ. Lille, CNRS, Centrale Lille, UMR 9189 -- CRIStAL -- Centre de Recherche en Informatique, Signal et Automatique de Lille, F-59000, Lille, France, France; Univ. Lille, CNRS, Centrale Lille, UMR 9189 -- CRIStAL -- Centre de Recherche en Informatique, Signal et Automatique de Lille, IMT Lille Douai, F-59000, Lille, France, France

Spiking neural networks (SNNs) are good candidates to produce ultra-energy-efficient hardware. However, the performance of these models is currently behind traditional methods. Introducing multi-layered SNNs is a promising way to reduce this gap. We propose in this paper a new threshold adaptation system which uses a timestamp objective at which neurons should fire. We show that our method leads to state-of-the-art classification rates on the MNIST dataset (98.60%) and the Faces/Motorbikes dataset (99.46%) with an unsupervised SNN followed by a linear SVM. We also investigate the sparsity level of the network by testing different inhibition policies and STDP rules.

8:50AM *Neural Population Coding for Effective Temporal Classification [#19925]*

Zihan Pan, Jibin Wu, Yansong Chua, Malu Zhang and Haizhou Li, National University of Singapore, Singapore; Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore, Singapore

Neural encoding plays an important role in faithfully describing the temporally rich patterns, instances of which, include human speech and environmental

sound. For tasks that involve classifying such spatio- temporal patterns with the Spiking Neural Networks (SNNs), how these patterns are encoded directly influence the difficulty of the task. In this paper, we compare several existing temporal and population coding schemes and evaluate them on both speech (TIDIGITS) and sound (RWCP) datasets. We show that, with population neural codings, the encoded patterns are linearly separable using the Support Vector Machine (SVM). We note that the population neural codings effectively project the temporal information onto the spatial domain, thus improving linear separability in the spatial dimension, achieving an accuracy of 95\% and 100\% for TIDIGITS and RWCP datasets classified using the SVM, respectively. This observation suggests that an effective neural coding scheme greatly simplifies the classification problem such that a simple linear classifier would suffice. The above datasets are then classified using the Tempotron, an SNN-based classifier. SNN classification results agree with the SVM findings that population neural codings help to improve classification accuracy. Hence, other than the learning algorithm, effective neural encoding is just as important for an SNN designed to recognize spatio-temporal patterns. It is an often neglected but powerful abstraction that deserves further study.

9:10AM *Competitive STDP-based Feature Representation Learning for Sound Event Classification [#19448]*

Jibin Wu, Yansong Chua, Malu Zhang and Haizhou Li, National University of Singapore, Singapore; Institute for Infocomm Research, A*STAR, Singapore

Humans are good at discriminating environmental sounds and associating them with opportunities or dangers. While the deep learning approach to sound event classification (SEC) is achieving human parity, unsolved problems remain, instances include high computational cost, requirement of massive labeled training data, and question of biological plausibility. Motivated by the human auditory system, we propose a biologically plausible SEC system, which integrates the auditory front-end, population coding, competitive spike-timing- dependent plasticity (STDP) based feature representation learning and supervised temporal classification into a unified spiking neural network (SNN) system. The proposed SEC system achieves a classification accuracy on the RWCP database that is on par with other competitive baseline systems. Furthermore, the STDP-based feature representation learning shows low intra-class variability and high inter-class variability in our experiments, which is highly desirable for pattern classification tasks.

1n: Other topics in artificial neural networks

Monday, July 15, 8:10AM-9:30AM, Room: Duna Salon II, Chair: Alexander Makarenko

8:10AM *Tensor Ring Restricted Boltzmann Machines [#20289]*

Maolin Wang, Chenbin Zhang, Yu Pan, Jing Xu and Zenglin Xu, SMILE Lab, School of Computer Science and Engineering, University of Electronic Science and Technology of China, China

Restricted Boltzmann Machines are important and useful generative models which learn a probability distribution from a set of vector inputs. Despite their success in a number of applications, standard RBMs designed for vectorized inputs are incapable of dealing with high-order data, since vectorization of high-order data may cause both modes collapsing and explosive parameter growth. To address this issue, we formulate a new tensor-input RBM model, which employs the tensor-ring (TR) decomposition structure to naturally represent the high-order relationship between the visual layer and the hidden layer. For convenience, we name the proposed model as TR-RBM. In particular, the tensor ring decomposition enjoys many good properties, such as the rank stableness, leading to better generalization performance compared with other low-rank decomposition methods. Moreover, TR-RBM can also reduce the complexity of RBM by reshaping of both visible and hidden layers into the tensor forms, leading a significant drop of parameter size. Experimental results in comparison with the classical RBMs and the Matrix-Product-Operator RBM have shown the promising performance of the proposed method in the tasks of feature extraction and denoising.

8:30AM *Multiple-Valued Artificial Neural Networks [#19527]*

Alexander Makarenko, Institute for Applied System Analysis at National Technical University of Ukraine "KPI", Ukraine

The generalizations of common single-valued artificial neural networks are proposed. The main proposition is to consider multiple-valued networks. Such neural networks can have multiple values of elements at given time moment. Neural networks with strong anticipation property are considered as the examples. Also some new research problems are proposed.

8:50AM *Convolutional Neural Network Architecture Design by the Tree Growth Algorithm Framework [#20310]*

Ivana Strumberger, Eva Tuba, Nebojsa Bacanin, Raka Jovanovic and Milan Tuba, Singidunum University, Serbia and Montenegro; Hamad bin Khalifa University, Qatar

This paper presents tree growth algorithm framework for designing convolutional neural network architecture. Convolutional neural networks are a special class of deep neural networks that typically consist of several convolution, pooling and fully connected layers. Convolutional neural networks have proved to be a robust method for tackling various image classification tasks. One of the most important challenges from this domain is to find the network architecture that has the best performance for the specific application. The performance of the network depends on the set of hyper-parameter values such as the number of convolutional and dense layers, the number of kernels per layer and kernel size. Optimization of hyper-parameters was performed by novel tree growth algorithm that belongs to the group of swarm intelligence metaheuristics. The robustness, performance and solutions quality of the proposed framework was validated against the well-known MNIST dataset. Conducted comparative analysis demonstrated that the proposed frameworks obtains promising results in this domain.

9:10AM *Encoding robust representation for graph generation [#20350]*

Dongmian Zou and Gilad Lerman, University of Minnesota, United States

Generative networks have made it possible to generate meaningful signals such as images and texts from simple noise. Recently, generative methods based on GAN and VAE were developed for graphs and graph signals. However, the mathematical properties of these methods are unclear, and training good generative models is difficult. This work proposes a graph generation model that uses a recent adaptation of Mallat's scattering transform to graphs. The proposed model is naturally composed of an encoder and a decoder. The encoder is a Gaussianized graph scattering transform, which is robust to signal and graph manipulation. The decoder is a simple fully connected network that is adapted to specific tasks, such as link prediction, signal generation on graphs and full graph and signal generation. The training of our proposed system is efficient since it is only applied to the decoder and the hardware requirements are moderate. Numerical results

demonstrate state- of-the-art performance of the proposed system for both link prediction and graph and signal generation.

2a: Supervised learning

Monday, July 15, 8:10AM-9:30AM, Room: Duna Salon III, Chair: Jacek Mandziuk

8:10AM *Who should bid higher, NS or WE, in a given Bridge deal? [#20098]*

Jacek Mandziuk and Jakub Suchan, Warsaw University of Technology, Faculty of Mathematics and Information Science, Poland

The paper proposes a neural model for a direct comparison of the two so-called Double Dummy Bridge Problem (DDBP) instances, along with a practical use-case for determining which pair, NS or WE, should propose the higher deal during a bidding phase in a Bridge game. The proposed system is composed of two identical subnetworks combined by a comparator layer placed on top of them. The base of each subnetwork is a shallow autoencoder (AE) which is further connected with a Multilayer Perceptron. The system is trained in two phases - an unsupervised one - used to create a meaningful feature-based input representation in AE compression layer, and a supervised one - meant for fine- tuning of the whole model. Training and test data are composed of pairs of Bridge deals in which the second deal in a pair is the first one rotated by 90 degrees. Since the task is to point which of the two deals promise a higher contract for the NS pair, due to deal rotation within a pair, the system effectively answers the title question ``Who should bid higher, NS or WE, in a given deal?". The proposed approach is experimentally compared with two other methods: one relying on a neural system solving the DDBP and the other one employing several estimators of hand strength used by experienced players. The results clearly indicate that both neural network approaches outperform the usage of human-scoring systems by a large margin, most notably in the trump (suit) contract.

8:30AM *A Count-sketch to Reduce Memory Consumption when Training a Model with Gradient Descent [#19170]*

Wissam Siblini, Frank Meyer and Pascale Kuntz, University of Nantes (LS2N) & Worldline, France; Orange Labs, France; University of Nantes (LS2N), France

Training machine learning models requires the storage of an ever-growing number n of parameters, which leads to memory management issues. In this paper, we propose a novel approach to accurately estimate the largest parameters sufficient for prediction tasks while saving memory. Inspired by the heavy hitter identification problem in data stream analysis, our count-sketch based strategy consists in independently storing t times all parameter values on different sets of r << n counters by randomly aggregating several values in the same counters. Each parameter is then approximated by the median of the values stored in its t associated counters. We conduct experiments on the popular use case of the linear regression problem for nine multi-label datasets inducing various model sizes (up to 11.7 GB of RAM). It shows that replacing the regular storage of all parameters with the proposed count-sketch storage preserves the quality of predictive performances while significantly reducing the use of memory (from 80% to 99.9%). A theoretical bound on the approximation error completes the experiments. In practice, this work gives a chance to several potentially interesting memory-hungry models to be implemented without having to resort to costly computational environments.

8:50AM *AX-DBN: An Approximate Computing Framework for the Design of Low-Power Discriminative Deep Belief Networks [#20401]*

Ian Colbert, Ken Kreutz-Delgado and Srinjoy Das, UC San Diego, United States

The power budget for embedded hardware implementations of Deep Learning algorithms can be extremely tight. To address implementation challenges in such domains, new design paradigms, like Approximate Computing, have drawn significant attention. Approximate Computing exploits the innate error-resilience of Deep Learning algorithms, a property that makes them amenable for deployment on low-power computing platforms. This paper describes an Approximate Computing design methodology, AX-DBN, for an architecture belonging to the class of stochastic Deep Learning algorithms known as Deep Belief Networks (DBNs). Specifically, we consider procedures for efficiently implementing the Discriminative Deep Belief Network (DDBN), a stochastic neural network which is used for classification tasks, extending Approximation Computing from the analysis of deterministic to stochastic neural networks. For the purpose of optimizing the DDBN for hardware implementations, we explore the use of: (a) Limited precision of neurons and functional approximations of activation functions; (b) Criticality analysis to identify the nodes in the network which can operate at reduced precision while allowing the network to maintain target accuracy levels; and (c) A greedy search methodology with incremental retraining to determine the optimal reduction in precision for all neurons to maximize power savings. Using the AX-DBN methodology proposed in this paper, we present experimental results across several network architectures that show significant power savings under a user-specified accuracy loss constraint with respect to ideal full precision implementations.

9:10AM *Dimensionality Reduction in Multilabel Classification with Neural Networks [#19679]*

Jacek Mandziuk and Adam Zychowski, Warsaw University of Technology, Poland

A new neural network method for Dimensionality Reduction (DR) of the input feature space in Multilabel Classification (MC) problems is proposed and experimentally evaluated in this paper. The method (abbreviated as TCART-MR) can be used in two possible scenarios: either as a stand-alone DR pre-processing phase, preceding subsequent application of any particular MC algorithm, or as a compact MC approach in which TCART-MR is applied twice - first to DR task and then to MC problem with reduced input space. Extensive experimental results proved statistically relevant advantage of TCART- MR over three state-of-the-art approaches in DR domain (in the context of MC), as well as its superiority over 10 state-of-the-art MC algorithms listed in a recent MC survey paper. The MC tests were performed on a set of 9 benchmark problems and 16 evaluation measures (leading to 144 experimental cases in total).

1a: Feedforward neural networks

Monday, July 15, 8:10AM-9:30AM, Room: Panorama I, Chair: Debasmit Das

8:10AM *Zero-shot Image Recognition Using Relational Matching, Adaptation and Calibration [#19040]*

Debasmit Das and C. S. George Lee, Purdue University, United States

Zero-shot learning (ZSL) for image classification focuses on recognizing novel categories that have no labeled data available for training. The learning is generally carried out with the help of mid-level semantic descriptors associated with each class. This semantic-descriptor space is generally shared by both seen and unseen categories. However, ZSL suffers from hubness, domain discrepancy and biased-ness towards seen classes. To tackle these problems, we propose a three-step approach to zero-shot learning. Firstly, a mapping is learned from the semantic-descriptor space to the image-feature space. This mapping learns to minimize both one-to-one and pairwise distances between semantic embeddings and the image features of the corresponding classes. Secondly, we propose test-time domain adaptation to adapt the semantic embedding of the unseen classes to the test data. This is achieved by finding correspondences between the semantic descriptors and the image features. Thirdly, we propose scaled calibration on the classification scores of the seen classes. This is necessary because the ZSL model is biased towards seen classes as the unseen classes are not used in the training. Finally, to validate the proposed three-step approach, we performed experiments on four benchmark datasets where the proposed method outperformed previous results. We also studied and analyzed the performance of each component of our proposed ZSL framework.

8:30AM *Non-negative Autoencoder with Simplified Random Neural Network [#19231]*

Yonghua Yin and Erol Gelenbe, Imperial College London, United Kingdom

A new shallow multi-layer auto-encoder that combines the spiking Random Neural Network (RNN) with the network architecture typically used in deep-learning, is proposed with a learning algorithm inspired by non-negative matrix factorization which satisfies the non-negative probability constraints of the RNN. Auto- encoders equipped with this learning algorithm are tested on typical images including the MNIST, Yale face and CIFAR-10 datasets, and also using 16 real- world datasets from different areas, exhibiting the desired high learning and recognition accuracy. Montecarlo simulations of the stochastic spiking behaviour of this RNN auto-encoder have also been carried out, showing that it can be implemented in a highly parallel manner to achieve substantial speed improvements.

8:50AM *The Cramming, Softening and Integrating Learning Algorithm with Parametric ReLU Activation Function for Binary Input/Output Problems [#19652]*

Yu-Han Tsai, Yu-Jie Jheng and Rua-Huan Tsaih, Dept. of Management Information Systems, National Chengchi University, Taiwan

Rare Artificial Neural Networks studies address simultaneously the challenges of (1) systematically adjusting the amount of used hidden layer nodes within the learning process, (2) adopting Parametric ReLU activation function instead of tanh function for fast learning, and (3) guaranteeing learning all training data. This study will address these challenges through deriving the CSI (Cramming, Softening and Integrating) learning algorithm for the single-hidden layer feed-forward neural networks with the binary input/output and making the technical justification. To further verify the proposed learning algorithm, this study conducts an empirical experiment using SPECT heart diagnosis data set from UCI Machine Learning repository. The learning algorithm is implemented via the advanced TensorFlow and GPU.

9:10AM *Mutual Information Generation for Improving Generalization and Interpretation in Neural Network [#19886]*

Ryotaro Kamimura, Tokai University, Japan

The present paper aims to show the importance of information augmentation by increasing the number of neurons. We suppose that, to cope with new input patterns, we should store information concerning given input patterns as much as possible, even if it is considered to be noises or harmful in improving performance. This is because the noises or unnecessary information may be changed into useful information for new situations. To realize this, we introduce mutual information between input and outputs, and information augmentation is supposed to be realized simply by increasing neurons and, correspondingly, dimensionality. In addition, all neurons are used as equally and differently as possible to make full use of the redundant number of neurons. The method was applied to the office equipment sales data set. The results confirmed that mutual information could be increased and, correspondingly, generalization performance could be increased. The reason for improved performance can be explained by the detailed representation of information contained in input patterns created by a large number of neurons. The detailed representation increases the possibility of obtaining necessary information for learning.

1l: Deep neural networks, Cellular Computational Networks

Monday, July 15, 8:10AM-9:30AM, Room: Panorama II, Chair: Nils Schaetti

8:10AM *Behaviors of Reservoir Computing Models for Textual Documents Classification [#19907]*

Nils Schaetti, University of Neuchatel, Switzerland

Reservoir Computing is a paradigm of recurrent neural network (RNN) models, attractive because of its ease of training and new neuromorphic optoelectronic implementations. Applied with success to time series prediction and speech recognition, few works have so far studied the behavior of these networks on natural language processing (NLP) tasks. Therefore, we decided to explore the ability of Echo State Network-based Reservoir Computing (ESN) models with additional embedding layers to classify text documents of the Reuters C50 data set based on authorship. We explored various learned representations such as word and character embedding and deep feature extractors. Our experiments demonstrate that ESN models can achieve state-of-the-art results on this task and are competitive with common models such as Support Vector Machines (SVM). Moreover, we show that these models compute documents as data streams and could then be able to handle other tasks such as event detection and text segmentation. The best performance is obtained by an ESN with a large reservoir of 1,500 neurons based on word vectors. We think that these results demonstrate the possibility of processing massive quantities of textual data in the future using Reservoir Computing-based systems.

8:30AM *Encoding of a Chaotic Attractor in a Reservoir Computer: A Directional Fiber Investigation [#19346]*

Sanjukta Krishnagopal, Garrett Katz, Michelle Girvan and James Reggia, University of Maryland, United States; Syracuse University, United States

In this work, we study the dynamical properties of a machine learning technique called reservoir computing in order to gain insight into how representations of chaotic signals are encoded through learning. We train the reservoir on individual chaotic Lorenz signals. The Lorenz system is characterized by a set of equations and known to have three fixed points, all of which are unstable in the chaotic regime of the strange attractor. Exploration of the fixed points of the reservoir whose outputs are trained allows us to understand whether inherent Lorenz dynamics are transposed onto reservoir dynamics during learning. We do so by using a novel fixed point finding technique called directional fibers. Directional fibers are mathematical objects that systematically locate fixed points in a high

dimensional space, and are found to be competitive and complementary with other traditional approaches. We find that the reservoir, after training of output weights, contains a higher dimensional projection of the Lorenz fixed points with matching stability, even though the training data did not include the fixed points. This tells us that the reservoir does indeed learn dynamical properties of the Lorenz attractor. We also find that the directional fiber also identifies additional fixed points in the reservoir space outside the projected Lorenz attractor region; these amplify perturbations during prediction and play a role in failure of long-term time series prediction.

8:50AM *Ensembling 3D CNN Framework for Video Recognition [#19148]*

Ruolin Huang, Hongbin Dong, Guisheng Yin and Qiang Fu, Harbin Engigeering University, China

Video-based behavior recognition is a challenging research topic. The three dimensional convolution neural network (3D CNN) is effectively adopted to capture features from videos directly. 3D CNN is extended by two-dimensional convolution neural network, in which a time dimension is added. 3D CNN is better than two-dimensional convolution network in expressing effective motion information, and it has certain advantages. In order to make better use of the valuable features extracted from the original video information, only stacked RGB frame data sets can be used as the input of network. Ensembling 3D CNN framework for video recognition is proposed in the paper. Firstly, the pre-training model of Sports-1M is initialized firstly, and a 3D convolution neural network based on multi-level feature fusion is constructed. . The final high-dimensional feature combination is obtained by fusing multiple convolution features. Then 3D convolutional neural network based on ensemble learning is proposed to increase motion information, enrich motion features and enhance the robustness of single feature representation. Three incomplete training data sets are obtained by Bagging algorithm. To get different networks, three data sets are employed to train three 3D convolution neural networks respectively, and the output of the three networks is integrated. The output features of the three networks are input into the SVM classifier through the Stacking algorithm and the final results are obtained. The integration effects of different ensemble methods are compared. The experimental results show that the method of this work can improve recognition accuracy on UCF-101 data set effectively.

9:10AM *Response Characterization for Auditing Cell Dynamics in Long Short-term Memory Networks [#19265]*

Ramin Hasani, Alexander Amini, Mathias Lechner, Felix Naser, Radu Grosu and Daniela Rus, Technische Universitat Wien (TU Wien), Austria; Massachusetts Institute of Technology (MIT), United States; Institute of Science and Technology (IST) Austria, Austria

In this paper, we introduce a novel method to interpret recurrent neural networks (RNNs), particularly long short-term memory networks (LSTMs) at the cellular level. We propose a systematic pipeline for interpreting individual hidden state dynamics within the network using response characterization methods. The ranked contribution of individual cells to the network's output is computed by analyzing a set of interpretable metrics of their decoupled step and sinusoidal responses. As a result, our method is able to uniquely identify neurons with insightful dynamics, quantify relationships between dynamical properties and test accuracy through ablation analysis, and interpret the impact of network capacity on a network's dynamical distribution. Finally, we demonstrate the generalizability and scalability of our method by evaluating a series of different benchmark sequential datasets.

Neural Network Models

Monday, July 15, 8:10AM-9:30AM, Room: Panorama III, Chair: Thar Baker

8:10AM *Simple 1-D Convolutional Networks for Resting-State fMRI Based Classification of Psychiatric Disorders [#20481]*

Ahmed Al Gazzar, Leonardo Cerliani, Guido Van Wingen and Rajat Mani Thomas, AMC, University of Amsterdam, Netherlands

Deep learning methods are increasingly being used with neuroimaging data like structural and function magnetic resonance imaging (MRI) to predict the diagnosis of neuropsychiatric and neurological disorders. For psychiatric disorders in particular, it is believed that one of the most promising modality is the resting-state functional MRI (rsfMRI), which captures the intrinsic connectivity between regions in the brain. Because rsfMRI data points are inherently high-dimensional (1M), it is impossible to process the entire input in its raw form. In this paper, we propose a very simple transformation of the rsfMRI images that captures all of the temporal dynamics of the signal but sub-samples its spatial extent. As a result, we use a very simple 1-D convolutional network which is fast to train, requires minimal preprocessing and performs at par with the state-of-the-art on the classification of Autism spectrum disorders.

8:30AM *Projectron - A Shallow and Interpretable Network for Classifying Medical Images [#19461]*

Aditya Sriram, Shivam Kalra and Hamid Tizhoosh, University of Waterloo, Canada

This paper introduces the "Projectron" as a new neural network architecture that uses Radon projections to both classify and represent medical images. The motivation is to build shallow networks which are more interpretable in the medical imaging domain. Radon transform is an established technique that can reconstruct images from parallel projections. The Projectron first applies global Radon transform to each image using equidistant angles and then feeds these transformations for encoding to a single layer of neurons followed by a layer of suitable kernels to facilitate a linear separation of projections. Finally, the Projectron provides the output of the encoding as an input to two more layers for final classification. We validate the Projectron on five publicly available datasets, a general dataset (namely MNIST) and four medical datasets (namely Emphysema, IDC, IRMA, and Pneumonia). The results are encouraging as we compared the Projectron's performance against MLPs with raw images and Radon projections as inputs, respectively. Experiments clearly demonstrate the potential of the proposed

8:50AM *A Fast Feature Extraction Algorithm for Image and Video Processing [#19608]*

Sadiq H. Abdulhussain, Abd Rahman Ramli, Basheera M. Mahmmod, M. Iqbal Saripan, S.A.R. Al-Haddad, Thar Baker, Wameedh N. Flayyih and Wissam A. Jassim, University of Baghdad, Iraq; Universiti Putra Malaysia, Malaysia; Liverpool John Moores University, United Kingdom; University of Dublin, Ireland

Medical images and videos are utilized to discover, diagnose and treat diseases. Managing, storing, and retrieving stored images effectively are considered important topics. The rapid growth of multimedia data, including medical images and videos, has caused a swift rise in data transmission volume and repository size. Multimedia data contains useful information; however, it consumes an enormous storage space. Therefore, high processing time for that sheer volume of data will be required. Image and video applications demand for reduction in computational cost (processing time) when extracting features. This paper introduces a novel method to compute transform coefficients (features) from images or video frames. These features are used to represent the local visual content of images and video frames. We compared the proposed method with the traditional approach of feature extraction using a standard image technique. Furthermore, the proposed method is employed for shot boundary detection (SBD) applications to detect transitions in video frames. The standard TRECVID 2005, 2006, and 2007 video datasets are used to evaluate the performance of the SBD applications. The achieved results show that the

proposed algorithm significantly reduces the computational cost in comparison to the traditional method.

9:10AM *Emotion helps Sentiment: A Multi-task Model for Sentiment and Emotion Analysis [#19685]*
Abhishek Kumar, Asif Ekbal, Daisuke Kawahra and Sadao Kurohashi, IIT Patna, India; Kyoto University, Japan

In this paper, we propose a two-layered multi-task attention based neural network that performs sentiment analysis through emotion analysis. The proposed approach is based on Bidirectional Long Short-Term Memory and uses Distributional Thesaurus as a source of external knowledge to improve the sentiment and emotion prediction. The proposed system has two levels of attention to hierarchically build a meaningful representation. We evaluate our system on the benchmark dataset of SemEval 2016 Task 6 and also compare it with the state-of-the-art systems on Stance Sentiment Emotion Corpus. Experimental results show that the proposed system improves the performance of sentiment analysis by 3.2 F-score points on SemEval 2016 Task 6 dataset. Our network also boosts the performance of emotion analysis by 5 F-score points on Stance Sentiment Emotion Corpus.

S01: Information Theory and Deep Learning

Monday, July 15, 8:10AM-9:30AM, Room: Panorama IV, Chair: Arturo Marban

8:10AM *Feature selection for orthogonal broad learning system based on mutual information [#19661]*
Liu Zhicheng, Chen Bao, Xie Bingxue, Huang Pingqiang and Zhu Ziqi, Wuhan University of Science and Technology, China

The broad learning system is a recently proposed neural network model. Different from the deep neural network model, the BLS has a flatted structure. Therefore, it can be trained efficiently and enjoyed much popularity on many applications. In this paper, we investigate the BLS from the perspective of information theory. By applying the newly proposed matrix-based R\'{e}nyi's α-entropy, we analysis the mutual information between the feature nodes and the output, and the results indicate that the contribution of different node varies from each other. Therefore, we propose a feature selection algorithm for orthogonal broad learning system based on mutual information. In the proposed algorithm, the mapping weights are chosen from a orthogonal base set to insure that the extracted feature is independent to each other. In addition, each node is valued based on the mutual information between the feature and the output, and a feature selection algorithm is given. We conducted extensive experiments on challenging datasets of classification and regression to demonstrate the effectiveness of the proposed algorithm.

8:30AM *A Low-Memory Learning Formulation for a Kernel-and-Range Network [#19479]*
Huiping Zhuang, Zhiping Lin and Kar-Ann Toh, Nanyang Technological University, Singapore; Yonsei University, Korea (South)

Recently, a learning method based on the kernel and the range space projections has been introduced. This method has been applied to learn the multilayer network analytically with interpretable relationships among the weight matrices. However, the learning method carries a high-memory demand during training. In this study, a low-memory formulation is proposed to address this issue of high-memory demand. The developed method is inspired by a recursive implementation of the Moore-Penrose inverse and is shown to be mathematically equivalent to the original batch learning. Next, we further improved our proposed low-memory formulation to annul the potential divergence caused by rounding errors. The regression and classification behaviors of the proposed learning method are demonstrated using both synthetic and benchmark datasets. Our experiments confirm that the proposed formulation consumes significantly lower memory.

8:50AM *Entropy-Constrained Training of Deep Neural Networks [#19375]*
Simon Wiedemann, Arturo Marban, Klaus-Robert Mueller and Wojciech Samek, Fraunhofer Heinrich Hertz Institute, Germany; Technical University of Berlin, Germany

Motivated by the Minimum Description Length (MDL) principle, we first derive an expression for the entropy of a neural network which measures its complexity explicitly in terms of its bit-size. Then, we formalize the problem of neural network compression as an entropy-constrained optimization objective. This objective generalizes many of the currently proposed compression techniques in the literature, in that pruning or reducing the cardinality of the weight elements can be seen as special cases of entropy reduction methods. Furthermore, we derive a continuous relaxation of the objective, which allows us to minimize it using gradient-based optimization techniques. Finally, we show that we can reach compression results, which are competitive with those obtained using state-of-the-art techniques, on different network architectures and data sets, e.g. achieving x71 compression gains on a VGG-like architecture.

9:10AM *Sparse Binary Compression: Towards Distributed Deep Learning with minimal Communication [#19378]*
Felix Sattler, Simon Wiedemann, Klaus-Robert Mueller and Wojciech Samek, Fraunhofer Heinrich Hertz Institute, Germany; Technical University of Berlin, Germany

Currently, progressively larger deep neural networks are trained on ever growing data corpora. In result, distributed training schemes are becoming increasingly relevant. A major issue in distributed training is the limited communication bandwidth between contributing nodes or prohibitive communication cost in general. To mitigate this problem we propose Sparse Binary Compression (SBC), a compression framework that allows for a drastic reduction of communication cost for distributed training. SBC combines existing techniques of communication delay and gradient sparsification with a novel binarization method and optimal weight update encoding to push compression gains to new limits. By doing so, our method also allows us to smoothly trade- off gradient sparsity and temporal sparsity to adapt to the requirements of the learning task. Our experiments show, that SBC can reduce the upstream communication on a variety of convolutional and recurrent neural network architectures by more than four orders of magnitude without significantly harming the convergence speed in terms of forward-backward passes. For instance, we can train ResNet50 on ImageNet in the same number of iterations to the baseline accuracy, using x3531 less bits or train it to a 1% lower accuracy using x37208 less bits. In the latter case, the total upstream communication required is cut from 125 terabytes to 3.35 gigabytes for every participating client.

Competition: Challenge UP: Multimodal Fall Detection
Monday, July 15, 8:10AM-9:30AM, Room: Panorama V, Chair: Lourdes Mart{\'i}nez-Villase{\=n}or, Le{\'o}n Palafox, Karina P{\'e}rez Hiram Ponce

Monday, July 15, 9:30AM-10:00AM

Special Lecture: Coffee Break
Monday, July 15, 9:30AM-10:00AM, Room: Pre-function area Intercontinental

Monday, July 15, 10:00AM-11:00AM

Plenary Talk: Isabelle Guyon, IRI France
Monday, July 15, 10:00AM-11:00AM, Room: Ballroom I + II +II, Chair: Hava Siegelmann,

Monday, July 15, 11:00AM-12:00PM

Plenary Talk: Ichiro Tsuda, Chubu University
Monday, July 15, 11:00AM-12:00PM, Room: Ballroom I + II +II, Chair: George Kampis,

Monday, July 15, 12:00PM-1:30PM

Special Lecture: Lunch Break
Monday, July 15, 12:00PM-1:30PM, Room: Various locations in the area

Monday, July 15, 1:30PM-3:30PM

1l: Deep neural networks, Cellular Computational Networks
Monday, July 15, 1:30PM-3:30PM, Room: Ballroom I, Chair: Changsheng Lu

1:30PM *Depth-Controllable Very Deep Super-Resolution Network [#19412]*
Dohyun Kim, Joongheon Kim, Junseok Kwon and Tae-Hyung Kim, Chung-Ang University, Korea (South); KT AI Tech Center, Korea (South)

Deep learning techniques not only have surpassed humans in several computer vision tasks, but also have achieved state-of-the-art performance on the super-resolution (SR) task, which enhances the resolution of reconstructed images from the observed low-resolution (LR) images. Very Deep Super-Resolution (VDSR) is a popular architecture with decent performance on the SR task. In general, the deeper the VDSR network, the better the performance. Say, on a computing device with limited computational resources, the SR task must support multiple resolutions or multiple output rates. For this multi-rate SR task, the state-of-the-art architectures require multiple networks with varying depths. However as the number of supported rates increases, so does the number of networks imposing incremental computational burden on the computing device. We propose a depth-controllable network and training principles for the multi-rate SR task. The proposed network is configured as a single network regardless of the number of supported rates. The inverse auxiliary loss and contiguous/progressive skip connections are presented to train the network end-to-end throughout the varying number of layers without biasing the performances of specific depths. With three data sets and three scaling factors, the proposed network is compared to the baseline network VDSR, along with the Super-Resolution Convolutional Neural Network (SRCNN).

Our network not only requires a single network at varying rates, but also performs as well as the baseline networks in terms of Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM). Our model outperforms the baseline networks when the depth of layers is more than eleven.

1:50PM *Sequencing the musical sections with deep learning [#19078]*

Xuange Cui, Mingxue Liao, Pin Lv and Changwen Zheng, Institute of Software, Chinese Academy of Sciences, China

Deep learning has become increasingly popular for sequence modeling in various domains. In this work, we address the musical order verification as a sequential pattern learning task.We present a more advanced self-supervised learning model, dubbed as the triple-wise similarity with recurrent neural networks (TSN-R) which is on the basis of state-of-the-art Similarity Embedding Network. We take triple-wise musical sections as the input instance, and leverage the temporal coherence as a supervisory signal. As a first step, we use the triplet Siamese network as the input layer, and assess the similarity of the triplet feature maps. After that, we choose the bidirectional LSTM to extract features along the time dimension. Finally, the subsequent fully-connected layers are used for order verification. The experiments show that our approach outperforms other competing models in terms of prediction accuracy. Additionally, this paper reveals the reasons for the good performance of our model through the 2d-visualization of features. All the source code, pre-trained models and the experiment results are available in our project page. https://github.com/ISCASTEAM/ Sequencing-the-musical-sections

2:10PM *Deeper Capsule Network for Complex Data [#19261]*

Yi Xiong, Guiping Su, Shiwei Ye, Yuan Sun and Yi Sun, University of Chinese Academy of Sciences, China; National Institute of Informatics, Japan

Capsule Network (CapsNet) addresses the problem of Convolutional Neural Network (CNN) by introducing dynamic routing between capsules. Our work further develops CapsNet in depth and performance. By introducing Convolutional Capsule Layer (Conv-Caps Layer), we deepen CapsNet, which greatly improves the performance. We also propose Capsule Pool (Caps-Pool), a new pooling operation, to reduce the number of parameters. This pooling operation preserves the full representation of features. In our experiment, our DeeperCaps model received so far the strongest result of CapsNet on the bigger dataset (Cifar-10). Our Caps-Pool reduces half of the parameters between layers while maintaining the performance. It also influences the output distribution to be intuitively more reasonable.

2:30PM *PointDoN: A Shape Pattern Aggregation Module for Deep Learning on Point Cloud [#19106]*

Shuxin Zhao, Chaochen Gu, Changsheng Lu, Ye Huang, Kaijie Wu and Xinping Guan, Shanghai Jiao Tong University, China

As point cloud is a typical and significant type of geometric 3D data, deep learning on the classification and segmentation of point cloud has received widely interests recently. However, the critical problems to process the irregularity of point cloud and feature extraction of shape pattern have not yet been fully explored. In this paper, a geometric deep learning architecture based on our PointDoN module is presented. Inspired by the Difference of Normals (DoN) in traditional point clouds processing, our PointDoN module is a feature aggregation module combining DoN shape pattern descriptor with both 3D coordinates and extra features (such as RGB colors). Our PointDoN-based architecture can be flexibly applied to multiple point cloud processing tasks such as 3D shape classification and scene semantic segmentation. Experiments demonstrate that PointDoN model achieves state-of-the-art results on multiple types of challenging benchmark datasets.

2:50PM *Learning Adaptive Weight Masking for Adversarial Examples [#19433]*

Yoshimasa Kubo, Michael Traynor, Thomas Trappenberg and Sageev Oore, Dalhousie University, Canada; Dalhousie University and Vector Institute for Artificial Intelligence, Canada

Adding small, well crafted perturbations to the pixel values of input images leads to adversarial examples, so called because these perturbed images can drastically affect the accuracy of machine learning classifiers. Defenses against such attacks are being studied, often with varying results. In this study, we introduce a model called the Stochastic-Gated Partially Binarized Network (SGBN), that incorporates binarization andinput-dependent stochasticity. In particular, a gate module learns the probability that individual weights in corresponding convolutional filters should be masked (turned on or off). The gate module itself consists of a shallow convolutional neural network, and its sigmoid outputs are stochastically binarized and pointwise multiplied with corresponding filters in the convolutional layer of the main network. We test and compare our model with several related approaches, and to try to gain an understanding of our model, we visualize activations of some of the gating network outputs and their corresponding filters.

3:10PM *Structured Pruning for Efficient ConvNets via Incremental Regularization [#20431]*

Huan Wang, Qiming Zhang, Yuehai Wang, Lu Yu and Haoji Hu, Zhejiang University, China; University of Sydney, Australia

Parameter pruning is a promising approach for CNN compression and acceleration by eliminating redundant model parameters with tolerable performance degrade. Despite its effectiveness, existing regularization-based parameter pruning methods usually drive weights towards zero with large and constant regularization factors, which neglects the fragility of the expressiveness of CNNs, and thus calls for a more gentle regularization scheme so that the networks can adapt during pruning. To achieve this, we propose a new and novel regularization-based pruning method, named IncReg, to incrementally assign different regularization factors to different weights based on their relative importance. Empirical analysis on CIFAR-10 dataset verifies the merits of IncReg. Further extensive experiments with popular CNNs on CIFAR-10 and ImageNet datasets show that IncReg achieves comparable to even better results compared with state-of-the-arts. Our source codes and trained models are available here: https://github.com/mingsun-tse/caffe_increg.

2e: Deep learning

Monday, July 15, 1:30PM-3:30PM, Room: Ballroom II, Chair: Hojung Lee

1:30PM *Local Critic Training of Deep Neural Networks [#19646]*

Hojung Lee and Jong-Seok Lee, Yonsei University, Korea (South)

This paper proposes a novel approach to train deep neural networks by unlocking the layer-wise dependency of backpropagation training. The approach employs additional modules called local critic networks besides the main network model to be trained, which are used to obtain error gradients without complete feedforward and backward propagation processes. We propose a cascaded learning strategy for these local networks. In addition, the approach is also useful from multi-model perspectives, including structural optimization of neural networks, computationally efficient progressive inference, and ensemble classification for performance improvement. Experimental results show the effectiveness of the proposed approach and suggest guidelines for determining appropriate algorithm parameters.

1:50PM *Stable Network Morphism [#19274]*
Tao Wei, Changhu Wang and Chang Wen Chen, State University of New York at Buffalo, United States; ByteDance AI Lab, China; The Chinese University of Hong Kong, Shenzhen, China

Deep neural networks perform better when they are deeper. Network morphism is one of the paradigms to construct deeper neural networks. It makes developing deeper neural networks building on existing ones possible by morphing a well-trained neural network into a new one with the network function completely preserved. The morphed network also has the potential to continue growing into a more powerful one as it has more parameters. Existing network morphism schemes include Net2Net and NetMorph. However, both of them suffer from significant initial performance drop when the morphed network is continually trained. Such unstability is very much undesired for a continual learning system. In this research, we first identify the reason for the unstability, which is due to the large amount of zeros padded into the parameters. Based on this observation, we propose an algorithm based on modified gradient descent to decompose the network morphism equation. As a result, the morphed parameters are all non-zeros and the continual training process become stable. Experimental results on benchmark datasets demonstrate the effectiveness of the proposed stable network morphism scheme.

2:10PM *Cross-Domain Car Detection Using Unsupervised Image-to-Image Translation: From Day to Night [#19615]*
Vinicius F. Arruda, Thiago M. Paixao, Rodrigo F. Berriel, Alberto F. De Souza, Claudine Badue, Nicu Sebe and Thiago Oliveira-Santos, Universidade Federal do Espirito Santo, Brazil; Instituto Federal do Espirito Santo, Brazil; University of Trento, Italy

Deep learning techniques have enabled the emergence of state-of-the-art models to address object detection tasks. However, these techniques are data-driven, delegating the accuracy to the training dataset which must resemble the images in the target task. The acquisition of a dataset involves annotating images, an arduous and expensive process, generally requiring time and manual effort. Thus, a challenging scenario arises when the target domain of application has no annotated dataset available, making tasks in such situation to lean on a training dataset of a different domain. Sharing this issue, object detection is a vital task for autonomous vehicles where the large amount of driving scenarios yields several domains of application requiring annotated data for the training process. In this work, a method for training a car detection system with annotated data from a source domain (day images) without requiring the image annotations of the target domain (night images) is presented. For that, a model based on Generative Adversarial Networks (GANs) is explored to enable the generation of an artificial dataset with its respective annotations. The artificial dataset (fake dataset) is created translating images from day- time domain to night-time domain. The fake dataset, which comprises annotated images of only the target domain (night images), is then used to train the car detector model. Experimental results showed that the proposed method achieved significant and consistent improvements, including the increasing by more than 10% of the detection performance when compared to the training with only the available annotated data (i.e., day images).

2:30PM *Reference-oriented Loss for Person Re-identification [#19653]*
Mingyang Yu, Zhigang Chang, Qin Zhou, Shibao Zheng and Tai Pang Wu, Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, China; Artificial Intelligence Center-City Brain, Alibaba Cloud, China; 1000 Video Technology Co. Limited, Suzhou, China

Deep metric learning methods are quite effective in exploring discriminative feature embeddings, among which triplet loss and its variants are widely utilized. However, in existing methods, the tightness information for intra-class samples is ignored, leading to large intra-class divergence and severe inter- class overlapping problem. To address this issue, a novel loss function called reference-oriented triplet loss is proposed in this paper. The proposed method introduces several reference images to guide training. More specifically, distances between the reference image and images of the same identity are required to be as similar as possible. By introducing reference images, images from the same class become much closer with each other and the inter-class overlapping problem is alleviated. Comparing to baseline batch hard triplet loss, the mAP accuracy increases by 3.75\%/5.69\% on person re-ID datasets Market1501 and DukeMTMC-Reid. Comparison results with state-of-the-art algorithms also demonstrate effectiveness of the proposed algorithm.

2:50PM *Double Transfer Learning for Breast Cancer Histopathologic Image Classification [#19840]*
Jonathan de Matos, Alceu de S. Britto Jr, Luiz S. Oliveira and Alessandro L. Koerich, Ecole de Technologie Superieure, Canada; Pontifical Catholic University of Parana, Brazil; Federal University of Parana, Brazil

This work proposes a classification approach for breast cancer histopathologic images (HI) that uses transfer learning to extract features from HI using an Inception-v3 CNN pre-trained with ImageNet dataset. We also use transfer learning on training a support vector machine (SVM) classifier on a tissue labeled colorectal cancer dataset aiming to filter the patches from a breast cancer HI and remove the irrelevant ones. We show that removing irrelevant patches before training a second SVM classifier, improves the accuracy for classifying malign and benign tumors on breast cancer images. We are able to improve the classification accuracy in 3.7% using the feature extraction transfer learning and an additional 0.7% using the irrelevant patch elimination. The proposed approach outperforms the state-of-art in three out of the four magnification factors of the breast cancer dataset.

3:10PM *Multiple Fake Classes GAN for Data Augmentation in Face Image Dataset [#20152]*
Adamu Ali-Gombe, Elyan Eyad and Jayne Chrisina, Robert Gordon University, United Kingdom; Oxford Brookes University, United Kingdom

Class-imbalanced datasets often contain one or more class that are under-represented in a dataset. In such a situation, learning algorithms are often biased toward the majority class instances. Therefore, some modification to the learning algorithm or the data itself is required before attempting a classification task. Data augmentation is one common approach used to improve the presence of the minority class instances and rebalance the dataset. However, simple augmentation techniques such as applying some affine transformation to the data, may not be sufficient in extreme cases, and often do not capture the variance present in the dataset. In this paper, we propose a new approach to generate more samples from minority class instances based on Generative Adversarial Neural Networks (GAN). We introduce a new Multiple Fake Class Generative Adversarial Networks (MFC-GAN) and generate additional samples to rebalance the dataset. We show that by introducing multiple fake class and oversampling, the model can generate the required minority samples. We evaluate our model on face generation task from attributes using a reduced number of samples in the minority class. Results obtained showed that MFC-GAN produces plausible minority samples that improve the classification performance compared with state-of-the-art AC-GAN generated samples.

8a: Applications of deep networks

Monday, July 15, 1:30PM-3:30PM, Room: Ballroom III, Chair: Wang Chen

1:30PM *Dog Identification using Soft Biometrics and Neural Networks [#19996]*

Kenneth Lai, Xinyuan Tu and Svetlana Yanushkevich, University of Calgary, Canada; Beijing Institute of Technology, China

This paper addresses the problem of biometric identification of animals, specifically dogs. We apply advanced machine learning models such as deep neural network on the photographs of pets in order to determine the pet identity. In this paper, we explore the possibility of using different types of "soft" biometrics, such as breed, height, or gender, in fusion with "hard" biometrics such as photographs of the pet's face. We apply the principle of transfer learning on different Convolutional Neural Networks, in order to create a network designed specifically for breed classification. The proposed network is able to achieve an accuracy of 90.80% and 91.29% when differentiating between the two dog breeds, for two different datasets. Without the use of "soft" biometrics, the identification rate of dogs is 78.09% but by using a decision network to incorporate "soft" biometrics, the identification rate can achieve an accuracy of 84.94%.

1:50PM *Adversarial Collaborative Auto-encoder for Top-N Recommendation [#19693]*

Feng Yuan, Lina Yao and Boualem Benatallah, University of New South Wales, Australia

Recently, deep learning-based recommendation models have been proved to have state-of-the-art recommendation accuracy. However, most of the existing work assume that user feedbacks are noise-free, on which the neural networks (NN) are trained. Although some methods apply man-made noises on the input data to train the networks more effectively (e.g. the collaborative denoising auto-encoder), the noises are randomly generated. To gain further improvements, we focus on boosting the overall recommendation performance through adversarial noises. We propose a general framework to adversarially train a NN-based item recommendation model. In particular, we select the collaborative auto-encoder model as an example and test our method on three public datasets. We show that our approach enhances both overall robustness and performance which outperforms competitive state-of-the-art item recommendation models.

2:10PM *Improving Route Choice Models by Incorporating Contextual Factors via Knowledge Distillation [#20456]*

Qun Liu, Supratik Mukhopadhyay, Ravindra Gudishala, Yimin Zhu, Sanaz Saeidi and Alimire Nabijiang, Louisiana State University, United States

Route Choice Models predict the route choices of travelers traversing an urban area. Most of the route choice models link route characteristics of alternative routes to those chosen by the drivers. The models play an important role in prediction of traffic levels on different routes and thus assist in development of efficient traffic management strategies that result in minimizing traffic delay and maximizing effective utilization of transport system. High fidelity route choice models are required to predict traffic levels with higher accuracy. Existing route choice models do not take into account dynamic contextual conditions such as the occurrence of an accident, the socio-cultural and economic background of drivers, other human behaviors, the dynamic personal risk level, etc. As a result, they can only make predictions at an aggregate level and for a fixed set of contextual factors. For higher fidelity, it is highly desirable to use a model that captures significance of subjective or contextual factors in route choice. This paper presents a novel approach for developing high- fidelity route choice models with increased predictive power by augmenting existing aggregate level baseline models with information on drivers' responses to contextual factors obtained from Stated Choice Experiments carried out in an Immersive Virtual Environment through the use of knowledge distillation.

2:30PM *Abstractive Summarization with Keyword and Generated Word Attention [#19057]*

Qianlong Wang and Jiangtao Ren, Sun Yat-sen University, China

The abstractive summarization is a important task in natural language processing field. In previous work, the sequence-to-sequence based models are widely used for abstractive summarization task. However, most of the current abstractive summarization models still suffer from two problems. One is that it is difficult for these models to learn an accurate source contextual representation from the redundancy and noisy source text at each decoding step. Another is the information loss problem, which is ignored in previous work. The inability of these models to effectively exploit previously generated words led to this problem. In order to address these two problems, in this paper, we propose a novel keyword and generated word attention model. Specifically, the proposed model first employs the hidden state of decoder to capture relevant keywords and previously generated words contextual at each time step. The model then utilizes obtained keywords and generated words contextual to create keywords-aware and generated words-aware source contextual, respectively. The keywords contextual contributes to learn an accurate source contextual representation, and the generated words contextual can alleviate the information loss problem. Experimental results on a popular Chinese social media dataset demonstrate that the proposed model outperforms baselines and achieves the state-of-the-art performance.

2:50PM *Utilizing Generative Adversarial Networks for Recommendation based on Ratings and Reviews [#19676]*

Wang Chen, Hai-Tao Zheng, Yang Wang, Wei Wang and Rui Zhang, Tsinghua-Southampton Web Science Laboratory Graduate School at Shenzhen, Tsinghua University, China; University of Melbourne, Australia

Many existing rating-based recommendation algorithms have achieved relative success. However, the real-world datasets are extremely sparse and most rating-based algorithms are still suffering from the data sparsity problem. Along with integer-valued ratings, we consider that the user-generated review is also an important user feedback. Furthermore, compared with the traditional recommendation algorithms which have the limited ability to learn the distributions of ratings and reviews simultaneously, the generative adversarial networks can learn better representations for data. In this paper, we propose Rating and Review Generative Adversarial Networks (RRGAN), an innovative framework for recommendation, in which the generative model and discriminative model play a minimax game. Specifically, the generative model predicts the ratings of top-N list for users or items based on reviews, while the discriminative model aims to distinguish the predicted ratings from real ratings. With the competition between these two models, RRGAN improves the ability of understanding users and items based on ratings and reviews. We introduce the user profiles, item representations and ratings into a matrix factorization model to predict the top-N list for the users. In addition, we study three different architectures to learn reasonable user profiles and item representations based on ratings and reviews to achieve better recommendations. To evaluate the performance of our model, we conduct the extensive experiments on three real-world amazon datasets in three parts, which are top-N recommendation analysis, case study and long-tail users analysis. The experimental results show that our method outperforms various state-of-the-art methods.

3:10PM *Gated Neural Network with Regularized Loss for Multi-label Text Classification [#19665]*

Yunlai Xu, Xiangying Ran, Wei Sun, Xiangyang Luo and Chongjun Wang, Nanjing University, China

Multi-label text classification is generally more difficult for its exponential output label space and more flexible input document considering its

corresponding number of labels. We observed that some long documents have only one or two labels while some short documents are related to much more labels. In this paper, we propose a dynamic Gated Neural Network architecture to simultaneously process the document through two parts, one to extract the most informative semantics and filter redundant information, the other to capture context semantics with most of the information in the document kept. And semantics from these two parts are dynamically controlled by a gate to perform subsequent classification. And to better the training we incorporate label dependencies into traditional binary cross-entropy loss by exploiting label co-occurrences. Experimental results on AAPD and RCV1-V2 datasets show that our proposed methods achieve state-of-art performance. Further analysis of experimental results demonstrate that the proposed methods not only capture enough feature information from both long and short documents assigned with various labels, but also exploit label dependency to regularize the proposed model to further improve its performance.

1b: Recurrent neural networks

Monday, July 15, 1:30PM-3:30PM, Room: Duna Salon I, Chair: Jinlei Xu

1:30PM *Context Gating with Short Temporal Information for Video Captioning [#19970]*

Jinlei Xu, Ting Xu, Xin Tian, Chunping Liu and Yi Ji, Soochow University, China

Video Captioning is a newly emerging task which automatically translates content in a video into a textual description. Similar to image captioning, most existing methods simply utilized extracted visual features to generate sentences. However, in video captioning temporal information is much more important for description. Though the short temporal information (STI) is always ignored. Meanwhile, the context of generated sentence seems not been mined enough. In this paper, we build a context gating mechanism with STI based on encoder-decoder (CG-ED) neural framework for video captioning. In our approach, based on the 2D feature space, we cut and recombine the whole 3D features to extract STI by temporal distribution. To balance the contributions of different context of sentences, context gating is designed . Our proposed model is evaluated on two large-scale datasets: Microsoft Research-Video to Text (MSR-VTT) and Microsoft Research Video Description Corpus(MSVD). Experimental results demonstrate that its precision of caption is higher than most of the state-of-the-art approaches.

1:50PM *Deep learning long-range information in undirected graphs with wave networks [#20288]*

Matthew Matlock, Arghya Datta, Na Le Dang, Kevin Jiang and S Joshua Swamidass, Washington University in Saint Louis, United States

Graph algorithms are key tools in many fields of science and technology. Some of these algorithms depend on propagating information between distant nodes in a graph. Recently, there have been a number of deep learning architectures proposed to learn on undirected graphs. However, most of these architectures aggregate information in the local neighborhood of a node, and therefore they may not be capable of efficiently propagating long-range information. To solve this problem we examine a recently proposed architecture, wave, which propagates information back and forth across an undirected graph in waves of nonlinear computation. We compare wave to graph convolution, an architecture based on local aggregation, and find that wave learns three different graph-based tasks with greater efficiency and accuracy. These three tasks include (1) labeling a path connecting two nodes in a graph, (2) solving a maze presented as an image, and (3) computing voltages in a circuit. These tasks range from trivial to very difficult, but wave can extrapolate from small training examples to much larger testing examples. These results show that wave may be able to efficiently solve a wide range of tasks that require long-range information propagation across undirected graphs. An implementation of the wave network, and example code for the maze task are included in the tflon deep learning toolkit (https://bitbucket.org/mkmatlock/tflon).

2:10PM *A Memory-Based STDP Rule for Stable Attractor Dynamics in Boolean Recurrent Neural Networks [#20311]*

Jeremie Cabessa and Alessandro Villa, University Paris 2, France; University of Lausanne, Switzerland

We consider a simplified Boolean model of the basal ganglia-thalamocortical network, and study the effect of a spike-timing-dependent plasticity (STDP) rule on the stabilization of its attractor dynamics. More precisely, we introduce an adaptive STDP rule which constantly updates its learning rate based on the attractors that the network encounters during a window of past time steps. This so-called network memory is assumed to be dynamic: its duration is step-wise increased every time a trigger input pattern is detected, and is decreased otherwise. In this context, we show that well-adjusted trigger inputs can fine tune the network memory and its associated STDP rule in such a way to drive the network into stable and rich attractor dynamics. We discuss how this feature might be related to reward learning processes in the neurobiological context.

2:30PM *Personalizing Session-based Recommendation with Dual Attentive Neural Network [#19949]*

Tianan Liang, Yuhua Li, Ruixuan Li, Xiwu Gu, Olivier Habimana and Yi Hu, Huazhong University of Science and Technology, China; Huazhong University of Science and Technology, Rwanda

Session-based recommendation which aims to recommend the next item in an anonymous session for users, becomes one of the most popular tasks in recommendation area. Traditional methods such as matrix factorization and item-to-item perform very poorly because they only take into account the last click of the session and ignore the information of the whole click sequence. On the other hand, Recurrent Neural Network (RNN) based methods have performed excellently in session-based recommendation. However, they only consider the user's sequential behavior in the current session or only use cross-session information to track user's interests over time, whereas the user preference is not emphasized in cross-session. Therefore, in this work, we design a novel neural network framework for personalized session-based recommendation, named Dual Attentive Neural Network (DANN). DANN considers user's main purpose of current session and user's personalized preference of cross-session. Specifically, in DANN we exploit a user-level attention mechanism to model user's personalized preference and capture user's main purpose in the current session via a session-level attention mechanism. The experimental results on two real-world datasets show that our DANN model outperforms other baseline models. Furthermore, we find that DANN achieves improvement when modeling user personalized preferences, which shows the advantage of modeling user's preference and user's purpose simultaneously.

2:50PM *Automatic Source Code Summarization with Extended Tree-LSTM [#19288]*

Yusuke Shido, Yasuaki Kobayashi, Akihiro Yamamoto, Atsushi Miyamoto and Tadayuki Matsumura, Graduate School of Informatics, Kyoto University, Japan; Center for Exploratory Research, Hitachi, Ltd., Japan

Neural machine translation models are used to automatically generate a document from given source code since this can be regarded as a machine translation task. Source code summarization is one of the components for automatic document generation, which generates a summary in natural language from given source code. This suggests that techniques used in neural machine translation, such as Long Short-Term Memory (LSTM), can be used for source code summarization. However, there is a considerable difference between source code and natural language: Source code is essentially structured, having loops and conditional branching, etc. Therefore, there is some obstacle to apply known machine translation

models to source code. Abstract syntax trees (ASTs) capture these structural properties and play an important role in recent machine learning studies on source code. Tree-LSTM is proposed as a generalization of LSTMs for tree-structured data. However, there is a critical issue when applying it to ASTs: It cannot handle a tree that contains nodes having an arbitrary number of children and their order simultaneously, which ASTs generally have such nodes. To address this issue, we propose an extension of Tree-LSTM, which we call Multi-way Tree-LSTM and apply it for source code summarization. As a result of computational experiments, our proposal achieved better results when compared with several state-of-the-art techniques.

3:10PM *Programming Style Analysis with Recurrent Neural Network to Automatic Pull Request Approval [#20375]*

Lucas Roque, Altino Dantas and Celso G. Camilo-Junior, Universidade Federal de Goias, Brazil

Although recognized as important, programming style is one aspect commonly neglected by developers. However, follow the pattern of programming presents in a project may be useful to understand and maintain the system. Usually, the companies build their own guidelines for coding bug fixes or new features. Nonetheless, developing this set of rules is not a simple task, and there are even inconsistencies in the specialized literature. Therefore, this paper proposes a new approach to programming style analysis, using a recurrent neural network (RNN) that learns the programming style presents in a project and determines whether a piece of code submitted to it, follows the project's pattern. A study on three real projects was conducted and demonstrated the promising of the approach, by revealing the RNN capability to recognize the programming style pattern of each project.

2a: Supervised learning

Monday, July 15, 1:30PM-3:30PM, Room: Duna Salon II, Chair: Teresa Ludermir

1:30PM *Analyzing the impact of data representations in classification problems using clustering [#20364]*

Felipe Farias, Teresa Ludermir, Carmelo Bastos-Filho and Flavio Oliveira, Universidade Federal de Pernambuco, Brazil; UNIVERSIDADE FEDERAL DE PERNAMBUCO, Brazil; Universidade de Pernambuco, Brazil; Instituto Federal de Educacao, Ciencia e Tecnologia de Pernambuco, Brazil

This work presents an investigation about how to better represent output data labels to be used in supervised training of classifiers. The posed hypothesis is that grouping cohesive patterns into clusters and assigning them sub-labels, may improve the classifier performance. We used 12 benchmark datasets to test our hypothesis. First, we create the clusters, and when appropriate, new sub-labels were generated, according to Fuzzy-CMeans and Silhouette score thresholds. After that, Multilayer Perceptrons were employed to model each dataset with cluster generated sub-labels, obtaining promising results. From results, we observed that in cases where the sub-labels were used, the accuracy increased with statistical significance with p=0.05 in 22 cases and remained statistically equivalent in 14 cases, presenting no decrease in accuracy.

1:50PM *k-Entropy Based Restricted Boltzmann Machines [#19063]*

Leandro Aparecido Passos, Marcos Cleison Santana, Thierry Moreira and Joao Paulo Papa, Federal University of Sao Carlos - UFSCar, Brazil; Sao Paulo State University - UNESP, Brazil

Restricted Boltzmann Machines achieved notorious popularity in the scientific community in the last decade due to outstanding results in a wide range of applications and also for providing the required mechanisms to build successful deep learning models, i.e., Deep Belief Networks and Deep Boltzmann Machines. However, their main bottleneck is related to the learning step, which is usually time-consuming. In this paper, we introduce a Sigmoid-like family of functions based on the Kaniadakis entropy formulation in the context of the RBM learning procedure. Experiments concerning binary image reconstruction are conducted in four public datasets to evaluate the robustness of the proposed approach. The results suggest that such a family of functions is suitable to increase the convergence rate when compared to standard functions employed by the research community.

2:10PM *Active Learning with Interpretable Predictor [#19162]*

Yusuke Taguchi, Keisuke Kameyama and Hideitsu Hino, University of Tsukuba, Japan; The Institute of Statistical Mathematics/RIKEN AIP, Japan

Active learning is a method of constructing a useful prediction model with the minimum number of annotations or labeling for a response variable. It is widely used as a modern experimental design method, particularly for problems with high annotation cost. An appropriate reason for the selection of the next experimental setting is required for experiments with high annotation cost, for example, situations requiring large-scale experiments or long-term experiments such as agricultural examinations. In conventional active learning, it is only known that the samples that can improve the prediction accuracy of a prediction model are selected. This is not a satisfactory explanation for the selection of the next experimental setting for approving an experiment. In this paper, we propose a novel active learning algorithm with the following two models: a model to predict a response variable and a model to predict the amount of decrease in test loss. A new sample is selected using a model that predicts the amount of decrease in test loss. It is possible to provide a reason for sample selection by employing a model that can evaluate variable importance, e.g., using a random forest as a model of predicting the decrease in test loss. We applied the proposed method to multiple datasets and showed that the prediction performance of the proposed method is comparable to those of existing methods and the computational time is superior to those of existing methods. In addition, we demonstrated that it is possible to provide suitable reasons for selecting a sample in the process of active learning.

2:30PM *Exploring Machine Learning and Deep Learning Frameworks for Task-Oriented Dialogue Act Classification [#20037]*

Tulika Saha, Saurabh Srivastava, Mauajama Firdaus, Sriparna Saha, Asif Ekbal and Pushpak Bhattacharyya, IIT Patna, India

Dialogue Act (DA) Classification plays a significant role in the understanding of an utterance in a dialogue. Components of Spoken Dialogue System (SDS) such as Natural Language Understanding (NLU) and Dialogue Management (DM) modules can significantly exploit the output of the DA classification. In this paper, we propose a task-oriented DA classifier based on both traditional supervised Machine Learning (ML) as well as Deep Learning (DL) techniques. The type and nature of dialogues basically depend on the domain and the DA itself. So, in order to make the model task-oriented, a new tag-set has been designed by studying the properties of the target domain. On the benchmark SwitchBoard (SWBD) and TRAINS corpus,

our proposed models have performed exceptionally well with the new tag-set. Experimental results indicate that our proposed models have achieved good accuracy on both the datasets and outperformed several state of the art approaches and the new tag-set is well suited for task-oriented applications.

2:50PM *Hierarchical Capsule Based Neural Network Architecture for Sequence Labeling [#20447]*

Saurabh Srivastava, Puneet Agarwal, Gautam Shroff and Lovekesh Vig, TCS Research, India

Sequence Labeling is one of the most prominent tasks in NLP. The traditional text classification models do not carry context from one sentence to another and hence may not perform well on these tasks. These models lack a hierarchical structure that can aid them in dissecting the input structure at different levels to allow flow of context between sentences. In this paper, we propose a hierarchical neural network comprising of BiLSTMs, Dilated Convolution operation, Capsules and Conditional Random Field (CRF) to understand the discourse/ abstract structure and predict next probable label by using label history. We have performed experiments on 3 publicly available datasets through which we have demonstrated that our model has achieved state-of-art performance on these datasets.

3:10PM *Guessing the Code: Learning Encoding Mappings Using the Back Propagation Algorithm [#20422]*

Amrutha Machireddy and Shayan Srinivasa Garani, Indian Institute of Science, India

Error correction codes such as low density parity check (LDPC) codes are popularly used to enhance the performance of digital communication systems. The current decoding framework relies on exchanging beliefs over a Tanner graph, which the encoder and decoder are aware of. However, this information may not be available readily, for example in covert communication. The main idea of this paper is to build a neural network to learn the encoder mappings in the absence of knowledge of the Tanner graph. We propose a scheme to learn the mappings using the back propagation algorithm. We investigate into the choice of different cost functions and the number of hidden neurons for learning the encoding function. The proposed scheme is capable of learning the parity check equations over a binary field towards identifying the validity of a codeword. Simulation results over synthetic data show that our algorithm is indeed capable of learning the encoder mappings and identifying the parity check equations.

2b: Unsupervised learning and clustering, (including PCA, and ICA)

Monday, July 15, 1:30PM-3:30PM, Room: Duna Salon III, Chair: Laura Muzzarelli

1:30PM *Multi-Hierarchy Attribute Relationship Mining Based Outlier Detection for Categorical Data [#19713]*

Xiaoyu Hu, Yijie Wang and Li Cheng, National University of Defense Technology, China

Outlier detection for categorical data is very important in many practical scenarios, such as intrusion detection, fraud detection, early detection of diseases, etc. However, there is no inherent difference measure for categorical data. The differences are hidden in complex attribute value relationships. Existing methods do not properly handle the internal relationship and external relationship of attributes, resulting in low accuracy of outlier detection. This paper proposes a novel unsupervised outlier detection method for categorical data based on Multi-Hierarchy Attribute Relationship Mining (MHARM). It detects outliers by mining the hierarchical and complex relationships between attribute values. MHARM first calculates the internal relationship. It processes each attribute independently via an information-theoretic difference to get an internal distance matrix. Then it handles different subhierarchy of external relationship. It divides attributes into two clusters, using mutual information as the correlation measure. For the external relationship of intra-cluster attributes, it iteratively updates an external distance matrix by using an entropy weighted Earth Mover's Distance (EMD) and the internal distance until convergence; for the external relationship of inter-cluster attributes, the joint entropy weighted sum is obtained to be the whole difference between objects. Finally, MHARM uses the sum of whole difference between objects as the outlier score, sorting it for outlier detection. Experiment results show that MHARM has an average auc value of 13.84% higher than the state-of-the-art methods and significantly reduced the detection volume multiples (dvM) for unearthing 90% outliers on the given nine data sets.

1:50PM *Unsupervised Representation Adversarial Learning Network: from Reconstruction to Generation [#19365]*

Yuqian Zhou, Kuangxiao Gu and Thomas Huang, ECE Department of UIUC, United States

A good representation for arbitrarily complicated data should have the capability of semantic generation, clustering and reconstruction. Previous research has already achieved impressive performance on either one. This paper aims at learning a disentangled representation effective for all of them in an unsupervised way. To achieve all the three tasks together, we learn the forward and inverse mapping between data and representation on the basis of a symmetric adversarial process. In theory, we minimize the upper bound of the two conditional entropy loss between the latent variables and the observations together to achieve the cycle consistency. The newly proposed RepGAN is tested on MNIST, fashionMNIST, CelebA, and SVHN datasets to perform unsupervised classification, generation and reconstruction tasks. The result demonstrates that RepGAN is able to learn a useful and competitive representation. To the author's knowledge, our work is the first one to achieve both a high unsupervised classification accuracy and low reconstruction error on MNIST.

2:10PM *Matrix Product Operator Restricted Boltzmann Machines [#20160]*

Cong Chen, Kim Batselier, Ching-yun Ko and Ngai Wong, The University of Hong Kong, Hong Kong; Delft University of Technology, Netherlands

A restricted Boltzmann machine (RBM) learns a probability distribution over its input samples and has numerous uses like dimensionality reduction, classification and generative modeling. Conventional RBMs accept vectorized data that dismiss potentially important structural information in the original tensor (multi-way) input. Matrix- variate and tensor-variate RBMs, named MvRBM and TvRBM, have been proposed but are all restrictive by model construction and have weak model expression power. This work presents the matrix product operator RBM (MPORBM) that utilizes a tensor network generalization of Mv/TvRBM, preserves input formats in both the visible and hidden layers, and results in higher expressive power. A novel training algorithm integrating contrastive divergence and an alternating optimization procedure is also developed. Numerical experiments compare the MPORBM with the traditional RBM and MvRBM for data classification and image completion and denoising tasks. The expressive power of the MPORBM as a function of the MPO-rank is also investigated.

2:30PM *Rank Selection in Non-negative Matrix Factorization: systematic comparison and a new MAD metric [#19395]*

Laura Muzzarelli, Susanne Weis, Simon B. Eickhoff and Kaustubh R. Patil, Forschungszentrum Juelich and HHU Duesseldorf, Germany

Non-Negative Matrix Factorization (NMF) is a powerful dimensionality reduction and factorization method that provides a part-based representation of the data. In the absence of a priori knowledge about the latent dimensionality of the data, it is necessary to select a rank of the reduced representation. Several rank selection methods have been proposed, but no

consensus exists on when a method is suitable to use. In this work, we propose a new metric for rank selection based on imputation cross-validation, and we systematically compare it against six other metrics while assessing the effects of data properties. Using synthetic datasets with different properties, our work critically evidences that most methods fail to identify the true rank. We show that properties of the data heavily impact the ability of different methods. Imputation-based metrics, including our new MADimput, provided the best accuracy irrespective of the data type, but no solution worked perfectly in all circumstances. One should therefore carefully assess characteristics of their dataset in order to identify the most suitable metric for rank selection.

2:50PM *Qualitative data clustering: a new Integer Linear Programming model [#19227]*

Luiz Henrique Nogueira Lorena, Marcos Goncalves Quiles, Luiz Antonio Nogueira Lorena, Andre C. P. L. F. de Carvalho and Juliana Garcia Cespedes, Federal University of Sao Paulo, Brazil; National Institute for Space Research, Brazil; University of Sao Paulo, Brazil

Qualitative data clustering is a fundamental data analysis task, with applications in many areas, like medicine, sociology, and economics. An appealing way to deal with this task is via Integer Linear Programming, as it avoids inappropriate inferences by the final user. This approach has two main advantages: the data are directly used, without the need of being converted to quantitative values, and the optimal number of clusters is automatically obtained by solving the optimization problem. However, it might create large and redundant models, which can limit the size of the problems it can be applied. Recently, models that are more compact and able to avoid some redundancy have been proposed in the literature. These models consume less memory and are faster to obtain the optimal solution set. In this study, a new model is introduced and compared with the state-of-the-art alternatives using datasets from different application domains. Empirical results show that the new model outperforms its predecessors, achieving the optimal solution set with lower computational time and memory consumption.

3:10PM *Attention-Guided Generative Adversarial Networks for Unsupervised Image-to-Image Translation [#19906]*

Hao Tang, Dan Xu, Nicu Sebe and Yan Yan, University of Trento, Italy; University of Oxford, England; Texas State University, United States

The state-of-the-art approaches in Generative Adversarial Networks (GANs) are able to learn a mapping function from one image domain to another with unpaired image data. However, these methods often produce artifacts and can only be able to convert low-level information, but fail to transfer high-level semantic part of images. The reason is mainly that generators do not have the ability to detect the most discriminative semantic part of images, which thus makes the generated images with low-quality. To handle the limitation, in this paper we propose a novel Attention-Guided Generative Adversarial Network (AGGAN), which can detect the most discriminative semantic object and minimize changes of unwanted part for semantic manipulation problems without using extra data and models. The attention-guided generators in AGGAN are able to produce attention masks via a built- in attention mechanism, and then fuse the input image with the attention mask to obtain a target image with high-quality. Moreover, we propose a novel attention-guided discriminator which only considers attended regions. The proposed AGGAN is trained by an end-to-end fashion with an adversarial loss, cycle-consistency loss, pixel loss and attention loss. Both qualitative and quantitative results demonstrate that our approach is effective to generate sharper and more accurate images than existing models.

1b: Recurrent neural networks

Monday, July 15, 1:30PM-3:30PM, Room: Panorama I, Chair: Tayfun Alpay

1:30PM *Question Answering with Hierarchical Attention Networks [#20465]*

Tayfun Alpay, Stefan Heinrich, Michael Nelskamp and Stefan Wermter, University of Hamburg, Germany

We investigate hierarchical attention networks for the task of question answering. For this purpose, we propose two different approaches: in the first, a document vector representation is built hierarchically from word-to-sentence level which is then used to infer the right answer. In the second, pointer sum attention is utilized to directly infer an answer from the attention values of the word and sentence representations. We evaluate our approach on the Children's Book Test, a cloze-style question answering dataset, and analyze the generated attention distributions. Our results show that, although a hierarchical approach does not offer much improvement over a shallow baseline, it does indeed offer a large performance boost when combining word and sentence attention with pointer sum attention.

1:50PM *SSA: A More Humanized Automatic Evaluation Method for Open Dialogue Generation [#19838]*

Zhiqiang Zhan, Zifeng Hou, Qichuan Yang, Jianyu Zhao, Yang Zhang and Changjian Hu, University of Chinese Academy of Sciences; Institute of Computing Technology, Chinese Academy of Sciences, China; Beihang University, China; Lenovo Research, China

Dialogue generation has been gaining ever-increasing attention, and various models have been proposed and adopted in many fields in recent years. How to evaluate their performance is critical. However, current evaluation metrics tend to be insufficient because of their simplicity and crudeness, resulting in weak correlation with human judgements. To solve this issue, we propose an automatic and comprehensive evaluation metric, which consists of three assessment criteria: Semantic Coherence, Syntactic Validity and Ability of Expression (SSA). The first two criteria are used to evaluate the generations from semantic and syntactic aspects respectively at the sentence level and the last one is to evaluate the overall performance at the model level. With two generative models, we conduct experiments on three datasets, including Twitter, Subtitle and Lenovo. Comparing with the previous metrics such as BLEU, METEOR and ROUGE, the correlation coefficient between SSA and human judgements is increased by 0.23-0.35, i.e. 324%-864% relative improvements. The experimental results demonstrate that SSA correlates more strongly with human judgements on the evaluation for open dialogue generation. Additionally, SSA is able to evaluate the semantic coherence and syntactic validity of generations exactly. More importantly, the evaluation models can be trained without human annotations. Thus, SSA is flexible and extensible to different datasets.

2:10PM *Multi-turn Intent Determination for Goal-oriented Dialogue systems [#20235]*

Waheed Ahmed Abro, Guilin Qi, Huan Gao, Muhammad Asif Khan and Zafar Ali, Southeast University, China

Intent determination is one of the main tasks of natural language understanding and dialogue systems, aiming to determine the intent of the user's input. Incorporating contextual information for intent determination has shown much promise. Recently, memory networks have been used to encode context from the dialogue history at each turn. However, these methods lack general mechanism for encoding intent patterns. In this paper, we investigate the incorporation of intent patterns extracted from regular expression for the multi-turn intent determination task. We propose a novel neural network model with two memories. The first memory encodes context from dialogue history at each turn whereas the second memory contains the information obtained from the regular expression patterns. We evaluate the model on Frames and Key-Value Retrieval datasets, the experimental results demonstrate that encoding intent patterns with memory networks significantly improves the multi-turn intent determination.

2:30PM *Multi-task Learning with Bidirectional Language Models for Text Classification [#19495]*
Qi Yang and Lin Shang, Nanjing University, China

Multi-task learning is an effective approach to extract task-invariant features by leveraging potential information among related tasks, which improves the performance of a single task. Most existing work simply divides the whole model into shared and private spaces. Unfortunately, there is no explicit mechanism to prevent the two spaces from merging information from each other. As a result, the shared space may be mixed with task-specific features, while the private space may extract some task- invariant features. To alleviate the problem mentioned, in this paper, we propose a bidirectional language models based multi-task learning method for text classification. More specifically, we add language modelling as an auxiliary task to the private part, aiming to enhance its ability to extract task-specific features. In addition, to promote the shared part to learn common features, a loss constraint via uniform label distribution is introduced to the shared part. Finally, put task- specific features and task-invariant features together in a weighted addition way to form the final representation, and it is then fed to the corresponding softmax layer. We do experiments on the FDU-MTL dataset which consists of 16 different text classification tasks. The experimental results show that our approach outperforms other typical methods.

2:50PM *Attention-based Multi-instance Neural Network for Medical Diagnosis from Incomplete and Low Quality Data [#19659]*
Zeyuan Wang, Josiah Poon, Sun Shiding and Simon Poon, The University of Sydney, Australia; Renmin University of China, China

One way to extract patterns from clinical records is to consider each patient record as a bag with various number of instances in the form of symptoms. Medical diagnosis is to discover informative ones first and then map them to one or more diseases. In many cases, patients are represented as vectors in some feature space and a classifier is applied after to generate diagnosis results. However, in many real-world cases, data is often of low-quality due to a variety of reasons, such as data consistency, integrity, completeness, accuracy, etc. In this paper, we propose a novel approach, attention based multi-instance neural network (AMI- Net), to make the single disease classification only based on the existing and valid information in the real-world outpatient records. In the context of a patient, it takes a bag of instances as input and output the bag label directly in end-to-end way. Embedding layer is adopted at the beginning, mapping instances into an embedding space which represents the individual patient condition. The correlations among instances and their importance for the final classification are captured by multi-head attention transformer, instance-level multi-instance pooling and bag-level multi-instance pooling. The proposed approach was test on two non- standardized and highly imbalanced datasets, one in the Traditional Chinese Medicine (TCM) domain and the other in the Western Medicine (WM) domain. Our preliminary results show that the proposed approach outperforms all baselines results by a significant margin.

3:10PM *Reduced-Gate Convolutional LSTM Architecture for Next-Frame Video Prediction Using Predictive Coding [#19159]*
Nelly Elsayed, Anthony S. Maida and Magdy Bayoumi, University of Louisiana at Lafayette, United States

Spatiotemporal sequence prediction is an important problem in deep learning. We study next-frame video prediction using a deep-learning-based predictive coding framework that uses convolutional, long short-term memory (convLSTM) modules. We introduce a novel reduced-gate convolutional LSTM architecture that achieves better next-frame prediction accuracy than the original convolutional LSTM while using a smaller parameter budget, thereby reducing training time and memory requirements. We tested our reduced gate modules within a predictive coding architecture on the gray-scale and RGB video datasets. We found that our reduced-gate model has a significant reduction of approximately 40 percent of the total number of training parameters and training time in comparison with the standard LSTM model which makes it attractive for hardware implementation especially on small devices.

1c: Self-organizing maps (including neural gas, etc.)

Monday, July 15, 1:30PM-3:30PM, Room: Panorama II, Chair: Lyes Khacef

1:30PM *Integer Self-Organizing Maps for Digital Hardware [#20091]*
Denis Kleyko, Evgeny Osipov, Daswin De Silva, Urban Wiklund and Damminda Alahakoon, Lulea University of Technology, Sweden; La Trobe University, Australia; Umea University, Sweden

The Self-Organizing Map algorithm has been proven and demonstrated to be a useful paradigm for unsupervised machine learning of two-dimensional projections of multi-dimensional data. The tri-state Self- Organizing Maps have been proposed as an accelerated resource-efficient alternative to the Self-Organizing Maps for implementation on field-programmable gate array (FPGA) hardware. This paper presents a generalization of the tri-state Self-Organizing Maps. The proposed generalization, which we call integer Self-Organizing Maps, requires only integer operations for weight updates. The presented experiments demonstrated that the integer Self-Organizing Maps achieve better accuracy in a classification task when compared to the original tri-state Self-Organizing Maps.

1:50PM *A Multi-Application, Scalable and Adaptable Hardware SOM Architecture [#20041]*
Mehdi Abadi, Slavisa Jovanovic, Khaled Ben Khalifa, Serge Weber and Mohamed Hedi Bedoui, UMR 7198, Institut Jean Lamour, Universite de Lorraine, Nancy, France; LR12ES06, Laboratoire de Technologie et Imagerie Medicale, Universite de Monastir, Monastir, Tunisia

In this work, a scalable and adaptable hardware SOM architecture allowing to execute multiple applications in parallel is presented. The proposed architecture allows to use simultaneously multiple SOM structures with different parameters in order to satisfy multiple applications with different needs. The application switching is done within a clock cycle at the neuron's level at run time only by analyzing the received input data. The proposed architecture was tested and validated in an image quantization experiment where 6 quantization applications with different parameters (from 6x6 to 15x15 SOMs with inputs varying from 3 to 12 elements) were performed simultaneously.

2:10PM *Self-organizing neurons: toward brain-inspired unsupervised learning [#19097]*
Lyes Khacef, Benoit Miramond, Diego Barrientos and Andres Upegui, Universite Cote d'Azur, CNRS, LEAT, France; InIT, hepia, University of Applied Sciences of Western Switzerland, Switzerland

During the last years, Deep Neural Networks have reached the highest performances in image classification. Nevertheless, such a success is mostly based on supervised and off-line learning: they require thus huge labeled datasets for learning, and once it is done, they cannot adapt to any change in the data from the environment. In the context of brain-inspired computing, we apply Kohonen- based Self-Organizing Maps for unsupervised learning without labels, and we explore original extensions such as the Dynamic SOM that enables continuous learning and the Pruning Cellular SOM that includes synaptic pruning in neuromorphic circuits. After presenting the three models and the experimental setup for MNIST classification, we compare different methods for automatic labeling based on very few labeled data (1% of the training dataset), and then we compare the performances of the three Kohonen-based Self-Organizing Maps with STDP-based Spiking Neural Networks in terms of accuracy, dynamicity and scalability.

2:30PM *A Semi-Supervised Self-Organizing Map with Adaptive Local Thresholds [#20380]*
Pedro Braga and Hansenclever Bassani, Universidade Federal de Pernambuco, Brazil

In the recent years, there is a growing interest in semi-supervised learning, since, in many learning tasks, there is a plentiful supply of unlabeled data, but insufficient labeled ones. Hence, Semi-Supervised learning models can benefit from both types of data to improve the obtained performance. Also, it is important to develop methods that are easy to parameterize in a way that is robust to the different characteristics of the data at hand. This article presents a new method based on Self-Organizing Map (SOM) for clustering and classification, called Adaptive Local Thresholds Semi-Supervised Self-Organizing Map (ALTSS-SOM). It can dynamically switch between two forms of learning at training time, according to the availability of labels, as in previous models, and can automatically adjust itself to the local variance observed in each data cluster. The results show that the ALTSS-SOM surpass the performance of other semi-supervised methods in terms of classification, and other pure clustering methods when there are no labels available, being also less sensitive than previous methods to the parameters values.

2:50PM *A Gaussian Process-based Self-Organizing Incremental Neural Network [#20369]*
Xiaoyu Wang, Giona Casiraghi, Yan Zhang and Jun-ichi Imura, Tokyo Institute of Technology, Japan; ETH Zurich, Switzerland

This paper proposes a Gaussian process-based self-organizing incremental neural network (GPINN) to address the density estimation problem of online unsupervised learning. First, we adopt Gaussian process models with adaptive kernels that map the distribution of the neighbors of each node to its link relationship. Second, combining GPINN and kernel density estimation, we derive the bandwidth matrix updating rule for adapting to the generated network. We theoretically analyze the advantages of the proposed approach in determining threshold regions over using distance measures. The experimental results on both synthetic data sets and real-world data sets show that our method achieves remarkable improvement in density estimation accuracy for large noisy data.

3:10PM *Distant Supervised Why-Question Generation with Passage Self-Matching Attention [#19529]*
Jiaxin Hu, Zhixu Li, Renshou Wu, Hongling Wang, An Liu, Jiajie Xu, Pengpeng Zhao and Lei Zhao, Soochow University, Neusoft Corporation, China; Soochow University, IFLYTEK Research, China; Soochow University, China

Question generation (QG) aims to create a fluent question from a passage and a target answer. State-of-the-art approaches are mainly based on encoder-decoder models to generate questions from the given passage and answer, which focus on using the information contained in a particular part of the passage for QG, but unaware of the clues hidden in other parts of the passage. Besides, the existing work on QG mainly focus on generating factoid questions, which are less suitable for generating non-factoid questions such as why-questions. In this paper, we propose to augment encoder-decoder framework with a pair-wise self-matching attention mechanism to dynamically collect inter-sentential evidence from the whole passage according to the current passage word and answer information. Besides, to let the model be more suitable for why-question generation, we also involve some causal features in the encoding process. Finally, to tackle the lack of why-question generation training data problem, we adopt a distant supervised method with an initial causal knowledge base to generate a large training data for why question generation. Extensive experiments on several data sets show that our model significantly outperforms state- of-the-art question generation models not only on why-question generation tasks, but also on other types of question generation tasks.

S31: Intelligent Vehicle and Transportation Systems and Other Applications

Monday, July 15, 1:30PM-3:30PM, Room: Panorama III, Chair: Tao Zheng

1:30PM *Removing Movable Objects from Grid Maps of Self-Driving Cars Using Deep Neural Networks [#20317]*
Ranik Guidolini, Raphael V. Carneiro, Claudine Badue, Thiago Oliveira-Santos and Alberto F. De Souza, Universidade Federal do Espirito Santo UFES, Brazil

We propose a technique for removing traces of movable objects from occupancy grid maps based on deep neural networks, dubbed enhanced occupancy grid map generation (E-OGM-G). In E-OGM-G, we capture camera images synchronized and aligned with LiDAR rays, semantically segment these images, and compute which laser rays of the LiDAR hit pixels segmented as belonging to movable objects. By clustering laser rays that are close together in a 2D projection, we are able to identify clusters that belong to movable objects and avoid using them in the process of generating the OGMs - this allows generating OGMs clean of movable objects. Clean OGMs are important for several aspects of self-driving cars' operation (i.e., localization). We tested E-OGM-G using data obtained in a real-world scenario - a 2.6 km stretch of a busy multi-lane urban road. Our results showed that E-OGM-G can achieve a precision of 81.19% considering the whole OGMs generated, of 89.76% considering a track in these OGMs of width of 12 m, and of 100.00% considering a track of width of 3.4 m. We then tested a self-driving car using the automatically cleaned OGMs. The self-driving car was able to properly localize itself and to autonomously drive itself in the world using the cleaned OGMs. These successful results showed that the proposed technique is effective in removing movable objects from static OGMs.

1:50PM *Traffic Light Recognition Using Deep Learning and Prior Maps for Autonomous Cars [#20432]*

Lucas C. Possatti, Ranik Guidolini, Vinicius B. Cardoso, Rodrigo F. Berriel, Thiago M. Paixao, Claudine Badue, Alberto F. De Souza and Thiago Oliveira-Santos, Universidade Federal do Espirito Santo, Brazil; Instituto Federal do Espirito Santo, Brazil

Autonomous terrestrial vehicles must be capable of perceiving traffic lights and recognizing their current states to share the streets with regular human drivers. Most of the time, human drivers can easily identify the relevant traffic lights. To deal with this issue, a common solution for autonomous cars is to integrate recognition with prior maps. However, additional solution is required for the detection and recognition of the traffic light. Deep learning techniques have showed great performance and power of generalization including traffic related problems. Motivated by the advances in deep learning, some recent works leveraged some state-of-the-art neural detectors to locate (and further recognize) traffic lights from 2D camera images. However, none of them combine the power of the deep learning-based detectors with prior maps to recognize the state of the relevant traffic lights. Based on that, this work proposes to integrate the power of deep learning-based detection with the prior maps used by our car platform IARA (acronym for Intelligent Autonomous Robotic Automobile) to recognize the relevant traffic lights of predefined routes. The process is divided in two phases: an offline phase for map construction and traffic lights annotation; and an online phase for traffic light recognition and identification of the relevant ones. The proposed system was evaluated on five test cases (routes) in the city of Vitoria, each case being composed of a video sequence and a prior map with the relevant traffic lights for the route. Results showed that the proposed technique is able to correctly identify the relevant traffic light along the trajectory.

2:10PM *Bio-Inspired Foveated Technique for Augmented-Range Vehicle Detection Using Deep Neural Networks [#20424]*

Pedro Azevedo, Sabrina Panceri, Ranik Guidolini, Vinicius B. Cardoso, Claudine Badue, Thiago Oliveira-Santos and Alberto F. De Souza, Universidade Federal do Espirito Santo, Brazil

We propose a bio-inspired foveated technique to detect cars in a long range camera view using a deep convolutional neural network (DCNN) for the IARA self-driving car. The DCNN receives as input (i) an image, which is captured by a camera installed on IARA's roof; and (ii) crops of the image, which are centered in the waypoints computed by IARA's path planner and whose sizes increase with the distance from IARA. We employ an overlap filter to discard detections of the same car in different crops of the same image based on the percentage of overlap of detections' bounding boxes. We evaluated the performance of the proposed augmented-range vehicle detection system (ARVDS) using the hardware and software infrastructure available in the IARA self-driving car. Using IARA, we captured thousands of images of real traffic situations containing cars in a long range. Experimental results show that ARVDS increases the Average Precision (AP) of long range car detection from 29.51% (using a single whole image) to 63.15%.

2:30PM *Attention-Driven Driving Maneuver Detection System [#20003]*

Xishuai Peng, Ava Zhao, Song Wang, Yi Lu Murphey and Yuanxiang Li, University of Michigan-Dearborn, United States; Shanghai Jiao Tong University, China

Driving Maneuver early Detection (DMD) is one of the most important tasks in Advanced Driver Assistance Systems (ADAS), it provides the early notification necessary for ADAS to predict dangerous circumstances and take appropriate actions. The end-to-end architectures such as Recurrent Neural Networks (RNNs) take advantage of deep networks to automatically learn non-linear discriminative features, which significantly boost the performances of DMD systems. However, due to the large number of parameters in the deep architectures, learning effective discriminative features requires millions of labeled images. Moreover the discriminative features are generally meaningless to human being, which makes the diagnose of end-to-end architectures extremely difficult. In this paper, we propose a novel DMD system, denoted as Attention-driven Driving Maneuver Detection system (ADMD), which uses drivers' attention as an intermediate concept (hint) to explore and understand the causal relation between driving surroundings and driving maneuvers. In the training phase, ADMD distills the knowledge from drivers' attention prediction model to provide initial search areas for learning effective features, which minimizes the labeled data requirements for training deep architectures and results in a fast optimization process. We compared the performances of ADMD with the state-of-the-art methods on 1953 miles (37 hours) of natural driving data collected from 7 drivers. The experimental results shows ADMD is capable of achieving better performances, which are mostly attributed to the proposed novel learning mechanism used to obtain meaningful interpretations of the driving surroundings that are closely related to the driver's intend maneuvers.

2:50PM *Generative Adversarial Network for Radar Signal Generation [#20214]*

Thomas Truong and Svetlana Yanushkevich, University of Calgary, Canada

A major obstacle in ultra-wideband radar based approaches for object detection concealed on human body is the difficulty in collecting high quality radar signal data. Generative adversarial networks (GAN) have shown promise in synthesizing data for image and audio processing. This paper proposes the design of a GAN for application in radar signal generation. Data collected using the Finite-Difference Time-Domain (FDTD) method on three concealed object classes (no object, large object, and small object) are used as training data. A GAN is trained to generate radar signal samples for each class. The proposed GAN is capable of synthesizing the radar signal data which is indistinguishable from the training data by qualitative analysis performed by human observers.

3:10PM *An Improved Recurrent Neural Network Language Model for Programming Language [#19237]*

Liwei Wu, Youhua Wu, Fei Li and Tao Zheng, Nanjing University, China

Language models are applied to the programming language. However, the existing language models may be confused with different tokens with the same name in different scopes and can generate the syntax error code. In this paper, we proposed a grammar language model to solve these two problems. The model is an improved recurrent neural network language model. The improved recurrent neural network language model has the scope-awared input feature and the grammar output mask. We evaluated our model and existing language models on a C99 code dataset. Our model gets a perplexity value of 2.91 and a top-1 accuracy rate of 74.23% which is much better than other models.

1a: Feedforward neural networks, 2k, 2m

Monday, July 15, 1:30PM-3:30PM, Room: Panorama IV, Chair: Gabriel Terejanu

1:30PM *Approximate Bayesian Neural Network Trained with Ensemble Kalman Filter [#19924]*
Chao Chen, Lin Xiao, Yuan Huang and Gabriel Terejanu, University of South Carolina, United States; University of North Carolina at Charlotte, United States

Neural networks have achieved significant success in many areas. Nevertheless, conventional neural networks lack uncertainty information, which plays an important role especially in critical-safety applications such as self-driving cars. When uncertainty is characterized using probability the main modeling approach is the construction of Bayesian neural networks. Obtaining the posterior distribution for these models is computationally intensive and analytical solutions are intractable. In this work, we propose a novel algorithm to infer the weights for Bayesian neural networks based on the ensemble Kalman filter. To evaluate the performance of the algorithm, we use ten regression datasets from University of California at Irvine machine learning repository, and a natural language dataset. The results suggest that EnKF can be used as a gradient-free alternative to training deep neural networks to capture prediction uncertainty.

1:50PM *Ensemble Attention For Text Recognition In Natural Images [#20462]*
Hongchao Gao, Yujia Li, Xi Wang, Jizhong Han and Ruixuan Li, IIE.AC.CN, China

Recognizing text from natural images is a challenging and hot research topic in computer vision, yet not completely solved. The recent methods regard this task as a sequence labeling problem. In this task, there is a strong correspondence between the position of the input image patches sequence and the output character sequence. However, most of the recent recognition systems rarely consider this local information of the input sequence when recognizing the current character. In contrast to this, we present a Local Restricted Attention (LRA) mechanism to encode the current vector by considering adjacent vectors of the input sequence. We propose an ensemble decoder block which combines LRA mechanism with a regular decoder mechanism. This block not only brings significant improvement of recognition results under shorter training time but also can be easily embedded in other recognition frameworks. In addition, we propose a scene text recognition network based on the ensemble decoder. The experimental performances show that the proposed model achieves the state-of-the-art on several benchmark datasets including IIIT- 5K, SVT, CUTE80, SVT-Perspective and ICDARs.

2:10PM *Multilayer Perceptron for Sparse Functional Data [#20267]*
Qiyao Wang, Shuai Zheng, Ahmed Farahat, Susumu Serita, Takashi Saeki and Chetan Gupta, Industrial AI Lab, Hitachi America, Ltd. R&D, United States

In this paper, we propose a novel algorithm for generalizing Multilayer Perceptron (MLP) to handle sparse functional data, wherein for a given subject there are multiple observations available over time and these observations are sparsely and irregularly distributed within the considered time range. The algorithm uses pooled observations across all the subjects to estimate a set of basis functions for the underlying correlation between time steps and then use these basis functions to build a sparse functional neuron that extracts features for each subject. We justify the validity of our algorithm through theoretical arguments. We also demonstrate the use of the proposed algorithm in solving three data challenges: the classification of synthetic curves, the prediction of patient's long-term survival, and the estimation of the remaining time to critical failures for turbofan engines. To show the superiority of our algorithm under sparse functional data scenarios, we compare the performance of our model with two alternative common practices, and demonstrate that our method outperforms baseline methods in all numerical studies.

2:30PM *AdaBoost with Neural Networks for Yield and Protein Prediction in Precision Agriculture [#19689]*
Amy Peerlinck, John Sheppard and Jacob Senecal, Montana State University, United States

Adaptive Boosting, or AdaBoost, is an algorithm aimed at improving the performance of ensembles of weak learners by weighing the data itself as well as the learners. Two versions of AdaBoost---AdaBoost-R and AdaBoost RΔ---are applied in this project, as well as a third novel algorithm combining ideas of the two existing versions, to the problem of predicting crop yield and protein content in support of precision agriculture. All three algorithms use Feedforward Neural Networks (FFNN) trained with backpropagation as the weak model. The models are applied to the field of Precision Agriculture, more specifically to predicting crop yield and protein of winter wheat in Montana. Data from four different fields were gathered as a result of on-farm experiments of different nitrogen rate applications using randomly stratified trials based on previous years' yield and protein. The three AdaBoost algorithms are compared to a simple FFNN with a single hidden layer. The results confirm previous findings in different fields, where ensemble methods outperform single models. The results are improved by 3 to 10 units for yield prediction, and by a small percentage for protein prediction.

2:50PM *Parallelizing Basis Pursuit Denoising [#19919]*
Cory Kromer-Edwards, Suely Oliveira and David Stewart, Dept of Computer Science, University of Iowa, United States; Dept of Mathematics, University of Iowa, United States

Basis Pursuit DeNoising (BPDN) is a version of sparse least squares, where there is an L1 penalty for the coefficients in addition to the sum of squares of the errors. To apply this technique to large data sets, parallel algorithms are needed. In this paper, we create a parallel version of a BPDN algorithm that can compute exact solutions for an interval of values for the L1 penalty parameter that exploits the sparsity of the solutions, unlike stochastic gradient descent-type algorithms.

3:10PM *Group k-Sparse Temporal Convolutional Neural Networks: Unsupervised Pretraining for Video Classification [#20243]*
Zoltan A. Milacski, Barnabas Poczos and Andras Lorincz, Faculty of Informatics, ELTE Eotvos Lorand University, Hungary; Machine Learning Department, Carnegie Mellon University, United States

In this paper we propose Group k-Sparse Temporal Convolutional Neural Networks for unsupervised pretraining using video data. Our work is the first to consider the recurrent extension of structured sparsity, thus enhancing representational power and explainability. We show that our architecture is able to outperform several state-of-the-art baselines on Rotated MNIST, Scanned CIFAR-10, COIL-100 and NEC Animal pretraining benchmarks for video classification using limited labeled data.

Competition: L2RPN: Learning to run a power network

Monday, July 15, 1:30PM-3:30PM, Room: Panorama V, Chair: Antoine Marot, Balthazar Donon, Benjamin Donnot Isabelle Guyon

Monday, July 15, 3:30PM-4:00PM

Special Lecture: Coffee Break
Monday, July 15, 3:30PM-4:00PM, Room: Pre-function area Intercontinental

Monday, July 15, 4:00PM-5:00PM

Plenary Talk: Erkki Oja, Aalto University, School of Science and Technology.
Monday, July 15, 4:00PM-5:00PM, Room: Ballroom I + II +II, Chair: Danilo Mandic,

Monday, July 15, 5:30PM-7:30PM

1l: Deep neural networks, Cellular Computational Networks
Monday, July 15, 5:30PM-7:30PM, Room: Ballroom I, Chair: Prof. S. Das

5:30PM *Directional Attention based Video Frame Prediction using Graph Convolutional Networks [#19890]*
Prateep Bhattacharjee and Sukhendu Das, Indian Institute of Technology Madras, India

This paper proposes a novel network architecture for video frame prediction based on Graph Convolutional Neural Networks (GCNN). Most recent methods often fail in situations where multiple close-by objects at different scales move in random directions with variable speeds. We overcome this by modeling the scene as a space-time graph with intermediate features from the pixels (or a local region) as vertices and the relationships among them as edges. Our main contribution lies within posing the frame generation problem with our proposed space- time graph, which enables the network to learn the spatial as well as temporal inter-pixel relationships independent of each other, thus making the system invariant to velocity differences among the moving objects present in the scene. Moreover, we also propose a novel directional attention mechanism for the graph based model to efficiently learn a significance score based on directional relationship between pixels in the original scene. We also show that the proposed model generalizes better on the much more challenging task of predicting semantic scene segmentation of future scenes, even without access to any raw RGB frames. We perform several proxy tasks such as comparison of the quality of the semantic segmentation produced on the generated frames and comparing the accuracies for the task of recognizing actions in case of the dataset consisting of human actions. We use the popular Cityscapes traffic scene segmentation dataset as well as UCF-101 and Penn Action containing human actions to quantitatively and qualitatively evaluate the proposed framework over the recent state-of-the-art.

5:50PM *Training Deep Neural Networks with Adversarially Augmented Features for Small-scale Training Datasets [#19134]*
Masato Ishii and Atsushi Sato, NEC, Japan

In this paper, we propose a novel method to train deep neural networks for small-scale training dataset. Since we focus on the situation in which the training dataset is limited, increasing the number of training data by data augmentation is the most straight-forward solution. Instead of augmenting data at an input layer as in typical data augmentation, we adversarially augment features at hidden layer by adding small perturbations to the original features extracted from training data. We call the augmented features as adversarial features. To effectively avoid overfitting of the trained network, the perturbation is designed to be adversarial, which means that they are designed to significantly change the output of the network. Moreover, to induce the adversarial feature to be more reasonable as real data, we adopt a coefficient layer to constrain the adversarial feature to be represented by linear combination of the original features. Experimental results on several benchmark datasets show that our method substantially improves the performance of the deep neural network.

6:10PM *DAGCN: Dual Attention Graph Convolutional Networks [#19706]*
Fengwen Chen, Shirui Pan, Jing Jiang, Huan Huo and Guodong Long, Centre for Artificial Intelligence, FEIT, University of Technology Sydney, Australia; Faculty of Information Technology, Monash University, Australia; School of software, FEIT, University of Technology Sydney, Australia

Graph convolutional networks (GCNs) have recently become one of the most powerful tools for graph analytics tasks in numerous applications, ranging from social networks and natural language processing to bioinformatics and chemoinformatics, thanks to their ability to capture the complex relationships between concepts. At present, the vast majority of GCNs use a neighborhood aggregation framework to learn a continuous and compact vector, then performing a pooling operation to generalize graph embedding for the classification task. These approaches have two disadvantages in the graph classification task: (1)when only the largest sub-graph structure (k-hop neighbor) is used for neighborhood aggregation, a large amount of early-stage information is lost during the graph convolution step; (2) simple average/sum pooling or max pooling utilized, which loses the characteristics of each node and the topology between nodes. In this paper, we propose a novel framework called, dual attention graph convolutional networks (DAGCN) to address these problems. DAGCN automatically learns the importance of neighbors at different hops using a novel attention graph convolution layer, and then employs a second attention component, a self-attention pooling layer, to generalize the graph representation from the various aspects of a matrix graph embedding. The dual attention network is trained in an end-to-end manner for the graph classification task. We compare our model with state-of-the-art graph kernels and other deep

learning methods. The experimental results show that our framework not only outperforms other baselines but also achieves a better rate of convergence.

6:30PM *Efficient Convolutional Neural Networks for Multi-Spectral Image Classification [#19045]*
Jacob Senecal, John Sheppard and Joseph Shaw, Montana State University, United States

While a great deal of research has been directed towards developing neural network architectures for RGB images, there is a relative dearth of research directed towards developing neural network architectures specifically for multi-spectral and hyper-spectral imagery. We have adapted recent developments in small efficient convolutional neural networks (CNNs), to create a small CNN architecture capable of being trained from scratch to classify 10 band multi-spectral images, using much fewer parameters than popular deep architectures, such as the ResNet or DenseNet architectures. We show that this network provides higher classification accuracy and greater sample efficiency than the same network using RGB images. Further, using a Bayesian version of our CNN architecture we show that a network that is capable of working with multi-spectral imagery significantly reduces the uncertainty associated with class predictions compared to using RGB images.

6:50PM *From Face Recognition to Facial Pareidolia: Analysing Hidden Neuron Activations in CNNs for Cross-Depiction Recognition [#19966]*
Asad Abbas and Stephan Chalup, The University of Newcastle, Australia

The imagination of non-existent faces in random patterns, clouds and rock formations is known as facial pareidolia. We show that facial pareidolia also occurs naturally in a standard Convolutional Neural Network (CNN) trained on face recognition. For achieving this we propose a new method to analyse CNNs that combines feature visualisation and dimensionality reduction methods to cluster the hidden neuron activations in convolutional layers into groups with discriminative roles. The main contributions of the present paper are 1.) an approach that uses a CNN trained on human face detection for facial pareidolia simulation without any additional training on a target image set of abstract facial patterns and 2.) a novel way of improving the generalisation capacity of a CNN for cross-depiction recognition and domain adaptation scenarios using features learned by hidden neurons.

7:10PM *Image Captioning Based On Sentence-Level And Word-Level Attention [#19749]*
Haiyang Wei, Zhixin Li, Canlong Zhang, Tao Zhou and Yu Quan, Guangxi Normal University, China

Existing attention models of image captioning typically extract only word-level attention information. i.e., the attention mechanism extracts local attention information from the image to generate the current word. We propose an image captioning approach based on self-attention to utilize image features more effectively. The self-attention mechanism can extract sentence-level attention information with richer visual representation from images. Furthermore, we propose a double attention model. The model combines sentence-level and word- level attention information to better simulate human perception system. We implement supervision and optimization in the intermediate stage of the model to solve over-fitting and information interference problems, and we apply reinforcement learning to two-stage training to optimize the evaluation metrics of the model. Finally, we evaluate our model on MSCOCO dataset. The experimental results show that our approach can generate more accurate and richer captions, and outperforms many state-of-the-art image captioning approaches on various evaluation metrics.

2e: Deep learning

Monday, July 15, 5:30PM-7:30PM, Room: Ballroom II, Chair: Andrew Skabar

5:30PM *Restricted Boltzmann Machines: an EigenCentrality-based Approach [#19109]*
Andrew Skabar, Department of Computer Science and Information Technology, La Trobe University, Australia

Restricted Boltzmann Machines (RBMs) are usually trained using contrastive divergence learning, which requires setting suitable values for a number of hyper-parameters, and requires a degree of practical know-how. This paper presents a radically different approach to RBMs, which we refer to as the Eigencentrality-based Restricted Boltzmann Machine (ECRBM). Rather than actually 'training' the RBM, the dataset is first represented as a bipartite graph consisting of a layer of attribute-value nodes and a layer of object nodes, with the representation chosen such that the components of an eigenvector calculated from the graph can be used to estimate the conditional probabilities of the attribute values. Matrix factorization is then used to replace this graph with a smaller graph, having the important property that eigenvectors calculated on this smaller graph preserve the interpretation of the eigenvector components as conditional probabilities. The model can be used to estimate and sample from any conditional or marginal distribution, and can therefore be applied to a diverse range of machine learning tasks including classification, regression, missing value imputation, outlier detection, and random vector generation. It can be applied to mixed-attribute datasets as easily as it can be applied to datasets containing variables of the same modality. We demonstrate the method by applying it to classification and random vector generation on a number of mixed-attribute datasets.

5:50PM *Adversarial Domain Adaptation via Category Transfer [#19337]*
Lusi Li, Haibo He, Jie Li and Guang Yang, University of Rhode Island, United States; Chongqing University of Science and Technology, China; Zhongnan University of Economics and Law, China

Adversarial domain adaptation has achieved some success in learning transferable feature representations and reducing distribution discrepancy between source and target domains. However, existing approaches mainly focus on alignment of global source and target distributions without considering complex structures in categories underlying different distributions, resulting in domain confusion and the mix of distinguishable structures. In this paper, we propose an adversarial domain adaptation via category transfer (ADACT) approach for unsupervised domain adaptation (UDA). ADACT first captures multi-category information through training source and target feature generators as well as a label predictor. Secondly, it uses multi-category domain critic networks to category-wisely estimate Wasserstein distances across domains. Then it learns category-invariant feature representations by finely grained matching different data distributions with the estimated Wasserstein distances. The adaptation can be achieved by the standard back-propagation training approach with this two-step iteration. The effectiveness of ADACT is demonstrated since it outperforms several state-of-the-art UDA methods on common domain adaptation datasets.

6:10PM *Deep Diffusion Autoencoders [#20156]*
Sara Dorado, Angela Fernandez and Jose R. Dorronsoro, Autonomous University of Madrid, Spain

Extending work by Mishne et al., we propose Deep Diffusion Autoencoders (DDA) that learn an encoder- decoder map using a composite loss function that simultaneously minimizes the reconstruction error at the output layer and

the distance to a Diffusion Map embedding in the bottleneck layer. These DDA are thus able to reconstruct new patterns from points in the embedding space in a way that preserves geometry of the sample and, as a consequence, our experiments show that they may provide a powerful tool for data augmentation.

6:30PM *Deep Multi-view Learning from Sequential Data without Correspondence [#19143]*

Tung Doan and Atsuhiro Takasu, SOKENDAI (The Graduate University for Advanced Studies), Japan; National Institute of Informatics, Japan

Multi-view representation learning has become an active research topic in machine learning and data mining. One underlying assumption of the conventional methods is that training data of the views must be equal in size and sample-wise matching. However, in many real-world applications, such as video analysis, text streaming, and signal processing, data for the views often come in form of sequences, which are different in length and misaligned. This results in failure of directly applying existing methods to such problems. In this paper, we firstly introduce a novel deep multi-view model that can implicitly discover sample correspondence while learning the representation. It can be shown that our method generalizes deep canonical correlation analysis - a popular multi-view learning method. We then extend our model by integrating the objective function with the reconstruction losses of autoencoders, forming a new variant of the proposed model. Extensively experimental results demonstrate superior performances of our models over competing methods.

6:50PM *Deep Q-Learning for Illumination and Rotation invariant Face Detection [#20347]*

Ariel Ruiz-Garcia, Vasile Palade, Ibrahim Almakky and Mark Elshaw, Coventry University, United Kingdom

The domain of automatic face detection is a challenging problem that has made great progress in the last two decades. Much of this progress is due to the continuous advancements in deep learning and computer vision. Nonetheless, contemporary state- of-the-art face detection models rely on exhaustive search or are unable to deal with changes in the data distribution. In this work, we propose a novel Deep Reinforcement Learning (DRL) approach for face detection on data with nonuniform conditions. More specifically, we address illumination and rotation invariance in face detection. Firstly, we train a Stacked Convolutional Autoencoder (SCAE) in a greedy layer-wise unsupervised fashion for illumination invariant feature extraction. We then train a deep Q-network on the illumination invariant features produced by the SCAE model, to learn an action-value policy that allows an agent to place a bounding box around a face. The proposed approach achieves state of-the-art recognition rates on images with varying degrees of illumination and images that contain faces with some degree of rotation.

7:10PM *Synthetic-to-Real Domain Adaptation for Object Instance Segmentation [#19338]*

Hui Zhang, Yonglin Tian, Kunfeng Wang, Haibo He and Fei-Yue Wang, Institute of Automation, Chinese Academy of Sciences, China; University of Science and Technology of China, China; University of Rhode Island, United States

Object instance segmentation can achieve preferable results, powered with sufficient labeled training data. However, it is time- consuming for manually labeling, leading to the lack of large-scale diversified datasets with accurate instance segmentation annotations. Exploiting the synthetic data is a very promising solution except for domain distribution mismatch between synthetic dataset and real dataset. In this paper, we propose a synthetic-to-real domain adaptation method for object instance segmentation. At first, this approach is trained to generate object detection and segmentation using annotated data from synthetic dataset. Then, a feature adaptation module (FAM) is applied to reduce data distribution mismatch between synthetic dataset and real dataset. The FAM performs domain adaptation from three different aspects: global-level base feature adaptation module, local-level instance feature adaptation module, and subtle-level mask feature adaptation module. It is implemented based on novel discriminator networks with adversarial learning. The three modules of FAM have positive effects on improving the performance when adapting from synthetic to real scenes. We evaluate the proposed approach on Cityscapes dataset by adapting from Virtual KITTI and SYNTHIA datasets. The results show that it achieves a significantly better performance over the state-of-the-art methods.

8a: Applications of deep networks

Monday, July 15, 5:30PM-7:30PM, Room: Ballroom III, Chair: Reda Al-Bahrani

5:30PM *Towards A Deep Learning Question-Answering Specialized Chatbot for Objective Structured Clinical Examinations [#20058]*

Julia El Zini, Yara Rizk, Mariette Awad and Jumana Antoun, American University of Beirut, Lebanon

Medical students undergo exams, called "Objective Structured Clinical Examinations" (OSCEs), to assess their medical competence in clinical tasks. In these OSCEs, a medical student interacts with a standardized patient, asking questions to complete a clinical assessment of the patient's medical case. In real OSCEs, standardized patients or "Actors" are recruited and trained to answer questions about symptoms mentioned in a script designed by the medical examiner. Developing a virtual conversational patient for OSCEs would lead to significant logistical savings. In this work, we develop a deep learning framework to improve the virtual patient's conversational skills. First, deep neural networks learned domain specific word embeddings. Then, long short-term memory networks derived sentence embeddings before a convolutional neural network model selected an answer to a given question from a script. Empirical results on a homegrown corpus showed that this framework outperformed other approaches, with an accuracy of 81%.

5:50PM *To Comprehend the New: On Measuring the Freshness of a Document [#20232]*

Tirthankar Ghosal, Abhishek Shukla, Asif Ekbal and Pushpak Bhattacharyya, IIT Patna, India; IIIT Kalyani, India

Detecting the novelty or freshness of an entire document is essential in this age of data duplication and semantic- level redundancy all across the web. Current techniques for the problem mostly root on handcrafted similarity and diver- gence based measures to classify a document as novel or non- novel. However, document-level novelty detection is relatively less explored in literature if compared to its sentence-level counterpart. In this work, we present a deep neural architecture to automatically predict the amount of new information contained in a document in the form of a novelty score. Along with, we offer a dataset of more than 7500 documents, annotated at the sentence-level to facilitate further research. Our approach which learns the notion of novelty and redundancy only from the data achieves significant performance improvement over the existing methods and adopted baselines (nearly 17% error reduction). Also, our approach complies with the Two-Stage theory of human recall essential to comprehend new information.

6:10PM *Peak Area Detection Network for Directly Learning Phase Regions from Raw X-ray Diffraction Patterns [#19901]*

Dipendra Jha, Aaron Gilad Kusne, Reda Al-Bahrani, Nam Nguyen, Wei-keng Liao, Alok Choudhary and Ankit Agrawal, Northwestern University, United States; National Institute of Standards and Technology, United States

X-ray diffraction (XRD) is a well-known technique used by scientists and engineers to determine the atomic- scale structures as a basis for understanding the composition-structure-property relationship of materials. The current approach for the analysis of XRD data is a multi-stage process requiring several intensive computations such as integration along 2\theta for conversion to 1D patterns (intensity-2\theta), background removal by polynomial fitting, and indexing against a large database of reference peaks. It impacts the decisions about the subsequent experiments of the materials under investigation and delays the overall process. In this paper, we focus on eliminating such multi-stage XRD analysis by directly learning the phase regions from the raw (2D) XRD image. We introduce a peak area detection network (PADNet) that directly learns to predict the phase regions using the raw XRD patterns without any need for explicit preprocessing and background removal. PADNet contains specially designed large symmetrical convolutional filters at the first layer to capture the peaks and automatically remove the background by computing the difference in intensity counts across different symmetries. We evaluate PADNet using two sets of XRD patterns collected from SLAC and Bruker D-8 for the Sn-Ti-Zn-O composition space; each set contains 177 experimental XRD patterns with their phase regions. We find that PADNet can successfully classify the XRD patterns independent of the presence of background noise and perform better than the current approach of extrapolating phase region labels based on 1D XRD patterns.

6:30PM *On the Discriminative Power of Learned vs. Hand-Crafted Features for Crowd Density Analysis [#20479]*

Mohamed Amine Marnissi, Hajer Fradi and Jean-Luc Dugelay, Laboratory of Advanced Technology and Intelligent Systems (LATIS) University of Sousse, Tunisia; EURECOM, France

Crowd density analysis is a crucial component in video surveillance mainly for security monitoring. This paper proposes a novel approach for crowd density classification, in which learned features substitute the commonly used handcrafted features. In particular, the approach consists of employing deep networks to extract useful crowd features that can further be manageable by a classifier. This process is favorable for crowd features extraction due to the large learning capability of deep networks compared to traditional methods based on handcrafted features. The proposed approach is evaluated on three challenging datasets, and the results demonstrate the effectiveness of learned features for crowd density classification. Furthermore, we include an extensive comparative study between different learned/hand-crafted features in order to investigate their discriminative power to handle such problems. Their performance is evaluated using different classifiers and strategies as well.

6:50PM *Emotion Intensity Estimation from Video Frames using Deep Hybrid Convolutional Neural Networks [#19700]*

Selvarajah Thuseethan, Sutharshan Rajasegarar and John Yearwood, PhD Student, Deakin University, Australia, Australia; Senior Lecturer, Deakin University, Australia, Australia; Professor, Deakin University, Australia, Australia

Detecting emotional states of human from videos is essential in order to automate the process of profiling human behaviour, which has applications in a variety of domains, such as social, medical and behavioural science. Considerable research has been carried out for binary classification of emotions using facial expressions. However, a challenge exists to automate the feature extraction process to recognise the various intensities or levels of emotions. The intensity information of emotions is essential for tasks such as sentiment analysis. In this work, we propose a metric-based intensity estimation mechanism for primary emotions, and a deep hybrid convolutional neural network-based approach to recognise the defined intensities of the primary emotions from spontaneous and posed sequences. Further, we extend the intensity estimation approach to detect the basic emotions. The frame level facial action coding system annotations and the intensities of action units associated with each primary emotion are considered for deriving the various intensity levels of emotions. The evaluation on benchmark datasets demonstrates that our proposed approach is capable of correctly classifying the various intensity levels of emotions as well as detecting them.

7:10PM *GANemotion: Increase Vitality of Characters in Videos by Generative Adversary Networks [#20002]*

Muhammad Hassan, Yutong Liu, Linghe Kong, Ziming Wang and Guihai Chen, Shanghai Jiao Tong University, China

Increasing the vitality of facial expression for characters in videos can greatly improve the user experience on entertainments, like in films, animations, news broadcasting, or even a static painting. Recent advances have been made on facial expression migration, especially by GANs for better performance, while most of them cannot be automatically applied in consecutive frames. In this paper, we propose GANemotion, an automatic facial expression migration architecture for real-time videos. Three main modules are designed for this purpose: audio emotion classification module, facial expression classification module, and facial expression migration module on GANs. The outputs of the first two modules will be the inputs of the third module, where the emotion type acquired from text analysis will be the target generation label and image from facial classification is the object for processing. Expression attention algorithm and standard AU library are specially designed for avoiding distortion and processing on non-real world characters. Experiments have been applied to each module and the whole architecture respectively. They show the feasibility of GANemotion with higher classification accuracy, and vivid generation, breaking the limits of non-real world characters and low illumination conditions.

1h: Spiking neural networks

Monday, July 15, 5:30PM-7:30PM, Room: Duna Salon I, Chair: Federico Corradi

5:30PM *A Spiking Network for Inference of Relations Trained with Neuromorphic Backpropagation [#19546]*

Johannes Christian Thiele, Olivier Bichler, Antoine Dupret, Sergio Solinas and Giacomo Indiveri, CEA/LIST, France; ETH Zurich and University of Zurich, Switzerland

The increasing need for intelligent sensors in a wide range of everyday objects requires the existence of low power information processing systems which can operate autonomously in their environment. In particular, merging and processing the outputs of different sensors efficiently is a necessary requirement for mobile agents with cognitive abilities. In this work, we present a multi-layer spiking neural network for inference of relations between stimuli patterns in dedicated neuromorphic systems. The system is trained with a new version of the backpropagation algorithm adapted to on-chip learning in neuromorphic hardware: Error gradients are encoded as spike signals which are propagated through symmetric synapses, using the same integrate-and-fire hardware infrastructure as used during forward propagation. We demonstrate the strength of the approach on an arithmetic relation inference task and on visual XOR on the MNIST dataset. Compared to previous, biologically-inspired implementations of networks for learning and inference

of relations, our approach is able to achieve better performance with less neurons. Our architecture is the first spiking neural network architecture with on-chip learning capabilities, which is able to perform relational inference on complex visual stimuli. These features make our system interesting for sensor fusion applications and embedded learning in autonomous neuromorphic agents.

5:50PM *A Spiking Neural Network with Distributed Keypoint Encoding for Robust Sound Recognition [#20001]*

Yanli Yao, Qiang Yu, Longbiao Wang and Jianwu Dang, Tianjin University, China

Compared to traditional artificial neural networks, spiking neural networks (SNNs) operate on an additional dimension of time which makes them more suitable for processing sound signals. However, two of the major challenges in sound recognition with SNNs are neural encoding and learning which demand more research efforts. In this paper, we propose a novel method by combining an improved local time-frequency encoding using key-points detection and biologically plausible tempotron spike learning for robust sound recognition. In the neural encoding part, local energy peaks, called key-points, are firstly extracted from local temporal and spectral regions in the spectrogram. The extracted key-points in each frequency channel are then distributed to multiple sub-channels according to their energy amplitudes with their temporal positions being retained. The resulted spatio-temporal spike patterns are then used as the inputs for spiking neural networks to learn and classify patterns of different categories. We use the RWCP database to evaluate the performance of our proposed system in mismatched environments. Our experimental results highlight that our proposed system, namely DKP-SNN, is effective and reliable for robust sound recognition, resulting in an improved recognition performance as compared to baseline methods.

6:10PM *eSPANNet: Evolving Spike Pattern Association Neural Network for Spike-based Supervised Incremental Learning and Its Application for Single-trial Brain Computer Interfaces [#20017]*

Kaushalya Kumarasinghe, Denise Taylor and Nikola Kasabov, Auckland University of Technology, New Zealand

Objective: Due to the non-stationarity and high trial to trial variability, online event prediction from biomedical signals is challenging. This is more significant when it is applied to neurological rehabilitation where the person incrementally learns to regain the control of movement. eSPANNet is a computational model inspired by the incremental learning for motor control in living nervous systems. It is inspired by the concept of 'population vectors' which have been experimentally proven by several computational neuroscience studies. In this paper, we present a proof-of-concept study on the proposed computational model. Our goal is to utilize the polychronization effect of Spiking Neural Networks to develop a better neural decoder for Brain-Computer Interfaces. Methods: The eSPANNet model contains a network of Spike Pattern Association Neurons, a spiking neuron model which is able to emit spikes at the desired time-point. Results: The proposed approach was experimentally validated using the finger flexion prediction dataset from the fourth BCI competition. The results show that eSPANNet results in 1) a higher classification accuracy, sensitivity and F1 score compared to several other multi-class classifiers and 2) a better approximation of the actual movement compared to several regression analysis based approaches. Conclusion and Significance: The novelty of our algorithm is the ability to learn which inputs to focus on in an online manner. We suggest that eSPANNet is a better BCI decoder due to its i) incremental and life-long learning, ii) compatibility with the neuromorphic platforms and, iii) ability to address the nonstationarity of brain data.

6:30PM *Intelligent Reservoir Generation for Liquid State Machines using Evolutionary Optimization [#19926]*

John J. M. Reynolds, James S. Plank and Catherine D. Schuman, University of Tennessee, Knoxville, United States; Oak Ridge National Laboratory, United States

Neuromorphic Computing is a burgeoning field of research. Many groups are exploring hardware architectures and theoretical ideas about spiking recurrent neural networks. The overarching goal is to exploit the low power promise of these neuromorphic systems. However, it is difficult to train spiking recurrent neural networks (SRNNs) to perform tasks and make efficient use of neuromorphic hardware. Reservoir Computing is an attractive methodology because it requires no tuning of weights for the reservoir itself. Yet, to find optimal reservoirs, manual tuning of hyperparameters such as hidden neurons, synaptic density, and natural structure is still required. Because of this, researchers often have to generate and evaluate many networks, which can result in non-trivial amounts of computation. This paper employs the reservoir computing technique (specifically liquid state machines) and genetic algorithms in order to develop useful networks that can be deployed on neuromorphic hardware. We build on past work in reservoir computing and genetic algorithms to demonstrate the power of combining these two techniques and the advantage it can provide over manually tuning reservoirs for use on classification tasks. We discuss the complexities of determining whether or not to use the genetic algorithms approach for liquid state machine generation.

6:50PM *ECG-based Heartbeat Classification in Neuromorphic Hardware [#19235]*

Federico Corradi, Pande Sandeep, Jan Stuijt, Ning Qiao, Siebren Schaafsma, Giacomo Indiveri and Francky Catthoor, Stichting IMEC Nederland, High Tech Campus 31, Eindhoven 5656 AE, Netherlands; Institute of Neuroinformatics, University of Zurich and ETH Zurich, Switzerland; IMEC Leuven, Kapeldreef 75, 3001 Heverlee, Belgium

Heart activity can be monitored by means of ElectroCardioGram (ECG) measure which is widely used to detect heart diseases due to its non-invasive nature. Trained cardiologists can detect anomalies by visual inspecting recordings of the ECG signals. However, arrhythmias occur intermittently especially in early stages and therefore they can be missed in routine check recordings. We propose a hardware setup that enables the always-on monitoring of ECG signals into wearables. The system exploits a fully event- driven approach for carrying arrhythmia detection and classification employing a bio-inspired Spiking Neural Network. The two staged SNN topology comprises a recurrent network of spiking neurons whose output is classified by a cluster of Leaky-Integrate-and-Fire (LIF) neurons that have been supervisely trained to distinguish 17 types of cardiac patterns. We introduce a method for compressing ECG signals into a stream of asynchronous digital events that are used to stimulate the recurrent SNN. Using ablative analysis, we demonstrate the impact of the recurrent SNN and we show an overall classification accuracy of 95% on the PhysioNet Arrhythmia Database provided by the Massachusetts Institute of Technology and Beth Israel Hospital (MIT/BIH). The proposed system has been implemented on an event-driven mixed-signal analog/digital neuromorphic processor. This work contributes to the realization of an energy-efficient, wearable, and accurate multi-class ECG classification system.

7:10PM *A Modular Approach to Construction of Spiking Neural Networks [#19158]*

Kyunghee Lee and Hongchi Shi, Pyeongtaek University, Korea (South); Texas State University, United States

In this paper, we propose a modular approach to construction of multi-layer spiking neural networks with a Coulomb energy function based learning algorithm for training each module. In this approach, a single-layer spiking neural network is constructed and trained with the Coulomb energy function

based learning algorithm. If the learning result is not sufficiently good, another layer is added, and the input is the output of the previous layer. The process continues until a desired learning result is achieved. The approach eliminates the need for advance determination of the number of hidden layers and the need for error-backpropagation training in multi-layer spiking neural networks. Experimental results of classifying a two-ring-shaped dataset and segmenting an aerial image show that our proposed modular multi-layer spiking neural network requires a simple learning algorithm and achieves better results compared with other approaches.

2a: Supervised learning

Monday, July 15, 5:30PM-7:30PM, Room: Duna Salon II, Chair: Vladimir Cherkassky

5:30PM *Group Learning for High-Dimensional Sparse Data [#20438]*

Vladimir Cherkassky, Hsiang-Han Chen and Han-Tai Shiao, University of Minnesota, Twin Cities, United States

We describe new methodology for supervised learning with sparse data, i.e., when the number of input features is (much) larger than the number of training samples (n). Under the proposed approach, all available (d) input features are split into several (t) subsets, effectively resulting in a larger number (t*n) of labeled training samples in lower-dimensional input space (of dimensionality d/t). This (modified) training data is then used to estimate a classifier for making predictions in lower-dimensional space. In this paper, standard SVM is used for training a classifier. During testing (prediction), a group of t predictions made by SVM classifier needs to be combined via intelligent post-processing rules, in order to make a prediction for a test input (in the original d-dimensional space). The novelty of our approach is in the design and empirical validation of these post-processing rules under Group Learning setting. We demonstrate that such post-processing rules effectively reflect general (common-sense) a priori knowledge (about application data). Specifically, we propose two different post-processing schemes and demonstrate their effectiveness for two real- life application domains, i.e., handwritten digit recognition and seizure prediction from iEEG signal. These empirical results show superior performance of the Group Learning approach for sparse data, under both balanced and unbalanced classification settings

5:50PM *Data complexity measures in feature selection [#19688]*

Lucas Okimoto and Ana Carolina Lorena, Universidade Federal de Sao Paulo, Brazil; Instituto Tecnologico de Aeronautica, Brazil

Feature selection (FS) is a pre-processing step often mandatory in data analysis by Machine Learning techniques. Its objective is to reduce data dimensionality by identifying and maintaining only the relevant features from a dataset. In this work we evaluate the use of complexity measures of classification problems in FS. These descriptors allow estimating the intrinsic difficulty of a classification problem by regarding on characteristics of the dataset available for learning. We propose a combined univariate-multivariate FS technique which employs two complexity measures: Fisher's maximum discriminant ratio and sum of intra-extra class distances. The results reveal that the complexity measures are indeed suitable for estimating feature importance in classification datasets. Large reductions in the numbers of features were obtained, while preserving, in general, the predictive accuracy of two strong classification techniques: Support Vector Machines and Random Forests.

6:10PM *Learning Minority Class prior to Minority Oversampling [#19632]*

Payel Sadhukhan, Indian Statistical Institute Kolkata, India

The success of minority oversampling in dealing with class-imbalanced dataset is well manifested by existing approaches. But that do not guarantee the true class of the synthetic minority points. We address the given context in this paper, Learning Minority Class prior to Minority Oversampling (LMCMO). To guarantee the class information of synthetic minority points, we estimate the minority spaces before generating the synthetic minority points. The performance efficiency of the LMCMO oversampled dataset is tested on C4.5 decision tree and Linear Support Vector Machine (SVM) classifier. Empirical evaluations on 21 datasets using four diversified metrics indicate substantial improvement in Linear SVM outcomes of the proposed method over four competing methods. A modest but still significant gain is achieved by our method over other methods on classification using C4.5 decision tree.

6:30PM *Selective Hypothesis Transfer for Lifelong Learning [#19915]*

Diana Benavides-Prado, Yun Sing Koh and Patricia Riddle, The University of Auckland, New Zealand

Selective transfer has been proposed as an alternative for transferring fragments of knowledge. Previous work showed that transferring selectively from a group of hypotheses helps to speed learning on a target task. Similarly, existing hypotheses could benefit by selective backward transfer of recent knowledge. This setting applies to supervised machine learning systems that observe a sequence of related tasks. We propose a novel scheme for bi-directional transfer between hypotheses learned sequentially using Support Vector Machines. Transfer occurs in two directions: forward and backward. During transfer forward, a new binary classification task is to be learned. Existing knowledge is used to reinforce the importance of subspaces on the target training data that are related to source support vectors. While this target task is learned, subspaces of shared knowledge between each source hypothesis and the target hypothesis are identified. Representations of these subspaces are learned and used to refine the sources by transferring backward. Albeit fundamental, the exploration of the problem of hypothesis refinement has been very limited. We define this problem and propose a solution. Our experiments show that a learning system can gain up to 5.5 units in mean classification accuracy of tasks learned sequentially using our scheme, within 26.6% of the number of iterations when these tasks are learned from scratch.

6:50PM *Are Traditional Neural Networks Well-Calibrated? [#20280]*

Ulf Johansson and Patrick Gabrielsson, Jonkoping University, Sweden; University of Boras, Sweden

Traditional neural networks are generally considered to be well-calibrated. Consequently, the established best practice is to not try to improve the calibration using general techniques like Platt scaling. In this paper, it is demonstrated, using 25 publicly available two-class data sets, that both single multilayer perceptrons and ensembles of multilayer perceptrons in fact often are poorly calibrated. Furthermore, from the experimental results, it is obvious that the calibration can be significantly improved by using either Platt scaling or Venn- Abers predictors. These results stand in sharp contrast to the standard recommendations for the use of neural networks as probabilistic classifiers. The empirical investigation also shows that for bagged ensembles, it is beneficiary to calibrate on the out-of-bag instances, despite the fact that this leads to using substantially smaller ensembles for the predictions. Finally, an outright comparison between Platt scaling and Venn-Abers predictors shows that the latter most often produced significantly better calibrations, especially when calibrated on out-of-bag instances.

7:10PM *Supervised Kernel Transform Learning [#19488]*

Jyoti Maggu and Angshul Majumdar, IIITD, India

This work introduces certain supervised formulations for transform learning. Transform learning is the analysis equivalent of dictionary learning. Four different types of supervision penalties are proposed. The first one is class-sparsity, which imposes common sparse support within representations of each class. The second one imposes similarity among intra-class features in terms of a low-rank constraint (high cosine similarity). The third penalty

enforces features of the same class to be nearby each other and features of different classes to be far apart. The final formulation is the well known label-consistency formulation which learns a linear map from the feature space to the class targets. For the first time, we show how transform learning (and its supervised versions can be kernelized). Finally this work also introduces stochastic regularization techniques like DropOut and DropConnect into the transform learning formulation. Experiments have been carried out on two different problems - computer vision and biomedial signal analysis. In both the problems, our method excels over all existing ones.

2f: Online learning

Monday, July 15, 5:30PM-7:30PM, Room: Duna Salon III, Chair: Pawel Wawrzynski

5:30PM *Efficient on-line learning with diagonal approximation of loss function Hessian [#19186]*

Pawel Wawrzynski, Warsaw University of Technology, Poland

The subject of this paper is stochastic optimization as a tool for on-line learning. New ingredients are introduced to Nesterov's Accelerated Gradient that increase efficiency of this algorithm and determine its parameters that are otherwise tuned manually: step-size and momentum decay factor. In this order a diagonal approximation of the Hessian of the loss function is estimated. In the experimental study the approach is applied to various types of neural networks, deep ones among others.

5:50PM *Pruned Sets for Multi-Label Stream Classification without True Labels [#20346]*

Joel Costa Junior, Elaine Faria, Jonathan Silva, Joao Gama and Ricardo Cerri, Departament of Computer Science - Federal University of Sao Carlos, Brazil; Federal University of Uberlandia, Brazil; Federal University of Mato Grosso do Sul, Brazil; Institute for Systems and Computer Engineering, Technology and Science, Portugal

In multi-label classification problems an example can be simultaneously classified into more than one class. This is also a challenging task in Data Streams (DS) classification, where unbounded and non-stationary distributed multi-label data contain multiple concepts that drift in different rates and patterns. In addition, the true labels of the examples may never become available and updating classification models in a supervised fashion is unfeasible. In this paper, we propose a Multi-Label Stream Classification (MLSC) method applying a Novelty Detection (ND) procedure task to update the classification model detecting any new patterns in the examples, which differ in some aspects from observed patterns, in an unsupervised fashion without any external feedback. Although ND is suitable for multiclass stream classification, it is still a not well-investigated task for multi-label problems. We improve an initial work proposed in [1] and extended it with a new Pruned Sets (PS) transformation strategy. The experiments showed that our method presents competitive performances over data sets with different concept drifts, and outperform, in some aspects, the baseline methods.

6:10PM *Sparse and online null proximal discriminant analysis for one class learning in large-scale datasets [#19819]*

Franck Dufrenois and Denis Hamad, Laboratoire d'Informatique du Signal et des Images de la Cote d'opale, France

Recently, null proximal discriminant analysis (NPDA) has been introduced in [dufrenois et al,2016] to identify anomalies in a target data set. However, the proposed algorithm involves high computational burden and memory requirement which limit its use to moderate sized static datasets. The goal of this paper is to propose an incremental and sparse NPDA which is able both to process sequentially a large volume of data and remove the redundant information. More precisely, our solution lies in computing the null projection direction by a recursive singular value decomposition in the kernel induced feature space. Sparsity of the solution is achieved by thresholding the null score between the already processed data and newly injected data. The combined action of these two steps introduces a significant gain in terms of time complexity and memory burden while offering comparable classification performance as the batch method. Experimental results on large-scale data sets confirm the effectiveness of the proposed algorithm.

6:30PM *Multi-Source Transfer Learning for Non-Stationary Environments [#19525]*

Honghui Du, Leandro Minku and Huiyu Zhou, University of Leicester, United Kingdom; University of Birmingham, United Kingdom

In data stream mining, predictive models typically suffer drops in predictive performance due to concept drift. As enough data representing the new concept must be collected for the new concept to be well learnt, the predictive performance of existing models usually takes some time to recover from concept drift. To speed up recovery from concept drift and improve predictive performance in data stream mining, this work proposes a novel approach called Multi-sourcE onLine TrAnsfer learning for Non-statIonary Environments (Melanie). Melanie is the first approach able to transfer knowledge between multiple data streaming sources in non-stationary environments. It creates several sub-classifiers to learn different aspects from different source and target concepts over time. The sub-classifiers that match the current target concept well are identified, and used to compose an ensemble for predicting examples from the target concept. We evaluate Melanie on several synthetic data streams containing different types of concept drift and on real world data streams. The results indicate that Melanie can deal with a variety drifts and improve predictive performance over existing data stream learning algorithms by making use of multiple sources.

6:50PM *GMM-VRD: A Gaussian Mixture Model for Dealing With Virtual and Real Concept Drifts [#19437]*

Gustavo Oliveira, Leandro Minku and Adriano Oliveira, Centro de Informatica, Brazil; School of Computer Science, United Kingdom

Concept drift is a change in the joint probability distribution of the problem. This term can be subdivided into two types: real drifts that affect the conditional probabilities p(y/x) or virtual drifts that affect the unconditional probability distribution p(x). Most existing work focuses on dealing with real concept drifts. However, virtual drifts can also cause degradation in predictive performance, requiring mechanisms to be tackled. Moreover, as virtual drifts frequently mean that part of the old knowledge remains useful, they require different strategies from real drifts to be effectively tackled. Motivated on this, we propose an approach called Gaussian Mixture Model for Dealing With Virtual and Real Concept Drifts (GMM-VRD), which updates and creates Gaussians to tackle virtual drifts and resets the system to deal with real drifts. The main results show that the proposed approach obtained the best results, in terms of average accuracy, in relation to the literature methods, which propose to solve that same problem. In terms of accuracy over time, the proposed approach showed lower degradation on concept drifts, which indicates that the proposed approach was efficient.

7:10PM *A Discretization-based Ensemble Learning Method for Classification in High-Speed Data Streams [#19585]*

Joao Bertini, University of Campinas, Brazil

Data stream mining has attracted much attention of the machine learning community in the last decade. Motivated by the upcoming issues associated with data stream applications, such as concept drift and the velocity into which data needs to be processed, several methods have been proposed in the literature, most of them resulting from adaptations of traditional

algorithms. Such methods are forced to satisfy hard requirements of restricted memory and processing time, while keeping track of the performance at the same time. In the classification context, ensembles are an effective and elegant way to handle this task. And mostly, the bottleneck of processing time and memory of an ensemble relies on the employed base learner and on the ensemble updating policy. This paper addresses both issues by proposing: 1) a fast base learning algorithm, which relies on discretizing every attribute range into disjoint intervals and associating, to each of them, a posterior probability relating it to a class; and 2) a static ensemble that comprises such base learners and handles concept drift without replacing base learners. Results comparing the proposed ensemble method to six ensemble approaches, on artificial and real data streams, showed it yields comparable results but with lower computational time; which makes the proposed ensemble an efficient alternative to high-speed data streams.

2e: Deep learning

Monday, July 15, 5:30PM-7:30PM, Room: Panorama I, Chair: Yi-Ling Liu

5:30PM *HDL: Hierarchical Deep Learning Model based Human Activity Recognition using Smartphone Sensors [#19656]*

Tongtong Su, Huazhi Sun, Chunmei Ma, Lifen Jiang and Tongtong Xu, School of Computer and Information Engineering, Tianjin Normal University, China

With the development and popularization of smartphones, human activity recognition methods based on contact perception are proposed. The smartphones which are embedded with various sensors can be used as a platform of mobile sensing for human activity recognition. In this paper, we propose an automated human activity recognition network HDL with smartphone motion sensor units. The HDL network combines DBLSTM (Deep Bidirectional Long Short-Term Memory) model and CNN (Convolutional neural network) model. The DBLSTM model is first used to model long sequence data and ultimately generate a bidirectional output vector in a abstract way. The DBLSTM model is good at dealing with serialization tasks but poor in the ability to extract features. Hence, the CNN model is then used to extract features from the abstract vector. Finally,the output layer employs a softmax function to classify human activities. We conduct experiments on the Public domain UCI dataset. The experimental results show that the proposed HDL network achieves reliable results with accuracy and F1 score as high as 97.95% and 97.27%. Compared with other networks based on the same smartphone dataset, the accuracy of HDL is higher than S-LSTM and Dropout CNN network by 2.14% and 6.97% respectively.

5:50PM *An MCTS-based Adversarial Training Method for Image Recognition [#19244]*

Yi-Ling Liu and Alessio Lomuscio, Imperial College London, United Kingdom

We present an adversarial training algorithm based on Monte Carlo Tree Search. We illustrate the robustness of the algorithm by studying its resistance to adversarial examples in the context of the MNIST and CIFAR10 datasets. For MNIST, after 2000 epochs the experimental results showed an average improvement of efficiency of 21.1% when compared to PGD. For CIFAR10, after 7000 epochs we obtained an average improvement of efficiency of 9.8% compared to PGD. We further compare the robustness of the algorithm against previous work against various attack methods. The results suggest that the adversarial training method here introduced is not only robust with respect to adversarial examples but also efficient during training.

6:10PM *A Deep Neural Network Model for Predicting User Behavior on Facebook [#20292]*

Hanen Ameur, Salma Jamoussi and Abdelmajid Ben Hamadou, Multimedia InfoRmation system and Advanced Computing Laboratory, Tunisia

On Facebook, the social media site, liking, commenting and sharing the posts or status of other users are usually considered to be the key mechanisms for exchanging opinions about different topics. Due to the non-availability of data and security constraints, only few research studies have analyzed such behavior. In this paper, we introduced a novel deep neural network model for user behavior prediction (like and comment). We presented an embedding representation method for the textual content of comments and posts based on the contextual recursive auto-encoders model. The users were represented using a deep joint auto-encoders model to fuse the users' like and comment information, and train the users' combined embedding representation. Then, the user behaviors towards a given post were embedded into the same feature space of users and posts, using the joint auto-encoders model. Thereafter, we used a fully connected layer for behavior prediction. To train and evaluate the effectiveness of the proposed method, we also constructed a large dataset collected from Facebook. The experimental results show that the proposed method could achieve better results than the previous alternative methods.

6:30PM *Analyzing Multi-Channel Networks for Gesture Recognition [#19976]*

Pradyumna Narayana, Ross Beveridge and Bruce Draper, Colorado State University, United States

Multi-channel architectures are becoming increasingly common, setting the state-of-the-art for performance in gesture recognition challenges. Unfortunately, we lack a clear explanation of why multi-channel architectures outperform single channel ones. This paper considers two hypotheses. The Bagging hypothesis says that multi-channel architectures succeed because they average the result of multiple unbiased weak estimators in the form of different channels. The Society of Experts (SoE) hypothesis suggests that multi-channel architectures succeed because the channels differentiate themselves, developing expertise with regard to different aspects of the data. To distinguish between these hypotheses, this paper reports on two experiments. The first measures the drop in individual channel performance when the input is degraded by removing high frequency, color, or motion information. The second looks at fusion weights relative to gesture properties. Both experiments support the SoE hypothesis, suggesting multi-channel architectures succeed because of channel specialization.

6:50PM *Image Captioning with Partially Rewarded Imitation Learning [#19336]*

Xintong Yu, Tszhang Guo, Kun Fu, Lei Li, Changshui Zhang and Jianwei Zhang, Tsinghua University, China; University of Hamburg, Germany

Current state-of-the-art image captioning algorithms have achieved great progress via reinforcement learning or generative adversarial nets, with hand-craft metrics such as CIDEr as the reward for the former and signals from adversarial discriminative networks for the latter. Despite the high scores on metrics or improvement in diversity gained from the application of these methods, they suffer from distinction with human-written sentences and drop of ratings on metrics respectively. In this paper, we propose a novel training objective for image captioning that consists of two parts representing explicit and implicit knowledge respectively. Optimizing the new reward partially with imitation learning, we devise an algorithm in which the caption generator is trained to maximize the combination of CIDEr and predictions from adversarial discriminator. Experiments on MSCOCO dataset demonstrate that the proposed method can integrate the strengths of state-of-the- arts, producing more human-like captions while maintaining comparable performance on traditional metrics.

7:10PM *Siamese Deep Dictionary Learning [#19643]*
Vanika Singhal, Angshul Majumdar, Mayank Vatsa and Richa Singh, IIITD, India

Researchers have explored the importance of Siamese networks in deep learning. With recent developments in deep learning and the effectiveness of deep dictionary learning, this research proposes the architecture of Siamese Deep Dictio- nary Learning. We first propose the architecture followed by solving the optimization problem. The experimental effectiveness is demonstrated on five different image databases pertaining to two classification problems: face verification and kinship verification. The experiments show that the proposed Siamese Deep Dictionary Learning yields comparable results compared to state-of-the-art algorithms on all five databases.

8a: Applications of deep networks

Monday, July 15, 5:30PM-7:30PM, Room: Panorama II, Chair: Jacek Mandziuk

5:30PM *DeepIQ: A Human-Inspired AI System for Solving IQ Test Problems [#19108]*
Jacek Mandziuk and Adam Zychowski, Warsaw University of Technology, Poland

This paper presents a neural network approach to solving the most common type of human IQ test problems - Raven's Progressive Matrices (RMs). The proposed DeepIQ system is composed of three modules: a deep autoencoder which is trained to learn a feature-based representation of various figure images used in IQ tests, an ensemble of shallow multilayer perceptrons applied to detection of feature differences, and a scoring module use for assessment of candidate answers. DeepIQ is able to learn the underlying principles of solving RMs (the importance of similarity of figures in shape, rotation, size or shading) in a domain-independent way, that allows its subsequent application to test instances constructed based on a different set of figures, never seen before, or another type of IQ problem, with no requirement for additional training. This transfer learning property is of paramount importance due to scarce availability of the real data, and is demonstrated in the paper on two different RM data sets, as well as two distinct types of IQ tasks (solving RMs and odd-one-out problems). Experimental results are promising, excelling human average scores by a large margin on the most challenging subset of RM instances and exceeding 90% accuracy in odd-one-out tests.

5:50PM *MIDS: End-to-End Personalized Response Generation in Untrimmed Multi-Role Dialogue [#19197]*
QIchuan Yang, Zhiqiang He, Zhiqiang Zhan, Jianyu Zhao, Yang Zhang and Changjian Hu, Beihang University, China; Chinese Academy of Sciences, Beihang University,Lenovo Ltd., China; Chinese Academy of Sciences, China; Lenovo Ltd., China

Multi-role dialogue is a challenging issue in nature language process (NLP), which needs not only to understand the sentences, but also to simulate the interaction among roles. However, existing methods treat all roles' speeches as one sequence and assume that only two speakers take turn to talk, which blurs the characteristics of roles and rarely happen in daily life. To address these issues, we propose a Multi-role Interposition Dialogue System (MIDS) which generates reasonable responses based on dialogue context and next speaker prediction. MIDS employs multiple role-defined encoders to understand each speaker, and an independent sequence model to predict the next speaker. The independent sequence model also works as a scheduler to integrate encoders with weights. Then, an attention-enhanced decoder generates responses based on dialogue context, speaker prediction and integrated encoders. Moreover, with the help of the unique speaker prediction, MIDS is able to generate diverse responses and join conversation actively when appropriate. Experimental results demonstrate that MIDS significantly improves the accuracy of speaker prediction and reduces the perplexity of generation over baselines. Furthermore, MIDS is able to interact with users without cue during real-life online conversations. This work marks a first step towards simulating multi- role dialogue generation.

6:10PM *Cyberthreat Detection from Twitter using Deep Neural Networks [#20231]*
Nuno Dionisio, Fernando Alves, Pedro M. Ferreira and Alysson Bessani, LASIGE, Faculty of Sciences, University of Lisbon, Portugal

--To be prepared against cyberattacks, most organizations resort to security information and event management systems to monitor their infrastructures. These systems depend on the timeliness and relevance of the latest updates, patches and threats provided by cyberthreat intelligence feeds. Open source intelligence platforms, namely social media networks such as Twitter, are capable of aggregating a vast amount of cybersecurity-related sources. To process such information streams, we require scalable and efficient tools capable of identifying and summarizing relevant information for specified assets. This paper presents the processing pipeline of a novel tool that uses deep neural networks to process cybersecurity information received from Twitter. A convolutional neural network identifies tweets containing security-related information relevant to assets in an IT infrastructure. Then, a bidirectional long short-term memory network extracts named entities from these tweets to form a security alert or to fill an indicator of compromise. The proposed pipeline achieves an average 94% true positive rate and 91% true negative rate for the classification task and an average F1-score of 92% for the named entity recognition task, across three case study infrastructures.

6:30PM *Evaluation of a Dual Convolutional Neural Network Architecture for Object-wise Anomaly Detection in Cluttered X-ray Security Imagery [#20461]*
Yona Falinie A. Gaus, Neelanjan Bhowmik, Samet Akcay, Guillen-Garcia Paolo M., Barker Jack W. and Breckon Toby P., Durham University, United Kingdom; Universidad Politecnica de Chiapas, Mexico

X-ray baggage security screening is widely used to maintain aviation and transport secure. Of particular interest is the focus on automated security X-ray analysis for particular classes of object such as electronics, electrical items and liquids. However, manual inspection of such items is challenging when dealing with potentially anomalous items. Here we present a dual convolutional neural network (CNN) architecture for automatic anomaly detection within complex security X-ray imagery. We leverage recent advances in region-based (R-CNN), mask-based CNN (Mask R-CNN) and detection architectures such as Reti- naNet to provide object localisation variants for specific object classes of interest. Subsequently, leveraging a range of established CNN object and fine-grained category classification approaches we formulate within object anomaly detection as a two-class problem (anomalous or benign). Whilst the best performing object localisation method is able to perform with 97.9% mean average precision (mAP) over a six-class X-ray object detection problem, subsequent two-class anomaly/benign classification is able to achieve 66% performance for within object anomaly detection. Overall, this performance illustrates both the challenge and promise of object-wise anomaly detection within the context of cluttered X-ray security imagery.

6:50PM *Single View Distortion Correction using Semantic Guidance [#20269]*
Szabolcs-Botond Lorincz, Szabolcs Pavel and Lehel Csato, Faculty of Mathemetics and Informatics, Babes-Bolyai University of Cluj-Napoca, Romania

Most distortion correction methods focus on simple forms of distortion, such as radial or linear distortions. These works undistort images either based on measurements in the presence of a calibration grid, or use multiple views to find point correspondences and predict distortion parameters. When possible distortions are more complex, e.g. in the case of a camera being placed behind a refractive surface such as glass, the standard method is to use a calibration grid. Considering a high variety of distortions, it is nonviable to conduct these measurements. In this work, we present a single view distortion correction method which is capable of undistorting images containing arbitrarily complex distortions by exploiting recent advancements in differentiable image sampling and in the usage of semantic information to augment various tasks. The results of this work show that our model is able to estimate and correct highly complex distortions, and that incorporating semantic information mitigates the process of image undistortion.

7:10PM *SpreadOut: A Kernel Weight Initializer for Convolutional Neural Networks [#20223]*
Matheus Hertzog, Ricardo Araujo and Ulisses Correa, Federal University of Pelotas, Brazil

Convolutional Neural Networks are based on the hierarchical extraction of features over many layers, by using convolutional filters, or kernels. Each kernel is coded by a set of weights, which represents a feature to be extracted from the input. On a given layer, it is undesirable to have redundant kernels and the training algorithm must learn to differentiate them. SpreadOut is a kernel weight initializer that differentiate kernels before training, so as to improve convergence rates. It does so by maximizing a distance metric calculated over all pairs of kernels on the same layer, after weights are initialized by a traditional technique. We show that SpreadOut improves convergence rates on several benchmark datasets when compared to traditional approaches

1g: Fuzzy Neural Networks

Monday, July 15, 5:30PM-7:30PM, Room: Panorama III, Chair: Jaishri Waghmare

5:30PM *Unbounded Recurrent Fuzzy Min-Max Neural Network for Pattern Classification [#19092]*
Jaishri Waghmare and Uday Kulkarni, SGGS Institute of Engineering and Technology, Nanded, India

A prominent stumbling block in use of the original fuzzy min-max neural network (FMN) and other FMN-based algorithms is their sensitivity to the expansion coefficient or hyperbox expansion parameter. Specifically, these algorithms need to be trained with its different values in the range [0,1], and it is adjusted appropriately to obtain 100 percent accuracy for training data set with a minimum number of hyperboxes. Hence, these algorithms execute multiple passes over the training data for different values of the expansion coefficient. Moreover, usually, its value has to be retuned for the addition of new patterns or classes to the existing classifier. Consequently, as expected these algorithms do not learn in a single pass. This paper proposes an Unbounded Recurrent FMN (URFMN). In the URFMN, several modifications are proposed such as removal of the expansion coefficient parameter, membership functions for different types of hyperboxes, and metamorphosis from feed-forward to the recurrent topology with a novel learning algorithm which works in two phases- offline and online phase. Moreover, during the online phase, it allows the addition of new patterns or classes without the need for retraining. Hence it is an online adaptive algorithm. Its performance is evaluated using the benchmark data sets.

5:50PM *Modulation of Activation Function in Triangular Recurrent Neural Networks for Time Series Modeling [#19682]*
Shyamala Sivakumar and Seshadri Sivakumar, Saint Mary's University, Canada; Pasumai EnergyTech LLC, United States

This paper introduces a novel method to dynamically vary the activation function slopes in recently developed upper and lower triangular recurrent neural networks (ULTRNN) to enhance their modeling capability. The ULTRNN employs a pair of triangular feedback weight matrices with block diagonal elements whose eigenvalues are constrained to lie on or close to the unit circle in the complex z-plane to maintain network and training stability. The activation function slopes of the ULTRNN state variables are dynamically varied by a second modulating network. The inputs to the modulating network are the state variables of the principal ULTRNNs and their inputs. The modulating network is trained simultaneously with the principal ULTRNN to compute the activation function slope for the latter's each state variable at each time step. Such dynamic variation of the activation function slopes selectively enhances the contribution of certain states while suppressing that of the others. A larger slope results in a longer time contribution of the corresponding state and helps model long-term dependencies. Conversely, a smaller slope results in a shorter time contribution and may be used to model controlled "forgetting". The proposed modulation technique enhances the ULTRNN's ability to effectively incorporate short-term memory and long-term dependencies. Simulation results show that with activation function modulation the ULTRNNs are able to autonomously replicate the outputs of sample chaotic dynamic system with good accuracy. This capability can be highly effective in modeling or characterizing the inherent process that generates the time series.

6:10PM *A Neural Field Model for Supervised and Unsupervised Learning of the MNIST Dataset [#19645]*
Michael Brady, AUCA, Kyrgyzstan

A biological model of cortical learning based on neural fields is elaborated and analyzed. In an effort to establish the method as a viable option in today's machine learning world, the method is evaluated based on the MNIST dataset. The approach relates to established work in dynamic systems and neural field models, with an innovation that prevents runaway feedback between neural fields. The method is advanced through four studies. The first three studies illustrate the mechanics of the method by way of supervised learning. Results are compared to analogous results achieved through other methods. The final study builds on the first three studies to illustrate and evaluate how unsupervised learning is accomplished. Concluding discussion considers advantages of the approach over other approaches, and how large neural field networks may be constructed and applied to temporal pattern learning in domains such as robotics.

6:30PM *FigureNet : A Deep Learning model for Question-Answering on Scientific Plots [#19291]*
Revanth Gangi Reddy, Rahul Ramesh, Ameet Deshpande and Mitesh M. Khapra, Indian Institute of Technology, Madras, India

Deep Learning has managed to push boundaries in a wide variety of tasks. One area of interest is to tackle problems in reasoning and understanding, with an aim to emulate human intelligence. In this work, we describe a deep learning model that addresses the reasoning task of question-answering on categorical plots. We introduce a novel architecture FigureNet, that learns to identify various plot elements, quantify the represented values and determine a relative ordering of these statistical values. We test our model on the FigureQA dataset which provides images and accompanying questions for scientific plots like bar graphs and pie charts, augmented with rich annotations. Our approach outperforms the state-of-the-art Relation Networks baseline by approximately 7% on this dataset, with a training time that is over an order of magnitude lesser.

6:50PM *Reconfiguration of Electric Power Distribution Networks using Unineuron and Nullneuron [#20325]*

Mariane Santana, Pyramo Costa, Maury Gouvea and Fabricio Lucas, Pontificia Universidade Catolica de Minas Gerais, Brazil

The reconfiguration of electric power distribution networks is an indispensable tool in the planning and operation of the distribution system. The presented work describes a multi-objective methodology for distribution networks reconfiguration, by using Unineurons and Nullneurons. The new configuration should minimize the total active power loss and minimize feeder loading without violating operational constraints. Tests with the algorithms were performed with IEEE 33 bus electrical system.The obtained results were compared with those from the literature, demonstrating that the developed methodology is promising and provides good results, taking into account all the physical and operational constraints of the problem.

7:10PM *RIT2FIS: A Recurrent Interval Type 2 Fuzzy Inference System and its Rule Base Estimation [#19245]*

Subhrajit Samanta, Andre Hartanto, Mahardhika Pratama, Suresh Sundaram and Narasimalu Srikanth, Nanyang Technological University, Singapore; Indian Institute of Science, Bengaluru, India

Two of the major challenges associated with time series modelling are handling uncertainty present in the data and tracing its dynamical behaviour. A Recurrent Interval Type 2 Fuzzy Inference System or RIT2FIS is presented in this paper. RIT2FIS adopts an interval type 2 fuzzy inference mechanism for superior handling of uncertainty. The memory neurons employed in its hidden and output layer, retain the temporal information, making RIT2FIS highly proficient in tracing system dynamics at a granular level. RIT2FIS also benefits from incorporating a k-means algorithm inspired approach to cluster the data in an unsupervised manner. An 'Elbow Method' is utilized next to determine the optimal clustering which is then employed as the optimal fuzzy rule base for RIT2FIS, eliminating the necessity of expert knowledge for fuzzy initiation. The antecedent and consequent parameters of RIT2FIS are updated using a gradient descent based backpropagation through time algorithm where the learning is made self- regulatory to avoid over-fitting and ensure generalization. Performance of RIT2FIS is evaluated against popular neuro-fuzzy methods on different benchmark and real world time series problems which distinctly indicates an improved accuracy and a parsimonious rule base.

S24: Evolving Machine Learning and Deep Learning Models for Computer Vision

Monday, July 15, 5:30PM-7:30PM, Room: Panorama IV, Chair: Li Zhang

5:30PM *Weather Based Photovoltaic Energy Generation Prediction Using LSTM Networks [#20092]*

Sahar Arshi, Li Zhang and Rebecca Strachan, Faculty of Engineering and Environment University of Northumbria, United Kingdom

Photovoltaic (PV) systems use the sunlight and convert it to electrical power. It is predicted that by 2023, 371,000 PV installations will be embedded in power networks in the UK. This may increase the risk of voltage rise which has adverse impacts on the power network. The balance maintenance is important for high security of the physical electrical systems and the operation economy. Therefore, the prediction of the output of PV systems is of great importance. The output of a PV system highly depends on local environmental conditions. These include sun radiation, temperature, and humidity. In this research, the importance of various weather factors are studied. The weather attributes are subsequently employed for the prediction of the solar panel power generation from a time-series database. Long-Short Term Memory networks are employed for obtaining the dependencies between various elements of the weather conditions and the PV energy metrics. Evaluation results indicate the efficiency of the deep networks for energy generation prediction.

5:50PM *Integrating Social Circles and Network Representation Learning for Item Recommendation [#19943]*

Yonghong Yu, Qiang Wang, Li Zhang, Can Wang, Sifan Wu, Boyu Qi and Xiaotian Wu, Nanjing University of Posts and Telecommunications, China; Northumbria University, United Kingdom; Griffith University, Australia

With the ever increasing popularity of social network services, social network platforms provide rich and additional information for recommendation algorithms. More and more researchers utilize the trust relationships of users to improve the performance of recommendation algorithms. However, most of the existing social-network-based recommendation algorithms ignore the following problems: (1) In different domains, users tend to trust different friends. (2) the performance of recommendation algorithms is limited by the coarse- grained trust relationships. In this paper, we propose a novel recommendation algorithm that integrates the social circles and the network representation learning for item recommendation. Specifically, we firstly infer the domain-specific social trust circles based on the original users' rating information and the social network information. Next, we adopt the network representation technique to embed the domain- specific social trust circle into a low-dimensional space, and then utilize the low-dimensional representations of users to infer the fine-grained trust relationships between users. Finally, we integrate the fine-gained trust relationships with the domain- specific matrix factorization model to learn the latent user and item feature vectors. Experimental results on real-world datasets show that our proposed approach outperforms the traditional social-network-based recommendation algorithms.

6:10PM *Evolving and Ensembling Deep CNN Architectures for Image Classification [#20188]*

Ben Fielding, Tom Lawrence and Li Zhang, Northumbria University, United Kingdom

Deep convolutional neural networks (CNNs) have traditionally been hand-designed owing to the complexity of their construction and the computational requirements of their training. Recently however, there has been an increase in research interest towards automatically designing deep CNNs for specific tasks. Ensembling has been shown to effectively increase the performance of deep CNNs, although usually with a duplication of work and therefore a large increase in computational resources required. In this paper we present Swarm Optimised Block Architecture Ensembles (SOBAE), a method for automatically designing and ensembling deep CNN models with a central weight repository to avoid work duplication. The models are trained and optimised together using particle swarm optimisation (PSO), with architecture convergence encouraged. At the conclusion of this combined process a base model nomination method is used to determine the best candidates for the ensemble. Two base model nomination methods are proposed, one using the local best particle positions from the PSO process, and one using the contents of the central weight repository. Once the base model pool has been created, the models inherit their parameters from the central weight repository and are then finetuned and ensembled in order to create a final system. We evaluate our system on the CIFAR-10 classification dataset and

demonstrate improved results over the single global best model suggested by the optimisation process, with a minor increase in resources required by the finetuning process. Our system achieves an error rate of 4.27% on the CIFAR-10 image classification task with only 36 hours of combined optimisation and training on a single NVIDIA GTX 1080Ti GPU.

6:30PM *Actively Semi-Supervised Deep Rule-based Classifier Applied to Adverse Driving Scenarios [#20197]*
Eduardo Soares, Plamen Angelov, Bruno Costa and Marcos Castro, Lancaster University, United Kingdom; Ford Motor Company, United States

This paper presents an actively semi-supervised multi-layer neuro-fuzzy modeling method, ASSDRB, to classify different lighting conditions for driving scenes. ASSDRB is composed of a massively parallel ensemble of AnYa type 0-order fuzzy rules. It uses a recursive learning algorithm to update its structure when new data items are provided and, therefore, is able to cope with nonstationarities. Different lighting conditions for driving situations are considered in the analysis, which is used by self-driving cars as a safety mechanism. Differently from mainstream Deep Neural Networks approaches, the ASSDRB is able to learn from unseen data. Experiments on different lighting conditions for driving scenes, demonstrated that the deep neuro-fuzzy modeling is an efficient framework for these challenging classification tasks. Classification accuracy is higher than those produced by alternative machine learning methods. The number of algebraic calculations for the present method are significantly smaller and, therefore, the method is significantly faster than common Deep Neural Networks approaches. Moreover, DRB produced transparent AnYa fuzzy rules, which are human interpretable.

6:50PM *Distant Pedestrian Detection in the Wild using Single Shot Detector with Deep Convolutional Generative Adversarial Networks [#20250]*
Ranjith Dinakaran, Li Zhang and Richard Jiang, Computer Science, Northumbria Univ, United Kingdom

In this work, we examine the feasibility of applying Deep Convolutional Generative Adversarial Networks (DCGANs) with Single Shot Detector (SSD) as data-processing technique to handle with the challenge of pedestrian detection in the wild. Specifically, we attempted to use in-fill completion (where a portion of the image is masked) to generate random transformations of images with portions missing to expand existing labelled datasets. In our work, GAN's been trained intensively on low resolution images, in order to neutralize the challenges of the pedestrian detection in the wild, and considered humans, and few other classes for detection in smart cities. The object detector experiment performed by training GAN model along with SSD provided a substantial improvement in the results. This approach presents a very interesting overview in the current state of art on GAN networks for object detection. We used Canadian Institute for Advanced Research (CIFAR), Caltech, KITTI data set for training and testing the network under different resolutions and the experimental results with comparison been showed between DCGAN cascaded with SSD and SSD itself.

7:10PM *Predicting Performance using Approximate State Space Model for Liquid State Machines [#20283]*
Ajinkya Gorad, Vivek Saraswat and Udayan Ganguly, Indian Institute of Technology Bombay, India

Liquid State Machine (LSM) is a brain-inspired architecture used for solving problems like speech recognition and time series prediction. LSM comprises of a randomly connected recurrent network of spiking neurons. This network propagates the non-linear neuronal and synaptic dynamics. Maass et al. have argued that the non-linear dynamics of LSM is essential for its performance as a universal computer. Lyapunov exponent (mu), used to characterize the non-linearity of the network, correlates well with LSM performance. We propose a complementary approach of approximating the LSM dynamics with a linear state space representation. The spike rates from this model are well correlated to the spike rates from LSM. Such equivalence allows the extraction of a memory metric (tau_M) from the state transition matrix. tau_M displays high correlation with performance. Further, high tau_M systems require fewer epochs to achieve a given accuracy. Being computationally cheap (1800x time efficient compared to LSM), the tau_M metric enables exploration of the vast parameter design space. We observe that the performance correlation of the tau_M surpasses that of Lyapunov exponent (mu), (2-4x improvement) in the high-performance regime over multiple datasets. In fact, while mu increases monotonically with network activity, the performance reaches a maxima at a specific activity described in literature as the edge of chaos. On the other hand, tau_M remains correlated with LSM performance. Hence, tau_M captures the useful memory of network activity that enables LSM performance. It also enables rapid design space exploration and fine-tuning of LSM parameters for high performance.

Panel Session: Funding Opportunities in Neural Networks and Biologically Inspired AI Research
Monday, July 15, 5:30PM-7:30PM, Room: Panorama V, Chair: Robert Kozma

Tuesday, July 16, 8:10AM-9:30AM

1l: Deep neural networks, Cellular Computational Networks
Tuesday, July 16, 8:10AM-9:30AM, Room: Ballroom I, Chair: Shiv Ram Dubey

8:10AM *A Performance Evaluation of Convolutional Neural Networks for Face Anti Spoofing [#19041]*
Chaitanya Nagpal and Shiv Ram Dubey, Indian Institute of Information Technology, Sri City, India

In the current era, biometric based access control is becoming more popular due to its simplicity and ease to use by the users. It reduces the manual work of identity recognition and facilitates the automatic processing. The face is one of the most important biometric visual information that can be easily captured without user cooperation in an uncontrolled environment. Precise detection of spoofed faces should be on the high priority to make face based identity recognition and access control robust against possible attacks. The recently evolved Convolutional Neural Network (CNN) based deep learning technique has proven as one of the excellent method to deal with the visual information very effectively. The CNN learns the hierarchical features at intermediate layers automatically from the data. Several CNN based methods such as Inception and ResNet have shown outstanding performance for image classification problem. This paper does a performance evaluation of CNNs for face anti-spoofing. The Inception and ResNet CNN architectures are used in this study. The results are computed over benchmark MSU Mobile Face Spoofing Database. The experiments are done by considering the different aspects such as the depth of the model, random weight initialization vs weight transfer, fine tuning vs training from scratch and different learning rate. The favorable results are obtained using these CNN architectures for face anti-spoofing in different settings.

8:30AM *Convolutional LSTM Network with Hierarchical Attention for Relation Classification in Clinical Texts [#19637]*

Li Tang, Fei Teng, Zheng Ma, Lufei Huang, Ming Xiao and Xuan Li, School of Information Science and Technology, Southwest Jiaotong University, China; The Third People's Hospital of Chengdu, China; School of Electrical Engineering, KTH Royal Institute of Technology, Sweden

Identifying relation from clinical texts is a complex and challenging task due to the specific biomedical knowledge. Existing methods for this work generally have the misclassification problem caused by sample class imbalance. In this paper, we propose a hierarchical attention-based convolutional long short-term memory (ConvLSTM) network model to solve this problem. We construct a sentence as multi- dimensional hierarchical sequence and directly learn local and global context information by a single-layer ConvLSTM network. Besides, a hierarchical attention-based pooling is built to capture the parts of a sentence that are relevant with the target semantic relation. Experiments on the 2010 i2b2/VA relation dataset show that our model outperforms several previous state-of-the- art models without relying on any external features.

8:50AM *Aggregation Connection Network For Tiny Face Detection [#19441]*

Chan Zhang, Tao Li, Song Guo, Ning Li, YingQi Gao and Kai Wang, Nankai University, China

Face detection has been greatly developed in recent years. Despite the remarkable progress, finding tiny faces in the wild is still a challenge due to the vastly scales, blur, occlusion and low resolution. This paper proposes an Aggregation Connection Network (ACN) which robustly solves these problems in tiny face detection. ACN utilizes the features from different convolution layers and performs superiorly on finding multi-scale faces in a single shot, especially for tiny faces. Specially, there are two novel modules in ACN that play significant roles: an aggregation connection module and a context module. First, by integrating efficient aggregation connection module, our ACN can effectively reduce the feature disappearance caused by image scaling. Second, the elaborately designed context module can make full use of the rich contextual cues without adding extra parameters. As a consequence, our ACN achieves state-of-the-art detection performance among several popular face detection benchmarks i.e. WIDER FACE, FDDB and Pascal Face.

9:10AM *Prediction Intervals With LSTM Networks Trained By Joint Supervision [#20262]*

Nicolas Cruz, Luis G Marin and Doris Saez, University of Chile, Chile

This paper presents an approach for prediction interval generation by training a LSTM neural network with a joint supervision Loss Function. The prediction interval model provides the expected value and the upper and lower bounds of the interval given a desired coverage probability. The prediction interval models based on LSTM networks are compared with the classical recurrent neural network approach and are tested using two case studies. The first case corresponds to the forecasting up to one day ahead of the demand profile of 20 dwellings from a town in the UK, and the second case corresponds to the net power from an energy community made up 30 dwellings with a 50% level of photovoltaic power penetration. By using LSTM networks as the backbone of the proposed architecture, high quality intervals are obtained with a narrower interval width compared with the classical recurrent neural network approach. Furthermore, the information provided by the prediction interval based on the LSTM network could be used to develop robust energy management systems that, for example, consider the worst-case scenario.

2e: Deep learning

Tuesday, July 16, 8:10AM-9:30AM, Room: Ballroom II, Chair: Manuel Roveri

8:10AM *Learning a Domain-Invariant Embedding for Unsupervised Person Re-identification [#20150]*

Nan Pu, Theodoros Georgiou, Erwin Bakker and Michael Lew, LIACS Media Lab,Leiden University, Netherlands

Person re-identification (Re-lD) aims at matching images of the same person where images are captured by non-overlapping camera views distributed at different locations. To solve this problem, most recent works require a large pre-labeled dataset for training a deep model. These methods are not always suitable for real- world applications, because the latter often lack labeled data. In order to tackle this drawback, we proposed a novel Domain-Invariant Embedding Network (DIEN) to learn a domain-invariant embedding (DIE) feature by introducing a multi-loss joint learning with Recurrent Top-Down Attention (RTDA) mechanism. Due to the improvement in traditional triplet loss, our proposed model can benefit from both source-domain (labeled) data and target-domain (unlabeled) data. Furthermore, the resulting DIE feature not only has improved class discrimination but also robustness to domain shift. We compared our method with recent competitive algorithms and also evaluated the effectiveness of the proposed modules.

8:30AM *Image Retrieval and Pattern Spotting using Siamese Neural Network [#19876]*

Kelly L. Wiggers, Alceu S. Britto Jr., Laurent Heutte, Alessandro L. Koerich and Luiz S. Oliveira, Pontifical Catholic University of Parana, Brazil; Normandie Univ, France; Ecole de Technologie Superieure, Canada; Federal University of Parana, Brazil

This paper presents a novel approach for image retrieval and pattern spotting in document image collections. The manual feature engineering is avoided by learning a similarity- based representation using a Siamese Neural Network trained on a previously prepared subset of image pairs from the Im- ageNet dataset. The learned representation is used to provide the similarity-based feature maps used to find relevant image candidates in the data collection given an image query. A robust experimental protocol based on the public Tobacco800 document image collection shows that the proposed method compares favor- ably against state-of-the-art document image retrieval methods, reaching 0.94 and 0.83 of mean average precision (mAP) for retrieval and pattern spotting (IoU=0.7), respectively. Besides, we have evaluated the proposed method considering feature maps of different sizes, showing the impact of reducing the number of features in the retrieval performance and time-consuming.

8:50AM *Abstractive Text Summarization with Multi-Head Attention [#19655]*

Jinpeng Li, Chuang Zhang, Xiaojun Chen, Yanan Cao, Pengcheng Liao and Peng Zhang, Institute of Information Engineering, Chinese Academy of Sciences. School of Cyber Security, University of Chinese Academy of Sciences, China; Institute of Information Engineering, Chinese Academy of Sciences, China

In this paper, we present a novel sequence-to-sequence architecture with multi-head attention for automatic summarization of long text. Summaries generated by previous abstractive methods have the problems of duplicate and missing original information commonly. To address these problems, we propose a multi-head attention summarization (MHAS) model, which uses multi-head attention mechanism to learn relevant information in different representation subspaces. The MHAS model can consider the previously predicted words when generating new words to avoid generating a summary of redundant repetition words. And it can learn the internal structure of the article by adding self-attention layer to the traditional encoder and decoder and make the model better preserve the original information. We also integrate the multi-head attention distribution into pointer network creatively to improve the performance of the model. Experiments are conducted on CNN/Daily Mail dataset, which is a long text English corpora. Experimental results show that our proposed model outperforms the previous extractive and abstractive models.

9:10AM *Learning Convolutional Neural Networks in presence of Concept Drift [#20303]*

Simone Disabato and Manuel Roveri, Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Italy

Designing adaptive machine learning systems able to operate in nonstationary conditions, also called concept drift, is a novel and promising research area. Convolutional Neural Networks (CNNs) have not been considered a viable solution for such adaptive systems due to the high computational load and the high number of images they require for the training. This paper introduces an adaptive mechanism for learning CNNs able to operate in presence of concept drift. Such an adaptive mechanism follows an ``active approach'', where the adaptation is triggered by the detection of a concept drift, and relies on the ``transfer learning'' paradigm to transfer (part of the) knowledge from the CNN operating before the concept drift to the one operating after. The effectiveness of the proposed solution has been evaluated on two types of CNNs and two real-world image benchmarks.

8a: Applications of deep networks

Tuesday, July 16, 8:10AM-9:30AM, Room: Ballroom III, Chair: Binyi Yin

8:10AM *Face Attribute Prediction in Live Video using Fusion of Features and Deep Neural Networks [#19703]*

Sudarsini Tekkam Gnanasekar and Svetlana Yanushkevich, University of Calgary, Canada

Face attribute analysis from live video is a valuable aide in biometric-based person identification. This is a challenging task due to variations in lighting, occlusion, pose and other variables. To address it, we propose an effective and robust approach: extract the face features using certain selected layers of the pre-trained Convolutional Neural Network (CNN) models such as AlexNet, GoogleNet and ResNet50. We focus on the intermediate CNN layers, since the reported experimental results suggest that the best results may not always be obtained when extracting deep features using the fully connected layers. Next, we train a linear SVM on the extracted features to perform the attribute classification. We also apply a feature level fusion by concatenating the features extracted from the intermediate layers of the aforementioned networks. Our approach applied on live video achieves an average accuracy of 89.40% using the fused features which is better than the results (between 86.6% and 87%) reported for the CNNs applied only on static images.

8:30AM *On the Influence of the Color Model for Image Boundary Detection Algorithms based on Convolutional Neural Networks [#19565]*

Tiago Jose dos Santos, Carlos Alexandre Barros de Mello, Cleber Zanchettin and Thiago Vinicius Machado de Souza, Universidade Federal de Pernambuco, Brazil

Image analysis and understanding are challenging tasks, usually having segmentation as a major step. Boundary detection is a type of segmentation which aims to highlight the boundaries of the objects in a scene. Models based on Convolutional Neural Networks (CNN) have presented promising results for boundary detection, where the input usually is the entire image or some patches, often described in the RGB color model. In this paper, we provide a qualitative analysis of boundary detection algorithms based on CNN but considering images in different color models. We have used the color models RGB, Lab, Luv, dRdGdB, YO1O2 and HSV for this analysis. The Holistically-Nested Edge Detection (HED) and Convolutional Encoder Decoder Network (CEDN) are the CNN's chosen due to their high performance. The benchmark BSDS is the boundary detection evaluator. Experiments show that the results of the edge detection process tend to be similar when training the CNN with weights randomly initialized, regardless of the color model used. For the HED architecture, the use of Lab and Luv color models has resulted in a significant improvement to the case of transfer learning and fine-tuning of weights.

8:50AM *Context-Aware Network for 3D Human Pose Estimation from Monocular RGB Image [#20270]*

Binyi Yin, Dongbo Zhang, Shuai Li, Aimin Hao and Hong Qin, Beihang University, China; Stony Brook University, United States

Convolutional Neural Network (CNN) has brought tremendous improvements in estimating 3D human pose from a monocular RGB image. However, the task of 3D human pose estimation still remains extremely challenging, especially when the task is geared towards estimating the depth of human body parts. Different from 2D human pose estimation, which focuses on the fusion of spatial information and context information, depth estimation demands more context information. Inspired by this, we build a Context-Aware Network (CAN) which can fully explore the context information to discover the underlying relationships among different body parts. The key ingredient of our network is High-Level Depth Estimation Module (HLDEM) designed to extract context information effectively. Additionally, multi-scale supervision is introduced in our network to extract context information at different scales. Experimental results show that our network achieves competitive performance compared with state-of-the-art methods on Human3.6M dataset.

9:10AM *Music Artist Classification with Convolutional Recurrent Neural Networks [#19893]*

Zain Nasrullah and Yue Zhao, Department of Computer Science, University of Toronto, Canada

Previous attempts at music artist classification use frame level audio features which summarize frequency content within short intervals of time. Comparatively, more recent music information retrieval tasks take advantage of temporal structure in audio spectrograms using deep convolutional and recurrent models. This paper revisits artist classification with this new framework and empirically explores the impacts of incorporating temporal structure in the feature representation. To this end, an established classification architecture, a Convolutional Recurrent Neural Network (CRNN), is applied to the artist20 music artist identification dataset under a

comprehensive set of conditions. These include audio clip length, which is a novel contribution in this work, and previously identified considerations such as dataset split and feature level. Our results improve upon baseline works, verify the influence of the producer effect on classification performance and demonstrate the trade-offs between audio length and training set size. The best performing model achieves an average F1 score of 0.937 across three independent trials which is a substantial improvement over the corresponding baseline under similar conditions. Additionally, to showcase the effectiveness of the CRNN's feature extraction capabilities, we visualize audio samples at the model's bottleneck layer demonstrating that learned representations segment into clusters belonging to their respective artists.

2c: Reinforcement learning and adaptive dynamic programming

Tuesday, July 16, 8:10AM-9:30AM, Room: Duna Salon I, Chair: Samuele Tosatto

8:10AM *Adversarial Imitation Learning via Random Search [#19367]*

MyungJae Shin and Joongheon Kim, Chung-Ang University, Korea (South)

Developing agents that can perform challenging complex tasks is the goal of reinforcement learning. The model-free reinforcement learning has been considered as a feasible solution. However, the state of the art research has been to develop increasingly complicated techniques. This increasing complexity makes the reconstruction difficult. Furthermore, the problem of reward dependency is still exists. As a result, research on imitation learning, which learns policy from a demonstration of experts, has begun to attract attention. Imitation learning directly learns policy based on data on the behavior of the experts without the explicit reward signal provided by the environment. However, imitation learning tries to optimize policies based on deep reinforcement learning such as trust region policy optimization. As a result, deep reinforcement learning based imitation learning also poses a crisis of reproducibility. The issue of complex model-free model has received considerable critical attention. A derivative-free optimization based reinforcement learning and the simplification on policies obtain competitive performance on the dynamic complex tasks. The simplified policies and derivative free methods make algorithm be simple. The reconfiguration of research demo becomes easy. In this paper, we propose an imitation learning method that takes advantage of the derivative-free optimization with simple linear policies. The proposed method performs simple random search in the parameter space of policies and shows computational efficiency. Experiments in this paper show that the proposed model, without a direct reward signal from the environment, obtains competitive performance on the MuJoCo locomotion tasks.

8:30AM *Accelerating the Deep Reinforcement Learning with Neural Network Compression [#19150]*

Hongjie Zhang, Zhuocheng He and Jing Li, University of Science and Technology of China, China

Acceleration in Deep Reinforcement Learning has attracted the attention of researchers. Many parallel training frameworks have been proposed to speed up the sampling and the training by running multiple agents simultaneously. However, the bottleneck of performance in a single agent is the neural network prediction which takes a lot of time as compared to the rapid environmental changes. Different from these parallel frameworks, we try to speed up the prediction of the agent to accelerate the entire training process. As far as we know, this is the first time to accelerate the training from this perspective. We propose a novel training framework NNC-DRL to accelerate the whole training process of deep reinforcement learning. NNC-DRL uses different neural networks to represent target policy and behavior policy. The behavior policy network is a smaller neural network derived from the target policy network, which could speed up the prediction. The target policy network will transfer its latest policy to the behavior policy network. The inconsistent policy distribution between behavior network and target network will degrade the convergence of the training. In NNC-DRL, we introduce the Important Sampling technique to estimate policy gradient, which could improve the convergence of the training. The experiments show that our approach NNC-DRL can speed up the whole training process by about 10-20% on Atari 2600 games with little performance loss.

8:50AM *Exploration Driven By an Optimistic Bellman Equation [#19157]*

Samuele Tosatto, Carlo D'Eramo, Joni Pajarinen, Marcello Restelli and Jan Peters, Technische Universitaet Darmstadt, Germany; Politecnico di Milano, Italy

Exploring high-dimensional state spaces and finding sparse rewards are central problems in reinforcement learning. Exploration strategies are frequently either naive (e.g., simplistic epsilon-greedy or Boltzmann policies), intractable (i.e., full Bayesian treatment of reinforcement learning) or rely heavily on heuristics. The lack of a tractable but principled exploration approach unnecessarily complicates the application of reinforcement learning to a broader range of problems. Efficient exploration can be accomplished by relying on the uncertainty of the state-action value function. To obtain the uncertainty, we maintain an ensemble of value function estimates and present an optimistic Bellman equation (OBE) for such ensembles. This OBE is derived from a relative entropy maximization principle and yields an implicit exploration bonus resulting in improved exploration during action selection. The implied exploration bonus can be seen as a well-principled type of intrinsic motivation and exhibits favorable theoretical properties. OBE can be applied to a wide range of algorithms. We propose two algorithms as an application of the principle: Optimistic Q-learning and Optimistic DQN which outperform comparison methods on standard benchmarks.

9:10AM *Event-triggered Adaptive Control for Discrete-Time Zero-Sum Games [#19578]*

Ziyang Wang, Qinglai Wei, Derong Liu and Yanhong Luo, University of Science and Technology Beijing, China; Chinese Academy of Sciences, China; Guangdong University of Technology, China; Northeastern University, China

In this paper, an event-triggered adaptive dynamic programming (ADP) method is developed for the discrete-time nonlinear two-player zero-sum games. First, an event-triggered ADP algorithm is presented to solve the Hamilton-Jacobi-Isaacs (HJI) equation. Then, a novel double event-triggered scheme is designed, the control inputs and the disturbance inputs will be updated only when the triggering conditions are satisfied. Therefore, the computational burden and the communication cost can be reduced. The algorithm is implemented by two neural networks, and the stability of the two-player system is proved. Finally, an example is employed to illustrate the effectiveness of the developed method.

2d: Semi-supervised learning

Tuesday, July 16, 8:10AM-9:30AM, Room: Duna Salon II, Chair: Suely Oliveira

8:10AM *Automatic Image Annotation based on Co-Training [#19139]*

Zhixin Li, Lan Lin, Canlong Zhang, Huifang Ma and Weizhong Zhao, Guangxi Normal University, China; Northwest Normal University, China; Central China Normal University, China

To learn a well-performed image annotation model, a large number of labeled samples are usually required. Although the unlabeled samples are readily available and abundant, it's a difficult task for humans to annotate large amounts of images manually. In this paper, we propose a novel semi-supervised approach based on co-training algorithm for automatic image annotation, which can utilize the labeled data and unlabeled data for the system simultaneously. Firstly, two different classifiers, namely the CNN (convolutional neural network) and the LDA-SVM, are constructed by all the labeled data. These two classifiers are in-dependently represented as different feature views. Then, the most confident data with relevant pseudo-labels are chosen and amalgamated with the whole labeled dataset. After that, the two classifiers are retrained with the new labeled dataset until a stop condition is reached. In each iteration process, the unlabeled samples are labeled by high confidence pseudo-labels that are estimated by an adaptive weighted fusion method. Finally, we conduct experiments on two datasets, namely, IAPR TC-2 and NUS-WIDE, and measure the performance of the model with standard criteria, including precision, recall, F- measure, N+ and mAP. The experimental results show that our approach has superior annotation performance and outperforms many state-of-the-art automatic image annotation approaches.

8:30AM *Metric Learning based Framework for Streaming Classification with Concept Evolution [#20213]*

Zhuoyi Wang, Hemeng Tao, Kong Zelun, Swarup Chandra and Latifur Khan, University of Texas at Dallas, United States

A primary challenge in label prediction over a stream of continuously occurring data instances is the emergence of instances belonging to unknown or novel classes. It is imperative to detect such novel-class instances quickly along the stream for a superior prediction performance. Existing techniques that perform novel class detection typically employ a clustering-based mechanism by observing that instances belonging to the same class (intra- class) are closer to each other (cohesion) than inter-class samples (separation). While this is generally true in low dimensional feature spaces, We observe that such a property is not intrinsic among instances in complex real-world high-dimensional feature space such as images and text. In this paper, we focus on addressing this key challenge that negatively affects prediction performance of a data stream classifier. Concretely, we develop a metric learning mechanism that transforms high-dimensional features into a latent feature space to make above property holds true. Unlike existing metric learning method which only focus on classification task, our approach address the novel class detection and stream classification simultaneously. We showcase a framework along the stream to achieve larger prediction performance compared to existing state-of-the-art detection techniques while using the least amount of labeled data during detection. Extensive experimental results on simulated and real-world stream demonstrate the effectiveness of our approach.

8:50AM *Interpretable Variational Autoencoders for Cognitive Models [#20248]*

Mariana Curi, Geoffrey Converse, Jeff Hajewski and Suely Oliveira, Unversity of Sao Paulo, Brazil; The University of Iowa, United States

One of the most used methodologies in the field of education assessment is Item Response Theory (IRT). In this work, we propose the use of a novel Variational Autoencoder (VAE) architecture for a multidimensional IRT model. Our approach combines the advantages of the IRT model while allowing us to model high latent trait dimensions, previously unattainable in prior work. Additionally, it has the advantage of interpretability in the domain of educational assessment. Our experiments show that, given enough data, the new model is competitive with the state-of-the-art methods with respect to predictive power and is much faster in runtime performance. In our experiments, we achieve competitive results on a sample size 20X larger in a runtime that is 40X faster than the state-of-the-art model.

S07: Advanced Machine Learning Methods for Big Graph Analytics

Tuesday, July 16, 8:10AM-9:30AM, Room: Duna Salon III, Chair: Shirui Pan

8:10AM *Feature-Dependent Graph Convolutional Autoencoders with Adversarial Training Methods [#19801]*

Di Wu, Ruiqi Hu, Yu Zheng, Jing Jiang, Nabin Sharma and Michael Blumenstein, University of Technology Sydney, Australia; Northwest A&F University, China

Graphs are ubiquitous for describing and modeling complicated data structures, and graph embedding is an effective solution to learn a mapping from a graph to a low-dimensional vector space while preserving relevant graph characteristics. Most existing graph embedding approaches either embed the topological information and node features separately or learn one regularized embedding with both sources of information, however, they mostly overlook the interdependency between structural characteristics and node features when processing the graph data into the models. Moreover, existing methods only reconstruct the structural characteristics, which are unable to fully leverage the interaction between the topology and the features associated with its nodes during the encoding-decoding procedure. To address the problem, we propose a framework using autoencoder for graph embedding (GED) and its variational version (VEGD). The contribution of our work is two-fold: 1) the proposed frameworks exploit a feature-dependent graph matrix (FGM) to naturally merge the structural characteristics and node features according to their interdependency; and 2) the Graph Convolutional Network (GCN) decoder of the proposed framework reconstructs both structural characteristics and node features, which naturally possesses the interaction between these two sources of information while learning the embedding. We conducted the experiments on three real-world graph datasets such as Cora, Citeseer and PubMed to evaluate our framework and algorithms, and the results outperform baseline methods on both link prediction and graph clustering tasks.

8:30AM *Community Detection with Indirect Neighbors based on Granular Computing in Social Networks [#19670]*

Naiyue Chen, Jie He, Xiang Wang, Zhiyuan Zhang, Ping Yang and Yanping Fu, School of Computer and Information Technology, Beijing Jiaotong University, China; CETC Big Data Research Institute Co.,Ltd., China; Signal and Communication Research Institute, China Academy of Railway Sciences, China; School of Electronic and Information Engineering, Beijing Jiaotong University,, China

Social relations exist in real life widely, which can be abstracted into graphs to show various social networks. Community detection in social networks has become an important and effective methodology to understand the structure and function of real world networks. In this paper, the algorithm we proposed fully considers the influence of direct nodes and indirect nodes and expresses the relationship between users comprehensively. The algorithm runs on the social network graph based on a new granular computing method. Experimental results on bench mark data show the superiority of the proposed algorithm compared to other well-known methods.

8:50AM *Deep Structure Learning for Rumor Detection on Twitter [#20148]*

Qi Huang, Chuan Zhou, Jia Wu, Mingwen Wang and Bin Wang, Institute of Information Engineering,Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences;, China; Institute of Information Engineering, Chinese Academy of Sciences, China; Department of Computing, Faculty of Science and Engineering, Macquarie University, Australia; School of Computer and Information Engineering, Jiangxi Normal University, China; Xiaomi AI Lab, China

With the development of social media and the popularity of mobile devices, it becomes increasingly easy to post rumors and spread rumors on social media. Widespread rumors may cause public panic and negative impact on individuals, which makes the automatic detection of rumors become necessary. Most existing methods for automatic rumor detection focus on modeling features related to contents, users and propagation patterns based on feature engineering, but few work consider the existence of graph structural information in the user behavior. In this paper, we propose a model that leverages graph convolutional networks to capture user behavior effectively for rumor detection. Our model is composed of three modules: 1) a user encoder that models user's attributes and behaviors based on graph convolutional networks to obtain user representation; 2) a propagation tree encoder, which encodes the structure of the rumor propagation tree as a vector with bridging the content semantics and propagation clues; 3) an integrator that integrates the output of the above modules to identify rumors. Experimental results on two public Twitter datasets show that our model achieves much better performance than the state-of-the-art methods.

9:10AM *Beyond the Power of Mere Repetition: Forms of Social Communication on Twitter through the Lens of Information Flows and Its Effect on Topic Evolution [#19284]*

Yunwei Zhao, Can Wang, Chi-Hung Chi, Willem-Jan van den Heuvel, Kwok-Yan Lam and Min Shu, CN-CERT, China; Griffith University, Australia; CSIRO, Australia; Tilburg University, Netherlands; Nanyang Technological University, Singapore

Understanding how people interact and exchange messages on social networks is significant for managing online contents and making predictions of future behaviors. Most existing research on the communication characteristics simply focuses on the user involvement. The current work largely neglects the content changes that imply how wide and deep the discussion in a topic goes, and to what degree people set forth their own views with the additional information supplemented. We are highly motivated to propose a theoretical framework to target those issues. In this paper, we define the communication modality constructs, and classify topics based on three dimensions: user involvement, information flow depth, and topic inter-relations, which substantially extend the traditional focus in user interaction analysis. The communication modality constructs comprise of (i) topic dialogicity, (ii) discussion intensiveness, and (iii) discussion extensibility. We introduce a quantitative model based on the topology of information flow graph, and use the information addition as well as the emotion attachment along the path to measure the pattern divergence between topic groups. Our model is empirically validated by using 78 million tweets, and experiments on Twitter demonstrate our contributions.

Neural Network Models

Tuesday, July 16, 8:10AM-9:30AM, Room: Panorama I, Chair: Zhihuan Yan

8:10AM *A Preprocessing Layer in Spiking Neural Networks - Structure, Parameters, Performance Criteria [#19450]*

Mikhail Kiselev and Andrey Lavrentyev, Chuvash State University, Russian Federation; Kaspersky Lab, Russian Federation

The subject of this article is the section of spiking neural network (SNN) closest to the source of input signals. The purpose of this layer is to provide input signal preprocessing and the extraction of primary informative features, promoting its higher-level analysis by subsequent SNN layers. The 1-layer and 2- layer WTA (winner takes all) architectures of this section are explored. Two independent criteria used to evaluate its necessity and efficiency in the given task are discussed. A variant of the STDP plasticity rule specially designed for the WTA preprocessing layer is described. An optimization procedure based on genetic algorithm and practical recommendations for its implementation are also included in the paper. The article contains three practical examples similar to real-world problems, which serve as illustrations for ideas presented in this work.

8:30AM *Evaluating the Stability of Recurrent Neural Models during Training with Eigenvalue Spectra Analysis [#20512]*

Priyadarshini Panda, Efstathia Soufleri and Kaushik Roy, Purdue University, United States

We analyze the stability of recurrent networks, specifically, reservoir computing models during training by evaluating the eigenvalue spectra of the reservoir dynamics. To circumvent the instability arising in examining a closed loop reservoir system with feedback, we propose to break the closed loop system. Essentially, we unroll the reservoir dynamics over time while incorporating the feedback effects that preserve the overall temporal integrity of the system. We evaluate our methodology for fixed point and time varying targets with least squares regression and FORCE training, respectively. Our analysis establishes eigenvalue spectra (which is, shrinking of spectral circle as training progresses) as a valid and effective metric to gauge the convergence of training as well as the convergence of the chaotic activity of the reservoir toward stable states.

8:50AM *Enhance knowledge graph embedding via fake triples [#19226]*

Zhihuan Yan, Rong Peng, Yaqian Wang and Weidong Li, Wuhan University, China

Embedding knowledge graphs (KGs) into continuous vector spaces is a focus of current research. Although previous works achieve great success, most of them are based on the closed-world assumption, where they only utilize knowledge explicitly exist in KGs but ignore implicit knowledge. This paper tries to improve the performance of KG embedding models by adding implicit knowledge to it. We consider that when an entity occurs only in head or tail of all triples in one KG, its connectivity with other entities is single and the inverse triples of it can provide as implicit knowledge to enrich its connectivity, and with this implicit knowledge KG embedding models can learn more accurate semantic of entities and relations in the KG. For this purpose, we introduce 'fake triples' (triples that theoretically should exist but not appear in knowledge graph explicitly) via dummy relations for zero in-degree and zero out-degree entities to enrich their connectivity and further improve the embedding models' performance. Extensive experiments on entity alignment task and linking predication task show that our approach achieves good results. On entity alignment task, the entity alignment model with fake triples (EA+F) obtains better results than a number of state-of-the-art entity alignment models. On linking predication task, our method achieves better Mean Rank on FB15K. Our work can prove that the performance of knowledge graph embedding may be promoted by elaborate analysis on dataset rather than designing complex models.

9:10AM *Neural Network Based Inverse System Identification from Small Data Sets [#19026]*
Chathura Wanigasekara, Akshya Swain, Sing Kiong Nguang and B. Gangadhara Prusty, The University of Auckland, New Zealand; University of New South Wales, Australia

Many applications in control, signal processing and manufacturing require the inverse model of a system. Identification of the inverse of a system (inverse modelling) is often an ill-posed problem and therefore a challenging task. The learning capability of the artificial neural network (ANN) has been exploited in the past to identify the inverse of a system. However, in certain applications, such as manufacturing, where the available data samples are less, the complexity of fitting an inverse model increases significantly. This results in small data learning problem. The present study solves this small data learning problem from the perspective of the inverse system identification using neural networks. Initially, the effectiveness of different combinations of various virtual sample generation (VSG) methods and machine learning tools are investigated to determine the optimum combination which gives the highest learning accuracy. Simulation results are included to demonstrate the effectiveness of ANN in identifying the inverse of various systems from small data.

2d: Semi-supervised learning

Tuesday, July 16, 8:10AM-9:30AM, Room: Panorama II, Chair: Min Peng

8:10AM *A Data Stratification Process for Instances Selection in Semi-Supervised Learning [#19684]*
Karliane M. O. Vale, Anne Magaly de P. Canuto, Cainan T. Alves, Arthur C. Gorgonio, Flavius L. Gorgonio, Amarildo J. F. Lucena and Araken M. Santos, Federal University of Rio Grande do Norte (UFRN), Brazil; Federal Rural University of Semi-Arido (UFERSA), Brazil

This paper presents a study in the field of semi-supervised learning and, more specifically, it proposes changes in the self-training algorithm in order to apply a data stratification method in the labeling process of this algorithm. Therefore, this work proposes a method, called FlexCon-CS, whose objective is to apply data stratification in the inclusion of new instances in the training data set. In this sense, the representativeness and class distribution will be maintained throughout the labeling process, with the same proportions of the initially labeled dataset. In order to evaluate this proposal, we performed experiments on 27 databases with different data distribution features. Each dataset was trained with four different classification algorithms, Naive Bayes, Decision Tree, ripper, and K-Nearest Neighbor classifiers. Moreover, the Friedman statistical test was applied to provide a statistically significant analysis of the obtained results. Our findings indicate that, in most cases, the proposed methods perform better than the original self-training method.

8:30AM *Unsupervised Domain Adaptation using Graph Transduction Games [#20296]*
Sebastiano Vascon, Sinem Aslan, Alessandro Torcinovich, Twan van Laarhoven, Elena Marchiori and Marcello Pelillo, Ca' Foscari University of Venice, Italy; Open University of the Netherlands, Netherlands; Radboud University Nijmegen, Netherlands

Unsupervised domain adaptation (UDA) amounts to assigning class labels to the unlabeled instances of a dataset from a target domain, using labeled instances of a dataset from a related source domain. In this paper, we propose to cast this problem in a game-theoretic setting as a non-cooperative game and introduce a fully automatized iterative algorithm for UDA based on graph transduction games (GTG). The main advantages of this approach are its principled foundation, guaranteed termination of the iterative algorithms to a Nash equilibrium (which corresponds to a consistent labeling condition) and soft labels quantifying the uncertainty of the label assignment process. We also investigate the beneficial effect of using pseudo-labels from linear classifiers to initialize the iterative process. The performance of the resulting methods is assessed on publicly available object recognition benchmark datasets involving both shallow and deep features. Results of experiments demonstrate the suitability of the proposed game-theoretic approach for solving UDA tasks.

8:50AM *Discriminative Regularization with Conditional Generative Adversarial Nets for Semi-Supervised Learning [#19317]*
Qianqian Xie, Min Peng, Jimin Huang, Bin Wang and Hua Wang, School of Computer Science, Wuhan University, China; Computer Science, Wuhan University, China; Xiaomi Incorporation, China; Victoria University, Australia

Existing generative adversarial networks (GANs) with manifold regularization for semi-supervised learning (SSL) have shown promising performance in image generation and semi-supervised learning (SSL), which penalize the smoothness of classifier over data manifold based on the smoothness assumption. However, the smoothness assumption is valid for data points in high density region while not hold for data points in low density region, thus they tend to misclassify boundary instances in low density region. In this paper, we propose a novel discriminative regularization method for semi-supervised learning with conditional generative adversarial nets (CGANs). In our method, the discriminative information from class conditional data distribution captured by CGANs is utilized to improve the discrimination of classifier. Different from regular manifold regularization, the discriminative regularization encourages the classifier invariance to local perturbations on the sub-manifold of each cluster, and distinct classification outputs for data points in different clusters. Moreover, our method can be easily implemented via the stochastic approximation without constructing the Laplacian graph or computing the Jacobian of classifier explicitly. Experimental results on benchmark datasets show that our method can achieve competitive performance against previous advanced methods.

9:10AM *Lifting 2d Human Pose to 3d : A Weakly Supervised Approach [#20454]*
Sandika Biswas, Sanjana Sinha, Kavya Gupta and Brojeshwar Bhowmick, TCS Research, Tata Consultancy Services, India

Estimating 3d human pose from monocular images is a challenging problem due to the variety and complexity of human poses and the inherent ambiguity in recovering depth from the single view. Recent deep learning based methods show promising results by using supervised learning on 3d pose annotated datasets. However, the lack of large-scale 3d annotated training data captured under in-the-wild settings makes the 3d pose estimation difficult for in-the-wild poses. Few approaches have utilized training images from both 3d and 2d pose datasets in a weakly-supervised manner for learning 3d poses in unconstrained settings. In this paper, we propose a method which can effectively predict 3d human pose from 2d pose using a deep neural network trained in a weakly-supervised manner on a combination of ground-truth 3d pose and ground- truth 2d pose. Our method uses re-projection error minimization as a constraint to predict the 3d locations of body joints, and this is crucial for training on data where the 3d ground-truth is not present. Since minimizing re-projection error alone may not guarantee an accurate 3d pose, we also use additional geometric constraints on skeleton pose to regularize the pose in 3d. We demonstrate the superior generalization ability of our method by cross-dataset validation

on a challenging 3d benchmark dataset MPI-INF-3DHP containing in the wild 3d poses.

1l: Deep neural networks, Cellular Computational Networks

Tuesday, July 16, 8:10AM-9:30AM, Room: Panorama III, Chair: Asim Iqbal

8:10AM *Decoding Neural Responses in Mouse Visual Cortex through a Deep Neural Network [#19491]*

Asim Iqbal, Phil Dong, Christopher Kim and Heeun Jang, UZH/ETH Zurich, Switzerland; Icahn School of Medicine at Mount Sinai, United States; National Institutes of Health, United States; Buck Institute for Research on Aging, United States

Finding a code to unravel the population of neural responses that leads to a distinct animal behavior has been a long- standing question in the field of neuroscience. With the recent advances in machine learning, it is shown that the hierarchically Deep Neural Networks (DNNs) perform optimally in decoding unique features out of complex datasets. In this study, we utilize the power of a DNN to explore the computational principles in the mammalian brain by exploiting the Neuropixel data from Allen Brain Institute. We decode the neural responses from mouse visual cortex to predict the presented stimuli to the animal for natural (bear, trees, cheetah, etc.) and artificial (drifted gratings, orientated bars, etc.) classes. Our results indicate that neurons in mouse visual cortex encode the features of natural and artificial objects in a distinct manner, and such neural code is consistent across animals. We investigate this by applying transfer learning to train a DNN on the neural responses of a single animal and test its generalized performance across multiple animals. Within a single animal, DNN is able to decode the neural responses with as much as 100% classification accuracy. Across animals, this accuracy is reduced to 91%. This study demonstrates the potential of utilizing the DNN models as a computational framework to understand the neural coding principles in the mammalian brain.

8:30AM *Bidirectional Learning for Robust Neural Networks [#19072]*

Sidney Pontes-Filho and Marcus Liwicki, Oslo Metropolitan University, Norway; Lulea University of Technology, Sweden

A multilayer perceptron can behave as a generative classifier by applying bidirectional learning (BL). It consists of training an undirected neural network to map input to output and vice-versa; therefore it can produce a classifier in one direction, and a generator in the opposite direction for the same data. The learning process of BL tries to reproduce the neuroplasticity stated in Hebbian theory using only backward propagation of errors. In this paper, two learning techniques are independently introduced which use BL for improving robustness to white noise static and adversarial examples. The first method is bidirectional propagation of errors, which the error propagation occurs in backward and forward directions. Motivated by the fact that its generative model receives as input a constant vector per class, we introduce as a second method the novel hybrid adversarial networks (HAN). Its generative model receives a random vector as input and its training is based on generative adversarial networks (GAN). To assess the performance of BL, we perform experiments using several architectures with fully and convolutional layers, with and without bias. Experimental results show that both methods improve robustness to white noise static and adversarial examples, and even increase accuracy, but have different behavior depending on the architecture and task, being more beneficial to use the one or the other. Nevertheless, HAN using a convolutional architecture with batch normalization presents outstanding robustness, reaching state-of-the-art accuracy on adversarial examples of hand-written digits.

8:50AM *Learning Syntactic and Dynamic Selective Encoding for Document Summarization [#19200]*

Haiyang Xu, Yahao He, Kun Han, Junwen Chen and Xiangang Li, Didi Chuxing Co., Ltd., China

Text summarization aims to generate a headline or a short summary consisting of the major information of the source text. Recent studies employ the sequence-to- sequence framework to encode the input with a neural network and generate abstractive summary. However, most studies feed the encoder with the semantic word embedding but ignore the syntactic information of the text. Further, although previous studies proposed the selective gate to control the information flow from the encoder to the decoder, it is static during the decoding and cannot differentiate the information based on the decoder states. In this paper, we propose a novel neural architecture for document summarization. Our approach has the following contributions: first, we incorporate syntactic information such as constituency parsing trees into the encoding sequence to learn both the semantic and syntactic information from the document, resulting in more accurate summary; second, we propose a dynamic gate network to select the salient information based on the context of the decoder state, which is essential to document summarization. The proposed model has been evaluated on CNN/Daily Mail summarization datasets and the experimental results show that the proposed approach outperforms baseline approaches.

9:10AM *Gaining Extra Supervision via Multi-task learning for Multi-Modal Video Question Answering [#19667]*

Junyeong Kim, Minuk Ma, Kyungsu Kim, Sungjin Kim and Chang D. Yoo, Korea Advanced Institute of Science and Technology, Korea (South); Samsung Electronics, Korea (South)

This paper proposes a method to gain extra supervision via multi-task learning for multi-modal video question answering. Multi-modal video question answering is an important task that aims at the joint understanding of vision and language. However, establishing large scale dataset for multi-modal video question answering is expensive and the existing benchmarks are relatively small to provide sufficient supervision. To overcome this challenge, this paper proposes a multi-task learning method which is composed of three main components: (1) multi-modal video question answering network that answers the question based on the both video and subtitle feature, (2) temporal retrieval network that predicts the time in the video clip where the question was generated from and (3) modality alignment network that solves metric learning problem to find correct association of video and subtitle modalities. By simultaneously solving related auxiliary tasks with hierarchically shared intermediate layers, the extra synergistic supervisions are provided. Motivated by curriculum learning, multi-task ratio scheduling is proposed to learn easier task earlier to set inductive bias at the beginning of the training. The experiments on publicly available dataset TVQA shows state-of-the-art results, and ablation studies are conducted to prove the statistical validity.

2a: Supervised learning

Tuesday, July 16, 8:10AM-9:30AM, Room: Panorama IV, Chair: Francesca Cipollini

8:10AM *Hybrid Model for Cavitation Noise Spectra Prediction [#19020]*
Francesca Cipollini, Miglianti Fabiana, Luca Oneto, Giorgio Tani and Michele Viviani, UNIGE, Italy

In the latest years, models combining physical knowledge of a phenomenon and statistical inference are becoming of much interest in many real world applications. In this context, ship propeller underwater radiated noise is an interesting field of application for these so-called hybrid models, especially when the propeller cavitates. Nowadays, model scale tests are considered the state-of-the-art technique to predict the cavitation noise spectra. Unfortunately, they are negatively affected by scale effects which could alter the onset of some interesting cavitating phenomena respect to the full scale propeller; as a consequence, for some ship operational conditions it is not trivial to correctly reproduce the cavitation pattern in model scale tests. Moreover, model scale tests are quite expensive and time-consuming; it is not feasible to include them in the early stage of the design. Nevertheless, data collected during these tests can be adopted in order to tune a data-driven model while the physical equation describing the occurring phenomenon can be used to refine the prediction. In this work, the authors propose a hybrid model for the prediction of ships propeller underwater radiated noise, able to exploit both the physical knowledge of the problem and the real data obtained from cavitation tunnel experiments performed on different propellers in different working conditions. Results on real data will support the validity and the effectiveness of the proposal.

8:30AM *Identifying Mislabeled Instances in Classification Datasets [#19751]*
Nicolas Mueller and Karla Markert, Fraunhofer AISEC, Germany

A key requirement for supervised machine learning is labeled training data, which is created by annotating unlabeled data with the appropriate class. Because this process can in many cases not be done by machines, labeling needs to be performed by human domain experts. This process tends to be expensive both in time and money, and is prone to errors. Additionally, reviewing an entire labeled dataset manually is often prohibitively costly, so many real world datasets contain mislabeled instances. To address this issue, we present in this paper a non- parametric end-to-end pipeline to find mislabeled instances in numerical, image and natural language datasets. We evaluate our system quantitatively by adding a small number of label noise to 29 datasets, and show that we find mislabeled instances with an average precision of more than 0.84 when reviewing our system's top 1% recommendation. We then apply our system to publicly available datasets and find mislabeled instances in CIFAR-100, Fashion-MNIST, and others. Finally, we publish the code and an applicable implementation of our approach.

8:50AM *Vulnerability of Covariate Shift Adaptation Against Malicious Poisoning Attacks [#19981]*
Muhammad Umer, Christopher Fredericson and Robi Polikar, Rowan University, United States

Adversarial machine learning has recently risen to prominence due to increased concerns over the vulnerability of machine learning algorithms to malicious attacks. While the impact of malicious poisoning attacks on some popular algorithms, such as deep neural networks, has been well researched, the vulnerability of other approaches has not yet been properly established. In this effort, we explore the vulnerability of unconstrained least squares importance fitting (uLSIF), an algorithm used for computing the importance ratio for covariate shift domain adaptation problems. The uLSIF algorithm is an accurate and efficient technique to compute the importance ratio; however, we show that the approach is susceptible to a poisoning attack, where an intelligent adversary - having full or partial access to the training data - can inject well crafted malicious samples into the training data, resulting in an incorrect estimation of the importance values. Through strategically designed synthetic as well as real world datasets, we demonstrate that importance ratio estimation through uLSIF algorithm can be easily compromised with the insertion of even modest number of attack points into the training data. We also show that incorrect estimation of importance values can then cripple the performance of a subsequent covariate shift adaptation.

9:10AM *Comparison of Probabilistic Models and Neural Networks on Prediction of Home Sensor Events [#19341]*
Flavia Dias Casagrande, Jim Toerresen and Evi Zouganeli, OsloMet - Oslo Metropolitan University, Norway; University of Oslo, Norway

We present results and comparative analysis on the prediction of sensor events in a smart home environment with a limited number of binary sensors. We apply two probabilistic methods, namely Sequence Prediction via Enhanced Episode Discovery - SPEED, and Active LeZi - ALZ, as well as Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) in order to predict the next sensor event in a sequence. Our dataset has been collected from a real home with one resident over a period of 30 weeks. The binary sensor events are converted to two different text sequences as dictated by SPEED and ALZ, which are also used as inputs for the LSTM networks. We compare the performance of the algorithms regarding the number of preceding sensor events required to predict the next one, the required amount of data for the model to reach peak accuracy and stability, and the execution time. In addition, we analyze these for two different sets of sensors. Our best implementation achieved a peak accuracy of 83% for a set with fifteen sensors including motion, magnetic and power sensors, and 87% for seven motion sensors.

Special Lecture: Doctoral Consortium
Tuesday, July 16, 8:10AM-9:30AM, Room: Panorama V, Speaker: Marcus Liwicki

Tuesday, July 16, 9:30AM-10:00AM

Special Lecture: Coffee Break
Tuesday, July 16, 9:30AM-10:00AM, Room: Pre-function area Intercontinental

Tuesday, July 16, 10:00AM-11:00AM

Plenary Talk: Lee Giles, Pennsylvania State University
Tuesday, July 16, 10:00AM-11:00AM, Room: Ballroom I + II +II, Chair: Robert Kozma,

Tuesday, July 16, 11:00AM-12:00PM

Plenary Talk: Wolf Singer, Ernst Strungmann Institute
Tuesday, July 16, 11:00AM-12:00PM, Room: Ballroom I + II +II, Chair: Barbara Hammer,

Tuesday, July 16, 12:00PM-1:30PM

Special Lecture: Lunch Break
Tuesday, July 16, 12:00PM-1:30PM, Room: Various locations in the area

Tuesday, July 16, 1:30PM-3:30PM

1l: Deep neural networks and artificial neural networks
Tuesday, July 16, 1:30PM-3:30PM, Room: Ballroom I, Chair: Balthazar Donon

1:30PM *Graph Neural Solver for Power Systems [#19349]*
Balthazar Donon, Benjamin Donnot, Isabelle Guyon and Marot Antoine, RTE R&D, UPSud/INRIA Universite Paris-Saclay, France; UPSud/INRIA Universite Paris-Saclay, France; RTE R&D, France

We propose a neural network architecture that emulates the behavior of a physics solver that solves electric- ity differential equations to compute electricity flow in power grids (so-called "load flow"). Load flow computation is a well studied and understood problem, but current methods (based on Newton-Raphson) are slow. With increasing usage expectations of the current infrastructure, it is important to find methods to accelerate computations. One avenue we are pursuing in this paper is to use proxies based on "graph neural networks". In contrast with previous neural network approaches, which could only handle fixed grid topologies, our novel graph-based method, trained on data from power grids of a given size, generalizes to larger or smaller ones. We experimentally demonstrate viability of the method on randomly connected artificial grids of size 30 nodes. We achieve better accuracy than the DC-approximation (a standard benchmark linearizing physical equations) on random power grids whose size range from 10 nodes to 110 nodes, the scale of real-world power grids. Our neural network learns to solve the load flow problem without overfitting to a specific instance of the problem.

1:50PM *Deep Domain Adaptation for Vulnerable Code Function Identification [#19347]*
Van Nguyen, Trung Le, Tue Le, Khanh Nguyen, Olivier DeVel, Paul Montague, Lizhen Qu and Dinh Phung, Monash University, Australia; Deakin University, Australia; Defence Science and Technology Group, Australia; Data61 Group, Australia

Due to the ubiquity of computer software, software vulnerability detection (SVD) has become crucial in the software industry and in the field of computer security. Two significant issues in SVD arise when using machine learning, namely: i) how to learn automatic features that can help improve the predictive performance of vulnerability detection and ii) how to overcome the scarcity of labeled vulnerabilities in projects that require the laborious labeling of code by software security experts. In this paper, we address these two crucial concerns by proposing a novel architecture which leverages deep domain adaptation with automatic feature learning for software vulnerability identification. Based on this architecture, we keep the principles and reapply the state-of-the-art deep domain adaptation methods to indicate that deep domain adaptation for SVD is plausible and promising. Moreover, we further propose a novel method named Semi-supervised Code Domain Adaptation Network (SCDAN) that can efficiently utilize and exploit information carried in unlabeled target data by considering them as the unlabeled portion in a semi-supervised learning context. The proposed SCDAN method enforces the clustering assumption, which is a key principle in semi-supervised learning. The experimental results using six real-world software project datasets show that our SCDAN method and the baselines using our architecture have better predictive performance by a wide margin compared with the Deep Code Network (VulDeePecker) method without domain adaptation. Also, the proposed SCDAN significantly outperforms the DIRT-T which to the best of

our knowledge is currently the-state-of-the-art method in deep domain adaptation and other baselines.

2:10PM *Language Modeling through Long-Term Memory Network [#20010]*
Anupiya Nugaliyadde, Kok Wai Wong, Ferdous Sohel and Hong Xie, Murdoch University, Australia

Recurrent Neural Networks (RNN), Long Short- Term Memory Networks (LSTM), and Memory Networks which contain memory are popularly used to learn patterns in sequential data. Sequential data has long sequences that hold relationships. RNN can handle long sequences but suffers from the vanishing and exploding gradient problems. While LSTM and other memory networks address this problem, they are not capable of handling long sequences (50 or more data points long sequence patterns). Language modelling requiring learning from longer sequences are affected by the need for more information in memory. This paper introduces Long Term Memory network (LTM), which can tackle the exploding and vanishing gradient problems and handles long sequences without forgetting. LTM is designed to scale data in the memory and gives a higher weight to the input in the sequence. LTM avoid overfitting by scaling the cell state after achieving the optimal results. The LTM is tested on Penn treebank dataset, and Text8 dataset and LTM achieves test perplexities of 83 and 82 respectively. 650 LTM cells achieved a test perplexity of 67 for Penn Treebank, and 600 cells achieved a test perplexity of 77 for Text8. LTM achieves state of the art results by only using ten hidden LTM cells for both datasets.

2:30PM *Exploiting Randomness in Deep Learning Algorithms [#20333]*
Seyed Hamed Fatemi Langroudi, Cory Merkel, Humza Syed and Dhireesha Kudithipudi, Rochester Institute of Technology, United States

The recent surge of interest to use deep neural networks for complex tasks has led to training complex networks with billions of parameters that use enormous amounts of training data. Performing the backpropagation algorithm on these deep networks is time consuming and requires a large amount of resources that is limited by the hardware. In order to move towards agile deep learning, we are motivated to exploiting randomness for deep neural networks as an alternative to the backpropagation algorithm. In this work we explore the effects of utilizing random weights in convolutional neural networks. This is done through random initialization of weights and freezing them while training on only the output layer. We also propose a novel weight distribution based on the sum of sinusoids for random convolutional neural networks. Our experiments show that by leaving the weights random in convolutional neural networks performance comparable to state-of-the-art results can be achieved for the MSTAR dataset.

2:50PM *A Model Based on Siamese Neural Network for Online Transaction Fraud Detection [#19385]*
Xinxin Zhou, Zhaohui Zhang, Lizhi Wang and Pengwei Wang, Donghua University, China

With the rapid development of Internet finance, the volume of online transactions increases gradually, but the risk of exposure is increasing, and fraud is emerging. Because of the characteristics of online transaction, such as large volume, high frequency and fast update speed. In addition the online transaction data has the problems of unbalanced positive and negative sample and sparse timing of transaction data. Most of the existing methods to solve the problem of data imbalance are sampled, but this method will change the distribution of the dataset, which is not conducive to improving the generalization ability of the model. There are some timing characteristics of online transaction data, and the common fraud detection model does not take the problem into account in the design of the model. Based on the problems, this paper puts forward the siamese neural network structure based on CNN and LSTM, uses the siamese neural network structure to solve the problem of sample imbalance in online transaction, uses the LSTM structure to make the model have memory function, remembers the user's transaction information, in order to better detect the fraudulent transaction. The model presented in this paper is verified in real B2C transaction data, and its precision and recall reach about 95% and 96%, respectively.

3:10PM *Gate-Layer Autoencoders with Application to Incomplete EEG Signal Recovery [#19303]*
Heba El-Fiqi, Kathryn Kasmarik, Anastasios Bezerianos, Kay Chen Tan and Hussein A. Abbass, UNSW-Canberra, Canberra, Australia; National University of Singapore, Singapore, Singapore; City University of Hong Kong, Kowloon, Hong Kong

Autoencoders (AE) have been used successfully as unsupervised learners for inferring latent information, learning hidden features and reducing the dimensionality of the data. In this paper, we propose a new AE architecture: Gate-Layer AE (GLAE). The novelty of GLAE lies in its ability to encourage learning of the relationships among different input variables, which affords it with an inherent ability to recover missing variables from the available ones and to act as a concurrent multi-function approximator. GLAE uses a network architecture that associates each input with a binary gate acting as a switch that turns on or off the flow to each input unit, while synchronising its action with data flow to the network. We test GLAE with different coding sizes, and compare its performance against the Classic AE, Denoising AE, and Variational AE. The evaluation uses Electroencephalograph (EEG) data with an aim to reconstruct the EEG signal when some data are missing. The results demonstrate GLAE's superior performance in reconstructing EEG signals with up to 25% missing data in an input stream.

2e: Deep learning

Tuesday, July 16, 1:30PM-3:30PM, Room: Ballroom II, Chair: Lesort Timothee

1:30PM *Learning Semantic Coherence for Machine Generated Spam Text Detection [#19674]*
Mengjiao Bao, Jianxin Li, Jian Zhang, Hao Peng and Xudong Liu, Beihang University, China

Using machine to generate text has attracted considerable attention recently. However, low quality text generated by machine will seriously impact the user experience due to the poor readability. Traditional methods for detecting machine generated text heavily depend on hand-crafted features. While most deep learning methods for general text classification tend to model the semantic representation of topics, and thus overlook the semantic coherence that is also useful for detecting machine generated text. In this paper, we propose an end-to-end neural architecture that learns semantic coherence of text sequences. We conduct experiments on both Chinese and English datasets with more than two million articles containing manually written and machine generated ones. Results show that our method is effective and achieves the state-of- the-art performance.

1:50PM *Generative Models from the perspective of Continual Learning [#19555]*
Lesort Timothee, Caselles-Dupre Hugo, Garcia-Ortiz Michael, Stoian Andrei and Filliat David, Ensta-Paristech, Thales, France; Ensta-Paristech, Softbank, France; Softbank, France; Thales, France; Ensta-Paristech, France

Which generative model is the most suitable for Continual Learning? This paper aims at evaluating and comparing generative models on disjoint sequential image generation tasks. We investigate how several models learn and forget, considering various strategies: rehearsal, regularization, generative replay and fine-tuning. We used two quantitative metrics to estimate the generation quality and memory ability. We experiment with sequential tasks on three commonly used benchmarks for Continual Learning (MNIST, Fashion MNIST and CIFAR10). We found that among all models, the original GAN performs best and among Continual Learning strategies,

generative replay outperforms all other methods. Even if we found satisfactory combinations on MNIST and Fashion MNIST, training generative models sequentially on CIFAR10 is particularly instable, and remains a challenge. Our code is available online (https://github.com/TLESORT/Generative_Continual_Learning).

2:10PM *Deep Networks with Adaptive-Nystrom Approximation [#20319]*

Luc Giffon, Stephane Ayache, Thierry Artieres and Hachem Kadri, Aix Marseille Universite, Universite de Toulon, CNRS, LIS, Marseille, France, France

Recent work has focused on combining kernel meth- ods and deep learning to exploit the best of the two approaches. Here, we introduce a new architecture of neural networks in which we replace the top dense layers of standard convolutional architectures with an approximation of a kernel function by relying on the Nystrom approximation. Our approach is easy and highly flexible. It is compatible with any kernel function and it allows exploiting multiple kernels. We show that our architecture has the same performance than standard architecture on datasets like SVHN and CIFAR100. One benefit of the method lies in its limited number of learnable parameters which makes it particularly suited for small training set sizes, e.g. from 5 to 20 samples per class.

2:30PM *Dynamic Unit Surgery for Deep Neural Network Compression and Acceleration [#20378]*

Minsam Kim and James Kwok, Hong Kong University of Science and Technology, Hong Kong

Successful deep neural network models tend to possess millions of parameters. Reducing the size of such models by pruning parameters has recently earned significant interest from the research community, allowing more compact models with similar performance level. While pruning parameters usually result in large sparse weight tensors which cannot easily lead to proportional improvement in computational efficiency, pruning filters or entire units allow readily available off-the-shelf libraries to harness the benefit of smaller architecture. One of the most well-known aspects of network pruning is that the final retained performance can be improved by making the process of pruning more gradual. Most existing techniques smooth the process by repeating the technique (multi-pass) at increasing pruning ratios, or by applying the method in a layer-wise fashion. In this paper, we introduce Dynamic Unit Surgery (DUS) that smooths the process in a novel way by using decaying mask values, instead of multi-pass or layer-wise treatment. While multi-pass schemes entirely discard network components pruned at the early stage, DUS allows recovery of such components. We empirically show that DUS achieves competitive performance against existing state-of-the-art pruning techniques in multiple image classification tasks. In CIFAR10, we prune VGG16 network to use 5% of the parameters and 23% of FLOPs while achieving 6.65% error rate with no degradation from the original network. We also explore the method's application to transfer learning environment for fine-grained image classification and report its competitiveness against state-of-the-art baseline.

2:50PM *Looking back at Labels: A Class based Domain Adaptation Technique [#19969]*

Vinod Kumar Kurmi and Vinay P Namboodiri, Indian Institute of Technology Kanpur, India

In this paper, we solve the problem of adapting classifiers across domains. We consider the problem of domain adaptation for multi-class classification where we are provided a labeled set of examples in a source dataset and we are provided a target dataset with no supervision. In this setting, we propose an adversarial discriminator based approach. While the approach based on adversarial discriminator has been previously proposed; in this paper, we present an informed adversarial discriminator. Our observation relies on the analysis that shows that if the discriminator has access to all the information available including the class structure present in the source dataset, then it can guide the transformation of features of the target set of classes to a more structure adapted space. Using this formulation, we obtain state-of-the-art results for the standard evaluation on benchmark datasets. We further provide detailed analysis which shows that using all the labeled information results in an improved domain adaptation.

3:10PM *Underwater Fish Detection with Weak Multi-Domain Supervision [#19534]*

Dmitry A. Konovalov, Alzayat Saleh, Michael Bradley, Mangalam Sankupellay, Simone Marini and Marcus Sheaves, James Cook University, Australia; National Research Council of Italy, Italy

Given a sufficiently large training dataset, it is relatively easy to train a modern convolution neural network (CNN) as a required image classifier. However, for the task of fish classification and/or fish detection, if a CNN was trained to detect or classify particular fish species in particular background habitats, the same CNN exhibits much lower accuracy when applied to new/unseen fish species and/or fish habitats. Therefore, in practice, the CNN needs to be continuously fine-tuned to improve its classification accuracy to handle new project-specific fish species or habitats. In this work we present a labelling-efficient method of training a CNN-based fish-detector (the Xception CNN was used as the base) on relatively small numbers (4,000) of project-domain underwater fish/no-fish images from 20 different habitats. Additionally, 17,000 of known negative (that is, missing fish) general-domain (VOC2012) above-water images were used. Two publicly available fish-domain datasets supplied additional 27,000 of above-water and underwater positive/fish images. By using this multi-domain collection of images, the trained Xception-based binary (fish/not-fish) classifier achieved 0.17% false-positives and 0.61% false-negatives on the project's 20,000 negative and 16,000 positive holdout test images, respectively. The area under the ROC curve (AUC) was 99.94%.

8a: Applications of deep networks

Tuesday, July 16, 1:30PM-3:30PM, Room: Ballroom III, Chair: Austin Okray

1:30PM *Music Classification using an Improved CRNN with Multi-Directional Spatial Dependencies in Both Time and Frequency Dimensions [#20443]*

Zhen Wang, Suresh Muknahallipatna, Maohong Fan, Austin Okray and Chao Lan, University of Wyoming, United States

In music classification tasks, Convolutional Recurrent Neural Network (CRNN) has achieved state-of-the-art performance on several data sets. However, the current CRNN technique only uses RNN to extract spatial dependency of music signal in its time dimension but not its frequency dimension. We hypothesize the latter can be additionally exploited to improve classification performance. In this paper, we propose an improved technique called CRNN in Time and Frequency dimensions (CRNN-TF), which captures spatial dependencies of music signal in both time and frequency dimensions in multiple directions. Experimental studies on three real-world music data sets show that CRNN-TF consistently outperforms CRNN and several other state-of-the-art deep learning-based music classifiers. Our results also suggest CRNN-TF is transferable on small music data sets via the fine-tuning technique.

1:50PM *A Multi-granularity Neural Neural Net work for Answer Sentence Selection [#19511]*

Zhang Chenggong, Zhang Weijuan, Zha Daren, Ren Pengjie and Mu Nan, State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China; School of Computer and Technology,Shandon University, China

In open-domain question answering system, the granularities of the answers vary with different types of questions.For example, for the questions asking about locations (Location type questions), their answers are usually short phrases.While for the questions asking about reasons (Description type questions), their answers are usually long clauses or sentences.This insight can be used to improve the performance of answer sentence selection, which is a crucial component of the open-domain QA system.In this paper, we propose a novel Multi-Granularity Neural Network (MGNN) model to better evaluate the semantic matching of questions and answers.First, MGNN has three classes of channels with each computing the similarity of question and answer pairs from one of the three granularity levels: clause level, phrase level and ngram level.Then, MGNN uses a parametrization weighting scheme which considers question types to combine these different granularity channels.We carry out experiments on a public available benchmark dataset for question answering.Empirical results show that our method outperforms state-of-the-art methods.

2:10PM *Generalized PatternAttribution for Neural Networks with Sigmoid Activations [#20307]*

Jiamei Sun and Alexander Binder, Singapore University of Technology and Design, Singapore

Explanation methods for deep neural networks (DNN) such as LRP, PatternAttribution, LIME, DeepLIFT have promoted the explanation of convolutional neural networks (CNN). Results are obtained mostly using ReLU activation function. In this paper, we investigate the performance of explanation methods on neural networks with sigmoid activations like sigmoid and tanh. PatternAttribution is a recent approach which allows learning explanation patterns from data. We show that the saturated zones of sigmoids pose difficulties for PatternAttribution. In order to solve these issues, as a first novelty, we generalize global explanations to piece-wise dependent explanations. In a second contribution, we learn the parameters of PatternAttribution in near-linear activation zones of the sigmoids, while replacing it with LRP in the saturated zones. Finally, we introduce and evaluate direct layer-wise Taylor approximation to show that LRP as a deep-Taylor-motivated approach outperforms the ad-hoc application of Taylor approximation. We show results on MNIST and also on LSTM and GRU networks used for two sentiment classification tasks as important application cases of models using sigmoids. Our results demonstrate that the proposed method, Piece-wise PAtternLRP (PPAP), outperforms PatternAttribution as well as LRP for networks with sigmoids, thus combining their strengths effectively.

2:30PM *Collaborative Multi-key Learning with an Anonymization Dataset for a Recommender System [#19049]*

Linh Nguyen and Tsukasa Ishigaki, Tohoku University, Japan

Balancing accuracy and privacy is an important tradeoff problem for information systems, including recommender systems. To achieve high performance, modern recommender systems tend to use as much information as possible. This trend is borne out by the increasing number of studies of hybrid methods that combine rating and auxiliary information. However, because of privacy concerns, in many cases, service providers can not require users to give their personal information. Therefore, numerous earlier reported methods only use item attributes for auxiliary information. To address these shortcomings, our manuscript provides a method to extract user profiles without using demographic data. Our model learns user and item latent variables through two separate deep neural networks and also learns implicit relations between users and items using the information and their ratings. Experiments verified that our model is a more effective recommender system than state-of-the-art baselines.

2:50PM *A Methodology Based on Deep Learning for the Classification of Power Quality Events Using Convolutional Network and Long Short-Term Memory [#20300]*

Wilson Rodrigues Junior, Fabbio Borges, Ricardo Rabelo, Bruno Lima and Jose Alencar, Federal University of Piaui (UFPI), Brazil; Federal Institute of Maranhao (IFMA), Brazil

The Electrical Power Quality (PQ) studies are commonly related to disturbances that alter the voltage sinusoidal features and/or current wave shapes. The classification approaches of electrical power quality disturbances found in the literature mainly consist of three steps: 1) signal analysis and feature extraction, 2) feature selection and 3) disturbances classification. However, there are some problems inherent in disturbances classification. The manual extraction of features is an imprecise and complex process, which can influence the results and, therefore, does not deal well with noisy signals. This article proposes an approach based on Deep Learning using the raw data, without pre-processing, manual extraction or manual feature selection of the PQ disturbances signals for the classification of fifteen electrical power quality disturbances. Therefore, a deep network is used, which consists of a hybrid architecture, composed by convolutional layers, a pooling layer, an LSTM layer, and batch normalization to execute the features extraction automatically. To adapt to the data type, it has adopted the 1-D convolution. The extracted features are used as input to completely connected layers, the last one being a SoftMax layer. The results are compared with the state of the art methods based on the three steps, showing that the proposed approach had satisfactory performance even with noisy data.

3:10PM *A Method based on Convolutional Neural Networks for Fingerprint Segmentation [#20286]*

Paulo Serafim, Aldisio Medeiros, Paulo Rego, Gilvan Maia, Fernando Trinta, Marcio Maia, Jose Macedo and Aloisio Lira, Federal University of Ceara, Brazil; Brazilian Federal Highway Police, Brazil

In forensic science, the resolution of crimes is associated with the identification of those involved. In the civil context, the security of automated processes depends on the identification of authorized people. In this sense, fingerprint-based recognition techniques stand out. A fundamental stage is the calculation of the degree of similarity between the samples presented, so the task of identifying a region of interest (ROI), excluding noisy regions, can improve the precision and reduce the computational cost. In this aspect, this work presents a technique of segmentation of the region of interest based on convolutional neural networks (CNN) without pre-processing steps. The new approach was evaluated in two different architectures from state of the art, presenting similarity indexes Distance of Hausdorff (5.92), Dice coefficient (97.28%) and Jaccard Similarity (96.77%) superior to the classic methods. The error rate (3.22%) was better than five segmentation techniques from state of the art and showed better results than another deep learning approach, presenting promising results to identify the region of interest with potential for application in systems based on biometric identification.

2t: Topics in machine learning

Tuesday, July 16, 1:30PM-3:30PM, Room: Duna Salon I, Chair: Khan Iftekharuddin

1:30PM *Compact Cluster-based Balanced Distribution Adaptation for Transfer Learning [#19991]*

Xu Zhang, Zuyu Zhang and Haeyoung Bae, Chongqing University of Posts and Telecommunications, China; Inha University, Korea (South)

Recently, Domain Adaptation has received great attention in addressing domain shift problem, where source domain and target domain follow different distribution. Existing methods often seek to minimize the marginal distribution and the conditional distribution across domains however treat them equally, which is not reasonable in real applications. Moreover, it is easy to misclassify on target data with large within-class scatter. In this paper, we propose a Compact Cluster-based Balanced Distribution Adaptation (CC-BDA) method for cross-domain classification. CC-BDA exploits the balanced domain adaptation by increasing the balance factor dynamically in each iteration until convergence. In addition, we construct compact clusters to reduce within-class scatter alongside the domain adaptation, which endows samples with the same label much closer from their means. Finally, we present analysis and performance evaluation over seven state-of-the-art baseline methods. The results show that CC-BDA significantly outperforms other adaptation methods in accuracy and execution time cost.

1:50PM *Combining Self-reported Confidences from Uncertain Annotators to Improve Label Quality [#20236]*

Christoph Sandrock, Marek Herde, Adrian Calma, Daniel Kottke and Bernhard Sick, University of Kassel, Germany

Class assignment (label) is not the only information that can be queried from annotators. Additional feedback, in form of confidences, quantifies the reliability of the provided label. Multiple annotators classifying the same sample with different levels of expertise posses different levels of confidences. In this article, we discuss different types of confidence inspired by real-world application and discuss how they relate to each other. The alpha-confidence is inspired by humans, the beta-confidence works with probabilistic outputs from intelligent machines (e.g., robots), and gamma-confidence models non-normalized confidences. We consider uncertain, benevolent annotators, thus the provided labels can be contradictory. To overcome this problem, we propose two fusion strategies that combine the confidences from multiple uncertain annotators to a single one. Numerical and graphical evaluation indicates superior performance of our strategy compared to related strategies, namely, confidence weighted majority vote, c-Certainty, and a maximum likelihood estimation.

2:10PM *Neural Regression Trees [#20345]*

Shahan Ali Memon, Wenbo Zhao, Bhiksha Raj and Rita Singh, Carnegie Mellon University, United States

Regression-via-Classification (RvC) is the process of converting a regression problem to a classification one. Current approaches for RvC use ad-hoc discretization strategies and are suboptimal. We propose a neural regression tree model for RvC. In this model, we employ a joint optimization framework where we learn optimal discretization thresholds while simultaneously optimizing the features for each node in the tree. We empirically show the validity of our model by testing it on two challenging regression tasks where we establish the state of the art.

2:30PM *Collaborative and Privacy-Preserving Machine Teaching via Consensus Optimization [#19896]*

Yufei Han, Yuzhe Ma, Christopher Gates, Kevin Roundy and Yun Shen, Symantec Research Labs, France; University of Wisconsin-Madison, United States; Symantec Research Labs, United States; Symantec Research Labs, United Kingdom

In this work, we define a collaborative and privacy-preserving machine teaching paradigm with multiple distributed teachers. We focus on consensus super teaching. It aims at organizing distributed teachers to jointly select a compact while informative training subset from data hosted by the teachers to make a learner learn better. The challenges arise from three perspectives. First, the state-of-the-art pool-based super teaching method applies mixed-integer non-linear programming (MINLP) which does not scale well to very large data sets. Second, it is desirable to restrict data access of the teachers to only their own data during the collaboration stage to mitigate privacy leaks. Finally, the teaching collaboration should be communication-efficient since large communication overheads can cause synchronization delays between teachers. To address these challenges, we formulate collaborative teaching as a consensus and privacy- preserving optimization process to minimize teaching risk. We theoretically demonstrate the necessity of collaboration between teachers for improving the learner's learning. Furthermore, we show that the proposed method enjoys a similar property as the Oracle property of adaptive Lasso. The empirical study illustrates that our teaching method can deliver significantly more accurate teaching results with high speed, while the non-collaborative MINLP-based super teaching becomes prohibitively expensive to compute.

2:50PM *A Proof of Local Convergence for the Adam Optimizer [#20268]*

Sebastian Bock and Martin Weiss, OTH Regensburg, Germany

Adaptive Moment Estimation (Adam) is a very popular training algorithm for deep neural networks, implemented in many machine learning frameworks. To the best of the authors knowledge no complete convergence analysis exists for Adam. The contribution of this paper is a method for the local convergence analysis in batch mode for a deterministic fixed training set, which gives necessary conditions for the hyperparameters of the Adam algorithm. Due to the local nature of the arguments the objective function can be non-convex but must be at least twice continuously differentiable.

3:10PM *Dimension Estimation and Topological Manifold Learning [#19673]*

Tasaki Hajime, Lenz Reiner and Chao Jinhui, Chuo University, Japan

We describe a framework which can be used to investigate the geometrical structure of datasets in high- dimensional spaces. Understanding these geometrical properties is essential in machine learning in general. An application which has received much attention recently is the investigation of adversarial examples which can be easily identified by humans but which are misleading neural networks. We will argue that these examples can be understood (and consequently avoided) by an investigation of the dimension and geometry of the classification hypersurface of the underlying neural network as a manifold using the tools introduced in this paper. We argue that a key concept in the analysis is the dimension of a manifold. Many machine learning methods use dimension reduction techniques to understand or classify the input data. We point out different definitions of a dimension of a manifold. We use tools from the theory of topological manifolds to introduce local, or intrinsic and global dimensions. The local dimension describes the local similarity between the dataset and a Euclidean space. The global dimension is the lowest dimension of a Euclidean space into which the dataset can be embedded. We introduce a framework for dimension estimation and topological manifold learning based on the measure ratio method to estimate the dimensions and structure of the data manifold. As an illustration we use images of handwritten digits and points of a Klein-bottle embedded in a five-dimensional space. We compare the results obtained by

the measure ratio method with the well-known local-principal component analysis estimation.

Neuroengineering

Tuesday, July 16, 1:30PM-3:30PM, Room: Duna Salon II, Chair: Hiroomi Hikawa

1:30PM *Neuromemristive Multi-Layer Random Projection Network with On-Device Learning [#19492]*

Abdullah Zyarah and Dhireesha Kudithipudi, Rochester Institute of Technology, United States

This paper proposes a neuromemristive multi-layer neural network with on-device learning. The proposed system is studied within the context of a feedforward multi-layer random projection network, where the core learning is modeled by a stochastic gradient descent simplified for memristor crossbar integration. Two random projection network topologies are explored for binomial and multinomial datasets. A detailed study on the resiliency of the networks in the presence of device failure is performed. The topology with softmax output layer exhibits stability and better resiliency in performance after experiencing a device failure. It is shown that this topology can regain full performance after experiencing 30% stuck-at-faults, with 2x increase in the hidden layer neurons.

1:50PM *Epilepsy detection using multiclass classifier based on spectral features [#19539]*

Jefferson Oliva and Joao Luis Rosa, University of Sao Paulo, Brazil

Epilepsy is the fourth most incident neurological disorder, for which approximately 50 million people are affected. The epilepsy is characterized by repeated seizures, occasioned by temporary electrical disturbances of the brain, and it is usually evidenced in four distinct events: pre-ictal, ictal, post-ictal, and interictal. This disorder can be diagnosed by electroencephalogram (EEG), which is the result of brain electrical activity monitoring. In the current work, we built multiclass classifiers in order to differentiate EEG segments among normal, interictal, and ictal classes. For this, the multitaper method was computed to generate power spectrum, spectrogram, and bispectrogram from each EEG segment, in which a total of 102 features were extracted. For classifier building, we used all-against- all approach to decompose the multiclass problem into binary subproblems, in which the following machine learning algorithms were applied: random forest, 1-nearest-neighbor, linear discriminant analysis (LDA), backpropagation based on multilayer perceptron (BP-MLP), k-means based on radial basis function network (KM-RBFN), and sequential minimal optimization. In the evaluation by ten-fold stratified cross-validation, the BP-MLP reached the highest accuracy (98.33\%). Afterward, an extremely significant statistical difference was found among classifiers. In a post hoc test application, it was proved that the KM-RBFN classifier reached an inferior performance compared to LDA and BP-MLP predictive models. All classifiers built in this work achieved promising results, in which accuracy above 90\% was obtained for classification of EEG segments.

2:10PM *Design Space Evaluation of a Memristor Crossbar Based Multilayer Perceptron for Image Processing [#19931]*

Chris Yakopcic, B. Rasitha Fernando and Tarek Taha, University of Dayton, United States

This paper describes a simulated memristor-based neuromorphic system that can be used for ex-situ training of a multi-layer perceptron algorithm. The presented programming technique can be used to map the weights required of a neural algorithm directly onto the grid of resistances in a memristor crossbar. Using this weight-to-crossbar mapping approach along with the dot product calculation circuit, neural algorithms can be easily implemented using this system. To show the effectiveness of this circuit, a Multilayer Perceptron is trained to perform Sobel edge detection. Following these simulations, an analysis was presented that shows how memristor programming accuracy and network size are related to output error; the results show that network size can be increased to reduce testing error. In some cases, the memristors in the circuit may be capable of operating with at lower precision if the network size is increased. This means that less precise (or lower resolution) memristor devices may be used to implement the proposed system. Furthermore, a power, timing, and energy analysis shows that this circuit has a computation throughput that allows it to process 4K UHD video in real time at approximately 337mW.

2:30PM *Nested Hardware Architecture for Self-Organizing Map [#20464]*

Hiroomi Hikawa, Kansai University, Japan

This paper proposes a hardware SOM architecture with nested structure. The proposed hardware SOM is made of a single module, and the module is made of four smaller modules, then the smaller module consists of four further smaller modules. With this simple recursive structure, size of SOM can be easily increased. In addition, the proposed SOM requires only one clock cycle to process a single input vector, and hence very high throughput operation is provided. Experimental results show that that the proposed SOM architecture provides he scalability without degrading the performance, and the proposed SOM achieved 33024 MCUPS of high throughput.

2:50PM *Cascaded Neural Network for Memristor based Neuromorphic Computing [#19204]*

Sheng-Yang Sun, Hui Xu, Jiwei Li, Haijun Liu and Qingjiang Li, National University of Defense Technology, China

Recent years, several memristor-based neuromorphic processing chips have been proposed. However, there is few architectures to consider the cascading problems, and the scalability is not strong while dealing some tasks. To address this issue, we present a memristor-based cascaded method with some basic computation unit, several neural network processing chips can be cascaded by this means to improve the processing capability of the dataset. Compared with VGGNet and GoogLeNet, the proposed cascaded framework can achieve 93.54% Fashion-MNIST accuracy under the 4.15M parameters. Extensive experiments are conducted show that the circuit simulation results can still provide a high recognition accuracy, the recognition accuracy loss after circuit simulation can be controlled at around 0.26%.

3:10PM *Hyperspectral Image Classification for Remote Sensing Using Low-Power Neuromorphic Hardware [#20074]*

Vivek Parmar, Jung-Ho Ahn and Manan Suri, Indian Insitute of Technology Delhi, India; NEPES Corporation, Korea (South); Indian Institute of Technology Delhi, India

In this paper, we present a novel feature extraction algorithm based approach for performing Hyperspectral Image Classification using a low-power Neuromorphic hardware. The application of interest for this study is HSI image classification for remote sensing. We demonstrate energy-efficient data processing pipeline optimized to use with on-edge neuromorphic hardware. The dataset used for the study is Salinas-A. We use the Brilliant USB stick with 4 NM500 chips for prototyping the application. Achieved recognition time is 18.4 us and energy consumption is 10 uJ with an accuracy of 97%.

8k: Signal processing, image processing, and multi-media

Tuesday, July 16, 1:30PM-3:30PM, Room: Duna Salon III, Chair: Nelson Enrique Yalta Soplin

1:30PM *Edge Focused Super-Resolution of Thermal Images [#19505]*

Yannick Zoetgnande, Jean-Louis Dillenseger and Javad Alirezaie, Universite Rennes 1, France; Ryerson Univeristy, Canada

In this work, a super-resolution method is proposed for indoor scenes captured by low-resolution thermal cameras. The proposed method is called Edge Focused Thermal Super-resolution (EFTS) which contains an edge extraction module enforcing the neural networks to focus on the edge of images. Utilizing edge information, our model, based on residual dense blocks, can perform super- resolution for thermal images, while enhancing the visual information of the edges. Experiments on benchmark datasets showed that our EFTS method achieves better performance in comparison to the state-of-the-art techniques.

1:50PM *Weakly-Supervised Deep Recurrent Neural Networks for Basic Dance Step Generation [#19803]*

Nelson Enrique Yalta Soplin, Shinji Watanabe, Kazuhiro Nakadai and Tetsuya Ogata, Waseda University, Japan; Johns Hopkins University, United States; Honda Research Institute Japan, Japan

Synthesizing human's movements such as dancing is a flourishing research field which has several applications in computer graphics. Recent studies have demonstrated the advantages of deep neural networks (DNNs) for achieving remarkable performance in motion and music tasks with little effort for feature pre-processing. However, applying DNNs for generating dance to a piece of music is nevertheless challenging, because of 1)DNNs need to generate large sequences while mapping the music input, 2)the DNN needs to constraint the motion beat to the music, and 3)DNNs require a considerable amount of hand-crafted data. In this study, we propose a weakly supervised deep recurrent method for real-time basic dance generation with audio power spectrum as input. The proposed model employs convolutional layers and a multilayered Long Short-Term memory (LSTM) to process the audio input. Then, another deep LSTM layer decodes the target dance sequence. Notably, this end-to-end approach has 1)an auto-conditioned decode configuration that reduces accumulation of feedback error of large dance sequence, 2)uses a contrastive cost function to regulate the mapping between the music and motion beat, and 3)trains with weak labels generated from the motion beat, reducing the amount of hand-crafted data. We evaluate the proposed network based on i)the similarities between generated and the baseline dancer motion with a cross entropy measure for large dance sequences, and ii)accurate timing between the music and motion beat with an F-measure. Experimental results revealed that, after training using a small dataset, the model generates basic dance steps with low cross entropy and maintains an F-measure score similar to that of a baseline dancer.

2:10PM *On Class Imbalance and Background Filtering in Visual Relationship Detection [#19547]*

Alessio Sarullo and Tingting Mu, University of Manchester, United Kingdom

In this paper we investigate the problems of class imbalance and irrelevant relationships in Visual Relationship Detection (VRD). State-of-the-art deep VRD models still struggle to predict uncommon classes, limiting their applicability. Moreover, many methods are incapable of properly filtering out background relationships while predicting relevant ones. Although these problems are very apparent, they have both been overlooked so far. We analyse why this is the case and propose modifications to both model and training to alleviate the aforementioned issues, as well as suggesting new measures to complement existing ones and give a more holistic picture of the efficacy of a model.

2:30PM *Boosted GAN with Semantically Interpretable Information for Image Inpainting [#19062]*

Li Ang, Qi Jianzhong, Zhang Rui and Kotagiri Ramamohanarao, The University of Melbourne, Australia

Image inpainting aims at restoring missing regions of corrupted images, which has many applications such as image restoration and object removal. However, current GAN-based inpainting models fail to explicitly consider the semantic consistency between restored images and original images. For example, given a male image with image region of one eye missing, current models may restore it with a female eye. This is due to the ambiguity of GAN-based inpainting models: these models can generate many possible restorations given a missing region. To address this limitation, our key insight is that semantically interpretable information (such as attribute and segmentation information) of input images (with missing regions) can provide essential guidance for the inpainting process. Based on this insight, we propose a boosted GAN with semantically interpretable information for image inpainting that consists of an inpainting network and a discriminative network. The inpainting network utilizes two auxiliary pretrained networks to discover the attribute and segmentation information of input images and incorporates them into the inpainting process to provide explicit semantic-level guidance. The discriminative network adopts a multi-level design that can enforce regularizations not only on overall realness but also on attribute and segmentation consistency with the original images. Experimental results show that our proposed model can preserve consistency on both attribute and segmentation level, and significantly outperforms the state-of-the-art models.

2:50PM *Visual Relationship Attention for Image Captioning [#19421]*

Zongjian Zhang, Qiang Wu, Yang Wang and Fang Chen, University of Technology Sydney, Australia

Visual attention mechanisms have been broadly used by image captioning models to attend to related visual information dynamically, allowing fine-grained image understanding and reasoning. However, they are only designed to discover the region-level alignment between visual features and the language feature. The exploration of higher-level visual relationship information between image regions, which is rarely researched in recent works, is beyond their capabilities. To fill this gap, we propose a novel visual relationship attention model based on the parallel attention mechanism under the learnt spatial constraints. It can extract relationship information from visual regions and language and then achieve the relationship-level alignment between them. Using combined visual relationship attention and visual region attention to attend to related visual relationships and regions respectively, our image captioning model can achieve state-of-the-art performances on the MSCOCO dataset. Both quantitative analysis and qualitative analysis demonstrate that our novel visual relationship attention model can capture related visual relationship and further improve the caption quality.

3:10PM *What's in a Word? Detecting Partisan Affiliation from Word Use in Congressional Speeches [#20327]*

Ulya Bayram, John Pestian, Daniel Santel and Ali Minai, University of Cincinnati and Cincinnati Children's Hospital, United States; Cincinnati Children's Hospital, United States; University of Cincinnati, United States

Politics is an area of broad interest to policy-makers, researchers, and the general public. The recent explosion in the availability of electronic data and advances in data analysis methods -- including techniques from machine learning -- have led to many studies attempting to extract political insight from this data. Speeches in the U.S. Congress represent an exceptionally rich dataset for this purpose, and these have been analyzed by many researchers using statistical and machine learning methods. In this paper, we analyze House of Representatives floor speeches from the 1981 - 2016 period, with the goal of inferring the partisan affiliation of the speakers from their use of

words. Previous studies with sophisticated machine learning models has suggested that this task can be accomplished with an accuracy in the 55 to 80% range, depending on the year. In this paper, we show that, in fact, very comparable results can be obtained using a much simpler linear classifier in word space, indicating that the use of words in partisan ways is not particularly complicated. Our results also confirm that, over the period of study, it has become steadily easier to infer partisan affiliation from political speeches in the United States. Finally, we make some observations about specific terms that Republicans and Democrats have favored over the years in service of partisan expression.

8a: Applications of deep networks

Tuesday, July 16, 1:30PM-3:30PM, Room: Panorama I, Chair: Alvaro S. Hervella

1:30PM *A Novel Neural Approach for News Reprint Prediction [#19760]*

Riheng Yao, Qiudan Li, Lei Wang and Daniel Dajun Zeng, Institute of Automation, Chinese Academy of Sciences;University of Chinese Academy of Sciences, China; Institute of Automation, Chinese Academy of Sciences, China; Beijing Wenge Technology Co., Ltd., China

News media has become a prevalent information spreading platform, where news sites can reprint news from other sites. To better understand the mechanism of news propagation, it is necessary to model reprint behavior and predict whether a news site will reprint a piece of news. Most existing works in news reprint analysis focus on analyzing the semantic of news content, little work has been done on integrating reprint relationship among sites and news content for reprint prediction from the perspective of sites. The challenge of improving prediction performance lies in how to effectively incorporate these two kinds of information to learn a more comprehensive reprint behavior model. In this paper, we propose an Integrated Neural Reprint Prediction (INRP) model that considers both reprint relationship and news content. It models the reprint relationships as a directed weighted graph and maps it into a latent space to learn sites representations. During news content modeling process, sites representations are embedded as attention guidance to build up more site-specific content representations. Finally, sites and news representations are jointly modeled to predict whether a piece of news will be reprinted by a site. We empirically evaluate the performance of the proposed model on a real world dataset. Experimental results show that taking both the reprint relationship and news content information into consideration could allow us make more accurate analysis of reprint patterns. The mined patterns could serve as a feedback channel for both corporations and management departments.

1:50PM *Self-Supervised Deep Learning for Retinal Vessel Segmentation Using Automatically Generated Labels from Multimodal Data [#20055]*

Alvaro S. Hervella, Jose Rouco, Jorge Novo and Marcos Ortega, Universidade da Coruna, Spain

This paper presents a novel approach that allows training convolutional neural networks for retinal vessel segmentation without manually annotated labels. In order to learn how to segment the retinal vessels, convolutional neural networks are typically trained with a set of pixel-level labels annotated by a clinical expert. This annotation is a tedious and error-prone task that limits the number of available training samples. To alleviate this problem, we propose the use of unlabeled multimodal data for learning about the retinal vasculature. Instead of using manually annotated labels, the networks learn to segment the retinal vessels from a complementary image modality where the vasculature is already highlighted. In this complementary modality, a vessel map can be easily constructed with simple image processing techniques. Then, a convolutional neural network is trained to learn the cross-modal mapping from the original modality to the automatically derived vessel maps. Using this strategy, the supervisory signal for training is automatically obtained from the unlabeled multimodal data. Thus, the number of training samples can be increased without any human annotation effort. Several experiments were conducted to evaluate the performance of the networks that were trained with the automatically derived labels, obtaining competitive results for retinal vessel segmentation in relevant public datasets. Furthermore, the results are promising towards including the presented approach in semi-supervised methods.

2:10PM *Deep Multimodal Reconstruction of Retinal Images Using Paired or Unpaired Data [#20220]*

Alvaro S. Hervella, Jose Rouco, Jorge Novo and Marcos Ortega, Universidade da Coruna, Spain

This paper explores the application of deep learning-based methods for the multimodal reconstruction of fluorescein angiography from retinography. The objective of this multimodal reconstruction is not only to estimate an invasive modality from a non-invasive one, but also to apply the learned models for transfer learning or domain adaption. Deep neural networks have demonstrated to be successful at learning the mapping between complementary image domains, using both paired or unpaired data. The paired data allows taking advantage of the rich information that is available from the pixelwise correspondence of paired images. However, this requires the pre-registration of the multimodal image pairs. In the case of the retinal images, the multimodal registration is a challenging task that may fail in complex scenarios, such as severe pathological cases or low quality samples. In contrast, the use of generative adversarial networks allows learning the mapping between image domains using unpaired data. This avoids the pre-registration of the images and allows including all the available data for training. In this work, we analyze both paired and unpaired deep learning-based approaches for the multimodal reconstruction of retinal images. The objective is to understand the implications of each alternative and the considerations for their future usage. For that purpose, we perform several experiments with the focus on producing a fair comparison between paired and unpaired approaches.

2:30PM *Adversarial Attacks on Remote User Authentication Using Behavioural Mouse Dynamics [#19711]*

Yi Xiang Marcus Tan, Alfonso Iacovazzi, Ivan Homoliak, Yuval Elovici and Alexander Binder, ST Engineering Electronics-SUTD Cyber Security Laboratory, Singapore

Mouse dynamics is a potential means of authenticating users. Typically, the authentication process is based on classical machine learning techniques, but recently, deep learning techniques have been introduced for this purpose. Although prior research has demonstrated how machine learning and deep learning algorithms can be bypassed by carefully crafted adversarial samples, there has been very little research performed on the topic of behavioural biometrics in the adversarial domain. In an attempt to address this gap, we built a set of attacks, which are applications of several generative approaches, to construct adversarial mouse trajectories that bypass authentication models. These generated mouse sequences will serve as the adversarial samples in the context of our experiments. We also present an analysis of the attack approaches we explored, explaining their limitations. In contrast to previous work, we consider the attacks in a more realistic and challenging setting in which an attacker has access to recorded user data but does not have access to the authentication model or its outputs. We explore three different attack strategies: 1) statistics-based, 2) imitation-based, and 3) surrogate-based; we show that they are able to evade the functionality of the authentication models, thereby impacting their robustness adversely. We show that imitation-based attacks often perform better than surrogate-based attacks, unless, however, the attacker can guess the architecture of the authentication model. In such cases, we propose a potential detection mechanism against surrogate-based attacks.

2:50PM *Predicting Parkinson's Disease using Latent Information extracted from Deep Neural Networks [#19909]*
Ilianna Kollia, Andreas-Georgios Stafylopatis and Stefanos Kollias, IBM Hellas, Greece; National Technical University of Athens, Greece; University of Lincoln, United Kingdom

This paper presents a new method for medical diagnosis of neurodegenerative diseases, such as Parkinson's, by extracting and using latent information from trained Deep convolutional, or convolutional-recurrent Neural Networks (DNNs). In particular, our approach adopts a combination of transfer learning, k-means clustering and k-Nearest Neighbour classification of deep neural network learned representations to provide enriched prediction of the disease based on MRI and/or DaT Scan data. A new loss function is introduced and used in the training of the DNNs, so as to perform adaptation of the generated learned representations between data from different medical environments. Results are presented using a recently published database of Parkinson's related information, which was generated and evaluated in a hospital environment.

3:10PM *Joint Graph Based Embedding and Feature Weighting for Image Classification [#20116]*
Ruifeng Zhu, Fadi Dornaika and Yassine Ruichek, Laboratory of Electronics, Information and Image(LE2I), CNRS,University of Bourgogne Franche-Comte, Belfort, France, France; Faculty of Computer Science, University of Basque Country San Sebastian, Spain, Spain

The graph-based embedding is an effective and useful method in reducing the dimension and extracting relevant data. This paper introduces a framework for classifying high dimensional data via a joint graph-based embedding and weighting method which could be used in semi-supervised or supervised learning. We design on effective optimization algorithm to solve the objective function. Experiments on image classification show that our proposed method can have a performance that is better than that of many state-of-the-art methods including linear and nonlinear methods.

2e: Deep learning

Tuesday, July 16, 1:30PM-3:30PM, Room: Panorama II, Chair: Ricardo Araujo

1:30PM *Combining Street-level and Aerial Images for Dengue Incidence Rate Estimation [#20173]*
Virginia Andersson, Cristian Cechinel and Ricardo Araujo, PPGC-UFPel, Brazil

The identification of urban locations with a high risk of diseases infections is a central aspect of public policies aiming at controlling these diseases. The presence of diseases, such as dengue fever, can be attributed to environmental factors in the urban scenario. Previous works have leveraged street-level imagery to provide estimates of dengue rates in an urban setting. In this paper, we apply Dense Deep Convolutional Neural Networks to both street-level and aerial imagery, providing evidence that aerial photography can provide better results than street-level images alone, while combining both leads to further improvements.

1:50PM *Vehicle Re-identification: an Efficient Baseline Using Triplet Embedding [#20382]*
Ratnesh Kumar, Edwin Weill, Farzin Aghdasi and Parthasarathy Sriram, NVIDIA, United States

In this paper we tackle the problem of vehicle re- identification in a camera network utilizing triplet embeddings. Re-identification is the problem of matching appearances of objects across different cameras. With the proliferation of surveillance cameras enabling smart and safer cities, there is an ever- increasing need to re-identify vehicles across cameras. Typical challenges arising in smart city scenarios include variations of viewpoints, illumination and self occlusions. Most successful approaches for re-identification involve (deep) learning an embedding space such that the vehicles of same identities are projected closer to one another, compared to the vehicles representing different identities. Popular loss functions for learning an embedding (space) include contrastive or triplet loss. In this paper we provide an extensive evaluation of these losses applied to vehicle re-identification and demonstrate that using the best practices for learning embeddings outperform most of the previous approaches proposed in the vehicle re-identification literature. Compared to most existing state-of-the-art approaches, our approach is simpler and more straightforward for training utilizing only identity-level annotations, along with one of the smallest published embedding dimensions for efficient inference. Furthermore in this work we introduce a formal evaluation of a triplet sampling variant (batch sample) into the re-identification literature.

2:10PM *ConvTimeNet: A Pre-trained Deep Convolutional Neural Network for Time Series Classification [#20439]*
Kathan Kashiparekh, Jyoti Narwariya, Pankaj Malhotra, Lovekesh Vig and Gautam Shroff, BITS-Pilani Goa Campus, Goa, India; TCS Research, New Delhi, India

Training deep neural networks often requires careful hyper-parameter tuning and significant computational resources. In this paper, we propose ConvTimeNet (CTN): an off-the-shelf deep convolutional neural network (CNN) trained on diverse univariate time series classification (TSC) source tasks. Once trained, CTN can be easily adapted to new TSC target tasks via a small amount of fine-tuning using labeled instances from the target tasks. We note that the length of convolutional filters is a key aspect when building a pre-trained model that can generalize to time series of different lengths across datasets. To achieve this, we incorporate filters of multiple lengths in all convolutional layers of CTN to capture temporal features at multiple time scales. We consider all 65 datasets with time series of lengths up to 512 points from the UCR TSC Benchmark for training and testing transferability of CTN: We train CTN on a randomly chosen subset of 24 datasets using a multi-head approach with a different softmax layer for each training dataset, and study generalizability and transferability of the learned filters on the remaining 41 TSC datasets. We observe significant gains in classification accuracy as well as computational efficiency when using pre-trained CTN as a starting point for subsequent task-specific fine-tuning compared to existing state-of-the-art TSC approaches.

2:30PM *Exploring Transferability in Deep Neural Networks with Functional Data Analysis and Spatial Statistics [#19869]*
Richard McAllister and John Sheppard, Montana State University, United States

Recent advances in machine learning have brought with them considerable attention in applying such methods to complex prediction problems. However, in extremely large dataspaces, a single neural network covering that space may not be effective, and generating large numbers of deep neural networks is not feasible. In this paper, we analyze deep networks trained from stacked autoencoders in a spatio-temporal application area to determine the extent to which knowledge can be transferred to similar regions. Our analysis applies methods from functional data analysis and spatial statistics to identify such correlation. We apply this work in the context of numerical weather prediction in analyzing large-scale data from Hurricane

Sandy. Results of our analysis indicate high likelihood that spatial correlation can be exploited if it can be identified prior to training.

2:50PM *Towards Optimizing Convolutional Neural Networks for Robotic Surgery Skill Evaluation [#20109]*

Dayvid Castro, Danilo Pereira, Cleber Zanchettin, David Macedo and Byron Bezerra, Federal University of Pernambuco, Brazil; University of Pernambuco, Brazil

In medicine courses, improve the skills of surgery students is an essential part of the program. For training the surgeon residents the institutions normally using a standard checklist to evaluate the student evolution. However, the checklist evaluation is susceptible to evaluator bias, inter-evaluator variability, besides being time-consuming. The automation of this process is an important evolution in medical training. An alternative to the instructor checklist is capturing and evaluation of kinematic data regarding the surgical motion. We propose a novel CNN architecture for automated robot-assisted skill assessment. We explore the use of the SELU activation function and a global mixed pooling approach based on the average and max-pooling layers. Finally, we examine two types of convolutional layers: real-value and quaternion-valued. The results suggest that our model presents a higher average accuracy across the three surgical sub-tasks of the JIGSAWS dataset.

3:10PM *Improving Universal Language Model Fine-Tuning using Attention Mechanism [#20204]*

Flavio Santos, Karina Guevara, David Macedo and Cleber Zanchettin, Universidade Federal de Pernambuco, Brazil

Inductive transfer learning is widespread in computer vision applications. However, in natural language processing (NLP) applications is still an under-explored area. The most common transfer learning method in NLP is the use of pre-trained word embeddings. The Universal Language Model Fine-Tuning (ULMFiT) is a recent approach which proposes to train a language model and transfer its knowledge to a final classifier. During the classification step, ULMFiT uses a max and average pooling layer to select the useful information of an embedding sequence. We propose to replace max and average pooling layers with a soft attention mechanism. The goal is to learn the most important information of the embedding sequence rather than assuming that they are max and average values. We evaluate the proposed approach in six datasets and achieve the best performance in all of them against literature approaches.

S03: Computational/Artificial Intelligence in Earth, Space, and Environmental Sciences

Tuesday, July 16, 1:30PM-3:30PM, Room: Panorama III, Chair: Vladimir Krasnopolsky

1:30PM *Classification of Stars using Stellar Spectra collected by the Sloan Digital Sky Survey [#19482]*

Michael Brice and Razvan Andonie, Central Washington University, United States

The classification of stellar spectra is a fundamental task in stellar astrophysics. There have been many explorations into the automated classification of stellar spectra but few that involve the Sloan Digital Sky Survey (SDSS). We use the SDSS dataset since it is the most important stellar spectra database available today. In our approach, we apply redshift corrections to the spectra and reduce the number of flux measurements by feature selection. Then we apply standard classifier methods: Random Forest and Support Vector Machine. We compare the accuracy of feature selection and classifier combinations for redshifted stellar spectra and rest stellar spectra. Even though redshifted stellar spectra create feature matrix discrepancies, classifiers utilizing redshifted stellar spectra perform with high accuracy. This creates a viable option for automated classification of stellar spectra without having to identify the redshift value.

1:50PM *Machine Learning Approaches for Predicting the 10.7 cm Radio Flux from Solar Magnetogram Data [#19557]*

Julio J. Valdes, Ljubomir Nikolic and Kenneth Tapping, National Research Council Canada, Canada; Natural Resources Canada, Canada

Using solar magetogram data, we explore potential of machine learning in space weather forecasting. In particular, unsupervised and supervised machine learning techniques are used to investigate the structure of magnetograms for 2006-2018, and their relation with the 10.7 cm solar radio flux. The similarity structure of the magnetograms is characterized with perceptionbased state of the art measures (the MSSIM index) and it was found that the data are contained in a space of intrinsically low dimension. The properties of these spaces were explored with methods preserving both local dissimilarity relationships, as well as conditional probability distributions within neighbourhoods. They reveal a clear relation with the intensity of the 10.7 cm flux. The flux was modeled using data driven supervised approaches in the form of model trees and convolutional neural networks. Models were found that allow prediction of the 10.7 cm radio flux with high accuracy. The results demonstrate significant potential which machine learning has in the space weather field.

2:10PM *A Deep Learning based architecture for rainfall estimation integrating heterogeneous data sources [#20255]*

Folino Gianluigi, Guarascio Massimo, Chiaravalloti Francesco and Gabriele Salvatore, ICAR-CNR, Italy; IRPI-CNR, Italy

Rain gauges are sensors providing direct measurement of precipitation intensity at individual point sites, and, usually, spatial interpolation methods are used to obtain an estimate of the precipitation field over the entire area of interest. Among them, Kriging with External Drift (KED) is a largely used and well-recognized method in this field. However, interpolation methods need to work with real-time data, and therefore can be hardly used in real-time scenarios. To overcome this issue, we propose a general machine learning framework, which can be trained offline, based on a deep learning architecture, also integrating information derived from remote sensing measurements such as weather radars and satellites. The framework allows to provide accurate estimations of the rainfall in the areas where no rain gauge data is available. Experimental results, conducted on real data concerning a southern region in Italy, provided by the Department of Civil Protection (DCP), show significant improvement in comparison with KED and other machine learning techniques.

2:30PM *Unsupervised Change Detection in Satellite Images Using Convolutional Neural Networks [#19124]*

Kevin Louis de Jong and Anna Sergeevna Bosman, University of Pretoria, South Africa

This paper proposes an efficient unsupervised method for detecting relevant changes between two temporally different images of the same scene. A convolutional neural network (CNN) for semantic segmentation is implemented to extract compressed image features, as well as to classify the detected changes into the correct semantic classes. A difference image is created using the feature map information generated by the CNN, without explicitly training on target difference images. Thus, the proposed change detection method is unsupervised, and can be performed using any CNN model pre-trained for semantic segmentation.

2:50PM *Deep Reinforcement Learning with Dual Targeting Algorithm [#20200]*

Naoki Kodama, Taku Harada and Kazuteru Miyazaki, Tokyo University of Science, Japan; National Institution for Academic Degrees and Quality Enhancement of Higher Education, Japan

Recently, deep reinforcement learning using the Deep Q-networks (DQN) algorithm has attracted attention, and extended methods continue to improve its learning performance. A multi-step DQN using an n-step TD method in the extended method contributes to faster learning. However, in the n-step TD methods, improvement in learning speed is better when using the intermediate prediction over the long-term prediction. Therefore, to further accelerate learning, methods that can use a long- term prediction effectively are required. A learning-accelerated DQN learns faster than DQN through a training neural network with bootstrap targets up to the next positive reward and 1-step bootstrap targets. It is, however, not possible for that method to use long-term prediction for tasks in which rewards are continuously observed. Furthermore, the use of two independent updates leads to instability with respect to the convergence of the neural network. We therefore propose a dual targeting algorithm that uses a single update with bootstrap targets up to the last reward in the next consecutive positive reward and 1-step bootstrap targets. The aim of the proposed method is to reduce instability in the convergence of the neural network by calculating the dual target from the same sampled experience. We apply the proposed method to a few classic control problems involving OpenAI Gym, compare it with DQN and multi-step DQN, and verify its effectiveness.

3:10PM *Fine-Grained Road Mining from Satellite Images with Bilateral Xception and DeepLab [#19272]*

Lele Cao, Activision Blizzard Group, Sweden

With the recent development of remote sensing and deep learning techniques, automatic and robust road extraction from satellite imaging data has become one of the most popular topics in both fields of Geographic Information System (GIS) and Computer Vision. Despite of the superior performance of Convolutional Neural Networks (DCNNs), a common problem of choosing between the classification and segmentation DCNNs still remains. By comparing two state-of-the-art baseline classification/segmentation DCNNs in several industrial application scenarios, we illustrate that their relative performance may vary, leading to different choices. Based on that observation, we propose a general fusion strategy that conveniently combines the strength of both classification and segmentation DCNNs using an end-to-end network architecture; this paradigm only requires pre-train segmentation/classification DCNNs once, which then can be reused in different road feature mining tasks. The task-specific experiments show that our fusion strategy guarantees superior results in all tested industrial scenarios.

2p: Feature selection, extraction, and aggregation

Tuesday, July 16, 1:30PM-3:30PM, Room: Panorama IV, Chair: Robi Pollikar

1:30PM *Feature Selection via Mutual Information: New Theoretical Insights [#19832]*

Mario Beraha, Alberto Maria Metelli, Matteo Papini, Andrea Tirinzoni and Marcello Restelli, Politecnico di Milano \ Universita degli Studi di Bologna, Italy; Politecnico di Milano, Italy

Mutual information has been successfully adopted in filter feature-selection methods to assess both the relevancy of a subset of features in predicting the target variable and the redundancy with respect to other variables. However, existing algorithms are mostly heuristic and do not offer any guarantee on the proposed solution. In this paper, we provide novel theoretical results showing that conditional mutual information naturally arises when bounding the ideal regression/classification errors achieved by different subsets of features. Leveraging on these insights, we propose a novel stopping condition for backward and forward greedy methods which ensures that the ideal prediction error using the selected feature subset remains bounded by a user-specified threshold. We provide numerical simulations to support our theoretical claims and compare to common heuristic methods.

1:50PM *Locality Preserving Projection via Deep Neural Network [#19191]*

Tianhang Long, Junbin Gao, Mingyan Yang, Yongli Hu and Baocai Yin, Beijing University of Technology, China; The University of Sydney, Australia; Xi'an Jiaotong University, China; Dalian University of Technology, China

Dimensionality reduction is an essential problem in data mining and machine learning fields. Locality Preserving Projection (LPP) is a well-known dimensionality reduction method which can preserve the neighborhood graph structure of data, and has achieved promising performance. However linear projection makes it difficult to analyze complex data with nonlinear structure.In order to deal with this issue, this paper proposes a novel nonlinear locality preserving projection method via deep neural network, termed as DNLPP, which replaces the linear projection with an appropriate deep neural network.Benefiting from the nonlinearity of neural networks and its powerful representation capability, the proposed method is more discriminative than the conventional LPP. In order to solve the new model, we propose an iterative optimization algorithm. Extensive experiments on several public datasets illustrate that the proposed method is overall superior to the other state-of-art dimensionality reduction methods.

2:10PM *Probabilistic Margin-Aware Multi-Label Feature Selection by Preserving Spatial Consistency [#20394]*

Yu Yin, Shuai An, Jun Wang, Jinmao Wei and Jianhua Ruan, College of Computer Science, Nankai University, China; Smart Supply Chain Y Bu, JD.com, China; College of Mathematics and Statistics Science,Ludong University, China; College of Computer Science, KLMDASR, Nankai University, China; Department of Computer Science, University of Texas at San Antonio, United States

Multi-label feature selection focuses on constructing a reduced feature space for discriminating multi-label instances. In consideration of the complex structures of label and feature spaces, a critical issue that explicitly determines selection performance is how to induce consistent information from both spaces to steer feature selection. Existing approaches tackle this issue in various spatial-aware views, without sufficient consideration of the negative effects of irrelevant features and imbalanced neighbors on inferring space structure. Inspired by the superiority of margin theory in assessing reliable space structure, we approach multi-label feature selection in the learning framework of preserving label-feature space consistency through probabilistic margin in this paper. In contrast to existing approaches, our model assesses the weighted margin based on the probabilistic nearest neighbors, and preserves consistent margin information in label and feature spaces. In this manner, label-feature space consistency is elegantly achieved, which conduces to effectively capturing discriminative features suitable for multi-label learning tasks and eliminating noisy features. Experimental results on multi-label data sets demonstrate the encouraging performance of the proposed model.

2:30PM *Efficient Estimation of Node Representations in Large Graphs using Linear Contexts [#20321]*
Tiago Pimentel, Rafael Castro, Adriano Veloso and Nivio Ziviani, Kunumi, Brazil; Universidade Federal de Minas Gerais, Brazil

Learning distributed representations in graphs has a rising interest in the neural network community. Recent works have proposed new methods for learning low dimensional embeddings of nodes and edges in graphs and networks. Several of these methods rely on the SkipGram algorithm to learn distributed representations, and they usually process a large number of multi-hop neighbors in order to produce the context from which node representations are learned. This is a limiting factor for these methods as graphs and networks keep growing in size. In this paper, we propose a simple alternate method which is as effective as previous methods, but being much faster at learning node representations. Our proposed method employs a restricted number of permutations over the immediate neighborhood of a node as context to generate its representation, thus avoiding long walks and large contexts while learning the representations. We present a thorough evaluation showing that our method outperforms state-of-the-art methods in six different datasets related to the problems of link prediction and node classification, being one to three orders of magnitude faster than baselines when generating node embeddings for very large graphs.

2:50PM *A Kernel Discriminant Information Approach to Non-linear Feature Selection [#19938]*
Hou Zejiang and Kung Sun-Yuan, Princeton University, United States

Feature selection has become a de facto tool for analyzing high-dimensional data, especially in bioinformatics. It is effective in improving learning algorithms' scalability and facilitating feature generalization or interpretability by removing noise and redundancy. Our focus is placed on the paradigm of supervised feature selection, which aims to find an optimal feature subset to best predict the target. We propose a nonlinear approach for finding a feature subset that achieves the highest inter-class separability in terms of the kernel Discriminant Information (KDI) measure. Theoretically, we prove the existence of good prediction hypotheses for feature subsets with high KDI value. We also establish the equivalency between maximizing the KDI statistic and minimizing a functional dependency measure of label variable on data. Moreover, we asymptotically prove the concentration property of the optimal feature subset found by maximizing the KDI measure. Practically, we provide an efficient gradient optimization algorithm for solving the KDI feature selection problem. We evaluate the proposed method based on 19 benchmark datasets in various domains, and demonstrates a noticeable improvement against state-of-the-art baselines on the majority of classification and regression tasks. Notably, our method is robust to the choice of hyper- parameters, works well with various downstream classifiers, has competitive computational complexity among the kernel based methods considered, and scales well the large-scale object recognition dataset, with generalization enhancement on CIFAR.

3:10PM *Distributed and Randomized Tensor Train Decomposition for Feature Extraction [#20320]*
Krzysztof Fonal and Rafal Zdunek, Wroclaw University of Science and Technology, Poland

Matrix factorization and their multi-linear extensions, known as tensor factorizations are widely known and useful methods in data analysis and machine learning for feature extraction and dimensionality reduction. Recently, new approaches to factorization models appeared -- tensor networks (TN) factorizations. They reduce storage, computational complexity, and aim to help with curse of dimensionality in decomposing multi-way data. Tensor train (TT) is one of the most popular TN models used in wide-range areas, such as quantum physics or chemistry. In this study, we improved TTs for classification tasks by combining the fundamental TT model with randomized decompositions and extending it to a distributed version according to the MapReduce paradigm. As a result, the proposed approach is not only scalable but also much faster than competing algorithms, and is able to perform large-scale dimensionality reduction, e.g. in classification tasks.

Competition: AutoCV Challenge
Tuesday, July 16, 1:30PM-3:30PM, Room: Panorama V, Chair: Yao Quanming, Wang Mengshuo, Hugo Jair Escalante, Isabelle Guyon Wei-Wei Tu

Tuesday, July 16, 3:30PM-4:00PM

Special Lecture: Coffee Break
Tuesday, July 16, 3:30PM-4:00PM, Room: Pre-function area Intercontinental

Tuesday, July 16, 4:00PM-5:00PM

Plenary Talk: Vera Kurkova, Institute of Computer science, Czech academy of sciences
Tuesday, July 16, 4:00PM-5:00PM, Room: Ballroom I + II +II, Chair: Irwin King,

Tuesday, July 16, 5:30PM-7:30PM

1n: Other topics in artificial neural networks

Tuesday, July 16, 5:30PM-7:30PM, Room: Ballroom I, Chair: Xiao Li

5:30PM *Fusion Strategies for Learning User Embeddings with Neural Networks [#19537]*

Philipp Blandfort, Tushar Karayil, Federico Raue, Joern Hees and Andreas Dengel, TUK and DFKI, Germany; DFKI, Germany

Growing amounts of online user data motivate the need for automated processing techniques. In case of user ratings, one interesting option is to use neural networks for learning to predict ratings given an item and a user. While training for prediction, such an approach at the same time learns to map each user to a vector, a so-called user embedding. Such embeddings can for example be valuable for estimating user similarity. However, there are various ways how item and user information can be combined in neural networks, and it is unclear how the way of combining affects the resulting embeddings. In this paper, we run an experiment on movie ratings data, where we analyze the effect on embedding quality caused by several fusion strategies in neural networks. For evaluating embedding quality, we propose a novel measure, Pair-Distance Correlation, which quantifies the condition that similar users should have similar embedding vectors. We find that the fusion strategy affects results in terms of both prediction performance and embedding quality. Surprisingly, we find that prediction performance not necessarily reflects embedding quality. This suggests that if embeddings are of interest, the common tendency to select models based on their prediction ability should be reconsidered.

5:50PM *Gated Sequential Recommendation with Dynamic Memory Network [#19267]*

Yunxiao Li, Jiaxing Song, Xiao Li and Weidong Liu, Computer science and Technology Department of Tsinghua University, China

Recommender systems provide users with a ranked list of items based on users' historical behavior. Existing methods usually embed users with their temporal dynamics, which may have lost their long-term, intrinsic preferences. RNN is widely adopted to model users' behaviors for its expressiveness on modeling sequential information. However, RNN-based methods are insufficient to capture users' intrinsic preferences due to short-term dependency issues. In this paper, we aim to make progress from two perspectives: 1. extract users' intrinsic preferences from their historical behaviors more effectively and robustly; 2. explore a fine-tuning way to combine users' intrinsic preferences and external dynamic interests. To do so, we propose a modification of Dynamic Memory Network with gating mechanism for generating a dynamic, precise representation for individual users. We also design an automatic and implicit method, and it utilizes the intrinsic preferences as a trigger signal to search the input sequences and retrieves the temporal dynamics of users in the memory module. As a result, a desirable combination of temporal dynamics and intrinsic preferences is achieved. Extensive experiments on three benchmark datasets have shown our model's superiority over the existing baseline method. Further analysis of the attention weights of each hop demonstrated the effectiveness of the proposed method.

6:10PM *Preempting Catastrophic Forgetting in Continual Learning Models by Anticipatory Regularization [#19508]*

Alaa El Khatib and Fakhri Karray, University of Waterloo, Canada

Neural networks trained on tasks sequentially tend to degrade in performance, on the average, the more tasks they see, as the representations learned for one task get progressively modified while learning subsequent tasks. This phenomenon - known as catastrophic forgetting - is a major obstacle on the road toward designing agents that can continually learn new concepts and tasks the way, say, humans do. A common approach to containing catastrophic forgetting is to use regularization to slow down learning on weights deemed important to previously learned tasks. We argue in this paper that, on their own, such post hoc measures to safeguard what has been learned can, even in their more sophisticated variants, paralyze the network and degrade its capacity to learn and counter forgetting as the number of tasks learned increases. We propose instead - or possibly in conjunction - that, in anticipation of future tasks, regularization be applied to drive the optimization of network weights toward reusable solutions. We show that one way to achieve this is through an auxiliary unsupervised reconstruction loss that encourages the learned representations not only to be useful for solving, say, the current classification task, but also to reflect the content of the data being processed - content that is generally richer than it is discriminative for any one task. We compare our approach to the recent elastic weight consolidation regularization approach, and show that, although we do not explicitly try to preserve important weights or pass on any information about the data distribution of learned tasks, our model is comparable in performance, and in some cases better.

6:30PM *Faster Training by Selecting Samples Using Embeddings [#19361]*

Santiago Gonzalez, Joshua Landgraf and Risto Miikkulainen, University of Texas at Austin, United States

Long training times have increasingly become a burden for researchers by slowing down the pace of innovation, with some models taking days or weeks to train. In this paper, a new, general technique is presented that aims to speed up the training process by using a thinned-down training dataset. By leveraging autoencoders and the unique properties of embedding spaces, we are able to filter training datasets to only include only the samples that matter the most. Through evaluation on a standard CIFAR-10 image classification task, this technique is shown to be effective. With this technique, training times can be reduced with a minimal loss in accuracy. Conversely, given a fixed training time budget, the technique was shown to improve accuracy by over 50%. This intelligent dataset sampling technique is a practical tool for achieving better results with large datasets and limited computational budgets.

6:50PM *Detecting Adversarial Perturbations Through Spatial Behavior in Activation Spaces [#20169]*

Ziv Katzir and Yuval Elovici, Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Israel

Although neural network-based classifiers outperform humans in a range of tasks, they are still prone to manipulation through adversarial perturbations. Prior research has resulted in the identification of effective defense mechanisms for many reported attack methods, however a defense against the C&W attack, as well as a holistic defense mechanism capable of countering multiple different attack methods, are still missing. All attack methods reported so far share a common goal. They aim to avoid detection by limiting the allowed perturbation magnitude, and still trigger incorrect classification. As a result, small perturbations cause classification to shift from one class to another. We coined the term activation spaces to refer to the hyperspaces formed by the activation values of the different network layers. We then use activation spaces to capture the differences in spatial dynamics between normal and adversarial examples, and form a novel adversarial example detector. We induce a set of k-nearest neighbor (k-NN)

classifiers, one per activation space, and leverage those classifiers to assign a sequence of class labels to each input of the neural network. We then calculate the likelihood of each observed label sequence and show that sequences associated with adversarial examples are far less likely than those of normal examples. We demonstrate the efficiency of our proposed detector against the C&W attack using two image classification datasets (MNIST, CIFAR-10) achieving an AUC of 0.97 for the CIFAR-10 dataset. We further show how our detector can be easily augmented with previously suggested defense methods to form a holistic multi-purpose defense mechanism.

7:10PM *$$L2Q: An Ultra-Low Loss Quantization Method for DNN Compression [#19298]*

Cheng Gong, Tao Li, Ye Lu, Cong Hao, Xiaofan Zhang, Deming Chen and Yao Chen, Nankai University, China; University of Illinois at Urbana-Champaign, United States; Advanced Digital Sciences Center, Singapore

Data quantization has been proved to be an effective method to compress deep neural networks (DNNs) by using less bits to represent the parameters and intermediate data. The bit width of the data directly affects the memory footprint, computing capability, and energy consumption during the computation of the DNN models. Although there have been numerous existing studies on data quantization, there is still no quantitative analysis of the existing quantization methods, which results in empirical quantization with unpredictable DNN accuracy loss. To address this problem, we propose an effective method, called ultra-low loss quantization (μL2Q), to provide DNN quantization schemes based on comprehensive quantitative data analysis. μL2Q builds the transformation of the original data to a data space with standard normal distribution, and then find the optimal parameters to minimize the loss of the quantization of a targeted bit width. In addition, we integrate the proposed μL2Q into a popular machine learning framework Caffe for convenient end-to-end DNN design and training. By comparing to the state-of-the-art DNN compression designs, μL2Q shows the greatest ability to maintain DNN accuracy after quantization. In the experiments, our proposed method can deliver 4.42\%, 16.70\%, 1.95\%, 8.26\% and 5.63\% accuracy improvements on Lenet-5, Cifarnet, VGG7-64 and Resnet-18 (Top1/5), respectively, compared to the state-of-the-art solutions with the same compression ratio.

2e: Deep learning

Tuesday, July 16, 5:30PM-7:30PM, Room: Ballroom II, Chair: Arijit Ukil

5:30PM *A Robust Embedding Method for Anomaly Detection on Attributed Networks [#19252]*

Zhang Le, Yuan Jun, Liu Zeyi, Pei Yang and Wang Lei, Institute of Information Engineering, Chinese Academy of Sciences, China

Anomalies detection is to spot the objects whose patterns singularly differ from the reference majority. Recently, attributed networks, which often contain node attributes and network structure, are widely used for real-life applications. Meanwhile, how to detect anomalies on attributed networks has caused a lot of attention. Most existing works on anomaly detection attempt to incorporate node attributes with the network structure. However, there may exist structurally irrelevant attributes in attributed networks, which may have adverse effects on the detection results. Besides, heterogeneity of node attributes and network structure may further make the detection of anomalies difficult. In order to overcome the above challenges, in this paper, we propose a Robust Embedding Method for Anomaly Detection on Attributed Networks, called REMAD. Methodologically, the proposed REMAD combines network embedding and residual analysis together. By performing network embedding on the network, REMAD obtains the representative attributes that are closely coherent with the network structure. Simultaneously, by adopting the residual analysis, REMAD characterizes and analyzes the residuals of attribute information to discover anomalies in the network. Experimental results on both synthetic and real-world datasets demonstrate the advantages of our proposed REMAD against the state-of-the-art anomaly detection methods.

5:50PM *DyReg-FResNet: Unsupervised Feature Space Amplified Dynamic Regularized Residual Network for Time Series Classification [#20075]*

Arijit Ukil, Soma Bandyopadhyay and Arpan Pal, Tata Consultancy Services, India

Time Series Classification (TSC) is a challenging problem owing to the practical constraints of lack of availability of training examples and insufficient sample points in the training instances. In order to ensure the construction of a robust trained model (under practical constraints) to address TSC, we propose DyReg-FResNet, which is a dynamically regularized Residual Network (ResNet), amplified by unsupervised feature space training. We generate signal processing, information theoretic and statistical features to augment the representation learning of the ResNet. The unsupervised features are capable of extracting the morphological and structural characteristics of the time series signals, whereas our proposed dynamic regularizer trades off by reducing substantial variance while not perturbing the bias. The regularization factor is a function of the signal dynamics (in effect, regularization factor is different for different training sets). DyReg-FResNet learns through residual mapping to minimize the exploding or vanishing gradient problems, amplified unsupervised features amplify the representation space by introducing lower level representation to guide the learning towards the basins of attraction of minima and dynamic regularizer minimizes the generalization error. We extensively experiment with publicly available UCR time series datasets. DyReg-FResNet demonstrates extremely superior performance by consistently outperforming the existing benchmark results as well as current state-of- the-art algorithms.

6:10PM *A Crowdsourcing based Human-in-the-Loop Framework for Denoising UUs in Relation Extraction Tasks [#19795]*

Mengting Li, Jing Yang, Wen Wu, Liang He, Yan Yang and Jian Jin, East China Normal University, China

In relation extraction tasks, distant supervision methods expand dataset by aligning entity pairs in different knowledge bases and completing the relations between two entities. However, these methods ignore the fact that sentences' labels generated by distant supervision methods with high confidence are often incorrect in the real world called Unknown Unknowns (UUs). To deal with this challenge, we propose a crowdsourcing based human-in-the-loop denoising framework which iteratively discovers UUs and corrects them by crowdsourcing to better extract relations. During each epoch of iterations, we choose one sentence bag and repeat two steps: Firstly, attention based Long Short-Term Memory network is applied as a selector to discover potential UUs. Secondly, these UUs are annotated by crowdsourcing with two answer collecting strategies and fed back into selector as positive samples. Until the accuracy of selector reaches a threshold, all annotated samples are added into relation classifier as cleaned train set and framework moves on to next epoch with new sentence bags. The experiments on the New York Times dataset and analysis of potential UUs demonstrate that our framework denoise the dataset and outperforms all the baselines on distant supervision relation extraction tasks.

6:30PM *Attention-based Adversarial Training for Seamless Nudity Censorship [#20360]*

Gabriel Simoes, Jonatas Wehrmann and Rodrigo C. Barros, PUCRS, Brazil

The amount of digital pornographic content over the Internet grows daily and accessing such a content has become increasingly easier. Hence, there is a

real need for mechanisms that can protect particularly-vulnerable audiences (e.g., children) from browsing the web. Recently, object detection methods based on deep neural networks such as CNNs have improved the effectiveness and efficiency of identifying and blocking pornographic content. Even though improvements in detecting intimate parts have been significant, the occlusion of the content is still primarily done by either blurring or removing regions of the image in an intrusive fashion. A recent study has addressed the problem of censoring the pornographic content in a non-intrusive way by generating the so- called seamless censorship via cycle-consistent generative adversarial networks. Such an approach has managed to automatically add bikinis to naked women without explicit supervision or paired training data. In this paper, we extend that method by designing a novel cycle-consistency framework that leverages sensitive information from an attention-based multi-label convolutional neural network. We evaluate the quality of our novel generative model by conducting a web survey with over 1000 opinions regarding the resulting images from our method and from baseline approaches. Results of the survey show that our method considerably improves the state-of-the-art on the seamless censorship task. Index Terms-- adversarial training, attention, convolutional neural networks, deep learning, GANs, pornography censorship.

6:50PM *Bagging Adversarial Neural Networks for Domain Adaptation in Non-Stationary EEG [#20039]*

Haider Raza and Spyridon Samothrakis, School of Computer Science and Electronics Engineering, University of Essex, United Kingdom

A major issue in bringing real-world applications of machine learning outside the laboratory is the difference in the data distributions between training and testing stages or domains. The diverging statistical properties in different domains can lead to decay the prediction performance. The technical term for a change in the distribution of features is covariate shift, which also happens to be a common challenge in electroencephalogram (EEG) based brain-computer interface (BCI); this is due to the presence of non-stationarities in the EEG signals. It is also the case that collecting and labelling samples is expensive, resulting in small datasets that are not in tune with the ``big data'' spirit that is the characteristic of the era. In this paper, we introduce a new method that handles domain adaptation in small datasets; the method combines elements of unsupervised domain adaptation with ensemble methods. We evaluate on real-world datasets corresponding to motor-imagery detection (BCI competition 2008 dataset 2A). The method produces state of the art results.

7:10PM *Quantum-Inspired Neural Architecture Search [#20215]*

Daniela Szwarcman, Daniel Civitarese and Marley Vellasco, PUC-Rio, IBM-Research, Brazil; IBM-Research, Brazil; PUC-Rio, Brazil

Deep neural networks have gained attention in the last decade as significant progress has been made in a variety of tasks thanks to these new architectures. Most of the time, hand- designed networks are responsible for this incredible success. However, this engineering process demands considerable time and expert knowledge, which leads to an increasing interest in automating the design of deep architectures. Several new algorithms have been proposed to address the neural architecture search problem, but many of them require significant computational resources. Quantum-inspired evolutionary algorithms (QIEA) have their roots on quantum computing principles and present promising results in respect to faster convergence. In this work, we propose Q-NAS (Quantum-inspired Neural Architecture Search): a quantum-inspired algorithm to search for deep neural architectures by assembling substructures and optimizing some numerical hyperparameters. We present the first results applying Q-NAS on the CIFAR-10 dataset using only 20 K80 GPUs for about 50 hours. The obtained networks are relatively small (less than 20 layers) compared to other state-of-the-art models and achieve promising accuracies with considerably less computational cost than other NAS algorithms.

8a: Applications of deep networks

Tuesday, July 16, 5:30PM-7:30PM, Room: Ballroom III, Chair: Tarek Taha

5:30PM *Image steganography using texture features and GANs [#19445]*

Jinjing Huang, Shaoyin Cheng, Songhao Lou and Fan Jiang, University of Science and Technology of China, China

As steganography is the main practice of hidden writing, many deep neural networks are proposed to conceal secret information into images, whose invisibility and security are unsatisfactory. In this paper, we present an encoder-decoder framework with an adversarial discriminator to conceal messages or images into natural images. The message is embedded into QR code first which significantly improves the fault-tolerance. Considering the mean squared error (MSE) is not conducive to perfectly learn the invisible perturbations of cover images, we introduce a texture-based loss that is helpful to hide information into the complex texture regions of an image, improving the invisibility of hidden information. In addition, we design a truncated layer to cope with stego image distortions caused by data type conversion and a moment layer to train our model with varisized images. Finally, our experiments demonstrate that the proposed model improves the security and visual quality of stego images.

5:50PM *Spatial-Temporal Attention Network for Malware Detection Using Micro-architecture Features [#19638]*

Fang Li, Jinrong Han, Ziyuan Zhu and Dan Meng, Institute of Information Engineering, Chinese Academy of Sciences;School of Cyber Security, University of Chinese Academy of Sciences, China; Institute of Information Engineering, Chinese Academy of Sciences, China

Malware detection is an imperative topic in computer security, since an evolutional malware will cause serious damage to computer system and user privacy information security. In recent years, some researchers began to utilize low-level hardware micro-architecture features to detect malware, because these micro-architecture features are difficult for malware evasion. However, these methods always adopt a long sample length and can hardly identify non-signature malware. This situation will inevitably affect the detection efficiency and effectiveness. To solve the above problems, we first select system call instruction as a trigger point to extract low-level features for avoiding blindly collecting unrelated data continuously. Specifically, we use the General-Purpose Registers (GPRs) as features for malware detection. Each register has specific functions and changes of its content contain the action information which thus can be used to detect illegal behaviours. To improve detection efficiency, we then propose a resampling method to well present the spatial and temporal properties of GPRs. Finally, a novel deep learning model is designed to highlight correlations among GPRs for accurate malware detection. Experimental results achieved 99% of Accuracy and zero False Positive rate (FPr) using only a short sample length and can also identify non-signature malware.

6:10PM *An Attention-based Hybrid LSTM-CNN Model for Arrhythmias Classification [#19473]*

Fan Liu, Xingshe Zhou, Tianben Wang, Jinli Cao, Zhu Wang, Hua Wang and Yanchun Zhang, Northwestern Polytechnical University, China; La Trobe University, Australia; Victoria University, Australia; Victoria Universiry, Australia

Electrocardiogram (ECG) signal based arrhythmias classification is an important task in healthcare field. Based on domain knowledge and observation results from large scale data, we find that accurately classifying different types of arrhythmias relies on three key characteristics of ECG: overall variation trends, local variation features and their relative location. However, these key factors are not yet well studied by existing methods. To tackle this problem, we design an attention- based hybrid LSTM-CNN model which is comprised of a stacked bidirectional LSTM (SB-LSTM) and a two-dimensional CNN (TD-CNN). Specifically, SB-LSTM and TD-CNN are utilized to extract the overall variation trends and local features of ECG, respectively. Furthermore, we add a trend attention gate (TAG) to SB-LSTM, meanwhile, add a feature attention mechanism (FAM) and a location attention mechanism (LAM) to TD-CNN. Thus, the effects of important trends and features at key locations in ECG can be enhanced, which is conducive to obtaining a better understanding of the fluctuation pattern of ECG. Experimental results on the MIT-BIH arrhythmias dataset indicate that our model outperforms three state-of-the-art methods, and achieve 99.3% of accuracy, 99.6% of sensitivity and 98.1% of specificity, respectively.

6:30PM *Pain Assessment From Facial Expression: Neonatal Convolutional Neural Network (N-CNN) [#20348]*

Ghada Zamzmi, Rahul Paul, Dmitry Goldgof, Rangachar Kasturi and Yu Sun, University of South Florida, United States

The current standard for assessing neonatal pain is discontinuous and suffers from inter-observer variations, which can result in delayed intervention and inconsistent treatment of pain. Therefore, it is critical to address the shortcomings of the current standard and develop continuous and less subjective pain assessment tools. Convolutional Neural Networks have gained much popularity in the last decades due to the wide range of its successful applications in medical image analysis, object recognition, and emotion recognition. In this paper, we propose a Neonatal Convolutional Neural Network, designed and trained end-to-end to detect neonatal pain. We evaluated the proposed network in two data sets of neonates and compared its performance to the performance of ResNet architecture in the same data sets. Our proposed method outperformed ResNet in recognizing neonates' pain and achieved around 91.00\% accuracy. While further research is needed, our preliminary results suggest that the presented network can be used for automatic pain assessment, and possibly similar applications. It also suggests that the automatic recognition of neonatal pain provides a viable and more efficient alternative to the current standard of pain assessment.

6:50PM *A Hierarchical Convolutional Neural Network for Malware Classification [#20312]*

Daniel Gibert, Carles Mateu and Jordi Planes, University of Lleida, Spain

Malware detection and classification is a challenging problem and an active area of research. Particular challenges include how to best treat and preprocess malicious executables in order to feed machine learning algorithms. Novel approaches in the literature treat an executable as a sequence of bytes or as a sequence of assembly language instructions. However, in those approaches the hierarchical structure of programs is not taken into consideration. An executable exhibits various levels of spatial correlation. Adjacent code instructions are correlated spatially but that is not necessarily the case. Function calls and jump commands transfer the control of the program to a different point in the instruction stream. Furthermore, these discontinuities are maintained when treating the binary as a sequence of byte values. In addition, functions might be arranged randomly if addresses are correctly reorganized. To address these issues we propose a Hierarchical Convolutional Network (HCN) for malware classification. It has two levels of convolutional blocks applied at the mnemonic-level and at the function-level, enabling us to extract n-gram like features from both levels when constructing the malware representation. We validate our HCN method on the dataset released for the Microsoft Malware Classification Challenge, outperforming almost every deep learning method in the literature.

7:10PM *Novel Ceiling Neuron Model and its Applications [#19105]*

Rama Murthy Garimella, Dileep Munugoti and Anil Rayala, Mahindra Ecole Centrale, India; IIT Guwahati, India; IIIT Hyderabad, India

In this research paper, a novel neuronal model named "Ceiling Neuron" has been proposed. Learning algorithm and proof of convergence of such a model are discussed. An activation function which mimics the ceiling behavior is proposed for Multi-Layer Neural Networks. The dynamic nature of the proposed activation function is discussed. An associative memory model based on Hopfield network is constructed using the proposed ceiling neuron and dynamics of the network are also studied

2t: Topics in machine learning

Tuesday, July 16, 5:30PM-7:30PM, Room: Duna Salon I, Chair: Tayo Obafemi-Ajayi

5:30PM *Visualizing Time Series Data with Temporal Matching Based t-SNE [#20452]*

Kwan-yeung Wong and Fu-lai Chung, Dept. of Computing, Hong Kong Polytechnic University, Hong Kong

Interpreting time series data has always been a hot research topic for various applications, especially when the dimensionality of time series dataset keeps growing almost prohibitively as technology advances. There exist several dimensionality reduction techniques attempting to address the related problems, but the inherent nature of time series datasets usually involves factors including time and amplitude shifting and scaling, which could impact the trustworthiness of the visualization results. t- distributed Stochastic Neighbor Embedding (t-SNE) is considered as a highly effective machine learning algorithm for visualization and it is in fact a nonlinear dimensionality reduction technique tailored made to embed high-dimensional data in a low-dimensional space of two to three dimensions only for proper visualization. In view of the key problem of adopting t-SNE to visualize time series data, we propose to introduce two temporal matching metrics, namely, dynamic time warping (DTW) and angular metric for shape similarity (AMSS), for t-SNE to enhance its time series data visualization ability. Both metrics provide a more robust similarity metric for time series data so that the embedding process in t-SNE can be made more effective, as demonstrated by various data visualization experiments.

5:50PM *Subword Semantic Hashing for Intent Classification on Small Datasets [#19329]*

Kumar Shridhar, Ayushman Dash, Amit Sahu, Gustav Grund Pihlgren, Pedro Alonso, Vinaychandran Pondenkandath, Gyorgy Kovacs, Foteini Simistira and Marcus Liwicki, Technical University Kaiserslautern, Germany; MindGarage, Germany; Lulea Technical University, Sweden; University of Fribourg, Switzerland

In this paper, we introduce the use of Semantic Hashing as embedding for the task of Intent Classification and achieve state-of-the-art performance on

three frequently used benchmarks. Intent Classification on a small dataset is a challenging task for data-hungry state-of-the-art Deep Learning based systems. Semantic Hashing is an attempt to overcome such a challenge and learn robust text classification. Current word embedding based methods are dependent on vocabularies. One of the major drawbacks of such methods is out-of-vocabulary terms, especially when having small training datasets and using a wider vocabulary. This is the case in Intent Classification for chatbots, where typically small datasets are extracted from internet communication. Two problems arise with the use of internet communication. First, such datasets miss a lot of terms in the vocabulary to use word embeddings efficiently. Second, users frequently make spelling errors. Typically, the models for intent classification are not trained with spelling errors and it is difficult to think about ways in which users will make mistakes. Models depending on a word vocabulary will always face such issues. An ideal classifier should handle spelling errors inherently. With Semantic Hashing, we overcome these challenges and achieve state-of-the-art results on three datasets: Chatbot, Ask Ubuntu, and Web Applications. Our benchmarks are available online: https://github.com/kumar-shridhar/Know-Your-Intent

6:10PM *A Methodology for Neural Network Architectural Tuning Using Activation Occurrence Maps [#20206]*

Rafael Garcia, Alexandre Xavier Falcao, Alexandru C. Telea, Bruno Castro da Silva, Jim Torresen and Joao Luiz Dihl Comba, Universidade Federal do Rio Grande do Sul, Brazil; Universidade de Campinas, Brazil; University of Groningen, Netherlands; University of Oslo, Norway

Finding the ideal number of layers and size for each layer is a key challenge in deep neural network design. Two approaches for such networks exist: filter learning and architecture learning. While the first one starts with a given architecture and optimizes model weights, the second one aims to find the best architecture. Recently, several visual analytics (VA) techniques have been proposed to understand the behavior of a network, but few VA techniques support designers in architectural decisions. We propose a hybrid methodology based on VA to improve the architecture of a pre-trained network by reducing/increasing the size and number of layers. We introduce Activation Occurrence Maps that show how likely each image position of a convolutional kernel's output activates for a given class, and Class Selectivity Maps, that show the selectiveness of different positions in a kernel's output for a given label. Both maps help in the decision to drop kernels that do not significantly add to the network's performance, increase the size of a layer having too few kernels, and add extra layers to the model. The user interacts from the first to the last layer, and the network is retrained after each layer modification. We validate our approach with experiments in models trained with two widely-known image classification datasets and show how our method helps to make design decisions to improve or to simplify the architectures of such models.

6:30PM *Stochastic Resonance Enables BPP/log* Complexity and Universal Approximation in Analog Recurrent Neural Networks [#19260]*

Emmett Redd, A. Steven Younger and Tayo Obafemi-Ajayi, Missouri State University, United States

Stochastic resonance (SR) is a natural process that without limit increases the precision of signal measurements in biological and physical sciences. Most artificial neural networks (NNs) are implemented on digital computers of fixed-precision. A NN accessing universal approximation and a computational complexity class more powerful that of a Turing machine needs analog signals utilizing SR's limitless precision increase. This paper links an analog recurrent (AR) NN theorem, SR, BPP/log* (a physically realizable, super-Turing computation class), and universal approximation so NNs following them can be made computationally more powerful. An optical neural network mimicking chaos indicates super-Turing computation has been achieved. Additional tests are needed which can verify super-Turing computation, show its superiority, and demonstrate its practical benefits. Truly powerful cognitively inspired computation needs to access the combination of ARNNs, SR, super-Turing mathematical complexity, and universal approximation.

6:50PM *Accelerate Mini-batch Machine Learning Training With Dynamic Batch Size Fitting [#19462]*

Liu Baohua, Shen Wenfeng, Li Peng and Zhu Xin, Shanghai University, China; The University of Aizu, Japan

Mini-batch Stochastic Gradient Descent (MGD) is one of the most widely used methods in Machine Learning (ML) model training. Typically, before a training process starts, researchers should manually set a fixed batch size, which is a hyper-parameter indicating the size of the random slice of the whole dataset that is trained in a single iteration. In this paper, we propose a light- weight dynamic batch size fitting algorithm based on online efficient evaluation, which has the ability of automatically tuning batch size during the train process to reach a best-so-far efficiency, but with little overhead. The experimental results have demonstrated that the algorithm is more effective compared with the commonly used fixed settings.

7:10PM *Online Estimation of Multiple Dynamic Graphs in Pattern Sequences [#19335]*

Jimmy Gaudreault, Arunabh Saxena and Hideaki Shimazaki, Polytechnique Montreal, Canada; Indian Institute of Technology Bombay, India; Kyoto University / Honda Research Institute Japan, Japan

Sequences of correlated binary patterns can represent many time-series data including text, movies, and biological signals. These patterns may be described by weighted combinations of a few dominant structures that underpin specific interactions among the binary elements. To extract the dominant correlation structures and their contributions to generating data in a time-dependent manner, we model the dynamics of binary patterns using the state-space model of an Ising-type network that is composed of multiple undirected graphs. We provide a sequential Bayes algorithm to estimate the dynamics of weights on the graphs while gaining the graph structures online. This model can uncover overlapping graphs underlying the data better than a traditional orthogonal decomposition method, and outperforms an original time-dependent Ising model. We assess the performance of the method by simulated data, and demonstrate that spontaneous activity of cultured hippocampal neurons is represented by dynamics of multiple graphs.

Neuroengineering and Bio-inspired Systems

Tuesday, July 16, 5:30PM-7:30PM, Room: Duna Salon II, Chair: Malte Schilling

5:30PM *Numerical Analysis on Wave Dynamics in a Spin-Wave Reservoir for Machine Learning [#20170]*
Ryosho Nakane, Gouhei Tanaka and Akira Hirose, The University of Tokyo, Japan

Reservoir computing is a computational framework which is originally based on software recurrent neural networks and recently achieved with physical systems as well. In our previous paper [Nakane et al., IEEE ACCESS vol. 6, p. 4462, 2018], we have proposed a spin-wave-based reservoir computing device with multiple input/output electrodes, and have demonstrated its high generalization ability in the estimation of input-signal parameters performed by the spin-wave-based reservoir computing. To successfully execute many types of estimation tasks with machine learning, it is necessary to investigate fundamental properties of spin-wave- based reservoir computing, particularly the relation between its input and output. From this background, the purposes of this work are to demonstrate a different estimation task with pulse input signals and to analyze the properties of spin waves which have important roles in the task. We first describe our approach to obtain spin waves with the features useful for reservoir computing by considering the fundamental properties of spin waves and feasible device technologies. Then, we investigate detailed characteristics of locally-excited spin waves in a garnet film by micromagnetics simulation. Using the resultant spin waves, we demonstrate a pulse interval estimation task and find that high diversity in the time-sequential signals generated by the spin-wave-based reservoir is effectively achieved. The spin- wave- based device is a highly promising hardware for next-generation machine- learning electronics.

5:50PM *Setup of a Recurrent Neural Network as a Body Model for Solving Inverse and Forward Kinematics as well as Dynamics for a Redundant Manipulator [#20222]*
Malte Schilling, Center of Excellence `Cognitive Interaction Technology', Bielefeld University, Germany

An internal model of the own body can be assumed a fundamental and evolutionary-early representation as it is present throughout the animal kingdom. Such functional models are, on the one hand, required in motor control, for example solving the inverse kinematic or dynamic task in goal-directed movements or a forward task in ballistic movements. On the other hand, such models are recruited in cognitive tasks as are planning ahead or observation of actions of a conspecific. Here, we present a functional internal body model that is based on the Mean of Multiple Computation principle. For the first time such a model is completely realized in a recurrent neural network as necessary normalization steps are integrated into the neural model itself. Secondly, a dynamic extension is applied to the model. It is shown how the neural network solves a series of inverse tasks. Furthermore, emerging representation in transformational layers are analyzed that show a form of prototypical population-coding as found in place or direction cells.

6:10PM *Unsupervised Feature Learning for Visual Place Recognition in Changing Environments [#20281]*
Dongye Zhao, Bailu Si and Fengzhen Tang, State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, China; School of Systems Science, Beijing Normal University, China

Visual place recognition in changing environments is a challenging and critical task for autonomous robot navigation. Deep convolutional neural networks (ConvNets) have recently been used as efficient feature extractors and obtained excellent performance in place recognition. However the success of ConvNets' learning highly relies on the availability of large datasets with millions of labeled images, the collection of which is a tedious and costly burden. Thus we develop an unsupervised learning method (the siamese VisNet) to autonomously learn invariant features in changing environments from unlabeled images. The siamese VisNet has two identical branches of sub-networks. With a Hebbian-type of learning rule incorporating a trace of previous activity patterns, the siamese VisNet learns features with increasing invariance in changing environments from layer to layer. Experiments conducted on multiple datasets demonstrate the robustness of the siamese VisNet against viewpoint changes, appearance changes, and joint viewpoint-appearance changes. In addition, the siamese VisNet, with lower complexity in architecture, outperforms the state-of-the-art place recognition ConvNets such as the CaffeNet and the PlaceNet. The proposed siamese VisNet constitutes a biologically plausible yet efficient method for unsupervised place recognition.

6:30PM *Transparent Machine Education of Neural Networks for Swarm Shepherding Using Curriculum Design [#19140]*
Alexander Gee and Hussein Abbass, University of New South Wales, Australia

Swarm control is a difficult problem due to the need to guide a large number of agents simultaneously. We cast the problem as a shepherding problem, similar to biological dogs guiding a group of sheep towards a goal. The shepherd needs to deal with complex and dynamic environments and make decisions in order to direct the swarm from one location to another. In this paper, we design a novel curriculum to teach an artificial intelligence empowered agent to shepherd in the presence of the large state space associated with the shepherding problem and in a transparent manner. The results show that a properly designed curriculum could indeed enhance the speed of learning and the complexity of learnt behaviours.

6:50PM *A QoS-oriented Scheduling and Autoscaling Framework for Deep Learning [#19960]*
Sikai Xing, Shiyou Qian, Bin Cheng, Jian Cao, Guangtao Xue, Jiadi Yu, Yanmin Zhu and Minglu Li, Shanghai Jiao Tong University, China

Deep learning is popular in many areas, but users must manually specify the resource configuration when submitting deep learning training jobs, usually over-provisioning resources. This kind of unreasonable resource configuration method results in slow training and low resource utilization. Therefore, it would be more convenient and efficient if users only need to specify the quality of service (QoS) for their jobs, and then the resources will be autoconfigured to meet the QoS. To satisfy this demand, we present a QoS-oriented scheduling and autoscaling framework that schedules and autoscales deep learning training jobs in the Kubernetes cluster. This paper focuses on the most important QoS requirement for deep learning training jobs: deadline. The goal of the framework is to guarantee that as many jobs as possible can be accomplished before their specified deadlines. To reach this goal, the framework schedules deep learning jobs by implementing a heuristic scheduling policy based on resource status and job deadline, and autoscales resource configuration by exploiting a characteristic of deep learning jobs: the predictability of training time. This predictability is used to predict whether a job can be accomplished before its deadline and estimate appropriate resource configuration if necessary. We implemented the framework by modifying the default scheduler of Kubernetes and conducted experiments to evaluate its performance. The experiment results show that our proposed framework can improve both the completion rate and resource utilization.

7:10PM *BCI and Multimodal Feedback Based Attention Regulation for Lower Limb Rehabilitation. [#19716]*
Jiaxing Wang, Weiqun Wang, Zeng-Guang Hou, Weiguo Shi, Xu Liang, Shixin Ren, Liang Peng and Yanjie Zhou, State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, China

Both motor and cognitive function rehabilitation benefits can be improved significantly by patients' active participation. However, post-stroke patients, especially with attention deficit disorders, can hardly engage in training for a longer time. In order to improve patients' attention focused on the training, an attention regulation system based on the brain-machine interface (BCI) and multimodal feedback is proposed for post-stroke lower limb rehabilitation. First, an interactive speed-tracking riding game is designed to increase the training challenge and patients' neural engagement. The character's riding

speed, which is synchronized with patients' actual cycling speed, is displayed on the screen in real time. And patients' attention can further be enhanced when they try their best to track the reference speed curve. Second, an attention classifier is designed and trained by using subjects' EEG signals, which are acquired if they are tracking the reference speed curve or not. This classifier is finally applied to monitor subject's attention. If the subject is recognized with inadequate attention, sharp voice (auditory feedback) and red screen (visual feedback) will be given by the designed game to remind the subject to focus on the training. The contrast experiment results show that subjects' performance indicated by speed tracking accuracy and muscle activation can be improved significantly by using the attention regulation system. Moreover, the phenomenon of prominent decrease in theta rhythm and increase in beta rhythm can be found, which is consistent with previous research and further validates the feasibility of the proposed system in attention enhancement.

8k: Signal processing, image processing, and multi-media

Tuesday, July 16, 5:30PM-7:30PM, Room: Duna Salon III, Chair: Hui Yu

5:30PM *A Super-Resolution Generative Adversarial Network with Simplified Gradient Penalty and Relativistic Discriminator [#19507]*

Hui Yu, Haitao Sa, Dafang Zou, Jiafa Mao and Weiguo Sheng, Zhejiang University of Technology, China; Junku (Shanghai) Information Technology Co.,Ltd., China; Hangzhou Normal University, China

Generative Adversarial Network (GAN) has been employed for single image super- resolution (SISR). However, unregularized GAN is difficult for training. This is due to gradient descent based GAN optimization is not easy to convergence, thus limiting its performance for image super-resolution. In this paper, a relativistic super-resolution GAN with a simplified gradient penalty (RSRGAN- GP) is proposed for single image super-resolution. In the proposed method, a compact residual network optimized by removing Batch-Normalization layers is employed as the generator to estimate photo-realistic images of 4X upscaling. Further, we introduce a residual network, which also has no Batch-Normalization layers as the conditional discriminator and adopt a simplified gradient regularization to penalize it for stabilizing the super-resolution GAN training, thus guaranteeing high-quality image reconstruction. Additionally, the super-resolution GAN is enhanced with a relativistic discriminator, which produces sharp and rich-detail images at no extra computational cost. The results on benchmark datasets show that our proposed method can effectively improve the visual quality of super-resolved images and achieves competitive performance compared with related works.

5:50PM *Unsupervised Synthesis of Anomalies in Videos: Transforming the Normal [#19897]*

Abhishek Joshi and Vinay P. Namboodiri, IIT Kanpur, India

Abnormal activity recognition requires detection of occurrence of anomalous events that suffer from a severe imbalance in data. In a video, normal is used to describe activities that conform to usual events while the irregular events which do not conform to the normal are referred to as abnormal. It is far more common to observe normal data than to obtain abnormal data in visual surveillance. In this paper, we propose an approach where we can obtain abnormal data by transforming normal data. This is a challenging task that is solved through a multi-stage pipeline approach. We utilize a number of techniques from unsupervised segmentation in order to synthesize new samples of data that are transformed from an existing set of normal examples. Further, this synthesis approach has useful applications as a data augmentation technique. An incrementally trained Bayesian convolutional neural network (CNN) is used to carefully select the set of abnormal samples that can be added. Finally through this synthesis approach we obtain a comparable set of abnormal samples that can be used for training the CNN for the classification of normal vs abnormal samples. We show that this method generalizes to multiple settings by evaluating it on two real world datasets and achieves improved performance over other probabilistic techniques that have been used in the past for this task.

6:10PM *Viewpoint-robust Person Re-identification via Deep Residual Equivariant Mapping and Fine-grained Features [#20221]*

Liang Yang, Xiao-yuan Jing, Fulin He, Fei Ma and Li Cheng, Wuhan University, China; Yunkang Technology co., Ltd., China

Existing person re-identification methods usually directly calculate the similarities of person pictures regardless of their viewpoints. However, matching persons under different viewpoints is difficult since it is intrinsically hard to directly learn a representation which is geometrically invariant to large viewpoint variations. In this paper, we explicitly take viewpoint information into account and propose a novel Deep Residual Equivariant Mapping and Fine-grained Features (DREMFF) approach for viewpoint- robust person re-identification. Specifically, DREMFF hypothesizes that there exists inherent mapping between different viewpoints of a person, and consequently, the global representation discrepancy of a person under different viewpoints will be bridged through equivariant mapping by adaptively adding residuals to original representation according corresponding angle deviation. What's more, based on attention mechanism, DREMFF extracts fine-grained features for each image from multiple salient regions as well as different scales. These captured information is capable of providing assistant decision- making at lower granularities. The mapped global features and the learned fine- grained features work collaboratively to enable viewpoint-robust person re- identification. Experiments on three challenging benchmarks consistently demonstrate the effectiveness of the proposed approach.

6:30PM *Two-stage Unsupervised Video Anomaly Detection using Low-rank based Unsupervised One-class Learning with Ridge Regression [#19905]*

Jingtao Hu, En Zhu, Siqi Wang, Siwei Wang, Xinwang Liu and Jianping Yin, National University of Defense Technology, China; Dongguan University of Technology, China

Video anomaly detection is a valuable but challenging task, especially in the field of surveillance videos for public safety. Almost all existing methods tackle the problem under the supervised setting and only a few attempts are conducted on the unsupervised learning. To avoid the cost of labeling training videos, this paper proposes to discriminate anomaly by a novel two-stage framework in a fully unsupervised manner. Unlike previous unsupervised approaches using local change detection to discover abnormality, our method enjoys the global information from video context by considering the pair-wise similarity of all video events. In this way, our method formulates video anomaly detection as an extension of unsupervised one-class learning, which has not been explored in the literature of video anomaly detection. Specifically, our method consists of two stages: The first stage of our kernel-based method, named Low-rank based Unsupervised One-class Learning with Ridge Regression (LR-UOCL-RR), reformulates the optimization goal of UOCL with ridge regression to avoid expensive computation, which enables our method to handle massive unlabeled data from videos. In the second stage, the estimated normal video events from the first stage are fed into the one-class support vector machine to refine the profile around normal events and enhance the performance. The experimental results conducted on two challenging video benchmarks indicate that our method is considerably superior, up to 15.7% AUC gain, to

the state-of-the-art methods in the unsupervised anomaly detection task and even better than several supervised approaches.

6:50PM *Deep Salient Object Detection with Fuzzy Superpixel Extraction and Controlled Filter Convolution [#19087]*

Yang Liu, Bo Wu and Bo Lang, Beihang University, China

Deep salient object detection (DSOD), which leverages the popular deep learning techniques, is a promising new branch of salient object detection (SOD). By training on large-scale public datasets, DSOD methods showed significant performance improvement while avoiding the involvement of manually designed visual features and prior knowledge of specific datasets. This paper proposes a novel superpixel-based DSOD method based on fuzzy superpixel extraction (FSE), a neural network-based differentiable superpixel extraction method, and controlled filter convolution (CFC), a modified convolution operation that accepts two input feature maps and can balance their influences without hand-picked coefficients. Different from other superpixel-based methods, by using FSE, the proposed method is able to include superpixel extraction in the training process, which optimizes the superpixel representations according to the datasets. Then, the CFC layers combine two different parts of the information possessed by the superpixels, which are intrasuperpixel features and intersuperpixel features, to generate a unified feature map. In the experiments conducted on 5 widely used public datasets, the proposed method significantly outperformed state-of-the-art models, which proved its effectiveness and generalization ability.

7:10PM *Prostate Segmentation using 2D Bridged U-net [#19872]*

Wanli Chen, Yue Zhang, Junjun He, Yu Qiao, Yifan Chen, Hongjian Shi, Xiaoying Tang and Ed X. Wu, Southern University of Science and Technology, China; The University of Hong Kong, Hong Kong; Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China; The University of Waikato, New Zealand

In this paper, we focus on three problems in deep learning based medical image segmentation. Firstly, U-net, as a popular model for medical image segmentation, is difficult to train when convolutional layers increase even though a deeper network usually has a better generalization ability because of more learnable parameters. Secondly, the exponential ReLU (ELU), as an alternative of ReLU, is not much different from ReLU when the network of interest gets deep. Thirdly, the Dice loss, as one of the pervasive loss functions for medical image segmentation, is not effective when the prediction is close to ground truth and will cause oscillation during training. To address the aforementioned three problems, we propose and validate a deeper network that can fit medical image datasets that are usually small in the sample size. Meanwhile, we propose a new loss function to accelerate the learning process and a combination of different activation functions to improve the network performance. Our experimental results suggest that our network is comparable or superior to state-of-the-art methods.

Computational Neuroscience

Tuesday, July 16, 5:30PM-7:30PM, Room: Panorama I, Chair: Ichiro Tsuda

5:30PM *Predictable Uncertainty-Aware Unsupervised Deep Anomaly Segmentation [#20412]*

Kazuki Sato, Kenta Hama, Takashi Matsubara and Kuniaki Uehara, Kobe University, Japan

Image-based anomaly segmentation is a fundamental topic for image analysis. For medical use, it supports treatments via refined diagnosis and growth rate evaluation of tumors and lesions. Especially, an unsupervised training is expected to generalize to unknown anomalies. Probabilistic models have been used for this purpose, whereby these models are trained to maximize the likelihood of known samples and detect anomalous samples by assigning low likelihoods. Recent studies have proposed a probabilistic model based on deep neural networks (DNNs) called AEs and they achieved significant performance thanks to their flexibility. However, AEs are sensitive to complex structure (e.g., ridges and grooves of a brain) rather than semantic anomalies (e.g., tumors and lesions). We decomposed the approximated log-likelihood into two terms; predictable uncertainty and normalized error. We found that the former represents the complexity of structure. Hence, we propose the normalized error as a novel uncertainty-sensitive score by removing the predictable uncertainty. We evaluated our score by experiments with head magnetic resonance imaging (MRI) datasets and demonstrate the robustness of the proposed normalized error to data complexity.

5:50PM *An undercomplete autoencoder to extract muscle synergies for motor intention detection [#20297]*

Domenico Buongiorno, Cristian Camardella, Giacomo Donato Cascarano, Luis Pelaez Murciego, Michele Barsotti, Irio De Feudis, Antonio Frisoli and Vitoantonio Bevilacqua, DEI - Polytechnic University of Bari, Bari / Apulian Bioengineering s.r.l. Modugno (BA), Italy; Percro Laboratory, Tecip Institute, Scuola Superiore Sant'Anna, Pisa, Italy

The growing interest in wearable robots for assistance and rehabilitation purposes opens the challenge for developing intuitive and natural control strategies. Among several human-machine interaction approaches, myoelectric control consists in decoding the motor intention from muscular activity (or EMG signals) with the aim at moving the assistive robotic device accordingly, thus establishing an intimate human-machine connection. In this scenario, bio- inspired approaches, e.g. synergy-based controllers, are reveling to be the most robust. In this work, the authors presented an undercomplete autoencoder (AE) to extract muscles synergies for motion intention detection. The proposed AE topology has been validate with EMG signals acquired from the main upper limb muscles during planar isometric reaching tasks performed in a virtual environment while wearing an exoskeleton. The presented AE have shown promising results in muscle synergy extraction comparing its performance with the Non-Negative Matrix Factorization algorithm, i.e. the most used approach in literature. The synergy activations extracted with the AE have been then used for estimating the moment applied at the shoulder and elbow joints. Comparing such estimation with the results of other synergy-based techniques already proposed in literature, it emerged that the proposed method achieves comparable performance.

6:10PM *Temporal Learning of Dynamics in Complex Neuron Models using Backpropagation [#20071]*

Christian Jarvers, Daniel Schmid and Heiko Neumann, Ulm University, Germany

One of the major challenges of computational cognitive neuroscience is to apply models of neural information processing to complex tasks. Hierarchical learning architectures like deep convolutional networks can be trained to solve tasks efficiently, but utilize simple mechanisms of activity integration and output generation. On the other hand, biologically plausible models of activation dynamics incorporate detailed mechanisms of changing membrane potentials and axonal firing properties. Making such elaborate models trainable requires learning of internal model parameters. Here, we propose to apply supervised learning and train a model of canonical cortical circuits via backpropagation through time. We train the model to settle to target equilibrium values, to generate oscillations, and to solve a contour completion task.

6:30PM *Transfer Entropy Based Connectivity Estimation of Spontaneously Firing Hippocampal Cultures on Multi Electrode Arrays [#20057]*
Nikesh Lama, Alan Hargreaves, Bob Stevens and T.M. McGinnity, Nottingham Trent University, United Kingdom

Accurate estimation of interactions in neuronal circuits is critical in understanding neural information processing and the neuronal dynamics of emergent networks. Transfer entropy(TE) is a model-free information theoretic measure of flow of information between two random processes. TE has recently gained much popularity due to its effectiveness in estimating effective connectivity among neurons from simulated networks. However, experimental recordings inherently lack the ground truth information of neural connectivity - making it difficult to identify true connections from spurious connections. To tackle this problem, we present a superimposition method where the estimated connectivity is superimposed onto the spatial firing density plot. The firing density plot, on itself, does not provide connectivity information but we assume that frequently firing channels are more likely to have more active interactions. A neuron's firing depends on the presynaptic inputs, and highly interconnected neurons tend to have more presynaptic inputs resulting in more frequent firing. The firing density plot is organised in a spatial layout, which corresponds substantially to the structural layout. We demonstrated that the connectivity estimated closely corresponds to the firing density plot when superimposed onto the firing density plot. To strengthen the arguement, we analyse the firing count and the connectivity inferred for each randomly sampled channel and found a positively correlated relationship.

6:50PM *AnxietyDecoder: An EEG-based Anxiety Predictor using a 3-D Convolutional Neural Network [#19344]*
Yi Wang, Brendan McCane, Neil McNaughton, Zhiyi Huang, Shabah Shadli and Phoebe Neo, University of Otago, New Zealand

In this paper, we propose and implement an EEG-based three-dimensional Convolutional Neural Network architecture, 'AnxietyDecoder', to predict anxious personality and decode its potential biomarkers from the participants. Since Goal-Conflict-Specific-Rhythmicity (GCSR) in the EEG is a sign of an anxiety-related system working, we first propose a two-dimensional Conflict-focused CNN (2-D CNN). It simulates the GCSR extraction process but with the advantages of automatic frequency band selection and functional contrast calculation optimization, thus providing more comprehensive trait anxiety predictions. Then, to generate more targeted hierarchical features from local spatio-temporal scale to global, we propose a three-dimensional Conflict- focused CNN (3-D CNN), which simultaneously integrates information in the temporal and brain-topology- related spatial dimensions. In addition, we embed Layer-wise Relevance Propagation (LRP) into our model to reveal the essential brain areas that are correlated to anxious personality. The experimental results show that the percentage variance accounted for by our 3-D CNN is 33%, which is almost four times higher than the previous GCSR theoretical model (7%). Meanwhile, it also outperforms the 2-D model (26%) and the t-test difference between the 3-D and 2-D models is significant ($t(4)=5.4962$, $p = 0.0053$). What's more, the reverse engineering results provide an interpretable way to understand the prediction decision-making and participants' anxiety personality. Our proposed AnxietyDecoder not only sets a new benchmark for EEG-based anxiety prediction but also reveals essential EEG components that contribute to the decision-making thus sheds some light on the anxiety biomarker research.

7:10PM *A Three-Modules Scenario in An Interpretation of Visual Hallucination in Dementia With Lewy Bodies and Preliminary Results of Computer Experiments [#19243]*
Shigetoshi Nara, Hiroshi Fujii, Hiromichi Tsukada and Ichiro Tsuda, Okayama University, Japan; Kyoto Sangyo University, Japan; Okinawa Institute of Science and Technology Graduate University, Japan; Chubu University, Japan

This paper reports a three-modules scenario in a computational interpretation of Recurrent Complex Visual Hallucination in Dementia with Lewy Bodies (RCVH- DLB). An algorithmic representation of the interpretation is proposed and preliminary results of computer experiments based on a recurrent neural network model consisting of three modules are shown. The three modules are considered to be visual module, memory module, and perceiving module, respectively. Intra- module and inter-module interactions via asymmetric synaptic connection strength are introduced, where each module consists of a rather large number of binary state neurons. Corresponding to certain inputs to visual module, veridical perception is conceptualized as dynamic convergence to particular states (perceiving attractors) of neuron activity patterns in high dimensional state space. The key idea is that, as a possible origin of hallucination, systematic reducing of neurons and pruning of synaptic connections following the observed data in physical measurements of grey matter loss and white matter atrophy in DLB patients obtained by using DTI-MR, MRI or SPECT are introduced. Preliminary experimental results suggest that these defects can cause destabilizing of perceiving attractors and mutual penetration of them which results in hallucination-like outputs in perceiving module can occur.

Neural Models of Perception, Cognition and Action

Tuesday, July 16, 5:30PM-7:30PM, Room: Panorama II, Chair: Hua Zheng

5:30PM *Retina-inspired Visual Module for Robot Navigation in Complex Environments [#20254]*
Hans Lehnert, Maria-Jose Escobar and Mauricio Araya, Department of Electronic Engineering, Universidad Tecnica Federico Santa Maria, Chile

Reinforcement learning (RL) has been widely used to implement autonomous navigation in artificial agents, where the goal is to learn a behaviuor which maximizes the reward, through interaction with the environment. Most of the recent architectures used in autonomous agents obtain information from the environment using visual modules implemented by convolutional neural networks, where the visual features resulting from learning are unknown or uncertain, which impose limitations considering the large number of parameters to be learned by the entire system. Research in retina physiology has been able to characterize it not as a single light-electrical transductor but as a complex device performing a variety of computations of the visual information, preparing the data for further stages of processing in the visual system. We propose an RL architecture that uses retina physiology knowledge to fed the convolutional neural network, avoiding the learning stage in the sensory input. The performance of the proposed architecture was evaluated using the DeepMind Lab environment simulating an agent moving inside two different maze scenarios. The results obtained reveal promising extension of the inclusion of biological-plausible mechanisms inside artificial intelligence applications.

5:50PM *Visual Cue Integration for Small Target Motion Detection in Natural Cluttered Backgrounds [#19188]*
Hongxin Wang, Jigen Peng, Qinbing Fu, Huatian Wang and Shigang Yue, University of Lincoln, United Kingdom; Guangzhou University, China

The robust detection of small targets against cluttered background is important for future artificial visual systems in searching and tracking applications. The insects' visual systems have demonstrated excellent ability to avoid predators, find prey or identify conspecifics - which always appear as small dim speckles in the visual field. Build a computational model of the insects' visual pathways could provide effective solutions to detect small

moving targets. Although a few visual system models have been proposed, they only make use of small-field visual features for motion detection and their detection results often contain a number of false positives. To address this issue, we develop a new visual system model for small target motion detection against cluttered moving backgrounds. Compared to the existing models, the small-field and wide-field visual features are separately extracted by two motion-sensitive neurons to detect small target motion and background motion. These two types of motion information are further integrated to filter out false positives. Extensive experiments showed that the proposed model can outperform the existing models in terms of detection rates.

6:10PM *A computational model of multi-sensory perception and its application to investigating the controversy around learning styles [#19630]*

A. Ravishankar Rao, Fairleigh Dickinson University, United States

The computational modeling of multi-sensory perception is challenging. It is important to understand and model temporal phenomena such as synchronization and synaptic integration over varying time intervals. Previously, we developed a sparse spatio-temporal model using networks of interconnected oscillators. This model achieved a binding of related object features in multiple sensory pathways. We apply this model to address a controversy in the field of psychology and education concerning the theory of learning styles. According to a popularly held view, learners fall into categories such as visual learners or auditory learners, and prefer to imbibe materials that match their learning styles. This view has been disputed as this categorization does not result in a demonstrable improvement in learning. We model preferences for a learning style through variations in synaptic learning efficacies. Hence a visual learner will possess higher synaptic efficacy in the visual pathway. The results of our simulation with the sparse spatio-temporal model show that there is little benefit to having or exploiting such differences in synaptic efficacy. The best system performance is achieved when all available sensory pathways are utilized rather than favoring a preferred pathway. Our model weighs in on this controversy, and supports the side that is against the categorization of students based on their learning styles. Since the idea of learning styles is quite prevalent in the literature and used in schools, it is important to disseminate appropriate research findings including those that utilize a computational neuroscience perspective, as in the current paper.

6:30PM *Neuro-Robotic Haptic Object Classification by Active Exploration on a Novel Dataset [#20190]*

Matthias Kerzel, Erik Strahl, Connor Gaede, Emil Gasanov and Stefan Wermter, University of Hamburg, Department of Informatics, Germany

We present an embodied neural model for haptic object classification by active haptic exploration with the humanoid robot NICO. When NICO's newly developed robotic hand closes around an object, multiple sensory readings from a tactile fingertip sensor, motor positions, and motor currents are recorded. We created a haptic dataset with 83200 haptic measurements, based on 100 samples of each of 16 different objects, every sample containing 52 measurements. First, we provide an analysis of neural classification models with regard to isolated haptic sensory channels for object classification. Based on this, we develop a series of neural models (MLP, CNN, LSTM) that integrate the haptic sensory channels to classify explored objects. As an initial baseline, our best model achieves a 66.6\% classification accuracy over 16 objects. We show that this result is due to the ability of the network to integrate the haptic data both over time domain and over different haptic sensory channels. Furthermore, we make the dataset publically available to address the issue of sparse haptic datasets for machine learning research.

6:50PM *Hierarchical Multi-dimensional Attention Model for Answer Selection [#20008]*

Wei Liu, Lei Zhang, Longxuan Ma, Pengfei Wang and Feng Zhang, School of Computer Science, Beijing University of Posts and Telecommunications, China; Graduate School, Beijing University of Posts and Telecommunications, China; Information Science Academy, China Electronics Technology Group Corporation, China

Answer selection is an important subtask of the question answering domain in natural language processing(NLP) applications. In this task, attention mechanism is a widely used technique which focuses on the context information and interrelationship between different words in the sentences to allocate different weight and enhance feature. However, the natural characteristics of words themselves are not fully excavated, thus the performance may be limited to a certain extent. In this paper, we propose a novel Hierarchical Multi-dimensional Attention (HMDA) model to address this issue. Especially, HMDA proposes a new kind of attention mechanism, word-attention, a true individual attention which can enhance the implied meaning of the word itself to extract features from word level which are more unique. Then HMDA uses global co-attention to better utilize word-attention and capture more common similar features. In order to utilize this attention-based semantic information on different granularities differently, HMDA designs a multi-layer structure which makes full use of all attention mechanisms by embedding attention features to model hierarchically. HMDA obtains various fine-grained information between question and candidate answers and avoids information loss. Empirically, we demonstrate that our proposed model can consistently outperform the state-of-the-art baselines under different evaluation metrics on all TrecQA, WikiQA and InsuranceQA datasets.

8l: Temporal data analysis, prediction, and forecasting; time series analysis

Tuesday, July 16, 5:30PM-7:30PM, Room: Panorama III, Chair: Nurilla Avazov Cheng Peng

5:30PM *CLEverReg: A CNN-LSTM based Linear Regression Technique for Temporal Fire Event Modelling [#20501]*

Syed Adnan Yusuf, Abdul Samad and David James Garrity, IntelliMon Pvt Ltd, United Kingdom; NED university of Engineering and Technology, Pakistan

The understanding of temperature rise events in enclosed compartments is an area of research with applications from industrial condition monitoring to household safety systems. This research originates from the need to understand and model fire behaviour to assist firefighters in their search & rescue activities in burning compartments. Fire spread in enclosed buildings pose many challenges and a timely warning of any catastrophic temperature changes is deemed critical for the personnel operating therein. The objective of this research was to utilise the real-time sensors temperature feedback to model and identify the occurrence of the so-called "flashover" phenomenon leading to a sudden increase in the temperature to 300+ degree Celsius. To achieve this, the research evaluated the effectiveness of a recurrent neural network for its ability to model and identify fire spread behaviour from events occurring earlier in the burning process. The proposed methodology achieved this by enhancing the learning process via a long short-term memory (LSTM) paradigm exploiting the spatial modelling capabilities of CNN. The underlying CNN-LSTM technique was modelled and evaluated against a diverse range of fire spread behaviours including multi-story, residential buildings and containers depicting single-floor, industrial/residential units. A sensitivity analysis against real-world and third-party fire-drill data showed the proposed model's resilience with body-mounted sensors and NIST dataset cases.

5:50PM *Deep Neural Network Ensembles for Time Series Classification [#19263]*

Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar and Pierre-Alain Muller, University of Haute-Alsace, France

Deep neural networks have revolutionized many fields such as computer vision and natural language processing. Inspired by this recent success, deep learning started to show promising results for Time Series Classification (TSC). However, neural networks are still behind the state-of-the-art TSC algorithms, that are currently composed of ensembles of 37 non deep learning based classifiers. We attribute this gap in performance due to the lack of neural network ensembles for TSC. Therefore in this paper, we show how an ensemble of 60 deep learning models can significantly improve upon the current state-of-the-art performance of neural networks for TSC, when evaluated over the UCR/UEA archive: the largest publicly available benchmark for time series analysis.Finally, we show how our proposed Neural Network Ensemble (NNE) is the first time series classifier to outperform COTE while reaching similar performance to the current state-of-the-art ensemble HIVE-COTE.

6:10PM *Periodic Neural Networks for Multivariate Time Series Analysis and Forecasting [#20342]*

Nurilla Avazov, Jiamou Liu and Bakhadyr Khoussainov, The University of Auckland, New Zealand

Designing systems that make accurate forecasts based on time dependent data is always a challenging and significant task. In this regard, a number of statistics and neural network-based models have been proposed for analysing and forecasting time series datasets. In this paper, we propose a novel machine learning model for handling and predicting multivariate time series data. In our proposed model we focus on supervised learning technique in which (1) some features of time series dataset exhibit periodic behaviour and (2) time t is considered as an input feature. Due to periodic nature of multivariate time series datasets, our model is a simple neural network where the inputs to the single output source are assumed to be in the form Asin(Bt+C)x as opposed to the standard form inputs Ax+B. We train our proposed model on various datasets and compare our model's performance with standard well-known models used in forecasting multivariate time series datasets. Our results show that our proposed model often outperforms other exiting models in terms of prediction accuracy. Moreover, our results show that the proposed model can handle time series data with missing values and also input data-values that are non-equidistant. We hope that the proposed model will be useful in fostering future research on designing accurate forecasting algorithms.

6:30PM *Adversarial attacks on deep neural networks for time series classification [#19532]*

Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar and Pierre-Alain Muller, University of Haute-Alsace, France

Time Series Classification (TSC) problems are encountered in many real life data mining tasks ranging from medicine and security to human activity recognition and food safety. With the recent success of deep neural networks in various domains such as computer vision and natural language processing, researchers started adopting these techniques for solving time series data mining problems. However, to the best of our knowledge, no previous work has considered the vulnerability of deep learning models to adversarial time series examples, which could potentially make them unreliable in situations where the decision taken by the classifier is crucial such as in medicine and security. For computer vision problems, such attacks have been shown to be very easy to perform by altering the image and adding an imperceptible amount of noise to trick the network into wrongly classifying the input image. Following this line of work, we propose to leverage existing adversarial attack mechanisms to add a special noise to the input time series in order to decrease the network's confidence when classifying instances at test time. Our results reveal that current state-of-the-art deep learning time series classifiers are vulnerable to adversarial attacks which can have major consequences in multiple domains such as food safety and quality assurance.

6:50PM *NAO Index Prediction using LSTM and ConvLSTM Networks Coupled with Discrete Wavelet Transform [#19772]*

Bin Mu, Jing Li, Shijin Yuan, Xiaodan Luo and Guokun Dai, Tongji University, China; Fudan University, China

The North Atlantic Oscillation (NAO) has a significant effect on the global weather and climate variation. Thus, it has widespread scientific research value to enhance prediction skill of the NAO events. The NAO is quantified by the NAO index which can be defined from sea level pressure (SLP), and has the characteristics of changeability and complexity. With the applications of deep learning approaches in the field of climatic prediction, deep neural networks provide the alternative to predict the NAO index beyond numerical models. In this paper, long short-term memory (LSTM) is used to predict the NAO index sequence. Furthermore, ConvLSTM is used to predict SLP grid data with capturing the temporal and spatial interdependencies. The predicted NAO index can be obtained by projecting output SLP on the NAO anomaly pattern. In order to improve the prediction accuracy, especially for extreme events, we adopt discrete wavelet transform (DWT) to preprocess data, thus DWT-LSTM and DWT- ConvLSTM models are proposed for the NAO index data and SLP grid data respectively. Preprocessing steps include discomposing the input data into the low-frequency and the high-frequency components using DWT along with considering the local time dependency. Observation data is selected as the benchmark, and our models are compared against multiple models, such as support vector regression (SVR), LSTM, gated recurrent unit (GRU) etc. The experimental results demonstrate that both DWT- LSTM and DWT-ConvLSTM perform better, particularly at peak values. In addition, our models are much closer to observation in multi-step forecasting compared with ensemble forecasts based on numerical models.

7:10PM *ENSO Forecasting over Multiple Time Horizons Using ConvLSTM Network and Rolling Mechanism [#19743]*

Bin Mu, Cheng Peng, Shijin Yuan and Lei Chen, Tongji University, China; Shanghai Central Meteorological Observatory, China

El Nino Southern Oscillation (ENSO) event is characterized by sea surface temperature (SST) anomalies in the tropical Pacific and mainly identified with Oceanic Nino Index (ONI). ENSO forecasting is very challenging owing to the existence of predictability barrier and chaos of climate variability. Recently, machine learning approaches have received considerable attention besides conventional numerical models for this task. However, these existing works mostly focus on investigating single ONI data, neglecting the spatial and temporal dependencies of SST data, and the skill of predictions reduce significantly beyond a lag time of 6 months. With the goal of capturing the spatial and temporal dependencies of SST simultaneously and improving the skill of prediction over longer time horizon, we propose ConvLSTM-RM model, which is a hybrid of convolutional LSTM and rolling mechanism, and use it to build an end-to-end trainable model for ENSO forecasting problem in this paper. Specifically, ENSO forecasting is formulated as a spatiotemporal sequence forecasting problem in which both the input and the output are SST sequences, and ONI can be acquired based on the output. Experiments on historical SST dataset demonstrate that ConvLSTM-RM outperforms seven well-known methods over multiple time horizons (6-, 9- and 12-month). We also apply this model to predict the latest ENSO event during 2015/2016. The results show that our model gets a relatively reliable prediction both from SST grid maps and ONI.

Neural Models of Perception, Cognition and Neurodynamics

Tuesday, July 16, 5:30PM-7:30PM, Room: Panorama IV, Chair: Huaping Liu

5:30PM *Zero-shot Object Detection for Indoor Robots [#19639]*

Abdalwhab Abdalwhab and Huaping Liu, Tsinghua University, China

Object detection is one of the most crucial tasks for robotic systems. Therefore, a plethora of researches have explored the traditional object detection, where the model must be trained on instances of all objects of interest. However, this is not a very realistic setting because of the huge number of classes in real-world settings and the high cost and time associated with annotating data. Thus, it is almost impossible to train a model in all possible classes. This led to the introduction of the concept of zero-shot object detection, in which a model is trained on some seen classes and then tries to use the obtained knowledge to be able to detect novel unseen classes. Despite the fact that, this setting is more realistic, but very few researchers have explored it. Furthermore, to the best of our knowledge, no previous work has explored it specifically for indoor robots. In fact, indoor environments can be more challenging, since the number of objects of interests is way larger compared to other applications like self- driving cars for instance. In this work, we explore zero shot object detection for indoor robots by embedding deep features and class labels to a shared semantic space. We use the famous SUN RGB-D dataset \cite{b10} and our own collected dataset for training and evaluation, and we propose two novel splits for SUN RGB-D dataset especially for zero-shot object detection.

5:50PM *Pinning Control for Synchronization of Drive-Response Memristive Neural Networks with Nonidentical Parameters [#19494]*

Yueheng Li, Biao Luo, Derong Liu, Zhe Dong and Zhanyu Yang, School of Automation and Electrical Engineering, University of Science and Technology Beijing, China; School of Automation, Central South University, China; School of Automation, Guangdong University of Technology, China; College of Electrical and Control Engineering, North China University of Technology, China; The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China

In this paper, the asymptotic synchronization for drive-response memristive neural networks(MNNs) with nonidentical parameters is investigated. Parameter inconformity is ubiquitous between drive and response systems due to environmental or internal influence. However, the majority of previous results were based on the well-matched MNNs. Thus, it is meaningful to study the synchronization problem of MNNs with nonidentical parameters. First, coefficient mismatches are dealt within the framework of set-valued maps and differential inclusions. Furthermore, in order to reduce the control cost, a pinning control strategy is adopted to drive two nonidentical MNNs to achieve asymptotic synchronization. And the sufficient stability conditions are given based on Lyapunov functional method. Finally, the effectiveness of proposed pinning controller is verified by a numerical example.

6:10PM *A novel hardware-efficient CPG model for a hexapod robot based on nonlinear dynamics of coupled asynchronous cellular automaton oscillators [#19758]*

Takeda Kentaro and Torikai Hiroyuki, Graduate School of Science and Engineering, Hosei University, Japan

This paper presents a novel central pattern generator (CPG) model based on nonlinear dynamics of asynchronous cellular automata. It is shown that the presented CPG model can exhibit various synchronization phenomena depending on parameter values. In order to evaluate usefulness of the presented CPG model, this paper focuses on controlling a hexapod robot. Based on intensive analyses of the synchronization phenomena, a parameter tuning method to realize a tripod gait of the hexapod robot is derived. Then the CPG model with a tuned parameter value is implemented in a field programmable gate array and it is shown that the CPG model can realize a tripod gait of the hexapod robot. Also, it is shown that the presented CPG model uses much fewer circuit elements and consumes much less power compared to a conventional CPG model.

6:30PM *Closed-loop Central Pattern Generator Control of Human Gaits in OpenSim Simulator [#19692]*

Andrii Shachykov, Oleksandr Shuliak and Patrick Henaff, Universite de Lorraine, CNRS, Inria, LORIA, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine; National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine; Universite de Lorraine, CNRS, Inria, LORIA, France

In this paper, a new neuro-musculoskeletal gait simulation platform is presented. This platform is developed to reproduce healthy or altered walking gaits. It is based on an original model of central pattern generator able to generate variable rhythmic signals for controlling biological human leg joints. Output signals of motoneurons are applied to excitation inputs of modelled muscles of the human lower limbs model. Eight central pattern generators control a musculoskeletal model made up of three joints per leg actuated by 44 Hill-type muscle models. Forward dynamics simulation in OpenSim show that it is possible to generate different stable walking gaits by changing parameters of controller. Further work is aimed on development of stable human standing by implementing reflexes.

6:50PM *Depersonalized Cross-Subject Vigilance Estimation with Adversarial Domain Generalization [#19827]*

Bo-Qun Ma, He Li, Yun Luo and Bao-Liang Lu, Shanghai Jiao Tong University, China

Subject variability is a major obstacle to vigilance estimation. The conventional subject-specific models fail to perform well on unknown subjects. The existing studies mainly focus on domain adaptation utilizing labeled/unlabeled subject-specific data. However, it is still expensive and inconvenient to collect task-specific data from unknown subjects in some real-world applications. In this paper, we introduce domain generalization methods for building vigilance estimation models without requiring any information from the unknown subjects. We first generalize the structure of Domain Adversarial Neural Network (DANN) into Domain Generalization (DG- DANN), and then propose a novel adversarial structure called Domain Residual Network (DResNet). We compare a popular domain generalization method, Domain-Invariant Component Analysis (DICA), with our proposed approach. In terms of the estimation accuracy and generalization ability, we designed two different settings for evaluation experiments on a public dataset called SEED-VIG. Experimental results indicate that our new model achieves comparable accuracy but more stable performance without using additional information from the unknown subjects in comparison with the state-of-the-art domain adaptation methods. Furthermore, domain generalization models also perform well on the tasks with multiple unknown subjects.

Panel Session: Deep Learning: Hype or Hallelujah?
Tuesday, July 16, 5:30PM-7:30PM, Room: Panorama V, Chair: University of Minnesota, USA Vladimir Cherkassky

Wednesday, July 17, 8:00AM-10:00AM

S11: Learning Representations for Structured Data

Wednesday, July 17, 8:00AM-10:00AM, Room: Ballroom I, Chair: Alessandro Sperduti

8:00AM *Large-Margin Multiple Kernel Learning for Discriminative Features Selection and Representation Learning [#19212]*

Babak Hosseini and Barbara Hammer, Bielefeld University-CITEC, Germany

Multiple kernel learning (MKL) algorithms combine different base kernels to obtain a more efficient representation in the feature space. Focusing on discriminative tasks, MKL has been used successfully for feature selection and finding the significant modalities of the data. In such applications, each base kernel represents one dimension of the data or is derived from one specific descriptor. Therefore, MKL finds an optimal weighting scheme for the given kernels to increase the classification accuracy. Nevertheless, the majority of the works in this area focus on only binary classification problems or aim for linear separation of the classes in the kernel space, which are not realistic assumptions for many real-world problems. In this paper, we propose a novel multi-class MKL framework which improves the state-of-the-art by enhancing the local separation of the classes in the feature space. Besides, by using a sparsity term, our large-margin multiple kernel algorithm (LMMK) performs discriminative feature selection by aiming to employ a small subset of the base kernels. Based on our empirical evaluations on different real-world datasets, LMMK provides a competitive classification accuracy compared with the state-of- the-art algorithms in MKL. Additionally, it learns a sparse set of non-zero kernel weights which leads to a more interpretable feature selection and representation learning.

8:20AM *Autoregressive Models for Sequences of Graphs [#20455]*

Daniele Zambon, Daniele Grattarola, Lorenzo Livi and Cesare Alippi, Universita della Svizzera italiana, Switzerland; University of Exeter, United Kingdom

This paper proposes an autoregressive (AR) model for sequences of graphs, which generalises traditional AR models. A first novelty consists in formalising the AR model for a very general family of graphs, characterised by a variable topology, and attributes associated with nodes and edges. A graph neural network (GNN) is also proposed to learn the AR function associated with the graph-generating process (GGP), and subsequently predict the next graph in a sequence. The proposed method is compared with four baselines on synthetic GGPs, denoting a significantly better performance on all considered problems.

8:40AM *Universal Readout for Graph Convolutional Neural Networks [#20249]*

Nicolo' Navarin, Dinh Van Tran and Alessandro Sperduti, University of Padova, Italy; University of Freiburg, Germany

Several machine learning problems can be naturally defined over graph data. Recently, many researchers have been focusing on the definition of neural networks for graphs. The core idea is to learn a hidden representation for the graph vertices, with a convolutive or recurrent mechanism. When considering discriminative tasks on graphs, such as classification or regression, one critical component to design is the readout function, i.e. the mapping from the set of vertex representations to a fixed-size vector (or the output). Different approaches have been presented in literature, but recent approaches tend to be complex, making the training of the whole network harder. In this paper, we frame the problem in the setting of learning over sets. Adopting recently proposed theorems over functions defined on sets, we propose a simple but powerful formulation for a readout layer that can encode or approximate arbitrarily well any continuous permutation-invariant function over sets. Experimental results on real-world graph datasets show that, compared to other approaches, the proposed readout architecture can improve the predictive performance of Graph Neural Networks while being computationally more efficient.

9:00AM *An Attention-Based Model for Learning Dynamic Interaction Networks [#19750]*

Sandro Cavallari, Vincent W Zheng, Hongyun Cai, Soujanya Poria and Erik Cambria, NTU, Singapore; ADSC, Singapore

In the physical world, complex systems are generally created as the composition of multiple primitive components that interact with each other rather than a single monolithic structure. Recently, spatio-temporal graphs received a reasonable amount of attention from the research community since they emerged as a natural representational tool able to capture the interactive and interrelated structure of a complex problem. To better understand the nature of complex systems, there is the need to define models that can easily explain the learned causal relationship. To this end, we propose an attentive model able to learn and project the relational structure into a fixed-size embedding. Such representation naturally captures the dynamic influence that each neighbors exert over a given vertex providing a valuable description of the problem setting. The proposed architecture has been extensively evaluated against strong baselines on toy as well as real-world tasks, such as prediction of household energy load and traffic congestion.

9:20AM *Bayesian Tensor Factorisation for Bottom-up Hidden Tree Markov Models [#20162]*

Daniele Castellana and Davide Bacciu, Universita' di Pisa, Italy

Bottom-Up Hidden Tree Markov Model is a highly expressive model for tree-structured data. Unfortunately, it cannot be used in practice due to the intractable size of its state-transition matrix. We propose a new approximation which lies on the Tucker factorisation of tensors. The probabilistic interpretation of such approximation allows us to define a new probabilistic model for tree-structured data. Hence, we define the new approximated model and we derive its learning algorithm. Then, we empirically assess the effective power of the new model evaluating it on two different tasks. In both cases, our model outperforms the other approximated model known in the literature.

9:40AM *A Novel End-to-End Multiple Tagging Model for Knowledge Extraction [#20164]*

Yunhua Song, Hongyun Bao, Zhineng Chen and Jianquan Ouyang, Xiangtan University, China; Institute of Automation Chinese Academy of Sciences, China

It is an emerging research topic in NLP to joint extraction of knowledge including entities and relations from unstructured text and representing them as meaningful triplets. Despite significant progresses made by recent deep neural network based solutions, these methods still confront the overlapping issue that different relational triplets may have overlapped entities in a sentence, and it is troublesome to address this issue by current solutions. In this paper, we propose a novel end2end multiple tagging model to address the overlapping issue and extract knowledge from unstructured text. Specifically, we devise a multiple tagging scheme that transforms the problem of joint entity and relation extraction into a multiple sequence tagging problem. By using GRU as the building block for encoding-decoding, the proposed model is capable of handling the triplet overlapping problem because the decoder layer allows one entity to take part in more than one triplet. The whole network is end-to-end trainable and outputs all triplets in a sentence directly. Experimental results on the NYT and KBP benchmarks demonstrate that the proposed model significantly improves the recall of triplet, and consequently, achieving the new state-of-the-art in the task of triplet extraction.

S12: Automatic Machine Learning and S13: Extreme Learning Machines (ELM)

Wednesday, July 17, 8:00AM-10:00AM, Room: Ballroom II, Chair: Donald Wunsch

8:00AM *RPR-BP: A Deep Reinforcement Learning Method for Automatic Hyperparameter Optimization [#19320]*

Jia Wu, SenPeng Chen and XiuYun Chen, University of Electronic Science and Technology of Chin, China

We introduce a new deep reinforcement learning architecture - RPR-BP to optimize hyperparameter for any machine learning model on a given data set. In this method, an agent constructed by a Long Short-Term Memory Network aims at maximizing the expected accuracy of a machine learning model on a validation set. At each iteration, it selects a set of hyperparameters and uses the accuracy of the model on the validation set as the reward signal to update its internal parameters. After multiple iterations, the agent learns how to improve its decisions. However, the computation of the reward requires significant time and leads to low sample efficiency. To speed up training, we employ a neural network to predict the reward. The training process for the agent and the prediction network is divided into three phases: Real-Predictive-Real (RPR). First, the agent and the prediction network are trained by the real experience; then, the agent is trained by the reward generated from the prediction network; finally, the agent is trained again by the real experience. In this way, we can speed up training and make the agent achieve a high accuracy. Besides, to reduce the variance, we propose a Bootstrap Pool (BP) to guide the exploration in the search space. The experiment was carried out by optimizing hyperparameters of two widely used machine learning models: Random Forest and XGBoost. Experimental results show that the proposed method outperforms random search, Bayesian optimization and Tree-structured Parzen Estimator in terms of accuracy, time efficiency and stability.

8:20AM *On the Performance of Differential Evolution for Hyperparameter Tuning [#20115]*

Mischa Schmidt, Shahd Safarani, Julia Gastinger, Tobias Jacobs, Sebastien Nicolas and Anett Schuelke, NEC Laboratories Europe GmbH, Germany

Automated hyperparameter tuning aspires to facilitate the application of machine learning for non-experts. In the literature, different optimization approaches are applied for that purpose. This paper investigates the performance of Differential Evolution for tuning hyperparameters of supervised learning algorithms for classification tasks. This empirical study involves a range of different machine learning algorithms and datasets with various characteristics to compare the performance of Differential Evolution with Sequential Model-based Algorithm Configuration (SMAC), a reference Bayesian Optimization approach. The results indicate that Differential Evolution outperforms SMAC for most datasets when tuning a given machine learning algorithm - particularly when breaking ties in a first-to-report fashion. Only for the tightest of computational budgets SMAC performs better. On small datasets, Differential Evolution outperforms SMAC by 19% (37% after tie-breaking). In a second experiment across a range of representative datasets taken from the literature, Differential Evolution scores 15% (23% after tie-breaking) more wins than SMAC.

8:40AM *FERNN: A Fast and Evolving Recurrent Neural Network Model for Streaming Data Classification [#19410]*

Monidipa Das, Mahardhika Pratama, Andri Ashfahani and Subhrajit Samanta, Nanyang Technological University (NTU), Singapore

With the recent explosion of data navigating in motion, there is a growing research interest for analyzing streaming data, and consequently, there are several recent works on data stream analytics. However, exploring the potentials of traditional recurrent neural network (RNN) in the context of streaming data classification is still a little investigated area. In this paper, we propose a novel variant of RNN, termed as FERNN, which features single-pass learning capability along with self-evolution property. The online learning capability makes FERNN fit for working on streaming data, whereas the self-organizing property makes the model adaptive to the rapidly changing environment. FERNN utilizes hyperplane activation in the hidden layer, which not only reduces the network parameters to a significant extent, but also triggers the model to work by default as per teacher forcing mechanism so that it automatically handles the vanishing/ exploding gradient issues in traditional RNN learning based on back-propagation- through-time policy. Moreover, unlike the majority of the existing autonomous learning models, FERNN is free from normal distribution assumption for streaming data, making it more flexible. The efficacy of FERNN is evaluated in terms of classifying six publicly available data streams, under the prequential test-then train protocol. Experimental results show encouraging performance of FERNN attaining state-of-the-art classification accuracy with fairly reduced computation cost.

9:00AM *Physical Activity Recognition Using Multi-Sensor Fusion and Extreme Learning Machines [#20351]*

Honggang Wang, WeiZhong Yan and Shaopeng Liu, GE Global Research, United States

Sensor based Physical Activity (PA) recognition is an imperative research topic due to its wide applications in many areas such as mobile healthcare, elderly care and personalized recommendation. However, PA recognition based on single sensor suffers from issues such as limited spatial coverage, imprecision and uncertainty. In contrast, fusion of heterogeneous sensor sources brings improved resolution, precision and robustness. In this paper, a multi-sensor fusion method based on variants of Extreme Learning Machine (ELM) is presented to improve the recognition performance in terms of speed and accuracy. The Kernel ELM, Weighted ELM, Regularized ELM are tasked directly to handle the multi-activity classification problem using the real- world test data with 46 test subjects. Their performance has been evaluated against previously published methods based on Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) using a two-step subject-wise cross validation scheme. The multi-sensor fusion based Kernel ELM has shown favorable performance with 88.4% average testing accuracy, outperforming SVM by 1.1%. And for real-time PA recognition, Kernel ELM is

30% faster than SVM. The ELMs trained with the two-step, subject-wise cross validation approach, combined with feature level data fusion also led to a better cross-person generalization than SVM and KNN.

9:20AM *Multi-Grained Cascade AdaBoost Extreme Learning Machine for Feature Representation [#19738]*

Hongwei Ge, Weiting Sun, Mingde Zhao, Kai Zhang, Liang Sun and Chao Yu, Dalian University of Technology, China; McGill University, Canada

Extreme learning machine (ELM) has been well recognized for characteristics such as less training parameters, fast training speed and strong generalization ability. Due to its high efficiency, researchers have embedded ELMs into deep learning frameworks to address the problems of high time-consumption and computational complexities that are encountered in the traditional deep neural networks. However, existing ELM-based deep learning algorithms usually neglect the spatial relationship of original data. In this paper, we propose a multi-grained cascade AdaBoost based weighted ELM algorithm (gcAWELM) for feature representation. We use AdaBoost based weighted ELM as a basic module to construct cascade structure for feature learning. Different ensemble ELMs trend to extract varied features. Moreover, multi-grained scanning is employed to exploit the spatial structure of the original data. The gcAWELM can determine the number of cascade levels adaptively and has simpler structure and fewer parameters compared with the traditional deep models. The results on image datasets with different scales show that the gcAWELM can achieve competitive performance for different learning tasks even with the same parameter settings.

9:40AM *Automatic Configuration of Deep Neural Networks with Parallel Efficient Global Optimization [#20111]*

Bas van Stein, Hao Wang and Thomas Back, University Leiden, Netherlands

Designing the architecture for an artificial neural network is a cumbersome task because of the numerous parameters to configure, including activation functions, layer types, and hyper-parameters. With the large number of parameters for most networks nowadays, it is intractable to find a good configuration for a given task by hand. In this paper the Mixed Integer Parallel Efficient Global Optimization (MIP-EGO) algorithm is proposed to automatically configure convolutional neural network architectures. It is shown that on several image classification tasks this approach is able to find competitive network architectures in terms of prediction accuracy, compared to the best hand-crafted ones in literature, when using only a fraction of the number of training epochs. Moreover, instead of the standard sequential evaluation in EGO, several candidate architectures are proposed and evaluated in parallel, which reduces the execution overhead significantly and leads to an efficient automation for deep neural network design.

S15: Machine Learning and Deep Learning Methods applied to Vision and Robotics (MLDLMVR)

Wednesday, July 17, 8:00AM-10:00AM, Room: Ballroom III, Chair: Jose Garcia-Rodriguez

8:00AM *Adversarial Action Data Augmentation for Similar Gesture Action Recognition [#20029]*

Di Wu, Junjun Chen, Nabin Sharma, Shirui Pan, Guodong Long and Michael Blumenstein, University of Technology Sydney, Australia; Beijing University of Chemical Technology, China; Monash University, Australia

Human gestures are unique for recognizing and describing human actions, and video-based human action recognition techniques are effective solutions to varies real-world applications, such as surveillance, video indexing, and human-computer interaction. Most existing video human action recognition approaches either using handcraft features from the frames or deep learning models such as convolutional neural networks (CNN) and recurrent neural networks (RNN); however, they have mostly overlooked the similar gestures between different actions when processing the frames into the models. The classifiers suffer from similar features extracted from similar gestures, which are unable to classify the actions in the video streams. In this paper, we propose a novel framework with generative adversarial networks (GAN) to generate the data augmentation for similar gesture action recognition. The contribution of our work is tri-fold: 1) we proposed a novel action data augmentation framework (ADAF) to enlarge the differences between the actions with very similar gestures; 2) the framework can boost the classification performance either on similar gesture action pairs or the whole dataset; 3) experiments conducted on both KTH and UCF101 datasets show that our data augmentation framework boost the performance on both similar gestures actions as well as the whole dataset compared with baseline methods such as 2DCNN and 3DCNN.

8:20AM *TactileGCN: A Graph Convolutional Network for Predicting Grasp Stability with Tactile Sensors [#19871]*

Alberto Garcia-Garcia, Brayan S. Zapata-Impata, Sergio Orts-Escolano, Pablo Gil and Jose Garcia-Rodriguez, University of Alicante, Spain

Tactile sensors provide useful contact data during the interaction with an object which can be used to accurately learn to determine the stability of a grasp. Most of the works in the literature represented tactile readings as plain feature vectors or matrix-like tactile images, using them to train machine learning models. In this work, we explore an alternative way of exploiting tactile information to predict grasp stability by leveraging graph-like representations of tactile data, which preserve the actual spatial arrangement of the sensor's taxels and their locality. In experimentation, we trained a Graph Neural Network to binary classify grasps as stable or slippery ones. To train such network and prove its predictive capabilities for the problem at hand, we captured a novel dataset of approximately 5000 three-fingered grasps across 41 objects for training and 1000 grasps with 10 unknown objects for testing. Our experiments prove that this novel approach can be effectively used to predict grasp stability.

8:40AM *Modulation Based Transfer Learning of Motivational Cues in Developmental Robotics [#20129]*

Alejandro Romero, Jose A. Becerra, Francisco Bellas and Richard J. Duro, Universidade da Coruna, Spain

The modeling of utility is an important problem in many fields, including reinforcement learning. However, when considering a developmental approach to open-ended learning a new aspect arises. In these settings, the efficiency of the modeling process becomes a key aspect, as these processes usually take place in real time and, to increase survivability, it is necessary for the robot to be able to produce utility models as fast as possible. In this paper, we address this issue by proposing a modulation based approach to the adaptation of the robot's experience, in the form of previously obtained ANN based utility models, to new situations. These previous utility models are perceptually recalled from a Long Term Memory and combined to produce an initial guess to the new utility model. After this, modulatory structures are created that lead to the fine adaptation of these initial guesses to the real utility model of the new situation. Some initial results of experiments using a real robot are presented to clarify the approach. Specifically, three realistic problems that a Baxter "cooking robot" must solve are faced with this modulating approach. With them, it is clearly shown the increase in efficiency of the utility model learning in real time.

9:00AM *Adaptive Model Learning of Neural Networks with UUB Stability for Robot Dynamic Estimation [#19319]*

Pedram Agand and Mahdi Aliyari Shoorehdeli, K. N. Toosi University of Technology, Iran

Since batch algorithms suffer from lack of proficiency in confronting model mismatches and disturbances, this contribution proposes an adaptive scheme based on continuous Lyapunov function for online robot dynamic identification. This paper suggests stable updating rules to drive neural networks inspiring from model reference adaptive paradigm. Network structure consists of three parallel self-driving neural networks which aim to estimate robot dynamic terms individually. Lyapunov candidate is selected to construct energy surface for a convex optimization framework. Learning rules are driven directly from Lyapunov functions to make the derivative negative. Finally, experimental results on 3- DOF Phantom Omni Haptic device demonstrate efficiency of the proposed method.

9:20AM *Multilevel Classification using a Taxonomy Applied to Recognizing Diptera Images [#19035]*

Javier Navarrete, Francisco Gomez-Donoso, Diego Viejo and Miguel Cazorla, Institute for Computer Research, University of Alicante, Spain

Most state-of-the-art machine learning classifiers currently perform a direct classification of categories. Nonethe- less, all the concepts are inherently connected by an inclusion dependency, creating a tree of relationships. For example, a car and a bus are elements included in the concept vehicle. Traditional approaches make use of finer grain concepts, thus discarding important data provided by these relational bounds. This paper describes different strategies to include this tree information in a deep learning pipeline. We compare our pro- posals, which produce multilevel classifications, with a direct classification baseline, outperforming it in each case. In addition, multilevel predictions could be further improved with different path correction strategies, which were also tested

9:40AM *Network Implosion: Effective Model Compression for ResNets via Static Layer Pruning and Retraining [#19270]*

Yasutoshi Ida and Yasuhiro Fujiwara, NTT Software Innovation Center, Japan

Residual Networks with convolutional layers are widely used in the field of machine learning. Since they effectively extract features from input data by stacking multiple layers, they can achieve high accuracy in many applications. However, the stacking of many layers raises their computation costs. To address this problem, we propose Network Implosion, it erases multiple layers from Residual Networks without degrading accuracy. Our key idea is to introduce a priority term that identifies the importance of a layer; we can select unimportant layers according to the priority and erase them after training. In addition, we retrain the networks to avoid critical drops in accuracy after layer erasure. A theoretical assessment reveals that our erasure and retraining scheme can erase layers without accuracy drop, and achieve higher accuracy than is possible with training from scratch. Our experiments show that Network Implosion can, for classification tasks on Cifar-10/100 and ImageNet, reduce the number of layers by 24.00 to 42.86 percent without any drop in accuracy.

S06: Deep and Generative Adversarial Learning

Wednesday, July 17, 8:00AM-10:00AM, Room: Duna Salon I, Chair: Ariel Ruiz-Garcia

8:00AM *Targeted Black-Box Adversarial Attack Method for Image Classification Models [#20081]*

Su Zheng, Jialin Chen and Lingli Wang, State Key Laboratory of ASIC & System, Fudan University, China

Deep neural networks (DNNs) are widely applied to image classification tasks. Due to the fact that these models are usually vulnerable, subtle perturbations of pixels may lead to classification errors, which poses a serious threat to the success of DNN applications. Moreover, perturbations of pixels can also corrupt other pattern recognition models such as Naive Bayes (NB), Decision Tree (DT) and Random Forest (RF). In this paper, a general method is proposed to carry out targeted black-box attacks for image classification models. The proposed method can achieve targeted fool rates (TFRs) of 0.873 and 0.781 on CIFAR-10 dataset with and without the access to the training set of the target model respectively. For cross-model attacks, the proposed method can still achieve a TFR of 0.630 on CIFAR-10. Furthermore, the proposed method is able to mount attacks for up to 100 classes on CIFAR-100 dataset with a TFR of 0.721, successfully handling 99 cases for each class. In our experiments, the proposed method shows higher performance and higher reliability than other black-box attack methods, with 0.123 greater maximum TFR and 0.602 greater minimum TFR than previous methods UPSET and ANGRI on CIFAR-10 in attacks trained on a single model.

8:20AM *Fine-grained Adversarial Image Inpainting with Super Resolution [#19282]*

Yang Li, Bitao Jiang, Yao Lu and Li Shen, Beijing Institute of Remote Sensing Information, China

Image inpainting refers to synthesizing plausible contents for images with missing regions. However, current methods often create blurry textures, distorted structures and loss of details, especially when the image has complex scenes or large missing regions. We propose a fine-grained adversarial image inpainting model with super resolution. It performs a coarse-to-fine inpainting procedure in two stages. The proposed generator first synthesizes initial predictions of the missing regions with a novel encoder-decoder structure. Then it refines the predicted missing regions by generating high-frequency details via super resolution. We evaluate the proposed from both pixel level and semantic level. Experiments demonstrate that the proposed can generate higher quality inpainting results than the baseline models in both metrics.

8:40AM *The Conditional Boundary Equilibrium Generative Adversarial Network and its Application to Facial Attributes [#20167]*

Marzouk Ahmed, Barros Pablo, Eppe Manfred and Wermter Stefan, University of Hamburg, Germany

We propose an extension of the Boundary Equilibrium GAN (BEGAN) neural network, named Conditional BEGAN (CBEGAN), as a general generative and transformational approach for data processing. As a novelty, the system is able of both data generation and transformation under conditional input. We evaluate our approach for conditional image generation and editing using five controllable attributes for images of faces from the CelebA dataset: age, smiling, cheekbones, eyeglasses and gender. We perform a set of objective quantitative experiments to evaluate the model's performance and a qualitative user study to evaluate how humans assess the generated and edited images. Both evaluations yield coinciding results which show that the generated facial attributes are recognizable in more than 80\% of all new testing samples.

9:00AM *Improving Prediction Accuracy in Building Performance Models Using Generative Adversarial Networks (GANs) [#20389]*

Chanachok Chokwitthaya, Edward Collier, Yimin Zhu and Supratik Mukhopadhyay, Louisiana State University, United States

Building performance discrepancies between building design and operation are one of the causes that lead many new designs fail to achieve their goals and objectives. One of main factors contributing to the discrepancy is

occupant behaviors. Occupants responding to a new design are influenced by several factors. Existing building performance models (BPMs) ignore or partially address those factors (called contextual factors) while developing BPMs. To potentially reduce the discrepancies and improve the prediction accuracy of BPMs, this paper proposes a computational framework for learning mixture models by using Generative Adversarial Networks (GANs) that appropriately combining existing BPMs with knowledge on occupant behaviors to contextual factors in new designs. Immersive virtual environments (IVEs) experiments are used to acquire data on such behaviors. Performance targets are used to guide appropriate combination of existing BPMs with knowledge on occupant behaviors. The resulting model obtained is called an augmented BPM. Two different experiments related to occupants lighting behaviors are shown as case study. The results reveal that augmented BPMs significantly outperformed existing BPMs with respect to achieving specified performance targets. The case study confirms the potential of the computational framework for improving prediction accuracy of BPMs during design.

9:20AM *Extracting Tables from Documents using Conditional Generative Adversarial Networks and Genetic Algorithms [#19739]*

Nataliya LeVine, Matthew Zeigenfuse and Mark Rowan, Swiss Re, United States; Swiss Re, Switzerland

Extracting information from tables in documents presents a significant challenge in many industries and in academic research. Existing methods which take a bottom-up approach of integrating lines into cells and rows or columns neglect the available prior information relating to table structure. Our proposed method takes a top-down approach, first using a generative adversarial network to map a table image into a standardised 'skeleton' table form denoting the approximate row and column borders without table content, then fitting renderings of candidate latent table structures to the skeleton structure using a distance measure optimised by a genetic algorithm.

9:40AM *Detection of Typical Pronunciation Errors in Non-native English Speech Using Convolutional Recurrent Neural Networks [#19552]*

Aleksandr Diment, Eemi Fagerlund, Adrian Benfield and Tuomas Virtanen, Tampere University, Finland

A machine learning method for the automatic detection of pronunciation errors made by non-native speakers of English is proposed. It consists of training word-specific binary classifiers on a collected dataset of isolated words with possible pronunciation errors, typical for Finnish native speakers. The classifiers predict whether the typical error is present in the given word utterance. They operate on sequences of acoustic features, extracted from consecutive frames of an audio recording of a word utterance. The proposed architecture includes a convolutional neural network, a recurrent neural network, or a combination of the two. The optimal topology and hyperparameters are obtained in a Bayesian optimisation setting using a tree- structured Parzen estimator. A dataset of 80 words uttered naturally by 120 speakers is collected. The performance of the proposed system, evaluated on a well-represented subset of the dataset, shows that it is capable of detecting pronunciation errors in most of the words (46/49) with high accuracy (mean accuracy gain over the zero rule 12.21 percent points).

8l: Temporal data analysis, prediction, and forecasting; time series analysis

Wednesday, July 17, 8:00AM-10:00AM, Room: Duna Salon II, Chair: Tom Gedeon

8:00AM *Domain Adaptation for sEMG-based Gesture Recognition with Recurrent Neural Networks [#20309]*

Istvan Ketyko, Ferenc Kovacs and Krisztian Varga, Member of technical staff, Hungary

Surface Electromyography (sEMG) is to record muscles' electrical activity from a restricted area of the skin by using electrodes. The sEMG-based gesture recognition is extremely sensitive of inter-session and inter-subject variances. We propose a model and a deep-learning-based domain adaptation method to approximate the domain shift for recognition accuracy enhancement. Analysis performed on sparse and High-Density (HD) sEMG public datasets validate that our approach outperforms state-of-the-art methods.

8:20AM *Competitive Feature Extraction for Activity Recognition based on Wavelet Transforms and Adaptive Pooling [#19174]*

Mubarak G. Abdu-Aguye and Walid Gomaa, Egypt-Japan University of Science and Technology, Egypt

Any application of machine learning requires feature extraction, whereupon the source data is processed to yield representations that are germane to obtaining the desired output. Traditional approaches to feature extraction include the estimation of statistical and structural properties of the data, the choice of which is mainly influenced by domain knowledge. However, Deep Learning has made it possible to learn the best features directly from the data itself, when sufficient data is available. For domains such as activity recognition where such plentiful data may be lacking, the application of deep learning may be limited. Therefore, methods which can yield deep learning-like performance without the need for massive amounts of data remain of interest. In this work we present a novel approach to feature extraction for sensor- generated activity recognition data. We first process the data using wavelet transforms, and subsequently use an adaptive pooling operator on the generated decomposition to obtain a compact, fixed-length representation of the data. Our experiments on seven different activity recognition datasets yield results comparable to those obtained from a deep neural network for all the considered datasets without the need for large amounts of data or any training overhead.

8:40AM *Generalized Alignment for Multimodal Physiological Signal Learning [#19933]*

Yuchi Liu, Yue Yao, Zhengjie Wang, Josephine Plested and Tom Gedeon, Australian National University, Australia

Revealing the correspondences and relationships between physiological signals is attractive for bioinformatics and human-computer interaction. Time alignment is a straightforward way to figure out correspondences between time sequential data. However, alignment between multimodal physiological signals is hard to achieve because the similarity metrics are difficult to define if the two physiological signals being investigated are non-linearly correlated, misaligned or quite different in morphology. In this paper, we propose a generalized time alignment method for multimodal physiological signals which (i) learns the feature extractions on physiological signals in a generalized way, and (ii) enables learned features to be in a coordinated space where the similarity between sub-components from two signals can be defined. Furthermore, we applied our alignment based multimodal feature fusion on an evaluation model to perform emotion recognition tasks on the DEAP multimodal physiological signal dataset. The experimental results show that the alignment based feature fusion outperforms the non-aligned feature fusion in most cases.

9:00AM *Dynamic Network Embedding by Semantic Evolution [#19313]*

Yujing Zhou, Weile Liu, Yang Pei, Lei Wang, Daren Zha and Tianshu Fu, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China, China

Network embedding, which aims to learn the low-dimensional representations of nodes, has attracted increasing attention in various fields such as social networks, paper citation networks and knowledge graphs. At present, most of the network embedding works are based on static networks, that is, the

evolution of networks over time is not taken into account. It is more realistic to consider temporal information in network embedding and it could also make the embedding get more abundant information. In this paper, we propose a dynamic network embedding model DynSEM with semantic evolution, to train node embeddings in a sequence of networks over time. The advantage of our method is that it presents an effective inheritance of historical information. Our method uses non-random initialization and orthogonal procrustes method to align the node embeddings into common space which makes node embedding able to inheritance information. In particular, in the common space, we train a model to capture the dynamics information of the networks and smooth temporal node embed- dings. We evaluate our method comparing it with other methods on three real-world datasets. The experimental results prove the effectiveness of dynamic network embeddings generated by DynSEM model.

9:20AM *Dealing with Limited Access to Data: Comparison of Deep Learning Approaches [#19079]*
Andreas Look and Stefan Riedelbauch, Phd Student, Germany; Professor, Germany

Nowadays deep learning is a tool, which is used in many different pattern recognition problems. For achieving strong results, which are reported in literature, often large amounts of data are necessary. In many use cases collecting or labeling data is very difficult or not possible at all. The aim of this paper is to compare three different algorithmic approaches in order to deal with limited access to labeled and unlabeled data. The drawbacks and benefits of each method are shown and compared to each other. A successful algorithmic approach, which is especially successful for one- and few-shot learning problems, is the usage of external data during the classification task. In this paper siamese neural networks will be investigated in order to evaluate how the usage of external data improves the accuracy. Another widely used approach is consistency regularization. Using consistency regularization state of the art results in semi-supervised learning (SSL) benchmarks are achieved. Virtual adversarial training (VAT) has shown strong results and is chosen as a representative algorithm for consistency regularization. The last approach is the usage of generative adversarial networks (GANs). In literature GANs are often used in order to create additional data and therefor to increase the generalization capability of the classification network. Furthermore the usage of unlabeled data for further performance improvement is considered. The use of unlabeled data is investigated for GANs and VAT.

9:40AM *Face Age Transformation with Progressive Residual Adversarial Autoencoder [#20435]*
Xuexiang Zhang, Ping Wei and Nanning Zheng, Xi'an Jiaotong University, Xi'an, China, China

Face age transformation is an important issue in many applications. While unidirectional and short-span face ageing has achieved remarkable progress, it remains a challenging problem to generate both younger-look and older-look face images over long age span. In this paper, we present a progressive residual adversarial autoencoder (PRAA) model for bidirectional and long-span face age transformation. Given an input face image, our model aims to synthesize face images of its younger looks (face rejuvenation) and older looks (face ageing). The PRAA contains adversarial generators and discriminators where age information is encoded as latent features. It adopts a residual-encoding strategy by which the original face images and the residual face images are jointly encoded. We adopt a progressive multi-scale method to train the network, by which our model can capture both the global structure and the local detail changes in face age transformation. We test our model on challenging data and the experimental results prove the strength of our method.

8: Other Applications

Wednesday, July 17, 8:00AM-10:00AM, Room: Duna Salon III, Chair: Vladimir Cherkassky

8:00AM *Deep Neural Networks for Network Routing [#20199]*
Joao Reis, Miguel Rocha, Truong Khoa Phan, David Griffin, Franck Le and Miguel Rio, University College London, United Kingdom; University of Minho, Portugal; IBM T.J. Watson Research Center, United States

In this work, we propose a Deep Learning (DL) based solution to the problem of routing traffic flows in computer networks. Routing decisions can be made in different ways depending on the desired objective and, based on that objective function, optimal solutions can be computed using a variety of techniques, e.g. with mixed integer linear programming. However, determining these solutions requires solving complex optimization problems and, thus, cannot be typically done at runtime. Instead, heuristics for these problems are often created but designing them is non-trivial in many cases. The routing framework proposed here presents an alternative to the design of heuristics, whilst still achieving good performance. This is done by building a DL model trained on the optimal decisions over flows from known traffic demands. To evaluate our solution, we focused on the problem of network congestion, even though a wide range of alternative objectives could be fitted into this framework. We ran experiments using two publicly available datasets of networks with real traffic demands and showed that our solution achieves close-to-optimal network congestion values.

8:20AM *Adaptive Edge Caching based on Popularity and Prediction for Mobile Networks [#19458]*
Li Li, Sarah Erfani, Chien Chan and Christopher Leckie, The University of Melbourne, Australia

Edge caching in mobile networks can improve users' experience, reduce latency and balance the network traffic load. However, edge caching requires suitable strategies for determining what files to pre-fetch at which cell and at what time. Due to the heterogeneity of users' content preferences and mobility, caching based only on popularity has limitations. Considering that cells located in different places have different predictability, in this paper, we propose an adaptive edge caching algorithm based on content popularity as well as the individual's prediction results to provide an optimal caching strategy, aiming to maximize the cache hit rate with acceptable file replacement cost. A heuristic optimization strategy based on genetic algorithms is presented, along with a prediction model based on an improved Markov model for each user according to the historical data. In the model, similar users are clustered based on their behavior patterns. We evaluate our algorithm on a simulation dataset as well as a 3-week real-life dataset from China Mobile. The results show that our optimal caching strategy can improve the cache hit rate compared with other methods, especially when the storage capacity is small and the similarity in content requests of users is low.

8:40AM *A Synchro-phasor Assisted Optimal Features Based Scheme for Fault Detection and Classification [#19866]*
Homanga Bharadhwaj, Avinash Kumar and Abheejeet Mohapatra, IIT Kanpur, India

A novel and efficient methodology for comprehensive fault detection and classification by using synchrophasor measurement based variations of a power system is proposed. Presently, Artificial Intelligence (AI) techniques have been used in power system protection owing to the greater degree of automation and robustness offered by AI. Evolutionary techniques like Genetic Algorithm (GA) are efficient optimization procedures mimicking the processes of biological evolution that have been shown to perform better than their gradient based counterparts in many problems. We propose a combined GA and Particle Swarm Optimization (PSO) approach to find the optimal features relevant to our fault detection process. As is evidenced by recent advances in multi-modal learning, it has been shown that this combined approach yields a more accurate feature optimization than that obtained by a single meta-heuristic. A systematic comparison of Artificial Neural Network (ANN) and Support Vector Machine (SVM) based methods

for fault classification using the identified optimal features is presented. The proposed algorithm can be effectively used for real time fault detection and also for performing postmortem analysis on signals. We demonstrate its effectiveness by simulation results on real world data from North American SynchroPhasor Initiative (NASPI) and signal variations from a test distribution system.

9:00AM *Methodology Based on ADABOOST Algorithm Combined with Neural Network for the Location of Voltage Sag Disturbance [#20301]*

Fabbio Borges, Ricardo Rabelo, Ricardo Fernandes and Marcel Araujo, Federal University of Piaui (UFPI), Brazil; Federal University of Sao Carlos (UFSCAR), Brazil; Federal Rural University of Pernambuco (UFRPE), Brazil

The correct location of the source of voltage sags is not a trivial task due to the short duration of these events and their rapid propagation in the distribution feeder. This paper proposes a method based on an ensemble method of the type Adaboost, with neural networks as base classifiers to determine the area where the voltage sag source is located. A voltage sag at a bus affects all other feeders, i.e. this disturbance is propagated in the whole system. The data management from smart meters installed in distribution feeders and decision support tools can become a viable alternative. In this sense, the smart meters could extract feature of the voltage sag and send it to the utility. At the utility, the AdaBoost performs location of the region by measuring the input's features similarity to samples from the training set. For this purpose, it was necessary to analyze the relevance of each feature extracted from smart meters'voltage signals to establish the structure which best represents the propagation of the disturbance in the system. The AdaBoost with Neural Network was tested in different scenarios of the 13-bus IEEE test feeder and was able to estimate the region with a good accuracy.

9:20AM *A Method for Voltage Sag Source Location Using Clustering Algorithm and Decision Rule Labeling [#20302]*

Jose Silva Filho, Fabbio Borges, Ricardo Rabelo and Ivan Silva, Federal University of Piaui (UFPI), Brazil

The voltage sag disturbance stands out as the most evident waveform change that is detected in electric networks, since the presence of these events in the network causes damages to the consumers. The first step in diagnosing the problem is to identify the location in the distribution system that is connected to the source causing the sinking disorder. This work presents a methodology based on clustering algorithm combined with decision rule to point out the region (cluster) that aggregates the place of origin. Clustering algorithm is responsible for analyzing the voltage signal data from different measurement nodes and separating these data into clusters. Then the Partial Decision Trees (PART) algorithm is responsible for defining the decision rule set that will confront the characteristics of each cluster and define which group aggregates the disturbance source location. For the clustering task, the k-means and fuzzy c-means clustering algorithms are evaluated and compared. The methodology was evaluated using the IEEE 34-bus test feeder system and the results show a hit rate higher than 90%.

9:40AM *Distantly Supervised Relation Extraction through a Trade-off Mechanism [#19163]*

Jun Ni, Yu Liu, Kai Wang, Zhehuan Zhao and Quan Z. Sheng, School of Software, Dalian University of Technology, China; Department of Computing, Macquarie University, Australia

Distantly supervised relation extraction can label large amounts of unstructured text without human annotations for training. However, distant supervision inevitably accompanies with the wrong labeling problem, which can deteriorate the performance of relation extraction. What's more, the entity-pair information, which can enrich instance information, is still underutilized. In the light of these issues, we propose TMNN, a novel Neural Network framework with a Trade-off Mechanism, which combines the feature of text and entity pair on the sentence level to predict relations. Our proposed trade-off mechanism is a probability generation module to dynamically adjust the weights of text and corresponding entity pair for each sentence. Experimental results on a widely used dataset show that the proposed method reduces the noisy labels and achieves substantial improvement over the state-of-the-art methods.

S10: Deep learning for brain data, S14: Evolutionary NN

Wednesday, July 17, 8:00AM-10:00AM, Room: Panorama I, Chair: Tetiana Aksenova

8:00AM *Decoding of Finger Activation from ECoG Data: a Comparative Study [#20139]*

Guillaume Jubien, Marie-Caroline Schaeffer, Stephane Bonnet and Tetiana Aksenova, Univ. Grenoble Alpes, CEA, LETI, CLINATEC, France; Univ. Grenoble Alpes, CEA, LETI, DTBS, SEIVI, LS2P, France

Motor Brain-Computer Interfaces (BCIs) are systems that allow severely motor- impaired patients to use their brain activity to interact with their environment. Electrocorticography (ECoG) arrays may be profitably used to develop safe and chronic motor BCI systems. BCI signal processing pipelines generally include neuronal signal pre-processing, feature extraction and classification/regression. The article presents a comparative study addressing the problem of neural feature classification in asynchronous multi-limb ECoG- driven BCIs. Several conventional classifiers often reported in the BCI literature were coupled with two preprocessing techniques and with a conventional feature extraction approach. They were compared to artificial neural network (ANN) end-to-end classifiers which mimic conventional BCI signal processing pipelines. Different initializations of ANNs were particularly studied. The comparison study was carried out using publicly available datasets (BCI competition IV).

8:20AM *Representation of White- and Black-Box Adversarial Examples in Deep Neural Networks and Humans: A Functional Magnetic Resonance Imaging Study [#20295]*

Chihye Han, Wonjun Yoon, Gihyun Kwon, Seungkyu Nam and Daeshik Kim, Korea Advanced Institute of Science and Technology, Korea (South); Hyundai Motor Company, Korea (South)

The recent success of brain-inspired deep neural networks (DNNs) in solving complex, high-level visual tasks has led to rising expectations for their potential to match the human visual system. However, DNNs exhibit idiosyncrasies that suggest their visual representation and processing might be substantially different from human vision. One limitation of DNNs is that they are vulnerable to adversarial examples, input images on which subtle, carefully designed noises are added to fool a machine classifier. The robustness of the human visual system against adversarial examples is potentially of great importance as it could uncover a key mechanistic feature that machine vision is yet to incorporate. In this study, we compare the visual representations of white- and black-box adversarial examples in DNNs and humans by leveraging functional magnetic resonance imaging (fMRI). We find a small but significant difference in representation patterns for different

(i.e. white- versus black-box) types of adversarial examples for both humans and DNNs. However, human performance on categorical judgment is not degraded by noise regardless of the type unlike DNN. These results suggest that adversarial examples may be differentially represented in the human visual system, but unable to affect the perceptual experience.

8:40AM *Improved Techniques for Building EEG Feature Filters [#19971]*

Yue Yao, Josephine Plested, Tom Gedeon, Yuchi Liu and Zhengjie Wang, Australian National University, Australia

Recent advances in the generative adversarial network (GAN) based image translation have shown its potential of being an image style transformer. Similarly, defined as a style transformer for physiological signals, a feature filter is used to filter privacy-related features while still keeping useful features. However, existing feature filter techniques have three problems: (1) the privacy-related features cannot be filtered out to the extent we need through a simple Conv-Deconv generator structure, and (2) the generator cannot control the semantics (maintain desired features) of given physiological signals. To address these problems, we utilize deeper neural networks and adopt techniques from domain adaptation. This includes semantic loss and a GAN based model structure with two generators, two discriminators and a classifier to form a game of five. Our results on the UCI EEG dataset demonstrate that our model can simultaneously (1) achieve the state-of-the-art accuracy removal for the privacy-related feature, (2) reduce the desired feature removal accuracy drop, and (3) make the filtered signals can be interpreted or visually checked.

9:00AM *Multi-Objective Autoencoder for Fault Detection and Diagnosis in Higher-Order Data [#19513]*

Ali Anaissi and Seid Miad Zandavi, The University of Sydney, Australia

We propose a multi-objective autoencoder method for fault detection and diagnosis in multi-way data based on the reconstruction error of autoencoder deep neural network (ADNN). Multi-way data analysis has become an essential tool for capturing underlying structures in higher-order data sets. Our method fuses data from multiple sources in one ADNN at which informative features are being extracted and utilized for anomaly detection. It also uses the generated anomaly scores to asses the severity of the anomalous data and localize it via a localization layer in the autoencoder. We evaluated our method on multi-way datasets in the area of structural health monitoring for damage detection purposes. Experimental results show that the proposed method can accurately detect structural damage. It was also able to estimate the different levels of damage severity, and capture damage locations in an unsupervised aspect. Compared to the state-of-the-art approaches, our proposed method shows better performance in terms of damage detection and localization.

9:20AM *A Prior Setting that Improves LDA in both Document Representation and Topic Extraction [#19616]*

Juncheng Ding and Wei Jin, University of North Texas, United States

Latent Dirichlet Allocation (LDA), as the most popular topic model, models documents as mixtures of topics and topics as mixtures of words. Topic mixture well represents documents while words mixture extracts meaningful topics from a corpus. The nature of LDA makes it a powerful tool in documents organizing and corpus summarizing. One limitation of LDA is that its performance depends heavily on the priors. Researchers show priors matters in LDA and propose methods to learn the priors for better modeling, regardless of using symmetric priors. However, LDA modeling ability does not necessarily consent with the performance of LDA in documents representation and topic extraction. In this paper, we propose a novel prior setting for LDA. The setting improves LDA in both documents representation and topic extraction performance. We experiment to compare our setting with symmetric priors and previously proposed priors that enhances modeling ability. Experiments on the topic quality show that LDA with our prior setting extracts better topics than LDA with other kinds of prior settings. We compare LDA document representation ability through tasks such as document clustering and document classification. These experiments demonstrate LDA with our proposed priors represents document better. Moreover, our analyses also reveal that better modeling does not necessarily lead to better performance in documents representation and topic extraction.

9:40AM *Optimization of a Convolutional Neural Network Using a Hybrid Algorithm [#19576]*

Chia-Ling Huang, Yan-Chih Shih, Chyh-Ming Lai, Vera Yuk Ying Chung, Wen-Bo Zhu, Wei-Chang Yeh and Xiangjian He, Department of Logistics and Shipping Management, Kainan University, Taiwan; Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan; Institute of Resources Management and Decision Science, Management College, National Defense University, Taiwan; School of Information Technology, University of Sydney, Australia; School of Automation, Foshan University, China; Integration and Collaboration Laboratory, Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan; Computer Vision and Recognition Laboratory, Research Centre for Innovation in IT Services and Applications, University of Technology, Sydney (UTS), Australia

In recent years, Convolutional Neural Networks (CNNs) have been widely used in image recognition due to their aptitude in large scale image processing. The CNN uses Back-propagation (BP) to train weights and biases, which in turn makes the error consistently smaller. The most common optimizers that uses a BP algorithm are Stochastic Gradient Decent (SGD), Adam, and Adadelta. These optimizers, however, have been proved to fall easily into the regional optimal solution. Little research has been conducted on the application of Soft Computing in CNN to fix the above problem, and most studies that have been conducted focus on Particle Swarm Optimization. Among them, the hybrid algorithm combined with SGD proposed by Albeahdili improves the image classification accuracy over that achieved by the original CNN. This study proposes the amalgamation of Improved Simplified Swarm Optimization (iSSO) with SGD, hence culminating in the iSSO-SGD which is intended train CNNs more efficiently to establish a better prediction model and improve the classification accuracy. The performance of the proposed iSSO-SGD can be affirmed through a comparison with the PSO-SGD, the Adam, Adadelta, rmsprop and momentum optimizers and their abilities in improving the accuracy of image classification.

2c: Reinforcement learning and adaptive dynamic programming

Wednesday, July 17, 8:00AM-10:00AM, Room: Panorama II, Chair: Chuxiong Sun

8:00AM *Efficient and Scalable Exploration via Estimation-Error [#19176]*

Chuxiong Sun, Rui Wang, Ruiying Li, Jiao Wu and XiaoHui Hu, Institute of Software Chinese Academy of Sciences(ISCAS),University of Chinese Academy of Sciences, China

Exploring efficiently in complex environments is still a challenging problem in reinforcement learning. Recent exploration algorithms based on ``optimism in the face of uncertainty" or intrinsic motivation achieved promising performance in sparse reward settings, but they often rely on additional structures which are hard to build in large scale problems. It renders them impractical and hinders the process of combining with reinforcement learning algorithms. Hence, the most state-of-the-art RL algorithms still use the naive action space noise as exploration strategy. In this paper, we model the uncertainty about environment through agent's ability to estimate the value across state and action space. Then, we parameterize the uncertainty by a neural network and regard it as a reward bonus signal to reward uncertain states. In this way, we generate an end-to-end bonus which can scale to complex environments with less computational cost. In order to prove the effectiveness of our method, we evaluate it on the challenging Atari 2600 games. We observed that our method achieves superior or comparable exploratory performance compared to action space noise in all environments, including environments whose rewards are sparse. The results demonstrate that our exploration method can motivate agent to explore effectively even in complex environments and it generally outperforms the naive action space noise.

8:20AM *A Human-Like Agent Based on a Hybrid of Reinforcement and Imitation Learning [#20026]*

Rousslan Fernand Julien Dossa, Xinyu Lian, Hirokazu Nomoto, Takashi Matsubara and Kuniaki Uehara, Graduate School of System Informatics, Kobe University, Japan; EQUOS RESEARCH Co., Ltd,, Japan

Reinforcement learning (RL) builds an effective agent that handles tasks in complex and uncertain environments by maximizing future reward. However, the efficiency is insufficient for practical use such as game AI and autonomous driving. An effective but selfish agent conflicts with other humans, and hence the demand of a human-like behavior arises. Imitation learning (IL) has been employed to trains an agent to mimic the actions of expert behaviors provided as training data. However, IL tends to build an agent limited in performance by the expert skill, and even worse, the agent exhibits an inconsistent behavior since IL is not goal-oriented. In this paper, we propose a training scheme by mixing RL and IL for both discrete and continuous action space problems. The proposed scheme builds an agent that achieves a performance higher than an agent trained by only IL and exhibits a more human-like behavior than agents trained by RL or IL, validated by human sensitivity.

8:40AM *Multi-Agent Deep Reinforcement Learning with Emergent Communication [#19388]*

David Simoes, Nuno Lau and Luis Paulo Reis, DETI/UA, IEETA, LIACC, Portugal; DETI/UA, IEETA, Portugal; LIACC, DEI/FEUP, Portugal

When compared with their single-agent counterpart, multi-agent systems have an additional set of challenges for reinforcement learning algorithms, including increased complexity, non-stationary environments, credit assignment, partial observability, and achieving coordination. Deep reinforcement learning has been shown to achieve successful policies through implicit coordination, but does not handle partial-observability. This paper describes a deep reinforcement learning algorithm, based on multi-agent actor-critic, that simultaneously learns action policies for each agent, and communication protocols that compensate for partial-observability and help enforce coordination. We also research the effects of noisy communication, where messages can be late, lost, noisy, or jumbled, and how that affects the learned policies. We show how agents are able to learn both highlevel policies and complex communication protocols for several different partially-observable environments. We also show how our proposal outperforms other communication-less state-of-the-art algorithms, even with noisy communication channels.

9:00AM *Parallel Transfer Learning in Multi-Agent Systems: What, when and how to transfer? [#19224]*

Adam Taylor, Ivana Dusparic, Maxime Gueriau and Siobhan Clarke, Trinity College Dublin, Ireland

Multi-agent Reinforcement Learning (RL) is frequently used in large-scale autonomous systems to learn the behaviours that best suit the system's operating environment. Learning can take a significant amount of time during which an RL system's performance is necessarily suboptimal. Transfer learning (TL), a method of reusing knowledge which has been gained in one task to improve the performance in another, has been used to speed up learning in single RL agent systems. TL requires learning on a source task to complete before transferring it to a target task, i.e., transfer is done offline. Parallel Transfer Learning (PTL), a technique which enables the source and target tasks to run concurrently, has been proposed to enable online transfers. However, the online selection of knowledge to be transferred, as well as online ways of integration of that knowledge on the receiving agents remain open issues. This paper proposes methods for selecting the knowledge to be transferred in PTL, frequency and size of transfers, and methods for knowledge integration into the target task. We evaluate the proposed approaches in three canonical RL examples: Cart Pole, Mountain Car, and Co-operative Predator Prey Pursuit. We show that PTL, similarly to RL, is highly sensitive to parameter selection and that suitable parameters differ per scenario.

9:20AM *Speeding Up Affordance Learning for Tool Use, Using Proprioceptive and Kinesthetic Inputs [#19228]*

Khuong Nguyen, Jaewook Yoo and Yoonsuck Choe, Texas A&M University, United States

End-to-end learning in deep reinforcement learning based on raw visual input has shown great promise in various tasks involving sensorimotor control. However, complex tasks such as tool use require recognition of affordance and a series of non-trivial subtasks such as reaching the tool, grasping the tool, and wielding the tool. In such tasks, end-to-end approaches with only pixel-wise images may fail to learn to perform the task or may take too long to converge. In this paper, inspired by the biological sensorimotor system, we explore the use of proprioceptive/kinesthetic inputs (internal inputs for body position and motion) as well as raw visual inputs (exteroception, external perception) for use in affordance learning for tool use tasks. We set up a reaching task in a simulated physics environment (MuJoCo), where the agent has to pick up a T-shaped tool to reach and drag a target object to a designated region in the environment. We used an Actor-Critic-based reinforcement learning algorithm called ACKTR and trained it using various input conditions to assess the utility of proprioceptive/kinesthetic inputs. Our results show that the inclusion of proprioceptive/kinesthetic inputs (position and velocity of the limb) greatly enhances the performance of the agent: higher success rate, and faster convergence to the solution. The lesson we learned is the important factor of the intertwined relationship of exteroceptive and proprioceptive in sensorimotor learning and that although end-to-end learning based on raw input may be appealing, separating the exteroceptive and proprioceptive/kinesthetic factors in the input to the learner, and providing the necessary internal inputs can lead to faster, more effective learning.

S18: Neuro-Inspired Computing with Nano-electronic Devices

Wednesday, July 17, 8:00AM-10:00AM, Room: Panorama III, Chair: Saibal Mukhopadhyay

8:00AM *FPCAS: In-Memory Foating Point Computations for Autonomous Systems [#20506]*
Sina Sayyah Ensan and Swaroop Ghosh, Pennsylvania State University, United States

Autonomous systems e.g., cars and drones generate vast amount of data from sensors that need to be processed in timely fashion to make accurate and safe decisions. Majority of these computations deal with Floating Point (FP) numbers. Conventional Von-Neumann computing paradigm suffers from overheads associated with data transfer. In-memory computing (IMC) can solve this challenge by processing the data locally. However, in-memory FP computing has not been investigated before. In this paper, we propose a FP arithmetic (adder/subtractor and multiplier) using Resistive RAM (ReRAM) crossbar based IMC. A novel shift circuitry is proposed to lower the shift overhead inherently present in the FP arithmetic. Simulation results show that the proposed single precision FP adder consumes 335 pJ and 322 pJ for NAND-NAND and NOR-NOR based implementation for addition/subtraction, respectively. The proposed adder/subtractor compared to MAGIC, improves latency, power. and energy by 828X, 3.2X, and 3.7X, respectively. Furthermore, the proposed multiplier reduces energy per operation by 1.13X and improves performance by 4.4X compared to ReVAMP.

8:20AM *Investigation of Neural Networks Using Synapse Arrays Based on Gated Schottky Diodes [#19992]*
Suhwan Lim, Dongseok Kwon, Sung-Tae Lee, Hyeongsu Kim, Jong-Ho Bae and Jong-Ho Lee, Seoul National University, Korea (South)

A synapse device array based on the gated Schottky diodes (GSDs) is fabricated. This GSD operates in reverse mode, so the synapse current is considerably low, helping to implement a low-power hardware-based neural networks (HNNs). The reverse Schottky diode current, which represents the synaptic weight, is modulated by applying program or erase pulses. In this GSD, the reverse diode current is used as the synapse current, but the forward diode current is cut off. This is an important feature to prevent the sneak path problem in the crossbar array. A synapse array consisting of fabricated 200 GSDs shows a variation of 0.34, 0.22, and 0.14 for three different synaptic weight states. By using this GSD array, we perform the vector-by-matrix multiplication, and evaluate the inference accuracy of MNIST. As a baseline accuracy for MNIST classification, a convolutional neural network similar to Lenet-5 is designed and gives an accuracy of 99.53%. Normalization method is applied to the weights trained in the network to map the weights into the conductance range of synapse device. The adaptive weight quantization is then applied to the normalized weights. We verify that the HNN using GSDs works well in comparison to the baseline network even in the presence of nonideal characteristics of synapse devices.

8:40AM *On Robustness of Spin-Orbit-Torque Based Stochastic Sigmoid Neurons for Spiking Neural Networks [#20326]*
Akhilesh Jaiswal, Amogh Agrawal, Indranil Chakraborty, Deboleena Roy and Kaushik Roy, Purdue University, United States

Nano-scale neuro-mimetic devices have recently gained wide research interest in the quest to enable brain-like energy-efficiency with cognitive computing abilities. Traditionally, neuromorphic devices have exploited deterministic nano-scale devices for emulating the intrinsic neuronal and synaptic behavior. However, of particular interest are stochastic neuromorphic devices owing to - 1) availability of nano-scale devices that are inherently stochastic based on intrinsic device physics 2) various neuroscience experiments have convincingly demonstrated that cortical neurons are stochastic in nature. In this paper, we focus on spin-orbit torque based Magnetic Tunnel Junction (SOT-MTJ) that exhibit stochastic sigmoid behavior with respect to the switching process. We first discuss the modeling framework that was used to study the effect of dimensional variations in SOT-MTJs and the resulting changes in the stochastic sigmoid behavior. Our model is based on the wellknown stochastic-Landau-Lifshitz-Gilbert-Slonczewski equation under mono-domain approximation. Subsequently, we abstract the sigmoid characteristic of the device into a behavioral model and study the effect of variations in sigmoid characteristics on a deep binary network. Our results show that the variations in the sigmoidal neuron behavior results in a minimal loss in accuracy (for CIFAR 10 dataset), additionally, the degradation in accuracy monotonically increases with increase in induced variations. This highlights the robustness of stochastic neural networks based on SOT-MTJs in presence of dimensional variations.

9:00AM *Improving Robustness of ReRAM-based Spiking Neural Network Accelerator with Stochastic Spike-timing-dependent-plasticity [#20239]*
Xueyuan She, Yun Long and Saibal Mukhopadhyay, Georgia Institute of Technology, United States

Spike-timing-dependent-plasticity (STDP) is an unsupervised learning algorithm for spiking neural network (SNN), which promises to achieve deeper understanding of human brain and more powerful artificial intelligence. While conventional computing system fails to simulate SNN efficiently, process-in-memory (PIM) based on devices such as ReRAM can be used in designing fast and efficient STDP based SNN accelerators, as it operates in high resemblance with biological neural network. However, the real-life implementation of such design still suffers from impact of input noise and device variation. In this work, we present a novel stochastic STDP algorithm that uses spiking frequency information to dynamically adjust synaptic behavior. The algorithm is tested in pattern recognition task with noisy input and shows accuracy improvement over deterministic STDP. In addition, we show that the new algorithm can be used for designing a robust ReRAM based SNN accelerator that has strong resilience to device variation.

9:20AM *Improving Noise Tolerance of Mixed Signal Neural Networks [#20497]*
Michael Klachko, Mohammad Mahmoodi and Dmitri Strukov, UCSB, United States

Mixed-signal hardware accelerators for deep learning achieve orders of magnitude better power efficiency than their digital counterparts. In the ultra-low power consumption regime, limited signal precision inherent to analog computation becomes a challenge. We perform a case study of a 6-layer convolutional neural network running on a mixed-signal accelerator and evaluate its sensitivity to hardware specific noise. We apply various methods to improve noise robustness of the network and demonstrate an effective way to optimize useful signal ranges through adaptive signal clipping. The resulting model is robust enough to achieve 80% classification accuracy on CIFAR-10 dataset with just 2mW power budget, while 10mW budget allows us to achieve 87% accuracy, which is within 1% of the software baseline. For comparison, the unoptimized version of the same model fails to converge at 2mW and achieves 61% accuracy at 10mW.

9:40AM *An Electronic Neuron with Input-Specific Spiking [#19986]*
Rebecca Lee and Alice Parker, University of Southern California, United States

Sensory information in the brain is encoded by intricate spiking patterns. Transient characteristics of individual action potentials may be used in neural encoding. Inspired by behaviors found in biological neurons, we have designed a neuromorphic circuit of a neuron that exhibits input-specific spiking. Our neuron circuit has multiple dendrites and uses prop- erties of localized high-voltage-activated (HVA) Ca2+ channels to generate action potentials with distinct shapes depending on the location of stimulation received at input synapses. Through circuit simulations, we show that our neuron circuit can encode differences in spatial locations of input stimuli through precise characteristics of output spikes, and spike shapes are tunable. We then show that neuron circuits, in conjunction with astrocytes, can be used to replace damaged neurons processing sensory inputs with healthy neurons augmented to process their own sensory inputs as well as the inputs originally intended for the damaged neurons, using spiking shapes to signal input sources.

S05: Deep Neural Audio Processing

Wednesday, July 17, 8:00AM-10:00AM, Room: Panorama IV, Chair: Leonardo Gabrielli

8:00AM *RNN-based speech synthesis using a continuous sinusoidal model [#19454]*

Mohammed Salah Al-Radhi, Tamas Gabor Csapo and Geza Nemeth, Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary

Recently in statistical parametric speech synthesis, we proposed a continuous sinusoidal model (CSM) using continuous F0 (contF0) in combination with Maximum Voiced Frequency (MVF), which was successfully giving state-of-the-art vocoders performance (e.g. similar to STRAIGHT) in synthesized speech. In this paper, we address the use of sequence-to-sequence modeling with recurrent neural networks (RNNs). Bidirectional long short-term memory (Bi-LSTM) is investigated and applied using our CSM to model contF0, MVF, and Mel-Generalized Cepstrum (MGC) for more natural sounding synthesized speech. For refining the output of the contF0 estimation, post-processing based on time-warping approach is applied to reduce the unwanted voiced component of the unvoiced speech sounds, resulting in an enhanced contF0 track. The overall conclusion is covered by objective evaluation and subjective listening test, showing that the proposed framework provides satisfactory results in terms of naturalness and intelligibility, and is comparable to the high-quality WORLD model based RNNs.

8:20AM *Processing Acoustic Data with Siamese Neural Networks for Enhanced Road Roughness Classification [#20025]*

Leonardo Gabrielli, Livio Ambrosini, Fabio Vesperini, Valeria Bruschi, Stefano Squartini and Luca Cattani, Universita' Politecnica delle Marche, Italy; ASK Industries SpA, Italy

In recent years, a lot of effort has been put in vehicle safety systems for manned and unmanned driving. Road conditions are crucial among the factors that influence the choice of the driving style and the safety systems. A few works based the detection of the road condition on acoustic sensors mounted on the vehicle using deep learning techniques. In this work we enhance the state of the art by introducing a Siamese Convolutional Neural Network architecture able to achieve improved results for the classification of the road surface roughness. A new dataset is recorded and the approach is tested, achieving a best overall F1- score of 95.6%, improving by 14% the results of the previous method.

8:40AM *Transfer Learning for Piano Sustain-Pedal Detection [#19340]*

Beici Liang, Gyorgy Fazekas and Mark Sandler, Queen Mary University of London, United Kingdom

Detecting piano pedalling techniques in polyphonic music remains a challenging task in music information retrieval. While other piano-related tasks, such as pitch estimation and onset detection, have seen improvement through applying deep learning methods, little work has been done to develop deep learning models to detect playing techniques. In this paper, we propose a transfer learning approach for the detection of sustain-pedal techniques, which are commonly used by pianists to enrich the sound. In the source task, a convolutional neural network (CNN) is trained for learning spectral and temporal contexts when the sustain pedal is pressed using a large dataset generated by a physical modelling virtual instrument. The CNN is designed and experimented through exploiting the knowledge of piano acoustics and physics. This can achieve an accuracy score of 0.98 in the validation results. In the target task, the knowledge learned from the synthesised data can be transferred to detect the sustain pedal in acoustic piano recordings. A concatenated feature vector using the activations of the trained convolutional layers is extracted from the recordings and classified into frame-wise pedal press or release. We demonstrate the effectiveness of our method in acoustic piano recordings of Chopin's music. From the cross-validation results, the proposed transfer learning method achieves an average F-measure of 0.89 and an overall performance of 0.84 obtained using the micro-averaged F-measure. These results outperform applying the pre-trained CNN model directly or the model with a fine-tuned last layer.

9:00AM *Cosine-similarity penalty to discriminate sound classes in weakly-supervised sound event detection [#19523]*

Thomas Pellegrini and Leo Cances, UPS - IRIT, France

The design of new methods and models when only weakly-labeled data are available is of paramount importance in order to reduce the costs of manual annotation and the considerable human effort associated with it. In this work, we address Sound Event Detection in the case where a weakly annotated dataset is available for training. The weak annotations provide tags of audio events but do not provide temporal boundaries. The objective is twofold: 1) audio tagging, i.e. multi-label classification at recording level, 2) sound event detection, i.e. localization of the event boundaries within the recordings. This work focuses mainly on the second objective. We explore an approach inspired by Multiple Instance Learning, in which we train a convolutional recurrent neural network to give predictions at frame-level, using a custom loss function based on the weak labels and the statistics of the frame-based predictions. Since some sound classes cannot be distinguished with this approach, we improve the method by penalizing similarity between the predictions of the positive classes during training. On the test set used in the DCASE 2018 challenge, consisting of 288 recordings and 10 sound classes, the addition of a penalty resulted in a localization F-score of 34.75%, and brought 10% relative improvement compared to not using the penalty. The approach also outperformed a false strong labeling baseline and an attention-based model. Our best model achieved a 26.20% F-score on the DCASE- 2018 official Eval subset close to the 10-system ensemble approach that ranked second in the challenge with a 29.9% F-score.

9:20AM *Representation Learning vs. Handcrafted Features for Music Genre Classification [#19878]*

Rodolfo M. Pereira, Yandre M. G. Costa, Rafael L. Aguiar, Alceu S. Britto Jr., Luiz E. S. Oliveira and Carlos N. Silla Jr., Pontifical Catholic University of Parana and Federal Institute of Parana - Pinhais, Brazil; State University of Maringa, Brazil; Pontifical Catholic University of Parana, Brazil; Federal University of Parana, Brazil

In this work we present a comprehensive set of experiments aiming to perform music genre classification using learned and handcrafted features plus the fusion of them. Handcrafted features were obtained from the audio signal itself, lyrics, chords and spectrogram images extracted from the audio. The rationale behind this investigation is based on the assumption that one can find some complementarity between classifiers created from these different resources. The experimental protocol was conducted on the Brazilian Music Dataset using the artist filter restriction and they confirm the power of non-handcrafted features to perform audio classification tasks. The experimental results have shown a significant complementarity among the handcrafted features for which the evaluated fusion strategies allowed an improvement in the classification accuracy up to 4 percent points. On the other hand, the fusion of learned and handcrafted features provided similar accuracy than the best individual CNN (0.7815).

9:40AM *Audio-based Recognition of Bipolar Disorder Utilising Capsule Networks [#19242]*
Shahin Amiriparian, Arsany Awad, Maurice Gerczuk, Lukas Stappen, Alice Baird, Sandra Ottl and Bjoern Schuller, University of Augsburg, Germany
Bipolar disorder (BD) is an acute mood condition, in which states can drastically shift from one extreme to another, considerably impacting an individual's wellbeing. Automatic recognition of a BD diagnosis can help patients to obtain medical treatment at an earlier stage and therefore have a better overall prognosis. With this in mind, in this study, we utilise a Capsule Neural Network (CapsNet) for audio-based classification of patients who were suffering from BD after a mania episode into three classes of Remission, Hypomania, and Mania. The CapsNet attempts to address the limitations of Convolutional Neural Networks (CNNs) by considering vital spatial hierarchies between the extracted images from audio files. We develop a framework around the CapsNet in order to analyse and classify audio signals. First, we create a spectrogram from short segments of speech recordings from individuals with a bipolar diagnosis. We then train the CapsNet on the spectrograms with 32 low-level and three high-level capsules, each for one of the BD classes. These capsules attempt both to form a meaningful representation of the input data and to learn the correct BD class. The output of each capsule represents an activity vector. The length of this vector encodes the presence of the corresponding type of BD in the input, and its orientation represents the properties of this specific instance of BD. We show that using our CapsNet framework, it is possible to achieve competitive results for the aforementioned task by reaching a UAR of 46.2% and 45.5% on the development and test partitions, respectively. Furthermore, the efficacy of our approach is compared with a sequence to sequence autoencoder and a CNN-based neural network.

Competition: AIML Contest 2019
Wednesday, July 17, 8:00AM-10:00AM, Room: Panorama V, Chair: Juan L. Castro-Garcia, Xiang Wu. Juyang Weng

Wednesday, July 17, 10:00AM-10:30AM

Special Lecture: Coffee Break
Wednesday, July 17, 10:00AM-10:30AM, Room: Pre-function area Intercontinental

Wednesday, July 17, 10:30AM-11:30AM

Plenary Talk: Nik Kasabov, KEDRI, Auckland University of Technology
Wednesday, July 17, 10:30AM-11:30AM, Room: Ballroom I + II +II, Chair: Marley Vellasco,

Wednesday, July 17, 11:30AM-12:30PM

Plenary Talk: Danil Prokhorov, Toyota R&D
Wednesday, July 17, 11:30AM-12:30PM, Room: Ballroom I + II +II, Chair: Asim Roy,

Wednesday, July 17, 12:30PM-2:00PM

Special Lecture: Lunch Break
Wednesday, July 17, 12:30PM-2:00PM, Room: Various locations in the area

Wednesday, July 17, 2:00PM-4:00PM

S09: Metrology of AI: blessing of dimensionality, tolerance and fits

Wednesday, July 17, 2:00PM-4:00PM, Room: Ballroom I, Chair: Danil Prokhorov

2:00PM *Do Fractional Norms and Quasinorms Help to Overcome the Curse of Dimensionality? [#19331]*

Evgeny M. Mirkes, Jeza Allohibi and Alexander N. Gorban, University of Leicester, Lobachevsky State University, United Kingdom; University of Leicester, United Kingdom

The curse of dimensionality causes well-known and widely discussed problems for machine learning methods. There is a hypothesis that usage of Manhattan distance and even fractional quasinorms lp (for p less than 1) can help to overcome the curse of dimensionality in classification problems. In this study, we systematically test this hypothesis for 37 binary classification problems on 25 databases.We confirm that fractional quasinorms have greater relative contrast or coefficient of variation than Euclidean norm l2, but we demonstrate also that the distance concentration shows qualitatively the same behaviour for all tested norms and quasinorms and the difference between them decays while dimension tends to infinity. Estimation of classification quality for kNN based on different norms and quasinorms shows that the greater relative contrast does not mean the better classifier performance and the worst performance for different databases was shown by the different norms (quasinorms). A systematic comparison shows that the difference in performance of kNN based on lp for p=2, 1, and 0.5 is statistically insignificant.

2:20PM *Practical Stochastic Separation Theorems for Product Distributions [#19556]*

Bogdan Grechuk, University of Leicester, United Kingdom

Stochastic separation theorems provide mathematical foundation for construction an extremely efficient mechanism for error correction in artificial intelligence systems. They imply that this mechanism works even if the number of points in the dataset is exponentially large in terms of dimension of the underlying space. However, in most such theorems in the literature, the bound for the size of the dataset in terms of dimension is either inexplicit or impractical for large but not extremely large dimensions (such as few hundreds or one thousand). In this work, we derive much less restrictive estimates for dataset size in terms of dimension, which still sufficient to guarantee Fisher separability with large probability, provided that data follow product distributions in the unit cube.

2:40PM *Toward Next Generation of Autonomous Systems with AI [#19912]*

Danil Prokhorov, Toyota, United States

I discuss a growing area of research in autonomous driving systems and overview what this means in terms of their testing. I also discuss implications for autonomous decision making systems of the future.

3:00PM *Estimating the effective dimension of large biological datasets using Fisher separability analysis [#19814]*

Luca Albergante, Jonathan Bac and Andrei Zinovyev, Institut Curie, France; Paris Diderot University, France

Modern large-scale datasets are frequently said to be high-dimensional. However, their data point clouds frequently possess structures, significantly decreasing their intrinsic dimensionality (ID) due to the presence of clusters, points being located close to low-dimensional varieties or fine-grained lumping. We introduce and test a dimensionality estimator, based on analysing the separability properties of data points, on several benchmarks and real biological datasets. We show that the introduced measure of ID has performance competitive with state-of- the-art measures, being efficient across a wide range of dimensions and performing better in the case of noisy samples. Moreover, it allows estimating the intrinsic dimension in situations where the intrinsic manifold assumption is not valid.

3:20PM *Kernel Stochastic Separation Theorems and Separability Characterizations of Kernel Classifiers [#20219]*

Ivan Y. Tyukin, Alexander N. Gorban, Bogdan Grechuk and Stephen Green, Univerity of Leicester, United Kingdom

In this work we provide generalizations and extensions of stochastic separation theorems to kernel classifiers. A general separability result for two random sets is also established. We show that despite feature maps corresponding to a given kernel function may be infinite-dimensional, kernel separability characterizations can be expressed in terms of finite-dimensional volume integrals. These integrals allow to determine and quantify separability properties of an arbitrary kernel function. The theory is illustrated with numerical examples.

3:40PM *Deep Learning of p73 Biomarker Expression in Rectal Cancer Patients [#19612]*

Tuan Pham, Chuanwen Fan, Hong Zhang and Xiao-Feng Sun, Linkoping University, Sweden; Orebro University, Sweden

By applying deep learning, we were able to compare p73 protein expression patterns of different tissue types including normal mucosa, primary tumor and lymph node metastasis in rectal cancer patients using immunohistochemical slides. The pair-wise pattern comparisons were automatedly carried out by considering color, edge, blobs, and other morphological information in the images. We discovered that when the pattern dissimilarity between primary tumor and lymph node metastasis is relatively low among other tissue pairs (primary tumor and distant normal, biopsy and distant normal, biopsy and primary tumor, biopsy and primary tumor, lymph node metastasis and distant normal, lymph node metastasis and biopsy), there was an implication of short-time survival. This original result suggests a novel application of advanced artificial intelligence in machine learning for clinical finding in rectal cancer and encourages relevant study of multiple biomarker expressions in cancer patients.

S22: Artificial Intelligence and Security (AISE)

Wednesday, July 17, 2:00PM-4:00PM, Room: Ballroom II, Chair: Francesco Mercaldo

2:00PM *Keystroke Analysis for User Identification using Deep Learning Networks [#20334]*

Mario Bernardi, Marta Cimitile, Fabio Martinelli and Francesco Mercaldo, Giustino Fortunato University, Italy; Unitelma Sapienza University, Italy; Institute for Informatics and Telematics, National Research Council of Italy (CNR), Italy

The current authentication systems based on password and pin code are not enough to guarantee attacks from malicious users. For this reason, in the last years, several studies are proposed with the aim to identify the users basing on their typing dynamics. In this paper, we propose a deep learning architecture aimed to discriminate between different users using a set of keystroke features. The idea behind the proposed method is to identify the users silently and continuously during their typing on a monitored system. To perform such user identification effectively, we propose a feature model that is able to capture the typing style that is specific to each given user. The proposed approach is evaluated on a large dataset derived by integrating two real-world datasets from existing studies. The merged dataset contains a total of 1530 different users each writing a set of different typing samples. A deep learning architecture with an increasing number of hidden layers and two different sets of features are tested with the aim to find the best configuration. The final best classifier scores a precision equal to 0.99, a recall equal to 0.987 and an accuracy equal to 0.97 using a neural network with 9 hidden layers. Finally, the performances obtained by using the deep learning approach are also compared with the performance of traditional machine learning algorithm, attesting the effectiveness of the deep learning-based classifiers in the domain of keystroke analysis.

2:20PM *NeuralAS: Deep Word-Based Spoofed URLs Detection Against Strong Similar Samples [#19132]*

Jing Ya, Tingwen Liu, Panpan Zhang, Jinqiao Shi, Li Guo and Zhaojun Gu, University of Chinese Academy of Sciences, China; Chinese Academy of Sciences, China; Civil Aviation University of China, China

Spoofed URLs are associated with various cyber crimes such as phishing and ransomware etc. Most existing detection approaches design a set of hand-crafted features and feed them to machine learning classifiers. However, designing such features is a time consuming and labor intensive process. This paper proposes an approach named NeuralAS (Neural Anti-Spoofing) by segmenting URLs into word sequences and detecting spoofed URLs with recurrent neural networks. As a result, NeuralAS can perform detection with high-abstract and poor-interpretable features learned automatically, and achieve accurate detection with contextual information in sequences. We also propose a novel method to construct indistinguishable data sets of strong similar samples, which can be used to evaluate the robustness of different approaches. Extensive experimental results show that NeuralAS works well on spoofed URLs detection, and has a significant effectiveness and robustness even on strong similar data sets.

2:40PM *TrustSign: Trusted Malware Signature Generation in Private Clouds Using Deep Feature Transfer Learning. [#19744]*

Daniel Nahmias, Aviad Cohen, Nir Nissim and Yuval Elovici, Ben-Gurion University, Israel

This paper presents TrustSign, a novel, trusted automatic malware signature generation method based on high-level deep features transferred from a VGG- 19 neural network model pre-trained on the ImageNet dataset. While traditional automatic malware signature generation techniques rely on static or dynamic analysis of the malware's executable, our method overcomes the limitations associated with these techniques by producing signatures based on the presence of the malicious process in the volatile memory. Signatures generated using TrustSign well represent the real malware behavior during runtime. By leveraging the cloud's virtualization technology, TrustSign analyzes the malicious process in a trusted manner, since the malware is unaware and cannot interfere with the inspection procedure. Additionally, by removing the dependency on the malware's executable, our method is capable of signing fileless malware. Thus, we focus our research on in-browser cryptojacking attacks, which current antivirus solutions have difficulty to detect. However, TrustSign is not limited to cryptojacking attacks, as our evaluation included various ransomware samples. TrustSign's signature generation process does not require feature engineering or any additional model training, and it is done in a completely unsupervised manner, obviating the need for a human expert. Therefore, our method has the advantage of dramatically reducing signature generation and distribution time. The results of our experimental evaluation demonstrate TrustSign's ability to generate signatures invariant to the process state over time. By using the signatures generated by TrustSign as input for various supervised classifiers, we achieved 99.5% classification accuracy.

3:00PM *Social Network Polluting Contents Detection through Deep Learning Techniques [#19517]*

Martinelli Fabio, Mercaldo Francesco and Santone Antonella, IIT-CNR, Italy; University of Molise, Italy

Nowadays social networks are widespread used not only to enable users to share comments with other users but also as tool from which is possible to extract knowledge. As a matter of fact, social networks are increasingly considered to understand the opinion trend about a politician or related to a certain event that occurred: in general social networks have been proved useful to understand the public opinion from both governments and companies. In addition, also from the end users point of view it is difficult to identify real contents. This is the reason why in last years we are witnessing a growing interest in tools for analyzing big data gathered from social networks in order to find common opinions. In this context, content polluters on social networks make the opinion mining process difficult to browse valuable contents. In this paper we propose a method aimed to discriminate between pollute and real information from a semantic point of view. We exploit a combination of word embedding and deep learning techniques to categorize semantic similarities between (pollute and real) linguistic sentences. We experiment the proposed method on a dataset composed of real-world sentences gathered from the Twitter social network obtaining interesting results in terms of precision and recall.

3:20PM *Cascade Learning for Mobile Malware Families Detection through Quality and Android Metrics [#19516]*

Fasano Fausto, Martinelli Fabio, Mercaldo Francesco and Santone Antonella, University of Molise, Italy; IIT-CNR, Italy

Considering the increasing diffusion of mobile devices, attackers started to explore the possibility to perpetrate attacks using mobile surfaces (i.e., smartphones and tablets). Unfortunately, common antimalware techniques are often ineffective to detect new threats with the current signature based approach mainly adopted. In this we propose a set of features with the aim to discriminate between malware and trusted mobile applications: in detail we design a cascade learner where the first classifier of the cascade performs a coarse-grain analysis (it discriminates between malware and trusted apps), while the second one performs a fine-grain analysis (it is aimed to identify the malware family). We obtain a precision equal to 0.947 and a recall equal to 0.962 in trusted samples identification, while a precision equal to 0.961 and a recall equal to 0.946 is obtained in malware detection. With regard to family identification, an average precision and recall of 0.961 is obtained across 12 malware families.

3:40PM *An Adversarial Perturbation Approach Against CNN-based Soft Biometrics Detection [#20376]*

Stefano Marrone and Carlo Sansone, University of Naples Federico II, Italy

The use of biometric-based authentication systems spread over daily life consumer electronics. Over the years, researchers' interest shifted from hard (such as fingerprints, voice and keystroke dynamics) to soft biometrics (such as age, ethnicity and gender), mainly by using the latter to improve the authentication systems effectiveness. While newer approaches are constantly being proposed by domain experts, in the last years Deep

Learning has raised in many computer vision tasks, also becoming the current state-of-art for several biometric approaches. However, since the automatic processing of data rich in sensitive information could expose users to privacy threats associated to their unfair use (i.e. gender or ethnicity), in the last years researchers started to focus on the development of defensive strategies in the view of a more secure and private AI. The aim of this work is to exploit Adversarial Perturbation, namely approaches able to mislead state-of-the-art CNNs by injecting a suitable small perturbation over the input image, to protect subjects against unwanted soft biometrics-based identification by automatic means. In particular, since ethnicity is one of the most critical soft biometrics, as a case of study we will focus on the generation of adversarial stickers that, once printed, can hide subjects ethnicity in a real-world scenario.

Deep Reinforcement Learning for Autonomous Driving

Wednesday, July 17, 2:00PM-4:00PM, Room: Ballroom III, Chair: Qichao Zhang

2:00PM *Deep Learning for System Trace Restoration [#20119]*

Ilia Sucholutsky, Apurva Narayan, Matthias Schonlau and Sebastian Fischmeister, University of Waterloo, Canada

Most real-world datasets, and particularly those collected from physical systems, are full of noise, packet loss, and other imperfections. However, most specification mining, anomaly detection and other such algorithms assume, or even require, perfect data quality to function properly. Such algorithms may work in lab conditions when given clean, controlled data, but will fail in the field when given imperfect data. We propose a method for accurately reconstructing discrete temporal or sequential system traces affected by data loss, using Long Short-Term Memory Networks (LSTMs). The model works by learning to predict the next event in a sequence of events, and uses its own output as an input to continue predicting future events. As a result, this method can be used for data restoration even with streamed data. Such a method can reconstruct even long sequence of missing events, and can also help validate and improve data quality for noisy data. The output of the model will be a close reconstruction of the true data, and can be fed to algorithms that rely on clean data. We demonstrate our method by reconstructing automotive CAN traces consisting of long sequences of discrete events. We show that given even small parts of a CAN trace, our LSTM model can predict future events with an accuracy of almost 90%, and can successfully reconstruct large portions of the original trace, greatly outperforming a Markov Model benchmark. We separately feed the original, lossy, and reconstructed traces into a specification mining framework to perform downstream analysis of the effect of our method on state- of-the-art models that use these traces for understanding the behavior of complex systems.

2:20PM *Clustering-enhanced PointCNN for Point Cloud Classification Learning [#19248]*

Yikuan Yu, Fei Li, Yu Zheng, Min Han and Xinyi Le, Shanghai Jiao Tong University, China; Beijing Institute of Electronic System Engineering,, China; Dalian University of Technology, China

3D shape feature learning plays a pivotal role in both industry and academia. PointCNN is one of excellent neural networks for 3D object databases classification. Instead of selecting representative points arbitrarily in PointCNN, clustering-enhanced PointCNN proposed in this paper can make representative points more logical and efficient for point cloud classification learning. The proposed clustering-based selection approach is able to distinguish more features and catch more details from 3D shapes. Both K-Means and Gaussian-Mixture-Model (GMM) clustering methods are applied during the point selection period. Both methods have been tested on several public data sets, which substantiates the superior classification accuracy with comparable training time.

2:40PM *Learning Private Neural Language Modeling with Attentive Aggregation [#19564]*

Shaoxiong Ji, Shirui Pan, Guodong Long, Xue Li, Jing Jiang and Zi Huang, The University of Queensland, Australia; Monash University, Australia; University of Technology Sydney, Australia

Mobile keyboard suggestion is typically regarded as a word-level language modeling problem. Centralized machine learning techniques require the collection of massive user data for training purposes, which may raise privacy concerns in relation to users' sensitive data. Federated learning (FL) provides a promising approach to learning private language modeling for intelligent personalized keyboard suggestions by training models on distributed clients rather than training them on a central server. To obtain a global model for prediction, existing FL algorithms simply average the client models and ignore the importance of each client during model aggregation. Furthermore, there is no optimization for learning a well- generalized global model on the central server. To solve these problems, we propose a novel model aggregation with an attention mechanism considering the contribution of client models to the global model, together with an optimization technique during server aggrega- tion. Our proposed attentive aggregation method minimizes the weighted distance between the server model and client models by iteratively updating parameters while attending to the distance between the server model and client models. Experiments on two popular language modeling datasets and a social media dataset show that our proposed method outperforms its counterparts in terms of perplexity and communication cost in most settings of comparison.

3:00PM *Model-Free Temporal Difference Learning for Non-Zero-Sum Games [#19422]*

Liming Wang, Yongliang Yang, Dawei Ding, Yixin Yin, Zhishan Guo and Donald Wunsch, University of Science and Technology Beijing, China; University of Central Florida, United States; Missouri University of Science and Technology, United States

In this paper, we consider the two-player non-zero-sum games problem for continuous- time linear dynamic systems. It is shown that the non-zero-sum games problem results in solving the coupled algebraic Riccati equations, which are nonlinear algebraic matrix equations. Compared with the algebraic Riccati equation of the linear dynamic systems with only one player, the coupled algebraic Riccati equations of non-zero- sum games with multi-player are more difficult to be solved directly. First, the policy iteration algorithm is introduced to find the Nash equilibrium of the non- zero-sum games, which is the sufficient and necessary condition to solve the coupled algebraic Riccati equations. However, the policy iteration algorithm is offline and requires complete knowledge of the system dynamics. To overcome the above issues, a novel online iterative algorithm, named integral temporal difference learning algorithm, is developed. Moreover, an equivalent compact form of the integral temporal difference learning algorithm is also presented. It is shown that the integral temporal difference learning algorithm can be implemented in an online fashion and requires only partial knowledge of the system dynamics. In addition, in each iteration step, the closed-loop stability using the integral temporal difference learning algorithm is analyzed. Finally, the simulation study shows the effectiveness of the presented algorithm.

3:20PM *Lane Change Decision-making through Deep Reinforcement Learning with Rule-based Constraints [#20518]*

Junjie Wang, Qichao Zhang, Dongbin Zhao and Yaran Chen, Institute of Automation, Chinese Academy of Sciences, China

Autonomous driving decision-making is a great challenge due to the complexity and uncertainty of the traffic environment. Combined with the rule-based constraints, a Deep Q-Network (DQN) based method is applied for autonomous driv- ing lane change decision-making task. Through the combination of high-level lateral decision-making and low-level rule-based trajectory modification, a safe and efficient lane change behavior can be achieved. With the setting of our state representation and reward function, the trained agent is able to take appropriate actions in a real-world-like simulator. The generated policy is evaluated on the simulator for 10 times, and the results demon- strate that the proposed rule-based DQN method outperforms the rule- based approach and the DQN method.

3:40PM *Model-Free Reinforcement Learning based Lateral Control for Lane Keeping [#20514]*

Qichao Zhang, Rui Luo, Dongbin Zhao, Chaomin Luo and Dianwei Qian, Institute of Automation, Chinese Academy of Sciences, China; North China Electric Power University, China; Department of Electrical and Computer Engineering, University of Detroit Mercy, United States; School of Control and Computer Engineering, North China Electric Power Unverisity, China

In this paper, the lateral control strategy for lane keeping task, which is an important module in the advanced assistant driver systems, is proposed based on the model-free reinforcement learning. Different from the model-based methods, our method only requires the generated data rather than the accurate system model. Furthermore, the lateral control strategy for driver model lane keeping is given, where driver controller and direct yaw controller (DYC) are working at the same time to maintain the vehicle stability. Note that the dynamic game theory is considered for this task, where the steering wheel controller for driver and the DYC compensated controller are obtained based on Nash game theory. Finally, we give simulation examples to prove the validity of the proposed schemes.

8n: Data mining and knowledge discovery

Wednesday, July 17, 2:00PM-4:00PM, Room: Duna Salon I, Chair: Erik Cambria

2:00PM *MMF: Attribute Interpretable Collaborative Filtering [#19130]*

Yixin Su, Sarah Monazam Erfani and Rui Zhang, The University of Melbourne, Australia

Collaborative filtering is one of the most popular techniques in designing recommendation systems, and its most representative model, matrix factorization, has been wildly used by researchers and the industry. However, this model suffers from the lack of interpretability and the item cold-start problem, which limit its reliability and practicability. In this paper, we propose an interpretable recommendation model called Multi-Matrix Factorization (MMF), which addresses these two limitations and achieves the state-of-the-art prediction accuracy by exploiting common attributes that are present in different items. In the model, predicted item ratings are regarded as weighted aggregations of attribute ratings generated by the inner product of the user latent vectors and the attribute latent vectors. MMF provides more fine grained analyses than matrix factorization in the following ways: attribute ratings with weights allow the understanding of how much each attribute contributes to the recommendation and hence provide interpretability; the common attributes can act as a link between existing and new items, which solves the item cold-start problem when no rating exists on an item. We evaluate the interpretability of MMF comprehensively, and conduct extensive experiments on real datasets to show that MMF outperforms state-of-the-art baselines in terms of accuracy.

2:20PM *Collecting Indicators of Compromise from Unstructured Text of Cybersecurity Articles using Neural-Based Sequence Labelling [#19774]*

Long Zi, Tan Lianzhi, Zhou Shengping, He Chaoyang and Liu Xin, Tencent, China

Indicators of Compromise (IOCs) are artifacts observed on a network or in an operating system that can be utilized to indicate a computer intrusion and detect cyber-attacks in an early stage. Thus, they exert an important role in the field of cybersecurity. However, state-of-the-art IOCs detection systems rely heavily on hand-crafted features with expert knowledge of cybersecurity, and require large-scale manually annotated corpora to train an IOC classifier. In this paper, we propose using an end-to-end neural-based sequence labelling model to identify IOCs automatically from cybersecurity articles without expert knowledge of cybersecurity. By using a multi-head selfattention module and contextual features, we find that the proposed model is capable of gathering contextual information from texts of cybersecurity articles and performs better in the task of IOC identification. Experiments show that the proposed model outperforms other sequence labelling models, achieving the average F1-score of 89.0% on English cybersecurity article test set, and approximately the average F1-score of 81.8% on Chinese test set.

2:40PM *LambdaGAN: Generative Adversarial Nets for Recommendation Task with Lambda Strategy [#19697]*

Yang Wang, Hai-tao Zheng, Wang Chen and Rui Zhang, Tsinghua-Southampton Web Science Laboratory Graduate School at Shenzhen, Tsinghua University, China, China; University of Melbourne, Australia

The Top-N recommendation task aims to recommend users the items they like most. Generative Adversarial Net (GAN) has achieved good results on recommendation, which learns user-item matrix by a generative adversarial training process. There are two main scenarios, pointwise scenarios and pairwise scenarios. However, for recommendation, GANs in pairwise scenarios perform not as well as these in pointwise scenarios. As pairwise rank is a position-independent algorithm, it does not consider Top-N ranking sufficiently. Recommendation task is position-dependent. Especially the Top-N item ranking accuracy is much more important than the ranking accuracy of the tail item. In this paper, we propose LambdaGAN for Top-N recommendation. LambdaGAN introduces lambda rank into generative adversarial training process in order to consider the ranking information of the item. The proposed model enables generative adversarial training in pairwise scenarios available for recommendation by optimizing the rank based metrics directly. Moreover, we adjust lambda function according to the characteristics of recommendation. Two new designed lambda functions are proposed. Experimental results show that LambdaGAN outperforms state-of-the-art algorithms including BPR, PRFM, LambdaFM and IRGAN in terms of four standard evaluation metrics on two widely used datasets, Movielens-100K and Netflix.

3:00PM *ST-RNet: A Time-aware Point-of-interest Recommendation Method based on Neural Network [#19945]*

Lu Gao, Yuhua Li, Ruixuan Li, Zhenlong Zhu, Xiwu Gu and Olivier Habimana, Huazhong University of Science and Technology, China; Huazhong University of Science and Technology, Rwanda

Point-of-interest (POI) recommendation is one of the most important services in the rapid growing location-based social networks (LBSNs). Good POI recommendation can help people explore the locations they haven't visited but are interested in, and help merchants find their target users. Time-aware POI recommendation aims to recommend unvisited POIs for a given user at a specified time in a day. However, previous methods, such as user-based collaborative filtering, lack the mining of the features of POIs and the learning of abstract spatio-temporal interactions. In this paper, we propose a novel time-aware POI recommendation method named ST-RNet (Spatio-Temporal Recommender Network) to address these shortages. ST-RNet works in the following fashion. Firstly, we analyze the crucial features in LBSNs to alleviate data sparsity problem and further measure the similarities between POIs. For subsequent network training, we then construct the embedding matrices with same dimension for users and POIs by POI-based Collaborative Filtering (PCF). Furthermore, the positive and negative check-in records are fed into a novel recommender neural network (RNet) to learn the embedding matrix of times and the abstract interactions between users, POIs and times. Finally, ST-RNet recommends the unvisited POIs most likely to be visited to a given user at a given time. The experimental results on Foursquare real-world dataset show that ST-RNet is effective on time-aware POI recommendation task and is capable of analyzing the hidden patterns behind spatio-temporal interactions.

3:20PM *Transfer Learning for Network Classification [#20421]*

Bowen Dong, Charu C Aggarwal and Philip S. Yu, University of Illinois at Chicago, United States; IBM T. J. Watson Research Center, United States

In this paper, we will study the problem of node classification in networks with the use of transfer learning. In many real applications, it may often be difficult to find a sufficient number of labels on the nodes for the classification process. When typical social networks often have a significant amount of content in the form of text, it is much harder to characterize the nodes in terms of pre-defined properties of interest (or class labels). This is because the nodes in the social network are autonomously created by different entities, who may not label them in a way, which is specific or friendly to particular kinds of applications. The lack of availability of labels on the nodes is analogous to the training data paucity which is often encountered in the classification problem. However, there is often copious availability of text collections, which describe a wide variety of subjects, including the desired properties of interest. Since social network nodes are also often associated with some amount of text content, such collections can be used as a transfer bridge in the learning process. In this paper, we will examine the use of transfer learning from these domains in order to perform collective classification. We will present experimental results, which show the effectiveness of our approach.

3:40PM *Personalized Point-of-Interest Recommendation on Ranking with Poisson Factorization [#19113]*

Yijun Su, Xiang Li, Wei Tang, Daren Zha, Ji Xiang and Neng Gao, Institute of Information Engineering, Chinese Academy of Sciences, China

The increasing prevalence of location-based social networks (LBSNs) poses a wonderful opportunity to build personalized point-of-interest (POI) recommendations, which aim at recommending a top-N ranked list of POIs to users according to their preferences. Although previous studies on collaborative filtering are widely applied for POI recommendation, there are two significant challenges have not been solved perfectly. (1) These approaches cannot effectively and efficiently exploit unobserved feedback and are also unable to learn useful information from it. (2) How to seamlessly integrate multiple types of context information into these models is still under exploration. To cope with the aforementioned challenges, we develop a new Personalized pairwise Ranking Framework based on Poisson Factor factorization (PRFPF) that follows the assumption that users' preferences for visited POIs are preferred over potential POIs, unvisited POIs are less preferred than potential POIs. The framework PRFPF is composed of two modules: candidate module and ranking module. Specifically, the candidate module is used to generate a series of potential POIs from unvisited POIs by incorporating multiple types of context information (e.g., social and geographical information). The ranking module learns the ultimate order of users' preference by leveraging the potential POIs. Experimental results evaluated on two large-scale real-world datasets show that our framework outperforms other state-of-the-art approaches in terms of various metrics.

S08: Dynamics, Applications, and Hardware Implementation of Reservoir Computing

Wednesday, July 17, 2:00PM-4:00PM, Room: Duna Salon II, Chair: Yoshihiko Horio

2:00PM *Chaotic Neural Network Reservoir [#19290]*

Yoshihiko Horio, Tohoku University, Japan

Simple structure and robust property of a reservoir neural network are preferable for hardware implementation of a high performance learning system, especially for time-series data processing. One of salient feature of the reservoir network is reproducibility or consistency of its responses to the same or similar inputs. This is usually guaranteed through the echo state property of the network by properly choosing synaptic weights among reservoir neurons. Another important feature is a variety of dynamics in the reservoir, which makes reservoir to process complex time-varying input signals. One way to increase the variety of dynamics is introducing chaotic behavior by destabilize the reservoir network by changing weight values. However, this will violate the echo state property, therefore, chaotic dynamics are usually avoided in the reservoir computing. In this paper, we propose a method to introduce high-dimensional chaotic dynamics into the reservoir network, but keeping its consistency. To achieve this, we use a chaotic neural network model in the reservoir network, while keeping the weight matrix in the reservoir network to satisfy the echo state property criteria. In order to show the consistency of the chaotic neural network reservoir, preliminary results for chaotic time-series predictions through the chaotic neural network reservoir are illustrated. In addition, we discuss the application of the chaotic neural network reservoir to a self-aware hardware system.

2:20PM *Reservoir Computing Based on Dynamics of Pseudo-Billiard System in Hypercube [#20372]*

Yuichi Katori, Hakaru Tamukoh and Takashi Morie, Future University Hakodate, Japan; Kyushu Institute of Technology, Japan

Reservoir computing (RC) is a framework for constructing recurrent neural networks with simple training rule and sparsely and randomly connected nonlinear units. The network (called reservoir) generates complex motion that can be used for many tasks including time series generation and prediction. We construct a reservoir based on the dynamics of the pseudo-billiard system that produce complex motion in a high- dimensional hypercube. In particular, we use the chaotic Boltzmann machine (CBM) whose units exhibit chaotic behavior in the hypercube. The units interact with each other in a time-domain manner through its binary state, and thus an efficient hardware implementation of the system is expected. In order to utilize the CBM as the reservoir, it is necessary to control its chaotic behavior for ensuring the echo state property of RC and establish encoding and decoding for input and output signal. For this purpose, we introduce a reference clock and analyze effects and properties of the reference input. We evaluate the proposed

model on the time series generation tasks and show that the model works properly on a broad range of parameter values. Our approach presents a novel mechanism for time-domain information processing and a fundamental technology for a brain like artificial intelligence system.

2:40PM *A Chaotic Boltzmann Machine Working as a Reservoir and Its Analog VLSI Implementation [#20163]*

Masatoshi Yamaguchi, Yuichi Katori, Daichi Kamimura, Hakaru Tamukoh and Takashi Morie, Kyushu Institute of Technology, Japan; Future University Hakodate, Japan

Reservoir computing attracts much interests because of its high computing ability especially for time-series prediction regardless of its simple structure and learning scheme. This paper proposes a reservoir computing hardware model using a chaotic Boltzmann machine (CBM) as the reservoir part, which can achieve complex motion in a dynamical system on a high-dimensional hypercube. The CBM uses analog nonlinear dynamics instead of stochastic operation as in the original Boltzmann machine model. To utilize CBMs as reservoir, chaotic operation has to be suppressed, and the echo state property should be satisfied. We modify the CBM model for simpler analog CMOS VLSI implementation, and propose to use it as a reservoir by adding an external reference clock signal, and verify its proper operation by numerical simulation. We also refine the CMOS VLSI circuit design from the previous design based on the proposed modified CBM model to improve the power consumption and calculation precision.

3:00PM *Short-term Prediction of Hyper Chaotic Flow Using Echo State Network [#20022]*

Aren Shinozaki, Kota Shiozawa, Kazuki Kajita, Takaya Miyano and Yoshihiko Horio, Ritsumeikan University, Japan; Tohoku University, Japan

We apply an echo state network with a reservoir consisting of 200 tanh neurons to sho to to to s to sho to sho to to sho to to to s to sho to sho to to to s to sho to to to sho to to sho to short-term prediction of a chaotic time series generated by the augmented Lorenz equations as a hyper chaotic flow model. The prediction error is analyzed on the b t th the the b t th the b t the b the b the b the basis of the Kolmogorov-Sinai entropy and the dimension of the chaotic attractor in comparison with those for chaotic flow models each with a single positive Lya Lyapunov exponent. We discuss the predictive performance of the reservoir as a universal simulator of chaotic attractors in terms of Ueda's view of chaos, i.e., random transitions between unstable periodic orbits in a chaotic attractor.

3:20PM *Analysis on Characteristics of Multi-Step Learning Echo State Networks for Nonlinear Time Series Prediction [#19193]*

Takanori Akiyama and Gouhei Tanaka, The University of Tokyo, Japan

Reservoir Computing (RC) is a framework based on recurrent neural networks for high-speed learning and has attracted much attention. RC has been applied to a variety of temporal recognition tasks. Especially, Jaeger showed that the Echo State Network (ESN), which is one of the RC models, was effective for chaotic time series prediction tasks. However, there are two inevitable problems in nonlinear time series prediction using the standard ESN. One is that its prediction ability reaches a saturation point as the reservoir size is increased. The other is that its prediction ability depends heavily on hyperparameter values. In this paper, we propose a multi-step learning ESN to solve these problems. The proposed system has multiple reservoirs and the prediction error of one ESN-based predictor is corrected by another subsequent predictor. We demonstrate the effectiveness of the proposed method in two nonlinear time series prediction tasks. Another experiment using Lyapunov exponents suggests that the performance of the proposed method is robust against changes in hyperparameter values. In addition, we clarify the characteristic of the proposed method with regard to nonlinearity and memory using simple function approximation tasks.

3:40PM *Quantitative Analysis of Dynamical Complexity in Cultured Neuronal Network Models for Reservoir Computing Applications [#20275]*

Satoshi Moriya, Hideaki Yamamoto, Ayumi Hirano-Iwata, Shigeru Kubota and Shigeo Sato, Tohoku University, Japan; Yamagata University, Japan

Reservoir computing is a machine learning paradigm that was proposed as a model of cortical information processing in the brain. It processes information using the spatiotemporal dynamics of a large-scale recurrent neural network and is expected to improve power efficiency and speed in neuromorphic computing systems. Previous theoretical investigation has shown that brain networks exhibit an intermediate state of full coherence and random firing, which is suitable for reservoir computing. However, how reservoir performance is influenced by connectivity, especially which revealed in recent connectomics analysis of brain networks, remains unclear. Here, we constructed modular networks of integrate-and-fire neurons and investigated the effect of modular structure and excitatory-inhibitory neuron ratio on network dynamics. The dynamics were evaluated based on the following three measures: synchronous bursting frequency, mean correlation, and functional complexity. We found that in a purely excitatory network, the complexity was independent of the modularity of the network. On the other hand, networks with inhibitory neurons exhibited complex network activity when the modularity was high. Our findings reveal a fundamental aspect of reservoir performance in brain networks, contributing to the design of bio-inspired reservoir computing systems.

8: Other Applications

Wednesday, July 17, 2:00PM-4:00PM, Room: Duna Salon III, Chair: Hui Liu

2:00PM *Ensemble Application of Transfer Learning and Sample Weighting for Stock Market Prediction [#19019]*

Simone Merello, Andrea Picasso Ratto, Luca Oneto and Erik Cambria, University of Genova, Italy; Nanyang Technological University, Singapore

Forecasting stock market behavior is an interesting and challenging problem. Regression of prices and classification of daily returns have been widely studied with the main goal of supplying forecasts useful in real trading scenarios. Unfortunately, the outcomes are not directly related with the maximization of the financial gain. Firstly, the optimal strategy requires to invest on the most performing asset every period and trading accordingly is not trivial given the predictions. Secondly, price fluctuations of different magnitude are often treated as equals even if during market trading losses or gains of different intensities are derived. In this paper, the problem of stock market forecasting is formulated as regression of market returns. This approach is able to estimate the amount of price change and thus the most performing assets. Price fluctuations of different magnitude are treated differently through the application of different weights on samples and the scarcity of data is addressed using transfer learning. Results on a real simulation of trading show how, given a finite amount of capital, the predictions can be used to invest in high performing stocks and, hence, achieve higher profits with less trades.

2:20PM *Stealing Knowledge from Protected Deep Neural Networks Using Composite Unlabeled Data [#20502]*

Itay Mosafi, Eli David and Nathan Netanyahu, Bar-Ilan University, Israel

As state-of-the-art deep neural networks are deployed at the core of more advanced AI-based products and services, the incentive for copying them (i.e., their intellectual properties) by rival adversaries is expected to increase considerably over time. The best way to extract and steal the knowledge of such networks is by querying them using a large dataset of random samples and recording the output of the network, followed by training a ``student'' network to mimic these outputs, while assuming nothing about the original network. The most effective way to protect against such a mimicking attack is to provide only the classification result, without confidence values associated with the softmax layer. In this paper, we present a novel method for generating composite images for attacking ``mentor'' neural nets using a ``student'' model. Our method assumes no information regarding the mentor's training dataset, its architecture, or weights. Furthermore, assuming no information regarding the mentor's softmax output values, we manage to successfully mimic its neural network and steal all of its knowledge. We further demonstrate that our student network (which copies the mentor) is impervious to watermarking protection methods, and thus would not be detected as a stolen model. Our results imply that all current neural networks are vulnerable, essentially, to mimicking attacks, even if they do not divulge anything but the most basic necessary output, and that the student model which mimics them cannot be easily detected and singled out as a stolen copy using currently available techniques.

2:40PM *Intranet User-Level Security Traffic Management with Deep Reinforcement Learning [#19787]*

Qiuqing Jin and Liming Wang, Institute of Information Engineering, University of Chinese Academy of Sciences, China; Institute of Information Engineering, Chinese Academy of Sciences, China

Insider threats gradually exert great influence in cybersecurity, causing a significant loss to organizations or companies. However, whatever the form of the threat is, insiders have to conduct the unauthorized activities through the communication traffic, such as controlling the victim systems and unauthorizedly requesting the resources. Moreover, as one of the most fundamental intranet resources, bandwidth is frequently targeted by insider attackers for sabotage to traffic communication and service delivery of the network. In this paper, we present a user-level full-lifecycle security management scheme for intranet traffic from anomaly detection to mitigation execution in an online manner. This scheme dynamically monitors abnormal users that deviate from normal behavior patterns through bidirectional Gated Recurrent Unit (bi-GRU) based online unsupervised log parser, then adaptively adjusts the traffic scheduling policy according to the adequate consideration of network security, network performance and user requirements by using deep Reinforcement Learning (RL) method for online decision-making. Extensive experimental evaluations show that our scheme can stably maintain the high performance of the traffic scheduling and effectively mitigate multifarious traffic threats. Our work is a valuable step towards designing self-adaptive intranets that learn to enhance the security management by themselves with high scalability and deployability.

3:00PM *Robust Neuro-adaptive Asymptotic Consensus for a Class of Uncertain Multi-agent systems: An Edge-based Paradigm [#19047]*

Dongdong Yue, Qi Li, Jinde Cao and Xuegang Tan, Southeast University, China

In this brief, we investigate the adaptive consensus problem for a class of matching uncertain multi-agent systems in the presence of unmodelled nonlinearities and unknown disturbances. A robust dynamic controller comprising a linear feedback term, a discontinuous feedback term and a neural network approximation term is designed, where edge-based adaptive coupling strengths are introduced for both feedback terms while adaptive weights are designed for the neural network term. Unlike most existing neural-network-based control algorithms, where the consensus errors are ultimately uniformly bounded, the asymptotic consensus of the uncertain MAS is theoretically proved based on graph theory, nonsmooth analysis and Lyapunov theoretic approach. A noticeable character of the proposed method is that no global information of the underlying network is needed, such as the eigenvalues of the Laplacian matrix. In the end, a numerical example is presented to testify the theoretical results.

3:20PM *Collaboration of Machines and Robots in Cyber Physical Systems based on Evolutionary Computation Approach [#20006]*

Fu-Shiung Hsieh, Chaoyang University of Technology, Taiwan

Due to the capability to effectively deal with changing demand and uncertainties, Cyber-Physical Systems (CPS) emerge as a paradigm for manufacturers to face challenges in the real business environment based on dynamic configuration of resources such as machines and robots. CPS relies on the development of an effective mechanism to make resources collaborate to meet the requirements of production processes. How to develop a mechanism for dynamic configuration of resources in CPS is an important issue. Motivated by this need, this paper aims to propose a methodology to support dynamic teaming between resources such as machines and robots in CPS. In this paper, we propose a method to achieve dynamic teaming of resources in CPS based on evolutionary computation approach and multi-agent system (MAS) architecture. To facilitate specification of production process and capabilities of resources, a modeling tool based on Petri nets is adopted. A problem is formulated based on Petri net models to meet the requirements of production processes. Two evolutionary computation algorithms, one based on Particle Swarm Optimization (PSO) approach and the other based on discrete Differential Evolution (DE) approach, have been developed to solve the problem. We illustrate and compare effectiveness of the proposed algorithms by examples.

3:40PM *A Novel Deep Learning-Based Encoder-Decoder Model for Remaining Useful Life Prediction [#19657]*

Hui Liu, Zhenyu Liu, Weiqiang Jia and Xianke Lin, State Key Laboratory of CAD&CG, Zhejiang University, China; Department of Mechanical Engineering, University of Ontario Institute of Technology, Canada

A novel encoder-decoder model based on deep neural networks is proposed for the prediction of remaining useful life (RUL) in this work. The proposed model consists of an encoder and a decoder. In the encoder, the Bi-directional Long Short-Term Memory Networks (Bi-LSTM) and Convolutional Neural Networks (CNN) are used to capture the long-term temporal dependencies and important local features from the sequential data, respectively. Besides, single 1*1 convolution filter in the last convolutional layer is used for dimensionality reduction. In the decoder, the fully connected networks are employed to decode the feature information to predict RUL. In addition, the proposed data-driven method can achieve end-to-end prediction, which does not need feature engineering. To evaluate the proposed model, experimental verification is carried out on a commonly used aero-engine C-MAPSS dataset. Compared with other state-of-the-art approaches on the same dataset, the effectiveness and superiority of the proposed framework are demonstrated. For example, the scoring function value of the second subset is reduced by up to 64.99% compared with the best existing result.

8a: Applications of deep networks

Wednesday, July 17, 2:00PM-4:00PM, Room: Panorama I, Chair: Donald Wunsch

2:00PM *Transfer Learning Using Ensemble Neural Networks for Organic Solar Cell Screening [#20448]*

Arindam Paul, Dipendra Jha, Reda Al-Bahrani, Wei-keng Liao, Alok Choudhary and Ankit Agrawal, Northwestern University, United States

Organic Solar Cells are a promising technology for solving the clean energy crisis in the world. However, generating candidate chemical compounds for solar cells is a time- consuming process requiring thousands of hours of laboratory analysis. For a solar cell, the most important property is the power conversion efficiency which is dependent on the highest occupied molecular orbitals (HOMO) values of the donor molecules. Recently, machine learning techniques have proved to be very useful in building predictive models for HOMO values of donor structures of Organic Photovoltaic Cells (OPVs). Since experimental datasets are limited in size, current machine learning models are trained on data derived from calculations based on density functional theory (DFT). Molecular line notations such as SMILES or InChI are popular input representations for describing the molecular structure of donor molecules. The two types of line representations encode different information, such as SMILES defines the bond types while InChi defines protonation. In this work, we present an ensemble deep neural network architecture, called SINet, which harnesses both the SMILES and InChI molecular representations to predict HOMO values and leverage the potential of transfer learning from a sizeable DFT-computed dataset- Harvard CEP to build more robust predictive models for relatively smaller HOPV datasets. Harvard CEP dataset contains molecular structures and properties for 2.3 million candidate donor structures for OPV while HOPV contains DFT-computed and experimental values of 350 and 243 molecules respectively. Our results demonstrate significant performance improvement from the use of transfer learning and leveraging both molecular representations.

2:20PM *MetODeep: A Deep Learning Approach for Prediction of Methionine Oxidation Sites in Proteins [#19899]*

Guillermo Lopez-Garcia, Jose M. Jerez, Daniel Urda and Francisco J. Veredas, Universidad de Malaga, Spain; Universidad de Cadiz, Spain

After being synthesized by ribosomes in the cells, proteins can suffer from post-translational modifications (PTM) that affect their functionality. One of the most studied PTMs is phosphorylation. Mass-spectrometry methods aimed at identifying phosphorylation sites in proteins are arduous and expensive. For these reasons, numerous studies propose the use of machine leaning techniques to predict this PTM. Like phosphorylation, methionine oxidation is another important PTM. Recently, we have proposed a machine learning approach that extracts a set of features from the primary and tertiary structure of the proteins to predict methionine oxidation sites. However, this work had an important limitation that impairs feature extraction, since the 3D structure of many proteins is not fully resolved. In this study, we present MetODeep, a deep learning approach to predict methionine oxidation. Unlike phosphorylation, for which datasets with several hundred thousands samples are available to train effective predictive models, methionine oxidation counts on small datasets, which could lead a deep neural network to experiment over-fitting issues. The recently evidenced existence of a cross-talk between phosphorylation and methionine oxidation, has motivated our transfer-learning approach. Thus, on the basis of a deep convolutional neural network (CNN) pre-trained with phosphorylation data, MetODeep is fine-tuned to predict methionine oxidation. The resulting CNN architecture allows us to omit manual feature extraction, since it accepts raw protein sequences as input data. The final model gives performance results (AUC 0.8267 (+-) 0.0174) that surpass state-of-art of computational models for the prediction of methionine oxidation.

2:40PM *Fully Automatic Dual-Guidewire Segmentation for Coronary Bifurcation Lesion [#19577]*

Yanjie Zhou, Xiaoliang Xie, Guibin Bian, Zengguang Hou, Yudong Wu, Shiqi Liu, Xiaohu Zhou and Jiaxing Wang, Institute of Automation, Chinese Academy of Sciences, China

Interventional therapy for coronary bifurcation lesion has always been an intractable problem in percutaneous coronary intervention (PCI). Dual-guidewire detection can greatly assist physicians in interventional therapy of bifurcated lesions. Nevertheless, this task often comes with the challenges of X-ray images with low signal noise ratio (SNR) as well as the thinner structure of the guidewire compared to other interventional tools. In this paper, a fully automatic detection method based on an improved U-Net and the modified focal loss is proposed for dual- guidewire segmentation in 2D X-ray fluoroscopy, which accomplishes accurate and robust segmentation. The main contributions of this paper are twofold: (1) the proposed method not only addresses the extreme foreground-background class imbalance generated by the slender guidewire structure, but also solve the problem of misclassified examples caused by the guidewire-like structures and contrast agents; (2) the running speed is about 8 frames per second, which reaches near-real-time processing speed. Furthermore, data augmentation algorithm and transfer learning are used to further improve the performance. The proposed method was verified on clinical 2D X-ray image sequences of 30 patients, in which F1-score reached 0.932. The experiment results indicated that our approach is promising for assisting bifurcation lesion surgery.

3:00PM *Spinal Stenosis Detection in MRI using Modular Coordinate Convolutional Attention Networks [#20024]*

Uddeshya Upadhyay, Badrinath Singhal and Meenakshi Singh, Indian Institute of Technology Bombay, India; Synapsica Technologies, India

Spinal stenosis is a condition in which a portion of spinal canal narrows and exerts pressure on nerves that travel through it causing pain and numbness that might require surgery. This narrowing can be caused by pathologies in bony structures (vertebrae) or soft tissue structures (intervertebral discs) that comprise the spine. Radiography, particularly Magnetic Resonance Imaging (MRI) is the modality of choice to evaluate stenosis and intervertebral disc pathology. Radiologists examine axial MRI scans at various levels along the spine to detect stenosis. Further, they evaluate the diameters of spinal canal and bulging in nearby discs which can indicate narrowing and compression on nerves. Hence measuring various diameters in a scan is a crucial step in diagnosis. However, affected regions occupy a very small fraction of the scan and there is virtually no room for error as a deviation of few pixels will also lead to discrepancies in measured and original lengths which makes it a very difficult and laborious task to measure the length of such intricate structures accurately. This paper proposes a novel deep learning based solution to tackle this problem. Proposed method attempts to solve it in two independent modules and makes prediction on the enlarged section of the scan which also makes it easier to measure various lengths. Human radiologists focus on certain parts of the scan rather than attending to the entire scan which largely consists of irrelevant background. Proposed modular approach is designed to mimic this attention mechanism. Both modules are built using coordinate convolutional networks, comparisons with baseline method empirically demonstrate superiority of the proposed approach.

3:20PM *JSAC: A Novel Framework to Detect Malicious JavaScript via CNNs over AST and CFG [#20132]*

Hongliang Liang, Yuxing Yang, Lu Sun and Lin Jiang, Beijing University of Posts and Telecommunications, China

JavaScript (JS) is a dominant programming language in web/mobile development, while it is also notoriously abused by attackers due to its powerful characteristics, e.g., dynamic, prototype-based and multi-paradigm, which foil most static and dynamic analysis approaches. To detect malicious JS instances, several machine learning-based methods have been developed recently. However, these methods took JS as a natural language instead of a programming one, which can not capture its syntactic and semantic features. In this paper, we present JSAC, a novel framework to detect JS malware. It combines deep learning and program analysis techniques to capture the syntactic and semantic features of JS programs. Specifically, to get a JS program's syntactic information, we build its abstract syntax tree and employ a tree-based convolutional neural network (CNN) to extract features from it. To get its semantic information, we construct its control flow graph and feed it to another graph-based CNN. Last, the features extracted from two CNNs are fused for final detection. Evaluation on a corpus of 69,523 JS files indicates that JSAC outperforms 4 other models with 98.73% F1-score in detecting JS malware.

3:40PM *Anomaly Detection for Visual Quality Control of 3D-Printed Products [#19806]*

Loek Tonnaer, Jiapeng Li, Vladimir Osin, Mike Holenderski and Vlado Menkovski, Eindhoven University of Technology, Netherlands; Signify, Netherlands

We present a method for detection of surface defects in images of 3D-printed products that enables automated visual quality control. The data characterising this problem is typically high-dimensional (high-resolution images), imbalanced (defects are relatively rare), and has few labelled examples. We approach these challenges by formulating the problem as probabilistic anomaly detection, where we use Variational Autoencoders (VAE) to estimate the probability density of non-faulty products. We train the VAE in an unsupervised manner on images of non-faulty products only. A successful model will then assign high likelihood to unseen images of non-faulty products, and lower likelihood to images displaying defects. We test this method on anomaly detection scenarios using the MNIST dataset, as well as on images of 3D-printed products. The demonstrated performance is related to the capability of the model to closely estimate the density distribution of the non-faulty (expected) data. For both datasets we present empirical results that the likelihood estimated with a convolutional VAE can separate the normal and anomalous data. Moreover we show how the reconstruction capabilities of VAEs are highly informative for human observers towards localising potential anomalies, which can aid the quality control process.

Machine Learning and Deep Learning

Wednesday, July 17, 2:00PM-4:00PM, Room: Panorama II, Chair: Spiros Georgakopoulos

2:00PM *Deep Rule-Based Aerial Scene Classifier using High-Level Ensemble Feature Descriptor [#19323]*

Xiaowei Gu and Plamen Angelov, Lancaster University, United Kingdom

In this paper, a new deep rule-based approach using high-level ensemble feature descriptor is proposed for aerial scene classification. By creating an ensemble of three pre-trained deep convolutional neural networks for feature extraction, the proposed approach is able to extract more discriminative representations from the local regions of aerial images. With a set of massively parallel IF...THEN rules built upon the prototypes identified through a self-organizing, nonparametric, transparent and highly human-interpretable learning process, the proposed approach is able to produce the state-of-the-art classification results on the unlabeled images outperforming the alternatives. Numerical examples on benchmark datasets demonstrate the strong performance of the proposed approach.

2:20PM *Tweet Act Classification : A Deep Learning based Classifier for Recognizing Speech Acts in Twitter [#20034]*

Tulika Saha, Sriparna Saha and Pushpak Bhattacharyya, IIT Patna, India

Speech act provides valuable insight for understanding the communicative intention and behaviour of a user utterance. This applies for conversation on any platform which includes social media network such as Twitter. This paper presents a deep learning based tweet act classifier (speech acts for Twitter) for understanding the content and intention of tweets and for discovering the valuable communications amongst the tweeters. A generic taxonomy of seven speech acts has been created for the task. A Convolution Neural Network (CNN) based Deep Learning (DL) architecture is integrated with different classifiers at the final layer such as softmax or linear Support Vector Machine (SVM) to develop a tweet act classifier. To make the classifier even more robust, various hand-crafted features are added to the system to boost its effectiveness towards improving the performance of the system. Experimental results indicate that the proposed model achieved good results with an average F1-score of 0.71 and outperformed state of the art approaches.

2:40PM *Chinese Clinical Named Entity Recognition with Word-Level Information Incorporating Dictionaries [#19808]*

Ningjie Lu, Jun Zheng, Wen Wu, Yan Yang, Kaiwei Chen and Wenxin Hu, East China Normal University, China; Shanghai Qiniu Information Technologies Co.,Ltd., China

Electronic Medical Records (EMRs) are the digital equivalent of paper records, which include treatment and medical history about a patient. At present, the main research goal of Chinese EMRS is to accurately recognize the body parts, drugs, illnesses and other information in the Chinese medical process. Implementing EMRs can boost both the quality and safety of patient care. In Chinese EMRs, how to accurately recognize named entities is important because it will be useful to predict the disease risk, therapeutic method and recovery probability. This paper proposes a novel deep learning framework, which uses character-word joint embedding and combines different feature information based on the dictionary. Compared with the predecessors, we incorporate word-level information based on the basic Bi-LSTM model. In addition, we propose an improved n-gram feature encoding method and compare it with PDET feature and PIET feature. Our experimental results demonstrate that our proposed model performs the best in predicting named entities in Chinese EMRs.

3:00PM *Multi-perspective Feature Generation Based on Attention Mechanism [#20470]*

Ma Longxuan and Zhang Lei, Beijing University of Posts and Telecommunications, China

Selecting features that represent a particular corpus is important to the success of many natural language processing tasks. However, the previous attention-based work only focused on feature augmentation in the lexical level, lacking the exploration of feature enhancement in the sentence level. In this paper, we exploit a novel feature generation method for information retrieval, denoted by Multi-perspective Feature Generation Network (MFGN).

We first add features at sentence level from multiple perspectives to enhance sentence representation, then the word level features make second interactive confirmation from two perspectives. After two-level processing for different granularity characteristics, the model complete the retrieval process. It is more in line with the cognitive process of human beings to grasp the overall meaning of sentences and then carefully study the potential meaning of specific words. In the process of feature generation and extraction, text interacts from multiple perspectives, takes into account the information of text itself and the information of text interaction in different levels and multiple granularities. Different kinds of characteristics are distilled according to specific tasks, enhancing the practicability and robustness of the model. MFGN relies solely on the text itself, requires no external knowledge or feature engineering. Experimental results show that our approach outperforms previous work on multiple well-known Retrieval-based question answering datasets. Through the analysis of the experiments, we prove that MFGN provides excellent retrieval and interpretative abilities.

3:20PM *Efficient Learning Rate Adaptation for Convolutional Neural Network Training [#20256]*
Spiros Georgakopoulos and Vassilis Plagianakos, Department of Computer Science, University of Thessaly, Greece, Greece

Convolutional Neural Networks (CNNs) have been established as substantial supervised methods for classification problems in many research fields. However, a large number of parameters have to be tuned to achieve high performance and good classification results. One of the most crucial parameter for the performance of a CNN is the learning rate (step) of the training algorithm. Although the heuristic search to tune the learning rate is a common practice, it is extremely time-consuming, considering the fact that CNNs require a significant amount of time for each training, due to their complex architectures and high number of weights. Approaches that integrate the adaptation of the initial learning rate in the optimization algorithm, manage to converge to high quality solutions and have been embraced by the research community. In this work, we propose an improvement of the recently proposed Adaptive Learning Rate algorithm (AdLR). The proposed learning rate adaptation algorithm (e-AdLR) exhibits excellent convergence properties and classification accuracy, while at the same time is fast and robust.

3:40PM *Fast segmentation for large and sparsely labeled coral images [#19934]*
Xi Yu, Ying Ma, Stephanie Farrington, John Reed, Bing Ouyang and Jose C Principe, University of Florida, United States; Florida Atlantic University, United States

Marine organism datasets often present sparse annotated labels and with many objects in cluttered background. Therefore, there are two challenges to do image segmentation on these sparsely labeled datasets: one is to obtain denser labeled training data and the other is to improve the speed of testing on large images. In this paper, we propose a label augmentation method to generate more labels for training based on the superpixel algorithm, and we also create coarse-to-fine approach to detect the coral areas quickly in the large images. Our experiments run on coral image dataset collected in Pulley Ridge, proving that this label augmentation and coarse-to-fine approach allows us to speed up the process of quantifying the percent of corals in large images while preserving accuracy.

2i: Support vector machines and kernel methods, 2: ML
Wednesday, July 17, 2:00PM-4:00PM, Room: Panorama III, Chair: Shigeo Abe

2:00PM *Flexible Kernel Selection in Multitask Support Vector Regression [#20185]*
Carlos Ruiz, Carlos Alaiz, Alejandro Catalina and Jose R. Dorronsoro, Autonomous University of Madrid, Spain

Multitask Learning (MTL) aims to solve several related problems at the same time exploiting the similarities between them. In particular, Support Vector Machines (SVMs) can be used for MTL, providing a model that is potentially more flexible than a classical SVM trained over the data of all the tasks, and which can use more information than a bunch of independent models trained over each one of the tasks. Nevertheless, a major drawback of these SVMs is the large number of hyperparameters if no reductionists assumptions are made, which prevents from using standard selection methods as grid or random searches. In particular, both the common kernel and the kernels associated to each one of the tasks have to be selected. In this paper, we propose an approach to set these kernels based on Gaussian Processes (GPs), whose Bayesian perspective allows to deal with several parameters naturally. In particular, a GP is trained for each task, and the resultant kernel parameters are transferred to the SVM-based MTL model. Several experiment in real-world datasets show empirically the usefulness of these approach and the advantages of the GP-based kernel selection method.

2:20PM *Analyzing Minimal Complexity Machines [#19083]*
Shigeo Abe, Kobe University, Japan

The minimal complexity machine (MCM) minimizes the maximum distance between training data and the separating hyperplane and is shown to generalize better than the conventional support vector machine. In this paper, we analyze the MCM and clarify the conditions that the solution of MCM is nonunique and unbounded. To resolve the unboundedness, we propose the minimal complexity linear programming support vector machine (MLP SVM), in which the minimization of the maximum distance between training data and the separating hyperplane is added to the linear programming support vector machine (LP SVM). By computer experiments we show that the solution of the MCM is unbounded under some conditions and that the MLP SVM generalizes better than the LP SVM for most of the two-class and multiclass problems.

2:40PM *A Multiple Kernel Machine with In-Situ Learning using Sparse Representation [#19855]*
Ali Pezeshki, Mahmood Azimi-Sadjadi and Christopher Robbiano, Colorado State University, United States

In this paper, a new multiple kernel machine with in-situ learning capability is introduced for pattern classification using sparse representation framework. A vector-valued score function for a given pattern is generated as a linear combination of multiple kernel functions, each of which measures the similarity between the pattern to be classified and the representative samples from every class. The representative samples for different classes are selected through the sparse approximations, by forcing the representation of the vector-valued score function to be sparse in a dictionary of kernel functions with respect to all data samples from different classes. After the sparse weight matrix is found, the vector- valued score function for a new pattern with unknown class label can be computed. An in-situ learning process is developed which allows for incremental updating of the system without impacting the previously learnt patterns. The proposed multiple kernel machine was then implemented and benchmarked against other multi-kernel methods for classifying underwater objects using low frequency sonar data.

3:00PM *Mixed Variational Inference [#19769]*
Nikolaos Gianniotis, Heidelberg Institute for Theoretical Studies gGmbH, Germany

The Laplace approximation has been one of the workhorses of Bayesian inference. It often delivers good approximations in practice despite the fact that it does not strictly take into account where the volume of posterior density lies. Variational approaches avoid this issue by explicitly minimising the Kullback-Leibler divergence DKL between a postulated posterior and the true (unnormalised) logarithmic posterior. However, they rely on a closed form DKL in order to update the variational parameters. To address this,

stochastic versions of variational inference have been devised that approximate the intractable DKL with a Monte Carlo average. This approximation allows calculating gradients with respect to the variational parameters. However, variational methods often postulate a factorised Gaussian approximating posterior. In doing so, they sacrifice a- posteriori correlations. In this work, we propose a method that combines the Laplace approximation with the variational approach. The advantages are that we maintain: applicability on non-conjugate models, posterior correlations and a reduced number of free variational parameters. Numerical experiments demonstrate improvement over the Laplace approximation and variational inference with factorised Gaussian posteriors.

3:20PM *An Approach to Cross-Lingual Voice Conversion [#19463]*

Sai Sirisha Rallabandi and Suryakanth V Gangashetty, IIIT-Hyderabad, India

The most prevalent multilingual Text-to-Speech (TTS) synthesis systems encounter an unnatural speaker shift at the language boundaries. This is observed when they are employed for code-mixed TTS synthesis. For the very fact that the collection of polyglot speech is non-trivial, many alternative approaches have been in focus. Cross-Lingual Voice Conversion (CLVC) has been one of those to generate speech with desired speaker and language identities. Our aim in this paper is to design a light-weighted CLVC framework between a pair of Mandarin-English speakers. CLVC is challenging when compared to traditional Voice Conversion (VC) because of its nature of accommodating unaligned corpus from the source and target speakers. We thus focus on generating a parallel corpus for CLVC and bridging the gap between speakers and languages. We perform a text-independent voice conversion with a three-layered conventional Neural Network (NN) for this purpose. The main contributions include i) Source similarity in both training and conversion stages of CLVC, ii) generation of a parallel corpus and iii) text independent and transcription free CLVC. We exploit two variants of a Neural Network in the proposed framework, i) an autoencoder to enable the source similarity and generation of parallel corpus, ii) a traditional DNN for feature mapping between the source and target. The subjective and objective evaluations show that the proposed method is indeed capable of performing a CLVC with an auto-encoded speech.

3:40PM *Twitter breaking news detector in the 2018 Brazilian presidential election using word embeddings and convolutional neural networks [#20189]*

Kenzo Sakiyama, Andre Bezerra Silva and Edson Takashi Matsubara, Federal University of Mato Grosso do Sul, Brazil

A breaking news event detector based on the time series of the number of positive, negative and neutral tweets obtained from a sentiment analysis classifier is proposed. The detector collects real-time tweets related to candidates and transforms them into word embeddings using the FastText algorithm. Using domain adaptation, the sentiment analysis classifier is trained based on a convolutional neural network (CNN) known as TextCNN. The number of positive, negative and neutral tweets in a time frame results in a time-series, which is monitored by an unsupervised time-series anomaly detector. The results show that the sentiment analysis classifier achieves an accuracy of 74% for the three classes, and the detector successfully detects significant breaking news in the 2018 Brazilian presidential election.

Neural Models of Perception, Cognition and Action

Wednesday, July 17, 2:00PM-4:00PM, Room: Panorama IV, Chair: Shengping Zhou

2:00PM *A Computational Model for a Multi-Goal Spatial Navigation Task inspired in Rodent Studies [#19917]*

Martin Llofriu, Pablo Scleidorovich, Gonzalo Tejera, Marco Contreras, Tatiana Pelc, Jean-Marc Fellous and Alfredo Weitzenfeld, University of South Florida, United States; Universidad de la Republica, Uruguay; Universidad Mayor, Chile; University of Arizona, United States

We present a biologically-inspired computational model of the rodent hippocampus based on recent studies of the hippocampus showing that its longitudinal axis is involved in complex spatial navigation. While both poles of the hippocampus, i.e. septal (dorsal) and temporal (ventral), encode spatial information; the septal area has traditionally been attributed more to navigation and action selection; whereas the temporal pole has been more involved with learning and motivation. In this work we hypothesize that the septal-temporal organization of the hippocampus axis also provides a multi-scale spatial representation that may be exploited during complex rodent navigation. To test this hypothesis, we developed a multi-scale model of the hippocampus evaluated it with a simulated rat on a multi-goal task, initially in a simplified environment, and then on a more complex environment where multiple obstacles are introduced. In addition to the hippocampus providing a spatial representation of the environment, the model includes an actor-critic framework for the motivated learning of the different tasks

2:20PM *Understanding Language Dependency on Emotional Speech using Siamese Network [#20290]*

Swaraj Kumar, Sandipan Dutta and Shaurya Chaturvedi, Netaji Subhas University of Technology, India

Emotion is made up of three components; physiological arousal, expressive behaviors, and conscious experience. Psychological theories like Social Constructionist theory believe that although emotions are a universal phenomenon, they are deeply influenced by the cultural background of the speaker. Since culture and language have an intimate connection between them, the language used by a particular social group would have a significant effect on emotions expressed in their verbal communication. We aim to study the susceptibility of emotions in different languages, which is expressed in speech. We introduce a novel Deep Convolution Siamese Network (DCSN) for determining the similarity between speech samples. The speech samples were obtained from emotional speech datasets spoken in different languages, namely English, Italian, and German. Two experiments were undertaken for studying the correlation between language and emotional speech. Similarity metric along with scatter plots were obtained, which establish that the DCSN is capable of identifying both languages and emotions in a speech sample. The results hence provide a conclusive proof for the model to be used for the study of different psychological aspects of emotional speech.

2:40PM *Condensed Convolution Neural Network by Attention over Self-attention for Stance Detection in Twitter [#19626]*

Shengping Zhou, Junjie Lin, Lianzhi Tan and Xin Liu, Tencent Technology Co., Ltd., China

In the era of Web 2.0, people have become accustomed to expressing their attitudes and exchanging opinions on social media sites such as Twitter. It is critical for security and business related applications to make sense of public opinions implied in users' texts. Stance detection aims to classify the stances users hold towards certain targets as FAVOR, AGAINST or NONE. In the literature, many efforts have been paid on neural network based stance detection to avoid hand-crafted features. As a widely used neural network structure, convolutional neural network (CNN) can mine and combine various local textual features for classifying stances with high training efficiency. However, global textual information is usually neglected in the convolution process. Besides, stance clues are often mixed with less informative words in noisy tweets, and it is hard for CNN to resolve and leverage long-distance semantic dependencies between words effectively. To address these issues, in this paper, we propose CCNN-ASA, the Condensed CNN by Attention over Self-Attention, to detect the stances of tweets. We first introduce self-

attention into CNN to adaptively enhance word embeddings with global textual information. To make stance clues close to each other and thus more salient, we further introduce attention- based condensation module which identifies stance-indicative words to condense tweets. Experiments on a benchmark dataset show that CCNN-ASA outperforms state-of- the-art methods in stance detection of tweets.

3:00PM *ChartNet: Visual Reasoning over Statistical Charts using MAC-Networks [#20046]*

Monika Sharma, Shikha Gupta, Arindam Chowdhury and Lovekesh Vig, TCS Research Delhi, India; Indian Institute of Technology, Mandi, India

Despite the improvements in perception accuracies brought about via deep learning, developing systems combining accurate visual perception with the ability to reason over the visual percepts remains extremely challenging. A particular application area of interest from an accessibility perspective is that of reasoning over statistical charts such as bar and pie charts. To this end, we formulate the problem of reasoning over statistical charts as a classification task using MAC-Networks to give answers from a predefined vocabulary of generic answers. Additionally, we enhance the capabilities of MAC-Networks to give chart-specific answers to open-ended questions by replacing the classification layer by a regression layer to localize the textual answers present over the images. We call our network ChartNet, and demonstrate its efficacy on predicting both in vocabulary and out of vocabulary answers. To test our methods, we generated our own dataset of statistical chart images and corresponding question answer pairs. Results show that ChartNet consistently outperform other state-of-the-art methods on reasoning over these questions and may be a viable candidate for applications containing images of statistical charts.

3:20PM *Executing Declarative Parallel Representations of Sequences with Temporal Pooling [#20423]*

Daniel Slack, Alistair Knott and Brendan McCane, Otago University, New Zealand

A Self-Organising Temporal Pooling (SOTP) network has been shown to be capable of forming declarative parallel representations of sequential events and chunking these events without supervision. However, such a network currently cannot take these declarative representations and execute the associated sequence; it is strictly a one-way sequence chunker and encoder. We present a modified model that can convert the lateral weights of the trained network into a primacy gradient usable by methods such as Competitive Queueing (CQ), thus granting the network a method for executing learned sequences. The resulting model has several benefits over traditional CQ. We further present an advanced method of executing sequences via SOTP itself, resulting in less error than CQ whilst being more flexible in replaying sequences from datasets with variable sequence lengths.

3:40PM *A Time-Frequency based Machine Learning System for Brain States Classification via EEG Signal Processing [#20207]*

Cosimo Ieracitano, Nadia Mammone, Alessia Bramanti, Silvia Marino, Amir Hussain and Francesco Carlo Morabito, University Mediterranea of Reggio Calabria, Italy; IRCCS Centro Neurolesi Bonino-Pulejo, Messina, Italy; National Research Council (CNR), Italy; Edinburgh Napier University, United Kingdom

In the last decades, the use of Machine Learning (ML) algorithms have been widely employed to aid clinicians in the difficult diagnosis of neurological disorders, such as Alzheimer's disease (AD). In this context, here, a data-driven ML system for classifying Electroencephalographic (EEG) segments (i.e. epochs) of patients affected by AD, Mild Cognitive Impairment (MCI) and Healthy Control (HC) individuals, is introduced. Specifically, the proposed ML system consists of evaluating the average Time-Frequency Map (aTFM) related to a 19-channels EEG epoch and extracting some statistical coefficients (i.e. mean, standard deviation, skewness, kurtosis and entropy) from the main five conventional EEG sub-bands (or EEG-rhythms: delta, theta, alpha1, alpha2, beta). Afterwards, the time-frequency features vector is fed into an Autoeconder (AE), a Multi-Layer Perceptron (MLP), a Logistic Regression (LR) and a Support Vector Machine (SVM) based classifier to perform the 2-ways EEG epoch- classification tasks: AD vs HC and AD vs MCI. The performances of the proposed approach have been evaluated on a dataset of 189 EEG signals (63 AD, 63 MCI and 63 HC), recorded during an eye-closed resting condition at IRCCS Centro Neurolesi Bonino Pulejo of Messina (Italy). Experimental results reported that the 1-hidden layer MLP (MLP1) outperformed all the other developed learning systems as well as recently proposed state-of-the-art methods, achieving accuracy rate up to 95.76% +\- 0.0045 and 86.84% +\- 0.0098 in AD vs HC and AD vs MCI classification, respectively.

Panel Session: NSF Career Award Winners in Intelligent and Adaptive Systems

Wednesday, July 17, 2:00PM-4:00PM, Room: Panorama V, Chair: NSF Anthony Kuh, Rowan University Robi Polikar and University of Rhode Island Haibo He

Wednesday, July 17, 4:00PM-4:30PM

Special Lecture: Coffee Break

Wednesday, July 17, 4:00PM-4:30PM, Room: Pre-function area Intercontinental

Wednesday, July 17, 4:30PM-5:30PM

Plenary Talk: Adam Miklosi, Eotvos Lorand University, Budapest

Wednesday, July 17, 4:30PM-5:30PM, Room: Ballroom I + II +II, Chair: Peter Erdi,

Wednesday, July 17, 7:30PM-11:00PM

: Banquet and Awards
Wednesday, July 17, 7:30PM-11:00PM, Room: Various locations in the area, Chair: C Jayne

Thursday, July 18, 8:00AM-9:40AM

Plenary Poster Session: Poster Session 1
Thursday, July 18, 8:00AM-9:40AM, Room: Ballroom I + II +II, Chair: Chrisina Jayne

P101 *A Deep Learning Algorithm for Fully Automatic Brain Tumor Segmentation [#19011]*
Yu Wang, Changsheng Li, Ting Zhu and Chongchong Yu, School of Computer and Information Engineering, Beijing Technology and Business University, China

Tumor segmentation is of great importance for diagnosis and prognosis of brain cancer in medical field. Many of the existing brain tumor segmentation methods are semiautomatic which need interventions of raters or specialists. In this paper an automatic method, named wide residual & pyramid pool network (WRN-PPNet), which can automatically segment glioma end to end is put forward. The main idea is described below. Firstly, substantial two-dimensional (2D) slices are obtained from three- dimensional (3D) MRI brain tumor images. Secondly, the 2D slices are normalized and put into the WRN-PPNet model, and the model will output the tumor segmentation results. Finally, dice coefficient (Dice), sensitivity coefficient (Sensitivity) and predictive positivity value (PPV) coefficient are used to evaluate the performance of WRN-PPNet quantitatively. The experimental results show that the proposed method is simple and robust compared with the other state-ofthe-art methods, and the average Dice, Sensitivity and PPV on the randomly selected test data can reach 0.94, 0.92 and 0.97 respectively.

P102 *Distributed Adaptive Dynamic Programming Algorithm for Office Energy Control with Multiple Batteries [#19021]*
Guang Shi, Chao Li, Bo Zhao, Qinglai Wei and Derong Liu, National Computer Network Emergency Response Technical Team/Coordination Center of China, China; School of Systems Science, Beijing Normal University, China; Institute of Automation, Chinese Academy of Sciences, China; Guangdong University of Technology, China

In this paper, the office energy coordination control problem with multiple batteries is solved by developing a distributed adaptive dynamic programming (ADP) algorithm. First, all the batteries are modeled using a unified battery by their average performance to transform the multi-input optimal control problem to a single-input one, which avoids high-dimensional optimization. Then, based on the optimal control law for the unified battery, only a single- input optimization problem is solved for one of all the batteries in each iteration. Finally, by achieving optimal control for all the batteries, the overall optimal control scheme for the office energy scheduling can be obtained, and therefore the energy cost from the power grid can be saved. Simulation analysis illustrates the effectiveness of the developed algorithm.

P103 *Learning Image Relations with Contrast Association Networks [#19028]*
Yao Lu, Zhirong Yang, Juho Kannala and Samuel Kaski, Australian National University, Australia; Norwegian University of Science and Technology, Norway; Aalto University, Finland

Inferring the relations between two images is an important class of tasks in computer vision. Examples of such tasks include computing optical flow and stereo disparity. We treat the relation inference tasks as a machine learning problem and tackle it with neural networks. A key to the problem is learning a representation of relations. We propose a new neural network module, contrast association unit (CAU), which explicitly models the relations between two sets of input variables. Due to the non-negativity of the weights in CAU, we adopt a multiplicative update algorithm for learning these weights. Experiments show that neural networks with CAUs are more effective in learning five fundamental image transformations than conventional neural networks.

P104 *KDSL: a Knowledge-Driven Supervised Learning Framework for Word Sense Disambiguation [#19031]*
Shi Yin, Yi Zhou, Chenguang Li, Shangfei Wang, Xiaoping Chen and Ruili Wang, School of Computer Science and Technology, University of Science and Technology of China, China; Shanghai Research Center for Brain Science and Brain Inspired Intelligence, China; Institute of Natural and Mathematical Sciences, Massey University (Albany Campus), New Zealand

We propose KDSL, a new word sense disambiguation (WSD) framework that utilizes knowledge to automatically generate sense-labeled data for supervised learning. First, from WordNet, we automatically construct a semantic knowledge base called DisDict, which provides refined feature words that highlight the differences among word senses, i.e., synsets. Second, we automatically generate new sense-labeled data by DisDict from unlabeled corpora. Third, these generated data, together with manually labeled data and unlabeled data, are fed to a neural framework conducting supervised and unsupervised learning jointly to model the semantic relations among synsets, feature words and their contexts. The experimental results show that KDSL outperforms several representative state-of-the-art methods on various major benchmarks. Interestingly, it performs relatively well even when manually labeled data is unavailable, thus provides a potential solution for similar tasks in a lack of manual annotations.

P105 *A Method of Pedestrian Fine-grained Attribute Detection and Recognition [#19038]*

Ma Xianqin, Yu Chongchong, Yang Xin, Chen Xiuxin, Chen Jianzhang and Zhou Lan, Beijing Technology and Business University, China; University of Illinois at Urbana Champaign, United States

The detection and recognition of pedestrians is one of the hottest research issues. This paper presents a method that employs the fusion of convolutional neural network (CNN) models based on multitask learning for multi-attributes, aiming to solve the problem of a low degree of accuracy in detecting and recognizing pedestrians' fine-grained attributes under complex circumstances. First, the model implements double detection for data preprocessing. Specifically, this paper uses the CNN model twice to detect pedestrians under complex circumstances. In the first detection, it conducts the coarse-grained detection for the whole pedestrian; in the second detection, based on the first, it conducts fine-grained detection of and then recognizes pedestrian subcomponents. Further, to address the problem of a low rate of correct recognition for the fine-grained detection of the attributes of pedestrian subcomponents, we use the concept of the fusion of multitask learning for multi-attributes, select the best recognized result for every attribute according to the fused results of different CNN classification models by the arrogance voting method, and then use a customized decision function to achieve a more accurate recognition of the coarse-grained attributes of pedestrians. Experimental results show that the proposed method achieved better performance than other recognition methods of pedestrians' attributes.

P106 *Short Text Topic Modeling with Flexible Word Patterns [#19058]*

Xiaobao Wu and Chunping Li, Tsinghua University, China

Since effective semantic representations are utilized in many practical applications, inferring discriminative and coherent latent topics from short texts is a critical and basic task. Traditional topic models like Probabilistic Latent Semantic Analysis (PLSA) and Latent Dirichlet Allocation (LDA) behave not well on short texts due to data sparsity problem. One novel model called Biterm Topic Model (BTM) which models unordered wordpairs (i.e., biterms) from whole corpus was proposed to solve this problem. However, both the performance and efficiency of BTM are reduced because of many irrelevant and useless biterms. In this paper, we propose a Multiterm Topic Model(MTM) for short text topic modeling. MTM extracts variable-length and more correlative word patterns (i.e., multiterms) from the whole corpus. By directly modeling the generative process of multiterms, MTM can infer the word distributions of each topic and the topic distribution of each short text to alleviate the sparsity problem in short texts modeling. With the the proper amount of flexible multiterms, learning process of MTM is enhanced. Through extensive experiments on two real-world short text collections, we show that MTM is more efficient and outperforms the baseline models in terms of topic coherence and text classification.

P107 *SOM-based Algorithm for Multi-armed Bandit Problem [#19067]*

Nobuhito Manome, Shuji Shinohara, Kouta Suzuki, Kosuke Tomonaga and Shunji Mitsuyoshi, SoftBank Robotics Corp./Graduate School of Engineering, The University of Tokyo, Japan; Graduate School of Engineering, The University of Tokyo, Japan

To satisfy a user facing a communication robot having various behavior options, it is necessary to output the behavior that is most suitable for the user faster. Such problems are formulated as a multi-armed bandit problem. The multi-armed bandit problem refers to the problem of maximizing gain in a situation with multiple arms where, by pulling a lever, a reward can be obtained from the arms with a certain probability, and the challenge is to select the arm that results in the maximum reward. Considering the behavior options of the communication robot as the arms, and the user satisfaction as the reward, under a condition with even more arms, it is desirable to select the arm that produces the maximum reward faster. This study proposes a new algorithm using a self-organizing map to solve the multi-armed bandit problem. Moreover, multiple numerical experiments have been conducted for the stochastic bandit problem, and it has been demonstrated that the proposed method is capable of selecting the arm with higher reward in a situation with even more arms faster, compared with the existing representative algorithms such as UCB1, UCB1-tuned, and Thompson sampling.

P108 *Text Classification Using Gated and Transposed Attention Networks [#19086]*

He Kang and Zhu Min, East China Normal University, China

Text classification models based on recurrent neural networks have poor parallel processing ability, as they need to be trained word by word. The model proposed in this paper has increased parallel computing capability compared to the one based on the recurrent neural networks. Text classification models based on convolution neural network struggle to obtain contextual information as they are limited by the size of the convolutional kernel. In this paper, we propose an attention network based on a gated control and transposed structure, which makes the network able to understand context and more efficiently extract important features from text. Experiments are carried out on six commonly used text classification datasets. The experimental results show that the model presented in this paper achieves the state of the art performance on four datasets and is very competitive on the other two. In addition, this paper also explores the effects of gate control and transposition on the whole model.

P109 *Adversarially Erased Learning for Person Re-identification by Fully Convolutional Networks [#19107]*

Shuangwei Liu, Yunzhou Zhang, Lin Qi, Sonya Coleman, Dermot Kerr and Shangdong Zhu, College of Information Science and Engineering, Northeastern University of China, China; Intelligent Systems Research Centre, University of Ulster, United Kingdom

The generalization ability of person re-identification model is subject to inadequate person data and occlusions. To relieve this dilemma, we propose a feature-level augmentation strategy, Adversarially Erased Learning Module (AELM), using two adversarial classifiers. Specifically, we utilize a classifier to identify discriminative regions and erase them to increase the variant of features. Meanwhile, we input the erased feature maps to another classifier to discover new body regions, which effectively resist occlusion of key parts. To easily perform end-to-end training for AELM, we propose a novel Identity model based on Fully Convolutional Networks (IFCN) to directly obtain body response heatmap during the forward pass by selecting corresponding class-specific feature map. Thus, the discriminative regions can be identified and erased in a convenient way. Moreover, to capture discriminative region for AELM, we present a Complementary Attention Module (CoAM) combined with channel and spatial attention to automatically focus on which feature types and positions are meaningful in the feature maps. In this paper, CoAM and AELM are cascaded into one module which is applied to the outputs of different convolutional layers to integrate mid- and high- level semantic features. Experimental results on three challenging benchmarks demonstrate the effectiveness of the proposed method.

P110 *Training a V1 Like Layer Using Gabor Filters in Convolutional Neural Networks [#19114]*

Jun Bai, Yi Zeng, Yuxuan Zhao and Feifei Zhao, Institute of Automation, Chinese Academy of Sciences, China

It is suggested from neural and cognitive science that the response of V1 neurons in primate visual streams behaves quite like Gabor filters. As an inspiration from primate neural circuits to artificial neural networks, we propose to replace the first layer as a series of Gabor filters in convolutional neural networks. To enhance the performance of the neural network, we introduce a lateral inhibitory mechanism in Gabor filters, enlightened from the research results of neural science. To improve performance, we explore a

parameter space and search the best suited parameters using cross validation. Experimental results demonstrate that the accuracy can basically match the results of the original convolutional neural networks. However, the adoption of Gabor filters can greatly reduce the time of training, as well as the memory and storage cost.

P111 *ShuffleNASNets: Efficient CNN models through modified Efficient Neural Architecture Search [#19117]*

Kevin Alexander Laube and Andreas Zell, Cognitive Systems Group, University of Tuebingen, Germany

Neural network architectures found by sophistic search algorithms achieve strikingly good test performance, surpassing most human-crafted network models by significant margins. Although computationally efficient, their design is often very complex, impairing execution speed. Additionally, finding models outside of the search space is not possible by design. While our space is still limited, we implement undiscoverable expert knowledge into the economic search algorithm Efficient Neural Architecture Search (ENAS), guided by the design principles and architecture of ShuffleNet V2. While maintaining baseline- like 2.85% test error on CIFAR-10, our ShuffleNASNets are significantly less complex, require fewer parameters, and are two times faster than the ENAS baseline in a classification task. These models also scale well to a low parameter space, achieving less than 5% test error with little regularization and only 236K parameters.

P112 *Parameter Reduction For Deep Neural Network Based Acoustic Models Using Sparsity Regularized Factorization Neurons [#19122]*

Hoon Chung, Euisok Chung, Jeon Gue Park and Ho-Young Jung, Electronics and Telecommunications Research Institute, Korea (South)

In this paper, we propose a Deep Neural Network (DNN) model parameter reduction technique for an efficient acoustic model. One of the most common DNN model parameter reduction techniques is to use low-rank matrix approximation. Although it can reduce a significant number of model parameters, there are two problems to be considered: one is the performance degradation, and the other is the appropriate rank selection. To solve these problems, retraining is carried out, and so-called explained variance is used. However, retraining takes additional time, and explained variance is not directly related to classification performance. Therefore, to mitigate these problems, we propose an approach that performs model parameter reduction simultaneously during model training from the aspect of minimizing classification error. The proposed method uses the product of three factorized matrices instead of a dense weight matrix, and applies sparsity constraint to make entries of the center diagonal matrix zero. After finishing training, a parameter reduced model can be obtained by discarding the left and right vectors corresponding to zero entries within the center diagonal matrix.

P113 *isAnon: Flow-Based Anonymity Network Traffic Identification Using Extreme Gradient Boosting [#19137]*

Zhenzhen Cai, Bo Jiang, Zhigang Lu, Junrong Liu and Pingchuan Ma, Institute of Information Engineering, Chinese Academy of Sciences, China

The abuse of anonymous communication technology brings serious challenges to network supervision. The valid identification of anonymity network traffic is a prerequisite and fundamentally important for preventing the violence of such techniques. However, due to the distinct characteristics of flow from anonymity networks including Tor, I2P, and JonDonym, existing studies don't take full advantage of these features, damaging the accuracy of identification. In this paper, we propose an effective anonymity network traffic identification model, called isAnon. Firstly, isAnon designs a novel hybrid feature selection algorithm by combining Modified Mutual Information and Random Forest (MMIRF) algorithm to filter out some irrelevant and redundant features quickly. Secondly, our proposed model applies a nested cross-validation scheme with an inner 5-fold cross-validation and an outer Monte Carlo cross-validation to prevent model overfitting. Finally, we use the Extreme Gradient Boosting (XGBoost) algorithm to identify Tor, I2P, and JonDonym networks for four scenarios. Comprehensive experimental results on several real-world anonymity network traffic datasets clearly show the effectiveness of our isAnon model compared with state-of-the-art baseline identification methods.

P114 *Label Distribution Feature Selection Based on Mutual Information in Fuzzy Rough Set Theory [#19138]*

Yingyao Wang and Jianhua Dai, Tianjin University, China; Hunan Normal University, China

As we all know, the multi-label learning is faced with the "dimension disaster", and the label distribution learning is actually faced with the same problem. Therefore, it is necessary to perform some appropriate pre- processing on the data before the label distribution learning, such as feature selection. Nowadays, researchers have proposed many feature selection algorithms for multi-label learning, while feature selection algorithms for label distribution learning are few. The difference between label distribution learning and multi-label learning is that in the existing multi-label classification methods, the importance of each label is considered to be the same, but in actual cases, different labels has different importance when they are describing the same instance, that is, there is an imbalance between the labels. Therefore, this paper proposed a feature selection algorithm based on information measures in fuzzy rough set theory for label distribution learning. The mutual information in fuzzy rough set theory is used to calculate the correlation between features and labels, and then a feature subset is selected by correlation maximization and redundancy minimization strategy. Finally, the experimental results demonstrate that the proposed feature selection algorithm is able to select valid subsets of features.

P115 *A new Spectral-Spatial Pseudo-3D Dense Network for Hyperspectral Image Classification [#19147]*

Ailin Li and Zhaowei Shang, Chongqing university, China

The recent research of hyperspectral images(HSIs) classification depicts that taking spectral-spatial features into account can considerably improve accuracy. Since HSI is a 3D cube datum, amounts of 3D network structures emerged to extract spectral- spatial features consecutively. In this paper, we present a simple spectral-spatial classification framework(SSP3DNet) based on a densely connected structure with Pseudo-3D block for HSIs. Firstly, a data augmentation strategy was implemented to solve the problem of limited and uneven training samples in the data preprocessing step. Secondly, the Pseudo-3D block can capture both spectral and spatial features simultaneously which is more economic by decreasing the number of parameters compared with traditional 3D convolution network. Then DenseNet learning framework is utilized to ease the training of networks as well as improving the classification performance. Especially, to prevent over-fitting, some tricks like early stopping, LeakyReLu, batch normalization, and dropout layers are used in our SSP3DNet which enable the SSP3DNet to obtain accuracy within 80 epochs. Experimental results on two well-known hyperspectral datasets show that the proposed SSP3DNet method achieves the best classification accuracy in comparison with lately traditional and deep-learning-based methods.

P116 *Clustering interval-valued data with automatic variables weighting [#19149]*

Sara Rodriguez and Francisco de Carvalho, Universidade Federal de Pernambuco - UFPE, Brazil

Over the past few years, Symbolic Data Analysis has gained popularity providing suitable methods for managing aggregated data represented by lists, intervals, histograms or even distributions. This paper proposes a partitioning clustering algorithm for interval-valued data based on the suitable adaptive Euclidean distance that takes into account the relevance of the variables according to the boundaries. The proposed distance changes at each algorithm iteration and is different from one cluster to another. The method provides a partition and a prototype for each cluster by optimizing an adequacy criterion that measures the fitting between groups and their

representatives. Experiments on synthetic and real interval-valued datasets corroborate the usefulness of the proposed algorithm.

P117 *On Correlation of Features Extracted by Deep Neural Networks [#19161]*

Babajide Ayinde, Tamer Inanc and Jacek Zurada, University of Louisville, United States

Redundancy in deep neural network (DNN) models has always been one of their most intriguing and important properties. DNNs have been shown to overparameterize or extract a lot of redundant features. In this work, we explore the impact of size (both width and depth), activation function, and weight initialization on the susceptibility of deep neural network models to extract redundant features. To estimate the number of redundant features in each layer, all the features of a given layer are hierarchically clustered according to their relative cosine distances in feature space and a set threshold. It is shown that both network size and activation function are the two most important components that foster the tendency of DNNs to extract redundant features. The concept is illustrated using deep multilayer perceptron and convolutional neural networks on MNIST digits recognition and CIFAR-10 dataset, respectively.

P118 *Learning Similarity: Feature-Aligning Network for Few-shot Action Recognition [#19168]*

Shaoqing Tan and Ruoyu Yang, Nanjing University, China

Deep learning structures have achieved impressive results in action recognition. However, most of deep models require extensive training on large scale datasets. Besides, Insufficient data can easily lead to overfitting. In this work, we propose a conceptually simple, flexible, and general approach for few-shot action recognition, where a model must learn to reliably classify an example having seen only few previous instances which belongs to the same class with it. Our method, called the Feature-Aligning Network (FAN), can be trained with a small amount of data. By applying "alignment" on representations from a pair of videos, we use a CNN to get an incorporating feature and learn a nonlinear similarity metric of it. In this way, FAN mainly focuses on capturing similarities and differences in the same type of feature maps between two feature map sets. We conduct standard few-shot classification experiments on UCF11, UCF101 and HMDB51 datasets, showing the ability of our model that it can quickly learn from few examples. Moreover, FAN is also applicable to action similarity labeling task, which is not only competitive, but also far simpler and more efficient than other approaches.

P119 *A Multiple Granularity Co-Reasoning Model for Multi-choice Reading Comprehension [#19172]*

Hang Miao, Ruifang Liu and Sheng Gao, Beijing University of Post and Telecommunications, China

We propose a multi-granularity co-reasoning model for multi-choice reading comprehension task, which aims to select the correct option based on the interaction between passage, question and candidate options. Firstly, we introduce a multiple granularity text matching module to interact passage with the question and each option. We take advantage of information extracted from diverse semantic spaces to conduct more extensive matching between text sequences. With this help, we could better match the passage against the question and each option to gather relevant information. Furthermore, we employ a multi-sentence co-reasoning module for sentence inference across multiple sentences. Specifically, we utilize 1D Convolutional Neural Network (1D-CNN) with different kernel sizes and self-attentive Recurrent Neural Network (RNN) to model the relationships of relevant sentences. This module could better synthesize and aggregate sentence-level evidence to make decisions. Experimental results demonstrate that our proposed model achieves state-of-the-art performance for single models on the RACE dataset.

P120 *A Deep Bidirectional Highway Long Short-Term Memory Network Approach to Chinese Semantic Role Labeling [#19177]*

Qi Xia, Chung-Hsing Yeh and Xiang-Yu Chen, Southeast University, China; Monash University, Australia

Existing approaches to Chinese semantic role labeling (SRL) mainly adopt deep long short- term memory (LSTM) neural networks to address the long-term dependencies problem. However, deep LSTM networks cannot address the vanishing gradient problem properly. In addition, the complexity of Chinese language, as a hieroglyphic language, decreases the performance of traditional SRL approaches to Chinese SRL. To address these problems, this paper proposes a new approach with a deep bidirectional highway LSTM network. The performance of the proposed approach is further improved by introducing the conditional random fields (CRFs) constraints and part-of-speech (POS) feature since POS tags are the classes of formal equivalents of words in linguistics. The experimental results on the commonly used Chinese Proposition Bank dataset show that the proposed approach outperforms existing approaches. With an easily acquired and reliable POS feature for practical applications, the proposed approach substantially improves Chinese SRL.

P121 *Mending is Better than Ending: Adapting Immutable Classifiers to Nonstationary Environments using Ensembles of Patches [#19179]*

Sebastian Kauschke, Lukas Fleckenstein and Johannes Fuernkranz, TU Darmstadt, Germany

In times of abundant streams of data, the only constant we can rely on is change itself. Given an existing classification model, it might be impractical or impossible to retrain when concept drift impairs the performance. In this paper, we propose an ensemble learning approach that fixes errors of a given model via patches: smaller, local models that are applied when the base model is likely to fail. These patches are added to the ensemble when the base learner shows decreased performance for one or multiple classes, which we discover by applying explicit drift detection. Each patch covers the drift of a subset of the classes. Via the drift detection method, the misclassification behavior of the base learner can be matched to an existing patch. This reduces the total amount of patches to the number of seen concept drifts, which is an advantage in the long run. Patches can be added to the ensemble and refined incrementally. The ensemble decision is conducted by a sequential classifier, which allows a high number of ensemble members that can be added incrementally. Experimental results show that ensemble patching performs well w.r.t. accuracy and adaptation speed.

P122 *ECG Segmentation by Neural Networks: Errors and Correction [#19185]*

Iana Sereda, Sergey Alekseev, Aleksandra Koneva, Roman Kataev and Grigory Osipov, Nizhny Novgorod State University, Russian Federation

In this study we examined the question of how error correction occurs in an ensemble of deep convolutional networks, trained for an important applied problem: segmentation of Electrocardiograms(ECG). We also explore the possibility of using the information about ensemble errors to evaluate a quality of data representation, built by the network. This possibility arises from the effect of distillation of outliers, which was demonstarted for the ensemble, described in this paper.

P123 *Seq2Seq Deep Learning Models for Microtext Normalization [#19199]*

Ranjan Satapathy, Yang Li, Sandro Cavallari and Erik Cambria, Nanyang Technological University, Singapore; Northwestern Polytechnical University, China

Microtext analysis is a crucial task for gauging social media opinion. In this paper, we compare four different deep learning encoder-decoder frameworks to handle microtext normalization problem. The frameworks have been evaluated on four different datasets in three different domains. To understand the impact of microtext normalization, we further integrate the framework into a sentiment classification task. This paper is the first of its kind to incorporate deep learning into a microtext normalization module and improve the sentiment analysis task. We show our models as a sequence to sequence character to word encoder-decoder model. We compare four deep learning models for microtext normalization task which further improve the accuracy of the sentiment analysis. Results show that the attentive LSTM and GRU cell both increase the sentiment analysis accuracy in the range of 4%-7% whereas LSTM and CNN with LSTM improve the accuracy in the range of 2%-4%

P124 *Generating Natural Video Descriptions using Semantic Gate [#19205]*

Hyungmin Lee and Il-Koo Kim, Samsung Electronics, Korea (South)

Video captioning task aims to generate a textual description of the situation in a video. It is challenging because of the nature of modality-difference between video and language. We present a novel method to bridge the gap between them by utilizing the semantic gate in two ways. First, we develop an activation mechanism to make a video description that captures the concept of the video. Next, we design a network that evaluates the similarity between visual and sentence feature. Semantic gate is used to transform sentence into a semantic embedding. We also conduct experiments to show that image and action classification task performance is transferred to video captioning task. Experimental results show that our proposed method has gained promising improvements compared to the baseline model. Consequently, our model demonstrated the effectiveness by achieving new best record on MSRVTT and MSVD dataset.

P125 *Patching Deep Neural Networks for Nonstationary Environments [#19207]*

Sebastian Kauschke, David Hermann Lehmann and Johannes Fuernkranz, TU Darmstadt, Germany

In this work we present neural network patching, an approach for adapting deep neural network models to nonstationary environments. Instead of creating or updating a network to accommodate concept drift, neural network patching leverages the inner layers of a previously trained network as well as its output to learn a patch that enhances the classification. It learns (i) a predictor that estimates whether the original network will misclassify an instance, and (ii) a patching network that fixes the misclassification. Neural network patching is based on the idea that the original network can still classify a majority of instances well, and that the inner feature representations encoded in the deep network aid the classifier to cope with unseen or changed inputs. We evaluated this technique on several datasets, comparing it to similar methods. Our finding is that neural network patching is adapting quickly to concept shifts, while also maintaining long-term learning capabilities similar to more complex methods that update the whole network.

P126 *Feature selection based on feature curve of subclass problem [#19209]*

Lei Liu, Bing Zhang, Shidong Wang, Shuangjie Li, Kaixiang Zhang and Shuqin Wang, College of Computer and Information Engineering, Tianjin Normal University, China

Feature selection is a key step to improve classification performance. Feature selection methods are divided into two types: filters and wrappers. Generally speaking, the filter methods use one score to judge the comprehensive classification ability of features for all classes. The higher the score is, the stronger the classification ability is. However, studies in many literature have indicated that only by selecting features with high scores often cannot achieve good effect. Therefore, this paper introduces a new feature selection method based on feature curve of subclass problem (referred to as IGRCFS) to find the features with high discriminal ability for each class, and then obtain the optimal feature subset. In order to verify the validity of the IGRCFS method, five kinds of existing feature selection methods are compared on eight datasets. The results demonstrate that the proposed method IGRCFS is effective.

P127 *Incremental Learning Based Subspace Modeling for Distributed Parameter Systems [#19219]*

Zhi Wang and Han-Xiong Li, City University of Hong Kong, China

In this paper, a novel incremental learning based subspace modeling method is developed for spatiotemporal modeling of distributed parameter systems (DPSs). First, the streaming snapshots are collected into small batches at a preset time interval in an online mode. The initial batch belongs to the first nominal subspace. Second, the dissimilarity analysis is further utilized to assign each new batch to one of the existing subspaces or a new subspace. Third, the local basis functions corresponding to the assigned subspace is updated or generated through incremental learning of the new batch data. Finally, all the local models are ensembled to approximate the system's dynamics over the whole time-space domain in real-time. The proposed method is tested on a hyperbolic advection system and a one-dimensional diffusion-reaction system. Results demonstrate that the proposed method is superior to the conventional global modeling, and achieves higher modeling accuracy for DPSs.

P128 *DNN-based Acoustic-to-Articulatory Inversion using Ultrasound Tongue Imaging [#19221]*

Dagoberto Porras, Alexander Sepulveda and Tamas Gabor Csapo, Universidad Industrial de Santander, Colombia; Budapest University of Technology and Economics, Hungary

Speech sounds are produced as the coordinated movement of the speaking organs. There are several available methods to model the relation of articulatory movements and the resulting speech signal. The reverse problem is often called as acoustic-to-articulatory inversion (AAI). In this paper we have implemented several different Deep Neural Networks (DNNs) to estimate the articulatory information from the acoustic signal. There are several previous works related to performing this task, but most of them are using ElectroMagnetic Articulography (EMA) for tracking the articulatory movement. Compared to EMA, Ultrasound Tongue Imaging (UTI) is a technique of higher cost-benefit if we take into account equipment cost, portability, safety and visualized structures. Seeing that, our goal is to train a DNN to obtain UT images, when using speech as input. We also test two approaches to represent the articulatory information: 1) the EigenTongue space and 2) the raw ultrasound image. As and objective quality measure for the reconstructed UT images, we use MSE, SSIM and CW-SSIM. Our experimental results show that CW-SSIM is the most useful error measure in the UTI context. We tested three different system configurations: a) simple DNN composed of 2 hidden layers with 64x64 pixels of an UTI file as target; b) the same simple DNN but with ultrasound images projected to the EigenTongue space as the target; c) and a more complex DNN composed of 5 hidden layers with UTI files projected to the EigenTongue space. In a subjective experiment the subjects found that the neural networks with two hidden layers were more suitable for this inversion task.

P129 *Two-Stream Convolution Neural Network with Video-stream for Action Recognition [#19281]*

Wei Dai, Yimin Chen, Chen Huang, Mingke Gao and Xinyu Zhang, School of Computer Engineering and Science,Shanghai University, China; China Electronics Technology Group Corporation, China

Recently, as the application of convolutional neural network in artificial intelligence is becoming increasingly diversified, a growing number of neural network methods is put forward. For example, 3D convolution and two-stream convolution method based on RGB and optical stream are applied to neural network. Convolutional neural network with 3D convolutional core is able to extract spatio-temporal features directly from a set of video sequences, used for action recognition. Although the 3D convolutional neural network can obtain partial spatio-temporal information, a new ConvNet architecture called CVDN (Combined Video-stream Deep Network) is proposed to extract more spatio- temporal features from video fragments so as to effectively utilize the temporal information in the dataset. We evaluate our method on the UCF-101 dataset and obtain a good result. The following is some details about our method: First, we use pre-trained ResNets models on Kinetics dataset to initialize our training models, training and extracting the video stream features from UCF-101 dataset. Then, optical flow graphs obtained from the UCF-101 dataset, which are the input of the optical stream, are used to extract the optical features. At length, two-stream features are combied and results are obtained after Softmax layer. When the linear fusion ratio of video stream features and optical stream features is 5:4, CVDN obtains good results. And the accuracy of our method with Resnet-101 achieves 92.2%.

P130 *Generative Adversarial Networks for Road Crack Image Segmentation [#19293]*

Ziping Gao, Bo Peng, Tianrui Li and Cong Gou, Southwest Jiaotong University, China

In this paper, we present a road crack segmentation method based on generative adversarial networks (GAN). Our GAN networks consist of two neural network models in terms of a generator and a discriminator, where two improved networks CU-Net and FU-Net are proposed based on U-Net. The U-Net, CU-Net and FU-Net are used as the generator, while two-class networks are used as the discriminator. The purpose of using the generator is to generate fake crack images which are very similar to real crack images. And the recognition is done by the discriminator to distinguish the real crack images from the fake crack images. After iterative training between the generator and the discriminator, the generator can generate fake crack images that are very similar to the real crack image. Finally, the generator can be used to segment the road crack images. Compared with the other state-of-the-art methods on three datasets, the proposed method achieves better performance. Specifically, the precision, recall and F1-score are 91.46, 73.40, and 77.33, respectively on one of the public datasets.

P131 *Dilated Convolutional Networks Incorporating Soft Entity Type Constraints for Distant Supervised Relation Extraction [#19301]*

Min Peng, Weilong Hu, Gang Tian, Bin Wang, Hua Wang and Gang Wang, Wuhan University, China; Xiaomi Inc, China; Victoria University, Australia

Although distant supervised relation extraction has been widely used, its performance is weakened by the wrong labeling problem. During eliminating noise, many previous work failed to grasp the trade-off between long dependencies and computational complexity while using neural networks, and ignored the potential noise in external knowledge. In this paper, we propose a neural model incorporating dilated convolutional neural networks with soft entity type constraints to jointly tackle the above two issues. Rather than using traditional convolutional neural networks or recurrent neural networks, we propose dilated convolutional networks as sentence encoder so as to capture long dependencies in large scale context and improve the robustness against local noise while keeping the efficiency of computation. Furthermore, we utilize entity types, which is considerd as external knowledge with noise, to denoise in relation classification by taking the constraints between entity types and relations into account and learning more precise entity types and attention weights simultaneously. Experimental results on the New York Times dataset show that, our model incorporating dilated convolutional networks with soft entity type constraints promotes distant supervised relation extraction and achieves state-of-art results compared with baselines.

P132 *A New Feature Selection Method based on Monarch Butterfly Optimization and Fisher Criterion [#19308]*

Xiaodong Qin, Xiabi Liu and Said Boumaraf, Beijing Institute of Technology, China; Beijing Institute of Technology, Algeria

This paper proposes an effective feature selection method based on monarch butterfly optimization and Fisher criterion. Fisher criterion is applied to evaluate the feature subsets, based on which the optimal feature subsets are searched by using monarch butterfly optimization algorithm. To combine these two components, a method is developed to binarize continuous solution vectors for deciding the feature selection. We conduct experiments on widely used UCI (University of California, Irvine) classification datasets to study the design of our algorithm and compare it with other state-of-the-art counterparts. The experimental results show that the proposed method is reasonable and effective, which achieves the best result of feature selection among the compared methods and has satisfactory efficiency.

P133 *A Position-aware Transformation Network for Aspect-level Sentiment Classification [#19318]*

Tao Jiang, Jiahai Wang, Youwei Song and Yanghui Rao, Sun Yat-sen University, China

This paper introduce a novel Position-aware Transformation Network (PTNet) for aspect-level sentiment classification. On the one hand, attention mechanisms have been employed to model the relationship between aspect and context. However, the position information of aspect words is rarely emphasized for sentiment prediction. The truth is that we should pay more attention to the word which is close to the aspect, since the words with closer distance may have a greater impact on the sentiment polarity of a sentence toward the aspect. On the other hand, existing approaches often adopt the average of aspect vectors or context vectors to calculate the attention weights, which may cause information loss if the aspect and context is not a single word. Therefore, this paper propose a position-aware layer and a context transformation layer in our model to solve the above two issues respectively. Moreover, several convolution kernels are also used to extract the n-gram information for prediction. We examine the performance of our model on three datasets: the first two are from SemEval2014 including the reviews of restaurants and laptops, and the third is a tweet collection. Experimental results show that our model consistently outperforms the stateof-the-art methods on all three datasets.

P134 *Impromptu Accompaniment of Pop Music using Coupled Latent Variable Model with Binary Regularizer [#19356]*

Bijue Jia, Jiancheng Lv, Yifan Pu and Xue Yang, Sichuan University, China

Symbolic music generation has long been an attractive topic in machine intelligence, which aims to automatically learn musical distribution from musical corpora and then to generate samples from the estimated distribution. As one of the most popular kinds, pop music is usually polyphonic and multi-track which makes difficulties in music generation. Furthermore, different from other tasks in machine intelligence such as image processing, the piano roll representation of music is discrete and binary, thus leading to a non-differentiable problem. In other words, most state-of- the-art models such as neural networks cannot be directly applied to achieve piano-roll-based symbolic music generation. To address these two issues, we propose a coupled latent variable model with binary regularizer. On the one hand, the proposed model employs a coupled mechanism to learn a latent variable that simultaneously captures the internal distribution of each track and the joint distribution of multiple tracks. On the other hand, we propose to reformulate the discrete and binary properties into a convex constraint in an

elegant way, thus obtaining a differentiable optimization problem and smoothly cooperating with neural networks. To show the effectiveness of our method, we carry out the experiment in impromptu accompaniment generation which is a popular application of music generation, and both the quantitative evaluation and human evaluation demonstrate promising performance of our model compared with some state-of-the-art models.

P135 *Correlation Filter Tracking Method via Metric Learning and Adaptive Multi-stage Appearance [#19363]*

Yan Hong, Jing Li, Yafu Xiao, Wenfan Zhang, Chengfang Song and Shan Xue, Wuhan University, China; Macquarie University, Australia

In the complex tracking environment such as background clutter, generally there are multiple peaks in the response map of the correlation filter. It is difficult to distinguish the real object from the interference; and using the fixed learning rate to update the appearance model, it is not only difficult to maintain the sample diversity, but also easy to introduce noise information. Aiming at this problem, this paper proposes a correlation filter tracking method via metric learning and adaptive multi-stage appearance. By introducing metric learning to discriminate candidate samples corresponding to multiple peaks in the response map, the influence of multi-peak response map on tracking results in complex environments such as background clutter is eliminated; the Gaussian mixture model is used to divide the object appearance samples into groups and assign corresponding weights according to the duration, and the redundant information is eliminated while maintaining the diversity of the appearance model samples. The experimental results on OTB100 and VOT2017 datasets show that the overall precision score obtained by the algorithm in this paper is 0.866. The overall success plot score is 0.628, and the expected average overlap score is 0.211, which is better than most existing tracking methods.

P136 *Unsupervised state representation learning with robotic priors: a robustness benchmark [#19377]*

TimothÓ Lesort, Mathieu Seurin, Xinrui Li, Natalia D z-Rodr_uez and David Filliat, ENSTA ParisTech & Thales, France; INRIA Lille, France; ENSTA ParisTech & INRIA Flowers, France

Our understanding of the world depends highly on our capacity to produce intuitive and simplified representations which can be easily used to solve problems. We reproduce this simplification process using a neural network to build a low dimensional state representation of the world from images acquired by a robot. As in Jonschkowski et al. 2015, we learn in an unsupervised way using prior knowledge about the world as loss functions called robotic priors and extend this approach to high dimension richer images to learn a 3D representation of the hand position of a robot from RGB images. We propose a quantitative evaluation metric of the learned representation that uses nearest neighbors in the state space and allows to assess its quality and show both the potential and limitations of robotic priors in realistic environments. We augment image size, add distractors and domain randomization, all crucial components to achieve transfer learning to real robots. Finally, we also contribute a new prior to improve the robustness of the representation. The applications of such low dimensional state representation range from easing reinforcement learning (RL) and knowledge transfer across tasks, to facilitating learning from raw data with more efficient and compact high level representations. The results show that the robotic prior approach is able to extract high level representation as the 3D position of an arm and organize it into a compact and coherent space of states in a challenging dataset.

P137 *Multiple Back Propagation Network and Metric Fusion for Person Re-identification [#19380]*

Si-Bao Chen, Feng Luo, Bin Luo, Chris Ding and Yi Liu, Anhui University, China; University of Texas at Arlington, United States; Peking University Shenzhen Institute, China

Person re-identification (Re-ID) is a research focus in pattern recognition, which is to identify a person from another camera view. Many researches have studied feature representations and metric distances of person images, which are robust to changes of view angle and illumination. In this paper, we propose a Multiple Back Propagation (MBP) network and Metric Fusion (MF) for person Re-ID. The proposed MBP network is based on DenseNet or ResNet. Each Dense-conv layer or Conv-ID block is linked by a MBP layer. Each MBP layer is divided into two sub-streams. One sub-stream is connected to softmax loss and the other sub-stream is transferred to a convolution layer followed by triplet loss. A Metric Fusion (MF) method with an optimized weighting scheme is proposed for deep feature fusion. Furthermore, we propose a new metric Re-ranking Euclidean distance joining metric fusion. Experiments on three large-scale person Re- ID benchmark datasets, including Market-1501, CUHK03 and DukeMTMC-reID, show that the proposed MBPMF method can achieve state-of-the-art performances.

P138 *SRAGAN: Generating Colour Landscape Photograph from Sketch [#19381]*

Si-Bao Chen, Peng-Cheng Wang, Bin Luo, Chris Ding and Jian Zhang, Anhui University, China; University of Texas at Arlington, United States; Peking University Shenzhen Institute, China

Generating sketch from colour landscape photograph is very easy while it is hard to generate colour photograph from landscape sketch. In this paper, a new automatic conversion network, named Sparse Residual Attention Generative Adversarial Networks (SRAGAN), is proposed to generate landscape colour photograph from sketch. Besides of generator adversarial loss, we not only adopt L1-regularized per-pix loss, but also combine L1-regularized perceptual loss together into our model. Due to the sparsity of L1-norm, it can preserve boundary edge information very well, which makes our model can handle well the conversion task of sketch-to-photo. In addition, we proposed a ResAttention block to our network structure, which combines the residual learning blocks with attention module. Experiments show that the landscape colour photographes generated by our SRAGAN looks more natural with bright colour and clear edge information. At the same time, we integrate two models so that we can generate winter-style and summer-style photographes from the same landscape sketch. Experiments demonstrate that our method outperforms many state-of-the-arts both in quantitative and in visual performance.

P139 *A Multi-Attentive Pyramidal Model for Visual Sentiment Analysis [#19401]*

Xiaohao He, Huijun Zhang, Ningyun Li, Ling Feng and Feng Zheng, Tsinghua University, China; Southern University of Science and Technology, China

Visual sentiment analysis aims to recognize emotions from visual contents. It is a very useful yet challenging task, especially when fine-grained emotions (such as love, joy, surprise, sadness, fear, anger, disgust, and anxiety) are analyzed. Existing methods based on convolutional neural networks learn sentiment representations based on global visual features, while ignoring the fact that both the local regions of the images and their relationships can have impact on sentiment representation learning. To address this limitation, in this paper, we propose a new Multi-Attentive Pyramidal model (MAP) for visual sentiment analysis. The model performs pyramidal segmentation and pooling upon the visual feature blocks obtained from a fully convolutional network, aiming to extract local visual features from multiple local regions at different scales of the global image. It then implants a self-attention mechanism to mine the associations between local visual features, and achieves the final sentiment representation. Extensive experiments on six benchmark datasets show the proposed MAP model outperforms the state-of-the-art methods in visual sentiment analysis.

P140 *Deep Feature Analysis in a Transfer Learning-based Approach for the Automatic Identification of Diabetic Macular Edema [#19415]*

Joaquim de Moura, Jorge Novo and Marcos Ortega, University of A Coruna, Spain

Diabetic Macular Edema (DME) is one of the most common causes of vision impairment and blindness in individuals with diabetes. Among the different imaging modalities, Optical Coherence Tomography (OCT) is a non-invasive ophthalmological imaging technique that is commonly used for the diagnosis, monitoring and treatment of DME. In this context, this paper proposes a new methodology for the automatic classification of DME using OCT images. Firstly, the method extracts a set of deep features from the target OCT images using a transfer learning-based approach. Then, the most relevant subset of deep features is selected using different feature selection strategies. Finally, a machine learning approach is applied to test the potential of the implemented method. The proposed methodology was validated using an OCT image dataset retrieved from 400 different patients, being 200 with DME and 200 normal cases. The proposed system achieved satisfactory results, reaching a best accuracy of 97.50%, using only 14.65% of the deep features in the classification of this ocular pathology, demonstrating also its competitive performance with respect to others approaches of the state-of-the-art.

P141 *Using Winning Lottery Tickets in Transfer Learning for Convolutional Neural Networks [#19417]*

Ryan Van Soelen and John Sheppard, Johns Hopkins University, United States; Montana State University, United States

Neural network pruning can be an effective method for creating more efficient networks without incurring a significant penalty in accuracy. It has been shown that the topology induced by pruning after training can be used to re-train a network from scratch on the same data set, with comparable or better performance. In the context of convolutional neural networks, we build on this work to show that not only can networks be pruned to 10\% of their original parameters, but that these sparse networks can also be re-trained on similar data sets with only a slight reduction in accuracy. We use the Lottery Ticket Hypothesis as the basis for our pruning method and discuss how this method can be an alternative to transfer learning, with positive initial results. This paper lays the groundwork for a transfer learning method that reduces the original network to its essential connections and does not require freezing entire layers.

P142 *Neural Networks Applied in the Prediction of Top Oil Temperature of Transformer [#19442]*

Wenxia Pan, Kun Zhao, Tianao Gao and Congchuang Gao, College of Energy and Electrical Engineering, Hohai University; Research Center for Renewable Energy Generation Engineering of Ministry of Education, Hohai University, China; College of Energy and Electrical Engineering, Hohai University, China; Jiangsu Guoxin Liyang Pumped Storage Power Generation Co., Ltd., China

Top oil temperature (TOT) is an important indicator reflecting the load capacity and insulation aging of the transformer. In order to predict the TOT accurately, this paper propose a transformer top oil temperature prediction method based on BP neural networks optimized by Adam. Firstly, we use the grey relational analysis method to calculate the correlation between other state variables of the transformer and the TOT, select state variables with larger correlation as the inputs and the TOT as the output to establish the neural networks prediction model (NNPM) of the TOT. Next NNPM of TOT is trained using historical data of transformer and Adam optimization algorithm. Then the case studying for historical data suggests that the prediction results of NNPM optimized by Adam of TOT are in accordance with measured results. Comparing with D Susa thermal circuit model and NNPM trained by SGD, the prediction accuracy of NNPM optimized by Adam is improved by 78.1% and 33.95% respectively. Finally, we choose different transformers to model and predict, and the results show that NNPM of TOT based on Adam has applicable ability to different transformers. The top oil temperature prediction method proposed in this paper provides a more accurate calculation basis for prediction of TOT of transformers and is of great significance for the safe and stable operation of the power transformers.

P143 *An End-to-End Joint Unsupervised Learning of Deep Model and Pseudo-Classes for Remote Sensing Scene Representation [#19446]*

Zhiqiang Gong, Ping Zhong, Weidong Hu, Fang Liu and BingWei Hui, National University of Defense Technology, China

This work develops a novel end-to-end deep unsupervised learning method based on convolutional neural network (CNN) with pseudo-classes for remote sensing scene representation. First, we introduce center points as the centers of the pseudo classes and the training samples can be allocated with pseudo labels based on the center points. Therefore, the CNN model, which is used to extract features from the scenes, can be trained supervised with the pseudo labels. Moreover, a pseudo-center loss is developed to decrease the variance between the samples and the corresponding pseudo center point. The pseudo-center loss is important since it can update both the center points with the training samples and the CNN model with the center points in the training process simultaneously. Finally, joint learning of the pseudo-center loss and the pseudo softmax loss which is formulated with the samples and the pseudo labels is developed for unsupervised remote sensing scene representation to obtain discriminative representations from the scenes. Experiments are conducted over two commonly used remote sensing scene datasets to validate the effectiveness of the proposed method and the experimental results show the superiority of the proposed method when compared with other state-of-the-art methods.

P144 *Bacteria shape classification by the use of region covariance and Convolutional Neural Network [#19459]*

Dawid Polap and Marcin Wozniak, Institute of Mathematics, Silesian University of Technology, Poland

In modern medical systems fast recognition of bacteria strain from microscopy image is very often done by the use of specialized computer systems. For these programs intelligent methodologies are very important tools. In this paper we present a model of bacteria recognition based on a composition of region covariance with Convolutional Neural Networks (CNN). In the first stage an input microscopy image is segmented by the use of region covariance model. Next these segments are forwarded to CNN for recognition of visible bacteria strains. Experiments were done for rod-shaped bacteria and spherical or nearly spherical shape bacteria. The results show high potential of the proposed methodology.

P145 *Latent Space Embedding for Unsupervised Feature Selection via Joint Dictionary Learning [#19465]*

Yang Fan, Jianhua Dai and Qilai Zhang, Tianjin University, China; Hunan Normal University, China

With the prevalence of unlabeled data, unsupervised feature selection is vital for comprehensive analysis of unlabeled high-dimensional data. Most existing unsupervised feature selection methods first generate cluster labels by specific techniques and then select features that can preserve cluster structure well. However, the selected features only reflect the distribution information of pseudo label space but ignore that of feature space. Instead, we propose a novel method, latent space embedding for unsupervised feature selection, which considers the common distribution of feature space and pseudo label space, spectral analysis and feature selection simultaneously. Inspired by the success of joint dictionary learning in cross-modality cases, we introduce a latent space shared by feature space and pseudo label space. By utilizing the mapping between feature space and latent space, features that can well maintain the common distribution of features and pseudo labels are selected. The $\ell_{2,1}$-norm minimization

constraint is added to the objective function to handle outliers and noise. Experimental results on benchmark datasets demonstrate that our algorithm outperforms the comparison methods in terms of clustering tasks.

P146 *LMLSTM: Extract Event-Oriented Keyphrase From News Stream [#19467]*

Lin Zhao, Longtao Huang, Liangjun Zang, Jizhong Han and Songlin Hu, Institute of Information Engineering,University of Chinese Academy of Sciences, China; Institute of Information Engineering, China

Keyphrase extraction, as a basis for many natural language processing and information retrieval tasks, can help people efficiently discover their interested information from vast streams of online documents. Previous methods are mostly proposed in general purpose, where keyphrases that represent the main topics are extracted. However, such keyphrases can hardly distinguish events from massive streams of long text documents that share similar topics and contain highly redundant information. In this paper, we address the task of keyphrase extraction for event-oriented retrieval. We propose a novel Long Short-Term Memory Network Language Model (LMLSTM) to extract event-oriented keyphrases that represent or related to a particular event. We conduct a series of experiments on a real-world dataset. The experimental results demonstrate the better performance of our approach than other state-of-the-art baselines.

P147 *Approximating Binarization in Neural Networks [#19485]*

Caglar Aytekin, Francesco Cricri, Jani Lainema, Emre Aksu and Miska Hannuksela, Nokia Technologies, Finland

Binarization of neural networks' activations may be a requirement for some applications. A typical example is end-to-end learned deep image compression systems where the encoder's output is requred to be a binary vector. Binarization is non-differentiable, therefore one needs to approximate it in order to train neural networks with stochastic gradient descent. In this paper, we investigate these training strategies and provide improvements over baselines. We find that during training, constraining the activations in a region that is far away from binary points leads to a better performance at test-time. The above finding provides a counter-intuitive result and leads to re-thinking the binarization approximation problem in neural networks.

P148 *Convolutional Recurrent Neural Networks for Text Classification [#19512]*

Ruishuang Wang, Zhao Li, Jian Cao, Tong Chen and Lei Wang, Big Data Engineering Technology Research Center of E-Government,Shandong, China; Qilu University of Technology(Shandong Academy of Sciences), Shandong Computer Science Center(National Supercomputer Center in Jinan), China

Text classification is an important task in natural language processing with wide applications. Traditional text classification methods manually extract the features which are later fed into the classifier for training. Recent researchers have employed convolutional neural networks or recurrent neural networks for text classification motivated by the noticeable success of deep learning. However, most of their models are based on single network. In this paper, we introduce a convolutional recurrent neural network for text classification, which enjoys both the advantages of convolutional neural networks for extracting local features from text and also those of recurrent neural networks (LSTM) in memory to connect the extracted features. We conduct extensive experiments on two Chinese data sets and five English data sets, and compare our method with several other classification methods. Experimental results show that the proposed method can achieve better accuracy on text classification tasks.

P149 *Improving the quality of enzyme prediction by using feature selection and dimensionality reduction [#19542]*

Luis Brito, Ana Lara, Luis Zarate and Cristiane Nobre, Pontifical Catholic University of Minas Gerais, Brazil

The coming of the post-genomic era has raised a growing demand for efficient procedures to identify protein functions, which can be accomplished by applying machine learning to the characteristics set extracted from the protein. This approach is called a feature-based approach and has been the focus of several works in Bioinformatics. In this work, we investigated the characteristics of proteins that improve the results found in the related works that used Support Vector Machine (SVM) in the enzymes class prediction. During the investigation, two approaches were evaluated: with and without dimensionality reduction, which was performed through the statistical technique Factor Analysis. Both approaches exceeded the results found in the related work, reaching F-measure averages of 85.10\% and 83.90\%, respectively, in the approachs with and without dimensionality reduction.

P150 *TCoD:A Traveling Companion Discovery Method Based on Clustering and Association Analysis [#19548]*

Ruihong Yao, Fei Wang and Shuhui Chen, National University of Defense Technology, China

Widely used mobile locating equipment, like phones, generates extensive spatio- temporal data every day. Since the data implicitly reflects behavior characteristics of moving objects, many applications focus on trajectory data mining while traveling companion discovery is one of the most fundamental techniques in these areas. Previous work based on time snapshot slicing have yielded some success on finding companions in regularly and frequently sampled trajectory data. However, it does not work well when data is sparse. This situation commonly appears in real world because of equipment failure or manual intervention. In this paper, we propose a novel Traveling Companion Discovery (TCoD) method that can discovery travelling companions even when time gaps between sample data are more than hours. TCoD combines density clustering and association analysis, while density clustering mine potential sets from perspective of location and association analysis identify related objects in potential sets. The evaluation on two real trajectory data sets shows that TCoD effectively overwhelms previous work with satisfying discovery of long-term companion patterns in sparse trajectory data.

P151 *Model Based on Deep Feature Extraction for Diagnosis of Alzheimer's Disease [#19554]*

Iago Silva, Gabriela Silva, Rodrigo Souza, Wellington Santos and Roberta Fagundes, University of Pernambuco, Brazil; Federal University of Pernambuco, Brazil

Alzheimer's disease (AD) is a neurodegenerative disease that results in loss of cognitive ability of the patient. Computational intelligence, more specifically Deep Learning, has been a powerful method for AD diagnosis. In this work we propose a model for AD diagnosis based on deep feature extraction for the classification using magnetic resonance imaging. This model aims to classify AD vs. HC (Healthy Controls). The database used in this project is the Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD), for validation of the proposed method. We select thirty slices from the upper region of the brain, above the eyes, for the apprenticeship in this work. The Convolutional Neural Network (CNN) architecture is designed in three convolutional layers to extract the best features of the selected region. After that, we put the selected attributes in a vector for learning and detection of patterns by another technique of computational intelligence. Finally, the data are partitioned with the 10- folds cross-validation method and trained with the Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN) algorithms with different parameters for evaluation. The results of accuracy are 0.8832, 0.9508 and 0.8745, for the algorithms mentioned above, respectively. According to a comparative analysis performed with other works of the literature, we can prove the efficiency and reliability of the model for the diagnosis of Alzheimer's disease.

P152 *A Composite Extended Nearest Neighbor Model for Day-Ahead Load Forecasting [#19562]*

Md. Rashedul Haq and Zhen Ni, South Dakota State Unviersity, United States

Day-ahead load forecasting is an important task for the reduction of electricity waste and efficient management of a smart grid. The electricity load profile data reveals the correlation of electricity load demand with weather condition, day type (working day or holiday), time of the day and season of the year. Thus the load forecasting problem has a high degree of complexity with consideration of those variables as input. To solve the problem of day-ahead short-term load forecasting (STLF), the proposed solution first classifies load profile data into different classes. To this end, a recent developed classification approach called extended nearest neighbor (ENN) algorithm is adopted. Then, a composite ENN model is proposed for dayahead load forecasting. The composite ENN model consists of three individual ENN models which are combined together by tuned weight factors for predicting final forecasting output. Unlike other statistical and computational intelligence approaches, the composite ENN model predicts electricity load demand from the maximum gain of intra-class coherence. By exploiting intraclass coherence from the generalized class-wise statistic of all available training samples, the composite ENN algorithm is able to learn from global distribution and therefore improve the accuracy of load forecasting. The proposed method is validated on two case study: (i) Australian National Energy Market Data and (ii) Brookings, South Dakota, USA Data. For case study 1, mean absolute percent error (MAPE) of composite ENN based load forecasting is decreased by 44.68% compared to composite kNN based load forecasting and mean absolute error (MAE) is decreased by 45.52%.

P153 *Intrusion Detection Method based on Information Gain and ReliefF Feature Selection [#19591]*

Zhang Yong, Ren Xuezhen and Zhang Jie, Liaoning Normal University, China

Traditional random forest has slow convergence in network intrusion detection and its learning performance is not perfect. In order to eliminate the redundant information in the original intrusion detection data, this paper proposes a random forest intrusion detection method based on the combination of information gain and ReliefF algorithm. The proposed method first uses the information gain to calculate the information gain value of each feature. Then, the ReliefF algorithm is used to calculate the weight of each feature. According to the information gain and the feature weight, the final feature subset is obtained. Finally, this paper uses a random forest classifier for classification. The experiment compares the three feature selection methods, including the proposed method, information gain based method and ReliefF-based method. The experimental results show that the precision, recall rate, and false positive rate of the proposed method are superior to those of the other two methods.

P154 *Noise-Aware Network Embedding for Multiplex Network [#19593]*

Xiaokai Chu, Xinxin Fan, Di Yao, Chenlin Zhang, Jianhui Huang and Jingping Bi, Institute of Computing Technology Chinese Academy of Sciences, University of Chinese Academy of Sciences, China; Institute of Computing Chinese Academy of Sciences, China; Institute of Computing Chinese Academy of Sciences, University of Chinese Academy of Sciences, China; National Key Laboratory for Novel Software Technology, Nanjing University, China

Network embedding aims at learning the latent representations of nodes while preserving the complex structure of the underlying graph. Real-world networks are usually related with each other via common nodes, the so-called multiplex network. To make the data mining work on the multiplex network more actionable, it become urgent and essential to transform it into low- dimension vector space. Recently, several works have been proposed to leverage the complementary information for embedding. However, they suffer from sacrificing distinct properties of the counterparts in different layers, as they preserve much noise information into embedding vectors. In this paper, we propose a Noise-Aware Network Embedding approach for Multiplex Network, namely NANE. Unlike previous works, NANE considers the roles of an identical node in different layers, and adopts a more robust and flexible strategy to rationally integrate the cross-layer information while keeping the unique characteristic of each layer. We perform extensive evaluations on several real-world datasets. The experimental results demonstrate that our NANE can achieve better performance on link prediction task and significantly outperform previous methods especially in noisy multiplex network scenarios.

P155 *A Hybrid Convolutional Approach for Parking Availability Prediction [#19606]*

Hadi Jomaa, Josif Grabocka and Lars Schmidt-thieme, Stiftung Universitat Hildesheim, Germany

Parking availability prediction is rapidly gaining in- terest within the community as an operationally cheap approach to identifying empty parking locations. Parking locations accom- modate multiple vehicles and are rarely completely occupied. This makes it difficult to predict occupied locations without the augmentation of external data, as the data becomes highly imbalanced. Existing forecasting models neither encapsulate the heterogeneous modes/types of parking data, nor can handle sparse measurements. The problem is formulated as a binary forecasting task, based on the parking occupancy information. In this paper, we propose a new convolutional hybrid model that is capable of capturing long term temporal dependencies and outperforming conventional time-series forecasting benchmarks on two types of parking data, namely on- and off-street parking. The performance of the proposed model is further boosted by integrating external features such as location identifiers, as well as local/global statistics. An extensive experimental evaluation proves that the proposed model is capable of handling sparse data by maintaining high precision and recall across different sparsity levels, which are controlled by empirically adjusting the occupancy cut-off threshold, as well as for multiple horizons, with an average F1 score improvement of 4.13% over strong off-the- shelf baselines.

P156 *Graph Convolutional Networks with Structural Attention Model for Aspect Based Sentiment Analysis [#19610]*

Junjie Chen, Hongxu Hou, Yatu Ji and Jing Gao, Inner Mongolia University, China; Inner Mongolia Agricultural University, China

With the amount of user-generated information on the Web, identifying the sentiment polarity of the given aspect provides more complete and in-depth results for businesses and customers. Aspect based sentiment analysis has gained increasing attention in decade years, but it remains a daunting task. Recently, approaches based on recurrent neural networks and convolutional neural networks have shown competitive results in this field. However, they don't take fully account of the entire text structure and the relation between words in a given document. In this paper, we propose a novel neural network method to address this problem, in which the text is treated as a graph and the aspect is the specific area of the graph. For the first time, we apply graph convolutional neural networks and structural attention model to aspect based sentiment analysis. Experiments on public-available datasets demonstrate the efficiency and effectiveness of our model.

P157 *Extracting Prerequisite Relations Among Concepts in Wikipedia [#19629]*

Yang Zhou and Kui Xiao, Hubei University, China

Extracting prerequisite relations among concepts is helpful for users to find out the background knowledge and determine reading order in a corpus. We investigate the problem and propose an effective method and multiple features to capture prerequisite relations between concepts in Wikipedia. Our experiments on eight datasets from both English and Chinese Wikipedia show that the proposed method outperforms existing prerequisite learning methods.

P158 *Cross-project Defect Prediction via ASTToken2Vec and BLSTM-based Neural Network [#19631]*

Hao Li, Xiaohong Li, Xiang Chen, Xiaofei Xie, Yanzhou Mu and Zhiyong Feng, Tianjin University, China; Nantong University, China; Nanyang Technological University, Singapore

Cross-project defect prediction (CPDP) as a means to focus quality assurance of software projects was under heavy investigation in recent years. However, existing CPDP approaches require the similar feature distribution between different projects. In this paper, we propose a novel CPDP approach via deep learning. In particular, we model each program module via simplified abstract syntax tree (S-AST). For each node in S-AST, only the project-independent node type is remained and other project-specific information (such as name of variable and method) is ignored, so that the modeling method is project-independent and suitable for CPDP issue. Then we extract token sequences from program modules modeled as S-AST. In addition, to construct meaningful vector representations for token sequences, we propose a novel unsupervised embedding method ASTToken2Vec, which learns semantic information from S-AST's natural structure. Finally, we use BLSTM (bi-directional long short-term memory) based neural network to automatically learn semantic features from vectorized token sequences and construct CPDP models. In our empirical studies, 10 real large-scale open source Java projects are chosen as our empirical subjects. Final results show that our proposed CPDP approach can perform significantly better than 5 state-of-the-art CPDP baselines in terms of AUC.

P159 *Event-Triggered $H_$ Tracking Control of Nonlinear Systems via Reinforcement Learning Method [#19636]*

Lili Cui, Wei Qu, Li Wang, Yanhong Luo and Zhanshan Wang, Shenyang Normal University, China; Northeastern University, China

In this paper, the H_{∞} tracking control problem of continuous-time nonlinear system is investigated. An event-triggered adaptive optimal tracking control scheme is developed based on the reinforcement learning method. By constructing an augmented system and introducing a discounted cost function, the corresponding event-driven tracking Hamilton-Jacobi-Isaacs (HJI) equation is derived, which provides the solution of the H_{∞} tracking control problem. Since the event-triggered tracking HJI equation is a nonlinear partial difference equation in essence, it is hard to be solved analytically. To overcome this difficulty, a novel reinforcement learning algorithm is proposed to learn the solution of derived event-triggered tracking HJI equation in which a critic network is employed to approximate the optimal cost function on-line without the requirement of initial admissible control policy. It is worthy to mention that the developed event-triggered tracking controller is updated only at the event-triggered instants, which can greatly decrease the controller update frequency and reduce the communication burden by contrast to the time- triggered one. The stability of closed-loop system is demonstrated by the Lyapunov theory. Simulation results validate the effectiveness of proposed event-triggered tracking control scheme.

P160 *A Unified Multi-output Semi-supervised Network for 3D Face Reconstruction [#19649]*

Pengrui Wang, Yi Tian, Wujun Che and Bo Xu, Institute of Automation, Chinese Academy of Sciences, Beijing, China, China

In this paper, we propose a method to reconstruct fine-grained 3D faces from single images base on a nearly unified multi-output regression network. The network estimates the facial shape, normal and appearance jointly in 2D UV map which preserves spatial adjacency relations among vertexes and provides semantic meaning of each vertex. Three contributions of the proposed method are: 1) we generate the UV map by as-rigid-as-possible parametrization to address the overlapping problem caused by cylindrical unwarp; 2) we directly estimate face normal rather than compute it from the estimated shape to let it catch geometric details from face texture; 3) we propose a post process strategy to generating more realistic faces and to employing the estimated normal. Experiments show that our network is able to learn a uniform appearance and predict more accurate shape from the proposed UV map. Additionally, the post process procedure can improve the quality of facial shapes and add geometric details from estimated normals.

P161 *Multi-Level Compare-Aggregate Model for Text Matching [#19683]*

Chunlin Xu, Hui Wang, Zhiwei Lin and Shengli Wu, University of Ulster, Northern Ireland

Text matching is important for a variety of natural language processing tasks, such as paraphrase identification and natural language inference. Recent studies have achieved very promising results under the compare-aggregate framework. A limitation of previous approaches following this framework is that they solely conduct matching at word level. In this paper, we propose a multi-level compare-aggregate model (MLCA), which matches each word in one text against the other text at three different levels, word level (word-by-word matching), phrase level (word-by-phrase matching) and sentence level (word-by-sentence matching). Then the results of these levels of matching are aggregated for making final matching decision. We evaluate our model on two different tasks: paraphrase identification and natural language inference. Experimental results show that our model achieves the state-of-the-art performance on both tasks.

P162 *DeepShapeSketch : Generating hand drawing sketches from 3D objects [#19694]*

Meijuan Ye, Shizhe Zhou and Hongbo Fu, College of Computer Science and Electronic Engineering, Hunan University, China; City University of Hong Kong, China

Freehand sketches are an important medium for expressing and communicating ideas. However creating a meaningful and understandable sketch drawing is not always an easy task especially for unskillful users.Existing methods for rendering 3D shape into line drawings such as Suggestive Contours, only consider the geometry-dependent and view-dependent information thus leads to over-regular or over-perfect results which doesn't look like a human freehand drawing.For this challenge we address the problem of producing freehand line drawing sketches from a 3D object under a given viewpoint automatically.The core solution here is a recurrent generative deep neural network, which learns a functional mapping from the suggestive contours of a 3D shape to a more abstract sketch representation.We drop the encoder of the generator, i.e., use only a decoder to achieve better stability of the sketch structure.Users can tune the level of freehand style of the generated sketches by changing a single parameter.Experiments show that our results are expressive enough to faithfully describe the input shape and at the same time be with the style of freehand drawings created by a real human.We also perform a comparative user study to verify the quality and style of generated sketch results over existing methods.We also retrain our network using several different mingled dataset to test the extendibility of our method for this particular application.As far as our knowledge this work is the first research effort to automate the generation of human-like freehand sketches directly from 3D shapes.

P163 *Author Disambiguation through Adversarial Network Representation Learning [#19712]*

Liwen Peng, Siqi Shen, Dongsheng Li, Jun Xu, Yongquan Fu and Huayou Su, National University of Defense Technology, China

Many persons share with the same name. Distinguishing different persons with the same name is important but challenging. Albeit much work has been proposed for author disambiguation, most of them do not adequately consider the heterogeneous relationships among authors and papers. In our work, ambiguous names and their related information, such as papers, conferences, titles, abstracts, etc., are constructed into a heterogeneous network which consists of different edge types. To fully incorporate all the information of the constructed network, we use Generative Adversarial Networks (GAN) to learn the network representation of the heterogeneous

network. Although GAN has been used in many fields such as image generation, it hasn't been used to obtain representations for the heterogeneous network. As far as we know, our work is the first work which use adversarial training to learn heterogeneous network representation. After the representations are learned, they are partitioned into different groups each representing distinct authors. After extensive experiments on three major author disambiguation datasets, we demonstrate that our method outperforms several state-of-the-art baselines in author disambiguation problem.

P164 *An End-to-end Network for Monocular Visual Odometry Based on Image Sequence [#19718]*

Mingwei Yao and Hongyan Quan, School of Computer Science and Software Engineering East China Normal University, China

Regarding depth and camera pose estimation tasks on the image sequence, we propose a monocular visual odometry frame named VONN on basis of convolutional neural network and recurrent neural network in this paper. The entire network frame is composed of two parts, i.e. depth estimator and camera 6-DoF position estimator. Besides, the paper also proposes a geometric consistency loss on this basis to train the network. By comparing VONN with other approaches in the accuracy of 6-DoF position estimation and accuracy of depth estimation, our method achieves better results and verifies the effectiveness of the frame and 3D geometric consistency loss proposed in this paper.

P165 *Network Search for Binary Networks [#19721]*

Jiajun Du, Yu Qin and Hongtao Lu, Shanghai Jiao Tong University, China

Neural networks with both high accuracy and small network size are urgently required for mobile phone applications. However, previous network search methods do not take network size into account. In this paper, we use reinforcement learning method to search for networks with both high accuracy and small network size. A RNN controller is used to generate networks and is updated via policy gradient. Gaussian policy is used to predict the number of channels in each convolutional layer. Our reward function contains both accuracy reward and parameter reward. We also use binary networks to further reduce network size. Without skip connections or branches, the network generated by our method is competitive with other methods on Cifar-10. Our network is much smaller than networks generated by other network search methods. Besides, our accuracy is higher than original binary network reported in BinaryConnect and is competitive with other real-valued networks.

P166 *A Semi-supervised Classification Using Gated Linear Model [#19724]*

Yanni Ren, Weite Li and Jinglu Hu, Graduate School of Information, Product and System, Waseda University, Japan

Semi-supervised learning aims to construct a classifier by making use of both labeled data and unlabeled data. This paper proposes a semi-supervised classification method using a gated linear model, based on the idea of effectively utilizing manifold information. A gating mechanism is firstly trained in a semi-supervised manner to capture manifold information which guides the generation of gate signals. Then the gated linear model is formulated into a linear regression form with the gate signals included. Secondly, a Laplacian regularized least squares (LapRLS) formulation is applied to optimize the linear regression form of the gated linear model. In this way, the gate signals are integrated into the kernel function, which is defined as the inner product of the regression vectors. Moreover, this kernel function is used as a better similarity function for graph construction. As a result, the manifold information is ingeniously incorporated into both kernel and graph Laplacian in the LapRLS. Experimental results exhibit the effectiveness of our proposed method.

P167 *Batch Mode Active Learning with Nonlocal Self-Similarity Prior for Semantic Segmentation [#19746]*

Yao Tan, Qinghua Hu and Zhibin Du, School of Computer Science and Technology,College of Intelligence and Computing,Tianjin University, China; China Automotive Technology & Research Center, China

Semantic segmentation is a task that heavily relies on the annotated data. The image annotation cost is very expensive. While active learning aims to select the most valuable samples by an iterative procedure. It can reduce the annotation cost and improve the performance of classification. In semantic segmentation, it's more common to select a batch of instances instead of a single instance at each iteration. In this paper, we propose a novel batch mode active learning algorithm for semantic segmentation. Different from the previous active learning algorithms for the image classification, we first introduce a new selecting criterion : the image prior of nonlocal self-similarity. It can measure the interaction between the pixels of the image. We combine the informativeness and representativeness with the criterion of nonlocal self-similarity to complete the selection of images at each iteration. In addition, we also use the model uncertainty to measure the information of samples. The model uncertainty is captured by using Monte Carlo Dropout in the semantic segmentation model. In this work, we evaluate our method on the CamVid and PASCAL VOC 2012 datasets. The importance of the nonlocal self-similarity is also assessed. The experiments demonstrate that our algorithm outperforms current state- of-the-art active learning methods over the segmentation performance.

P168 *Multi-Satellite Resource Scheduling Based on Deep Neural Network [#19753]*

Huan Meng, Changde Li, Weizhi Lu, Yuhan Dong, Zhipeng Zhao and Bin Wu, Tianjin University, China; Beijing Institute of Satellite Information Engineering, China

Resource scheduling is one of the main problems for multi-satellite Tracking, Telemetry and Command (TT&C) networks. Traditional multi-resource joint scheduling algorithms are with long solution time, low efficiency, high computational cost, and simple description on the system. Deep Neural Network (DNN) provides a possible new way to solve those problems, but it is difficult to handle correlations among the input data. This motivates our work to solve the strong correlation problem based on the accumulated historical data, and thus enables DNN for TT&C resource scheduling. By discretizing the data, multiple constraints and related attributes are transformed into different flags, and some binary bits of the data are used to reflect the constraint relationship. Then, we can use DNN model and construct an intelligent TT&C resource scheduling system to handle multiple constraints and data attributes (such as priorities among tasks and others). This improves the efficiency of TT&C resources utilization and automation. Effectiveness of the proposed model is verified by simulations.

P169 *A Feature Learning Siamese Model for Intelligent Control of the Dynamic Range Compressor [#19759]*

Di Sheng and Gyorgy Fazekas, Queen Mary University of London, United Kingdom

In this paper, a siamese DNN model is proposed to learn the characteristics of the audio dynamic range compressor (DRC). This facilitates an intelligent control system that uses audio examples to configure the DRC, a widely used non-linear audio signal conditioning technique in the areas of music production, speech communication and broadcasting. Several alternative siamese DNN architectures are proposed to learn feature embeddings that can characterise subtle effects due to dynamic range compression. These models are compared with each other as well as handcrafted features proposed in previous work. The evaluation of the relations between the hyperparameters of DNN and DRC parameters are also provided. The best model is able to produce a universal feature embedding that is capable of predicting multiple DRC parameters simultaneously, which is a significant

improvement from our previous research. The feature embedding shows better performance than handcrafted audio features when predicting DRC parameters for both mono-instrument audio loops and polyphonic music pieces.

P170 *A Novel Recommender System using Hidden Bayesian Probabilistic Model based Collaborative Filtering [#19778]*

Xin Dai, Fanzhang Li, Xiaopei Li and Helan Liang, Soochow University, China

For the problems of data sparseness and cold start of goods in the existing recommendation algorithm, in this paper, we propose a new method based on the hidden Bayesian method to predict user preferences. Our approach is to use the variational Bayesian non-negative matrix factorization on observable rating matrices (users- items), which can predict the filling of the scoring matrix and cluster the users. On this basis, the user's hidden information and pre- rating are obtained, and the pre-rating is corrected in combination with the item attribute information by improved naive Bayes classifier. Experimental results show that 1) this method does not require additional clustering algorithms, which saves execution time. 2) Compared with the classical matrix factorization and similarity algorithm, our method solves the cold start problem of new items and the greatly improved the accuracy of recommendation.

P171 *Improving Sentence Representations with Local and Global Attention for Classification [#19780]*

Zesheng Liu, Xu Bai, Tian Cai, Chanjuan Chen, Wang Zhang and Lei Jiang, University of Chinese Academy of Sciences. Institute of Information Engineering, Chinese Academy of Sciences, China; Institute of Information Engineering, Chinese Academy of Sciences, China; China National Machinery Industry Corporation, China

Representation learning is a key issue for text classification tasks. Few existing representation models are able to learn sufficient text information, including local semantic information and global structure information. This paper focuses on how to generate better semantic and structure representations to obtain better sentence representation with them. In detail, we propose a hierarchical local and global attention network to learn sentence representation automatically. We generate semantic and structure representations respectively with local attention. Global attention is used to get the final representation. The final representation obtained is used for training and prediction. Experimental results show that our method achieves ideal results in several text classification tasks, including sentiment analysis, subjectivity classification and question type classification. The specific accuracies are 81.6%(MR), 93.6%(SUBJ), 49.4%(SST-5) and 95.6%(TREC).

P172 *EEG-Based Motor Imagery Classification with Deep Multi-Task Learning [#19781]*

Yaguang Song, Danli Wang, Kang Yue, Nan Zheng and Zuo-Jun Shen, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China; University of California, Berkeley, United States

In the past decade, Electroencephalogram (EEG) has been applied in many fields, such as Motor Imagery (MI) and Emotion Recognition. Traditionally, for classification tasks based on EEG, researchers would extract features from raw signals manually which is often time consuming and requires adequate domain knowledge. Besides that, features manually extracted and selected may not generalize well due to the limitation of human. Convolutional Neural Networks (CNNs) plays an important role in the wave of deep learning and achieve amazing results in many areas. One of the most attractive features of deep learning for EEG-based tasks is the end-to-end learning. Features are learned from raw signals automatically and the feature extractor and classifier are optimized simultaneously. There are some researchers applying deep learning methods to EEG analysis and achieving promising performances. However, supervised deep learning methods often requires large-scale annotated dataset, which is almost impossible to acquire in EEG-based tasks. This problem limits the further improvements of deep learning models for classification based on EEG. In this paper, we propose a novel deep learning method DMTL-BCI based on Multi-Task Learning framework for EEG-based classification task. The proposed model consists of three modules, the representation module, the reconstruction module and the classification module. Our model is proposed to improve the classification performance with limited EEG data. Experimental results on benchmark dataset, BCI Competition IV dataset 2a, show that our proposed method outperforms the state-of-the-art method by 3.3%, which demonstrates the effectiveness of our model.

P173 *Scene Recognition via Object-to-Scene Class Conversion: End-to-End Training [#19788]*

Hongje Seong, Junhyuk Hyun, Hyunbae Chang, Suhyeon Lee, Suhan Woo and Euntai Kim, Yonsei University, Korea (South)

When a person recognize the scene of an image, contextual understanding from its environmental elements is necessary. These environmental elements are variant and require comprehensive understanding of various situations. Especially, objects are frequently used as environmental elements related with scene. In this paper, we suggest a score level Class Conversion Matrix (CCM) for scene recognition with a great focus on relationship between objects and scene. A lot of existing methods have already build scene recognition systems with consideration of close relationship between object and scenes. However, most of these methods are using the object features directly without any conversions or reconstructions, and it lack confirmation whether these object features are helpful to recognize scenes correctly. To solve this problem, CCM, a matrix converting object feature to scene feature, is suggested. Moreover, CCM can be implemented with neural network layer and end-to-end trainable. Extensive experiments on Places 2 dataset demonstrate the effectiveness of our approach, when it is applied to the existing deep convolutional neural network architectures. The code is available at https://github.com/Hongje/Class_Conversion_Matrix-Places365

P174 *Learning "What" and "Where": An Interpretable Neural Encoding Model [#19793]*

Haibao Wang, Lijie Huang, Changde Du and Huiguang He, Research Center for Brain-Inspired Intelligence, CASIA, China

Neural encoding modeling aims to reveal how brain processes perceived information by establishing a quantitative relationship between stimuli and evoked brain activities. In the field of visual neuroscience, many studies have been dedicated to building the neural encoding model for primary visual cortex and demonstrate that the population receptive field (pRF) models can be used to explain how neurons in primary visual cortex work. However, these models rely on either the inflexible prior assumptions imposed on the spatial characteristics of pRF or the clumsy parameter estimation methods which requiring too much manual adjustment. Suffering from these issues, current methods yield dissatisfactory performance on mimicking brain activity. In this paper, we address the problems under a novel "what and where" neural encoding framework. Basing on deep neural network (DNN) and the separability of the spatial ("where") and visual feature ("what") dimensions, the proposed method is not only powerful in extracting nonlinear features from images, but also rich in interpretability. Owing to two forms of regularization: sparsity and smoothness, receptive fields are estimated automatically for each voxel without prior assumptions on shape, which gets rid of the shortcomings of previous methods. Extensive empirical evaluations on publicly available fMRI dataset show that the proposed method has superior performance gains over several existing methods.

P175 *FSC-CapsNet: Fractionally-Strided Convolutional Capsule Network for complex data [#19799]*

Jian-wei Liu, Feng Gao, Run-kun Lu, Yuan-feng Lian, Dian-zhong Wang, Xiong-lin Luo and Chu-ran Wang, Department of Automation China University of Petroleum Beijing, Beijing, China, China; Department of Automation, China University of Petroleum , Beijing Campus (CUP), China; College of Information Science and Engineering, China University of Petroleum, Beijing Campus (CUP), China; Beijing Institute of Space Mechanics & Electricity, China; Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China, China

Recently, a novel neural network called CapsNet has attracted the attention of many researchers. It is a great attempt to overcome the drawback of convolutional neural networks (CNNs) and achieves state-of-the-art performance on some simple datasets like MNIST. However, this network architecture is built specifically for MNIST and gets a poor performance on more complex datasets like CIFAR-10. To address this problem, aiming at complex data, we propose a new CapsNet architecture called Fractionally-Strided Convolutional Capsule Network (FSC-CapsNet). We modify both network structures of the encoder and decoder of CapsNet. For the purpose of extracting better features, we increase the number of convolutional layers before capsule layer in the encoder and improve the reconstruction performance by adopting two fractionally-strided convolutional layers in the decoder. In addition, no pooling layers are used in our architecture. To assess the performance of our proposed network on complex data, we conduct experiments with a single model without using any ensembled methods and data augmentation techniques on five real-world datasets, which are of higher dimensionality and larger size than MNIST. The experimental results demonstrate that our proposed method achieves better performance and improves the reconstruction performance compared with the normal CapsNet.

P176 *A New Knowledge Distillation for Incremental Object Detection [#19804]*

Li Chen, Chunyan Yu and Lvcai Chen, Fuzhou University, China

Nowadays, the Convolutional Neural Network is successfully applied to the images object detection. When new classes of object emerges, it is popular to adapt the convolutional neural network based detection model through a retraining process with the new classes of samples. Unfortunately, the adapted model can only detect the new classes of objects, but cannot identify the old classes of objects, which is called catastrophic forgetting, also occurring in incremental classification tasks. Knowledge distillation has achieved good results in incremental learning for classification tasks. Due to the dual tasks within object detection, object classification and location at the same time, a straightforward migration of knowledge distillation method cannot provide a satisfactory result in incremental learning for object detection tasks. Hence, this paper propose a new knowledge distillation for incremental object detection, which introduces a new object detection distillation loss, a loss not only for classification results but also for location results of the predicted bounding boxes, not only for all final detected regions of interest but also for all intermediate regions proposal. Furthermore, to avoid forgetting learned knowledge from old datasets, this paper not only employs hint learning to retain the characteristic information of the initial model, but also innovatively uses confidence loss to extract the confidence information of the initial model. A series of experiment results on the PASCAL VOC 2007 dataset verify the effectiveness of the proposed method.

P177 *Evaluation of Heart Disease Diagnosis Approach using ECG Images [#19810]*

Marcos Aurelio A. Ferreira Junior, Mateus Valentim Gurgel, Leandro B. Marinho, Navar Medeiros M. Nascimento, Suane Pires. P. da Silva, Shara Shami A. Alves, Geraldo Luis Bezerra Ramalho and Pedro Pedrosa Reboucas Filho, Instituto Federal do Ceara, Brazil; Federal University of Ceara, Brazil

Among illnesses, heart diseases are accounted for as one of the most responsible for deaths. Precise and fast diagnoses increase the patient's chances to receive treatment time. A non- invasive and low-cost way to diagnose it is by using Electrocardiogram (ECG). In this paper, we propose a way to diagnosis two types of heart arrhythmia, by using the ECG record as an image. To access the performance of our system, five feature extraction methods well-known in literature are used along with five different classifiers are tested. We were able to identify heart disorders with over 96.00% of accuracy, using a vanilla neural-network, Multilayer Perceptron (MLP), and Local Binary Patterns (LBP) from ECG images. This investigation has shown promising results from a medical point-of-view.

P178 *Multimodal Data Enhanced Representation Learning for Knowledge Graphs [#19826]*

Zikang Wang, Linjing Li, Qiudan Li and Daniel Zeng, The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences; School of Computer and Control Engineering, University of Chinese Academy of Sciences, China; The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China

Knowledge graph, or knowledge base, plays an important role in a variety of applications in the field of artificial intelligence. In both research and application of knowledge graph, knowledge representation learning is one of the fundamental tasks. Existing representation learning approaches are mainly based on structural knowledge between entities and relations, while knowledge among entities per se is largely ignored. Though a few approaches integrated entity knowledge while learning representations, these methods lack the flexibility to apply to multimodalities. To tackle this problem, in this paper, we propose a new representation learning method, TransAE, by combining multimodal autoencoder with TransE model, where TransE is a simple and effective representation learning method for knowledge graphs. In TransAE, the hidden layer of autoencoder is used as the representation of entities in the TransE model, thus it encodes not only the structural knowledge, but also the multimodal knowledge, such as visual and textural knowledge, into the final representation. Compared with traditional methods based on only structural knowledge, TransAE can significantly improve the performance in the sense of link prediction and triplet classification. Also, TransAE has the ability to learn representations for entities out of knowledge base in zero-shot. Experiments on various tasks demonstrate the effectiveness of our proposed TransAE method.

P179 *Integrating Dual User Network Embedding with Matrix Factorization for Social Recommender Systems [#19828]*

Liying Chen, Honglei Zhang and Jun Wu, Beijing Jiaotong University, China

To address the data sparsity problem faced by recommender systems, social network among users is often utilized to complement rating data for improving the recommendation performance. One of current trends is to combine the idea of matrix factorization (MF) for predicting ratings with the idea of graph embedding (GE) for analyzing social network towards recommendation tasks. Despite enjoying many advantages, the existing integrated models have two critical limitations. First, such models are designed to work with either explicit or implicit social network, but little is

known in taking both into account. Second, the users' embeddings learned by GE are fed to the downstream MF, but not reverse, which is sub-optimal because rating information is not considered for learning the users' embeddings. In this paper, we propose a novel social recommendation algorithm which exploits both explicit and implicit social networks towards the task of rating prediction. In Particular, we seamlessly integrate MF model and GE model within a unified optimization framework, in which MF and GE tasks can be reinforced each other during the learning process. Our encouraging experimental results on three real-world benchmarks validate the superiority of the proposed approach to state-of-the-art social recommendation methods.

P180 *View-Invariant Gait Recognition Based on Deterministic Learning and Knowledge Fusion [#19836]*

Muqing Deng, Haonan Yang, Jiuwen Cao and Xiaoreng Feng, The Chinese University of Hong Kong, Hong Kong; Hangzhou Dianzi University, China; The University of Hong Kong, Hong Kong

Deformation of gait silhouettes caused by different view angles heavily affects the performance of gait recognition. In this paper, a new method based on deterministic learning and knowledge fusion is proposed to eliminate the effect of view angle for efficient view-invariant gait recognition. First, the binarized walking silhouettes are characterized with three kinds of time-varying width parameters. The nonlinear dynamics underlying different individuals' width parameters is effectively approximated by radial basis function (RBF) neural networks through deterministic learning algorithm. The extracted gait dynamics captures the spatio-temporal characteristics of human walking, represents the dynamics of gait motion, and is shown to be insensitive to the variance across various view angles. The learned knowledge of gait dynamics is stored in constant RBF networks and used as the gait pattern. Second, in order to handle the problem of view change no matter the variation is small or large, the learned knowledge of gait dynamics from different views is fused by constructing a deep convolutional and recurrent neural network (CRNN) model for later human identification task. This knowledge fusion strategy can take advantage of the encoded local characteristics extracted from the CNN and the long-term dependencies captured by the RNN. Experimental results show that promising recognition accuracy can be achieved.

P181 *Deeper Monocular Depth Prediction via Long and Short Skip Connection [#19847]*

Zhaokai Wang, Limin Xiao, Rongbin Xu, Shubin Su, Shupan Li and Song Yao, Beihang University, China

This paper presents a fully convolutional neural network to tackle the mapping between single view RGB images and depth maps. To regress the depth maps from monocular images,we leverage deep short skip connections in residual learning for extracting features rather than using hand-crafted features. We further propose long skip connections in up-sampling stage to reuse the feature maps which is proved to enhance the result experimentally. To show the impact of loss functions in monocular depth map predictions, we train our model with kind of loss functions and compare the results qualitatively and quantitatively. The proposed model outperforms all current state-of- the-art results with less training data as well as less than half of training epochs in two standard benchmark data sets without any post-processing procedures or other refinement steps.

P182 *Recurrent Layer Aggregation using LSTM [#19852]*

Yu Qin, Jiajun Du, Xinyao Wang and Hongtao Lu, Shanghai JiaoTong University, China

Standard convolutional neural networks assemble multiple convolutional layers to extract high-level features. Recent efforts keep designing deeper and wider architectures. Even with skip connections applied to combine different layers, the useful low-level features are not effectively utilized. Some deep layer aggregation methods have been proposed to aggregate features of all levels, using a simple linear combination or complex non-linear transformation. In this paper, we treat convolutional features as a sequence, and propose our Recurrent Aggregation of Convolutional Neural Network (CNN-RA). Our aggregation method splits a standard CNN into blocks and maps their feature matrices to a sequence of vectors of the same length. LSTM is employed to connect to the sequence and better fuse features across blocks. Our proposed CNN-RA can be directly appended to any standard CNN without any modifications. Experiments show remarkable improvements of CNN-RA over the original architectures across datasets.

P183 *Recurrent Network and Multi-arm Bandit Methods for Multi-task Learning without Task Specification [#19012]*

Thy Nguyen and Tayo Obafemi-Ajayi, Missouri State University, United States

This paper addresses the problem of multi-task learning (MTL) in settings where the task assignment is not known. We propose two mechanisms for the problem of inference of task's parameter without task specification: parameter adaptation and parameter selection methods. In parameter adaptation, the model's parameter is iteratively updated using a recurrent neural network (RNN) learner as the mechanism to adapt to different tasks. For the parameter selection model, a parameter matrix is learned beforehand with the task known apriori. During testing, a bandit algorithm is utilized to determine the appropriate parameter vector for the model on the fly. We explored two different scenarios in MTL without task specification, continuous learning and reset learning. In continuous learning, the model has to adjust its parameter continuously to a number of different task without knowing when task changes. Whereas in reset learning, the parameter is reset to an initial value to aid transition to different tasks. Results on three real benchmark datasets demonstrate the comparative performance of both models with respect to multiple RNN configurations, MTL algorithms and bandit selection policies.

S25: Artificial Intelligence in Health and Medicine: from Theory to Applications

Thursday, July 18, 8:00AM-9:40AM, Room: Duna Salon I, Chair: Hissam Tawfik

8:00AM *Neural Networks for Lung Cancer Detection through Radiomic Features [#19520]*

Luca Brunese, Francesco Mercaldo, Alfonso Reginelli and Antonella Santone, University of Molise, Italy; IIT-CNR, Italy; University of Campania, Italy

Lung cancer is one of primary premature death causes. As a matter of fact, people die more for lung cancer than prostate, colon or breast. To assist medics and radiologists in the diagnosis formulation, in this paper we propose a neural network-based method aimed to discriminate between different lung cancer types. We exploit a set of 30 radiomic feature directly obtained from magnetic resonance, tuning the neural network model when the momentum and the loss functions are varying with the aim to find the best model in terms of features and network parameters. We evaluate the effectiveness of the proposed method on a dataset of 2000 MRI labelled through medical reports, obtaining a precision equal to 0.918 and a recall equal to 0.923 in T1a lung cancer detection while a precision equal to 0.931 and a recall equal to 0.918 is obtained in T2b lung cancer detection.

8:20AM *An Object Detection by using Adaptive Structural Learning of Deep Belief Network [#19594]*

Shin Kamada and Takumi Ichimura, Hiroshima City University, Japan; Prefectural University of Hiroshima, Japan

Deep learning forms a hierarchical network structure for representation of multiple input features. The adaptive structural learning method of Deep Belief Network (DBN) can realize a high classification capability while

searching the optimal network structure during the training. The method can find the optimal number of hidden neurons for given input data in a Restricted Boltzmann Machine (RBM) by neuron generation-annihilation algorithm. Moreover, it can generate a new hidden layer in DBN by the layer generation algorithm to actualize a deep data representation. The proposed method showed higher classification accuracy for image benchmark data sets than several deep learning methods including well-known CNN methods. In this paper, a new object detection method for the DBN architecture is proposed for localization and category of objects. The method is a task for finding semantic objects in images as Bounding Box (B-Box). To investigate the effectiveness of the proposed method, the adaptive structural learning of DBN and the object detection were evaluated on the Chest X- ray image benchmark data set (CXR8), which is one of the most commonly accessible radio- logical examination for many lung diseases. The proposed method showed higher performance for both classification (more than 94.5% classification for test data) and localization (more than 90.4% detection for test data) than the other CNN methods.

8:40AM *Machine Learning to Identify Gender via Hair Elements [#19518]*

Pasquale Avino, Francesco Mercaldo, Vittoria Nardone, Ivan Notardonato and Antonella Santone, University of Molise, Italy; IIT-CNR, Italy; University of Sannio, Italy

Currently, the gender is mainly inferred through bone and dental analyses. However, such sort of analysis is useless when bones are not available. Hair is a stable substance that, depending on its length, is capable of retaining years of information. In this paper we propose a machine learning based approach aimed to identify gender through hair elements. Preliminary results have even indicated that the method is promising also for forensics analysis since it is able to identify the gender with a high accuracy.

9:00AM *Heartbeat Anomaly Detection using Adversarial Oversampling [#20112]*

Jefferson Lima, David Macedo and Cleber Zanchettin, Centro de Informatica - Universidade Federal de Pernambuco, Brazil

Cardiovascular diseases are one of the most common causes of death in the world. Prevention, knowledge of previous cases in the family, and early detection is the best strategy to reduce this fact. Different machine learning approaches to automatic diagnostic are being proposed to this task. As in most health problems, the imbalance between examples and classes is predominant in this problem and affects the performance of the automated solution. In this paper, we address the classification of heartbeats images in different cardiovascular diseases. We propose a two-dimensional Convolutional Neural Network for classification after using a InfoGAN architecture for generating synthetic images to unbalanced classes. We call this proposal Adversarial Oversampling and compare it with the classical oversampling methods as SMOTE, ADASYN, and RandomOversampling. The results show that the proposed approach improves the classifier performance for the minority classes without harming the performance in the balanced classes.

9:20AM *Development of a Simulation Experiment to Investigate In-Flight Startle using Fuzzy Cognitive Maps and Pupillometry [#20521]*

Abiodun Brimmo Yusuf, Ah-Lian Kor and Hissam Tawfik, Leeds Beckett University, United Kingdom

Loss of control in-flight (LOC-I), following loss of situational awareness and startle has been identified as a leading cause of aviation-based fatalities in recent decades. This has led to significant effort toward improving safety records; particularly in the fields of flight crew training; and in-flight support technologies that aid better decision making and management of reactions to a startling occurrence. One way to achieve quality decision making in the cockpit is by providing adequate cueing and response activating mechanisms carefully designed to aid human information processing. These response performances, especially in the context of reactionary management of startle in flight, could be honed through simulator-based training. This paper discusses the setup of such a simulation experiment and the methods driving the work.

S29: Biologically Inspired Learning for Cognitive Robotics

Thursday, July 18, 8:00AM-9:40AM, Room: Duna Salon II, Chair: Artur Istvan Karoly

8:00AM *OCSVM-based Evaluation Method for Generative Neural Networks [#19426]*

Artur Istvan Karoly, Marta Takacs and Peter Galambos, Obuda University, Hungary

Humanity has desired to create machines that can comprehend fine arts for a long while. Recently, several articles report neural network models that managed to create something which can be considered as art, such as paintings or music. Naturally, these proposals are also leveraged in the field of robotics and autonomous vehicles for tasks such as training environment generation and efficient exploration in reinforcement learning. The common feature of all these approaches is the utilization of generative neural networks. These generative models are a trendy and actively researched area of machine learning methods. However, it is hard to find a general and objective metric for the evaluation of such models. In this paper, a One-Class Support Vector Machine-based solution is proposed. The basic concept is demonstrated via experiments on a set of GANs. The presented method can be extended and refined in further, more exhaustive studies to serve as an alternative for the current state-of-the-art procedures.

8:20AM *Confidence Identification Based on the Combination of Verbal and Non-Verbal factors in Human Robot Interaction [#20103]*

Wei-Fen Hsieh, Youdi Li, Erina Kasano, Shimokawara Eri-Sato and Toru Yamaguchi, Tokyo Metropolitan University, Japan

Not only verbal information but also Non-verbal information is an essential factor in Human-Robot Interaction (HRI). In order to understand Human partner's perception, confidence status detection plays an important role for the robot to be capable to provide further information in the conversation. Moreover, the confident/uncertain feature can be utilized on robot expression to smooth the interaction and gain social skill. On another aspect, certain/uncertain expression style makes it possible to form extraversion/introversion personality which might spice up the interaction. This paper presented the concept of confident expression in HRI and an experiment to analyze both verbal and non-verbal features for confident expression. The confidence conditions were defined as 5-level status according to the questionnaire answer: most confident, comparative certainty, neutral, relative uncertain, most unconfident. The results showed that verbal and nonverbal features are possible to be classified into most confident/uncertain condition with the accuracy of 73.2143 % analyzed by multinomial logistic regression model via Weka. However, the ambiguous differences between the expression style are difficult to be classified. The

accuracy of predicting comparative certainty and relative uncertain status can merely reach 66.67%. The classification of comparative certainty and relative uncertainty still need more features to increase prediction accuracy.

8:40AM *Stepwise Acquisition of Dialogue Act Through Human-Robot Interaction [#20137]*
Akane Matsushima, Ryosuke Kanajiri, Yusuke Hattori, Chie Fukada and Natsuki Oka, Kyoto Institute of Technology, Japan

Dialogue act (DA) represents the meaning of an utterance at the illocutionary force level (Aus force level (Austin 1962) such as questions, requests, and greetings. Since DAs take charge of the most fundamental part of communication, we believe that the elucidation of DA learning mechanism is important for cognitive science and artificial intelligence, especially for cognitive robotics. The purpose of this study is to verify that scaffolding takes place when a human teaches a robot, and to let and to let a robot learn to estimate DAs and to make a response based on them step by step utili step utilizing scaffolding provided by a human. To realize that, it is necessary for the robot the robot to detect changes in utterance and rewards given by the partner and continue continue learning accordingly. Experimental results demonstrated that participants who continued interaction for a sufficiently long time often gave scaffolding for the ro the robot. Although the number of experiments is still insufficient to obtain a defini definite conclusion, we observed that 1) the robot quickly learned to respond to DAs in most most cases if the participants only spoke utterances that match the situation, 2) in the the case of participants who built scaffolding differently from what we assumed, learning did not proceed quickly, and 3) the robot could learn to estimate DAs almost exactly if the participants kept interaction for a sufficiently long time even if the scaffolding was unexpected.

9:00AM *Curious Meta-Controller: Adaptive Alternation between Model-Based and Model-Free Control in Deep Reinforcement Learning [#20322]*
Muhammad Burhan Hafez, Cornelius Weber, Matthias Kerzel and Stefan Wermter, University of Hamburg, Germany

Recent success in deep reinforcement learning for continuous control has been dominated by model-free approaches which, unlike model-based approaches, do not suffer from representational limitations in making assumptions about the world dynamics and model errors inevitable in complex domains. However, they require a lot of experiences compared to model-based approaches that are typically more sample-efficient. We propose to combine the benefits of the two approaches by presenting an integrated approach called Curious Meta-Controller. Our approach alternates adaptively between model-based and model-free control using a curiosity feedback based on the learning progress of a neural model of the dynamics in a learned latent space. We demonstrate that our approach can significantly improve the sample efficiency and achieve near-optimal performance on learning robotic reaching and grasping tasks from raw-pixel input in both dense and sparse reward settings.

9:20AM *Spatial Map Learning with Self-Organizing Adaptive Recurrent Incremental Network [#20187]*
Wei Hong Chin, Naoyuki Kubota, Chu Kiong Loo, Zhaojie Ju and Honghai Liu, Tokyo Metropolitan University, Japan; University of Malaya, Malaysia; University of Portsmouth, United Kingdom

Biological information inspires the advancement of a navigational mechanism for autonomous robots to help people explore and map real-world environments. However, the robot's ability to constantly acquire environmental information in real-world, dynamic environments has remained a challenge for many years. In this paper, we propose a self-organizing adaptive recurrent incremental network that models human episodic memory to learn spatiotemporal representations from novel sensory data. The proposed method termed as SOARIN consists of two main learning process that is active learning and episodic memory playback. For active learning (robot exploration), SOARIN quickly learns and adapts incoming novel sensory data as episodic neurons via competitive Hebbian Learning. Episodic neurons are connecting with each other and gradually forms a spatial map that can be used for robot localization. Episodic memory playback is triggered whenever the robot is in an inactive mode (charging or hibernating). During playback, SOARIN gradually integrates knowledge and experience into more consolidate spatial map structures that can overcome the catastrophic forgetting. The proposed method is analyzed and evaluated in term of map learning and localization through a series of real robot experiments in real-world indoor environments.

S30: Exploring Uncertainties in Big Data by Neural Fuzzy Systems

Thursday, July 18, 8:00AM-9:40AM, Room: Duna Salon III, Chair: Jie Lu

8:00AM *Unsupervised Domain Adaptation with Sphere Retracting Transformation [#19271]*
Zhen Fang, Jie Lu, Feng Liu and Guangquan Zhang, Centre for Artificial Intelligence FEIT, Univervisity of Technology Sydney, Australia

Unsupervised domain adaptation aims to leverage the knowledge in training data (source domain) to improve the performance of tasks in the remaining unlabeled data (target domain) by mitigating the effect of the distribution discrepancy. Existing approaches resolve this problem mainly by 1) mapping data into a latent space where the distribution discrepancy between two domains is reduced; or 2) reducing the domain shift by weighting the source domain. However, most of these approaches share a common issue that they neglect inter-class margins while matching distributions, which has a significant impact on classification performance. In this paper, we analyze the issue from the theoretical aspect and propose a novel unsupervised domain adaptation approach: Sphere Retracting Transformation (SRT), which reduces the distribution discrepancy and increases inter-class margins. We implement SRT, according to our theoretical analysis by (1) assigning class-specific weights for data in the source domain, and (2) minimizing the intra-class variations. Experiments confirm that the SRT approach outperforms several competitive approaches for standard domain adaptation benchmarks.

8:20AM *Cross-domain Recommendation with Semantic Correlation in Tagging Systems [#19580]*
Qian Zhang, Peng Hao, Jie Lu and Guangquan Zhang, University of Technology Sydney, Australia

The tagging system provides users with a platform to express their preferences as they annotate terms or keywords to items. Tag information is a bridge between two domains for transferring knowledge and helping to alleviate the data sparsity problem, which is a crucial and challenging problem in most recommender systems. Existing methods incorporate correlations extracted from overlapping tags at a lexical level in cross-domain recommendation, but they neglect semantical relationships between different tags, which impairs prediction accuracy in the target domain. To solve this challenging problem, we propose a cross-domain recommendation method with semantic correlation in tagging systems. This method automatically captures the semantic relationships between non- identical tags and applies them to the recommendation. The word2vec technique is used to learn the latent representations of tags. Semantically equivalent tags are then grouped to form a joint embedding space comprised of tag clusters. This embedding space serves as the bridge between domains. By mapping users and items from both the source and target domains into the same embedding space, similar users or items across domains can be identified. Thus, the recommendation in a sparse target domain is improved by transferring knowledge through correlated users and items. Experimental results with three datasets on six cross-domain recommendation tasks demonstrate that

the proposed method exploits the semantic links from tags in two domains and outperforms five benchmarks in prediction accuracy. The results indicate that transferring knowledge through tags semantics is feasible and effective.

8:40AM *A Hybrid Incremental Regression Neural Network for Uncertain Data Streams [#19129]*

Hang Yu, Jie Lu, Jialu Xu and Guangquan Zhang, University of Technology Sydney, Australia; Shanghai University, China

The design of classical regression algorithms was based on the assumption that all the required data is obtained at one time. With the emergence of big data, however, data is increasingly displayed in sequence form, such as in data streams, and can be read only once in a specific order. Many incremental regression algorithms which process data in a sequential manner have been proposed, but the accuracy of these algorithms deteriorates when the value of the data is uncertain. This paper proposes a hybrid incremental regression neural network based on self-organizing incremental neural network and incremental fuzzy support vector regression. In our proposed network, the neurons of the regression neural network are obtained by an improved self-organized incremental neural network (SOINN). This enables the regression neural network structure to self-organize as the number of neurons increases. An incremental fuzzy support vector regression (IFSVR) algorithm is then used to modify the parameters of the regression neural network. By combining the improved SOINN and IFSVR algorithms, our proposed hybrid incremental regression neural network is able to learn an accurate regression model from large uncertain data. Experiments on both artificial and real-world datasets indicate that our proposed hybrid incremental regression neural network achieves superior performance compared to other incremental regression algorithms.

9:00AM *RsyGAN: Generative Adversarial Network for Recommender Systems [#20451]*

Ruiping Yin, Kan Li, Jie Lu and Guangquan Zhang, School of Computer Science and Technology, Beijing Institute of Technology, China; Centre for Artificial Intelligence, University of Technology Sydney, Australia

Many recommender systems rely on the information of user-item interactions to generate recommendations. In real applications, the interaction matrix is usually very sparse, as a result, the model cannot be optimised stably with different initial parameters and the recommendation performance is unsatisfactory. Many works attempted to solve this problem, however, the parameters in their models may not be trained effectively due to the sparse nature of the dataset which results in a lower quality local optimum. In this paper, we propose a generative network for making user recommendations and a discriminative network to guide the training process. An adversarial training strategy is also applied to train the model. Under the guidance of a discriminative network, the generative network converges to an optimal solution and achieves better recommendation performance on a sparse dataset. We also show that the proposed method significantly improves the precision of the recommendation performance on several datasets.

Deep Learning and Applications

Thursday, July 18, 8:00AM-9:40AM, Room: Panorama I, Chair: Athanasios Davvetas

8:00AM *Evidence Transfer for Improving Clustering Tasks Using External Categorical Evidence [#19014]*

Athanasios Davvetas, Iraklis Angelos Klampanos and Vangelis Karkaletsis, National Centre for Scientific Research "Demokritos", Greece

In this paper we introduce evidence transfer for clustering, a deep learning method that can incrementally manipulate the latent representations of an autoencoder, according to external categorical evidence, in order to improve a clustering outcome. By evidence transfer we define the process by which the categorical outcome of an external, auxiliary task is exploited to improve a primary task, in this case representation learning for clustering. Our proposed method makes no assumptions regarding the categorical evidence presented, nor the structure of the latent space. We compare our method, against the baseline solution by performing k-means clustering before and after its deployment. Experiments with three different kinds of evidence show that our method effectively manipulates the latent representations when introduced with real corresponding evidence, while remaining robust when presented with low quality evidence.

8:20AM *Effortless Deep Training for Traffic Sign Detection Using Templates and Arbitrary Natural Images [#19586]*

Lucas Tabelini Torres, Thiago M. Paixao, Rodrigo F. Berriel, Alberto F. De Souza, Claudine Badue, Nicu Sebe and Thiago Oliveira-Santos, Universidade Federal do Espirito Santo, Brazil; Instituto Federal do Espirito Santo, Brazil; University of Trento, Italy

Deep learning has been successfully applied to several problems related to autonomous driving. Often, these solutions rely on large networks that require databases of real image samples of the problem (i.e., real world) for proper training. The acquisition of such real-world data sets is not always possible in the autonomous driving context, and sometimes their annotation is not feasible (e.g., takes too long or is too expensive). Moreover, in many tasks, there is an intrinsic data imbalance that most learning-based methods struggle to cope with. It turns out that traffic sign detection is a problem in which these three issues are seen altogether. In this work, we propose a novel database generation method that requires only (i) arbitrary natural images, i.e., requires no real image from the domain of interest, and (ii) templates of the traffic signs, i.e., templates synthetically created to illustrate the appearance of the category of a traffic sign. The effortlessly generated training database is shown to be effective for the training of a deep detector (such as Faster R-CNN) on German traffic signs, achieving 95.66% of mAP on average. In addition, the proposed method is able to detect traffic signs with an average precision, recall and F1-score of about 94%, 91% and 93%, respectively. The experiments surprisingly show that detectors can be trained with simple data generation methods and without problem domain data for the background, which is in the opposite direction of the common sense for deep learning.

8:40AM *A Distant Supervised Relation Extraction Model with Two Denoising Strategies [#20145]*

Zikai Zhou, Yi Cai, Jingyun Xu, Jiayuan Xie, Qing Li and Haoran Xie, South China University of Technology, China; Guangdong University of Technology, China; The Hong Kong Polytechnic University, Hong Kong; The Education University of Hong Kong, Hong Kong

Distant supervised relation extraction has been an effective way to find relational facts from text. However, distant supervised method inevitably accompanies with wrongly labeled sentences. Noisy sentences lead to poor performance of relation extraction models. Though existing piecewise convolutional neural network model with sentence-level attention (PCNN+ATT) is an effective way to reduce the effect of noisy sentences, it still has two limitations. On one hand, it adopts a PCNN module as sentence encoder, which only captures local contextual features of words and might lose important information. On the other hand, it neglects the fact that not all words contribute equally to the semantics of sentences. To address these two issues, we propose a hierarchical attention-based bidirectional GRU (HA-BiGRU) model. For the first limitation, our model utilizes a BiGRU module in place of PCNN, so as to extract global contextual information. For the second limitation, our model combines word-level and sentence-level attention mechanisms, which help get accurate sentence representations. To further

alleviate the wrongly labeling problem, we first calculate the co-occurrence probabilities (CP) between the shortest dependency path (SDP) and the relation labels. Based on these co-occurrence probabilities, two denoising strategies are proposed to reduce noise interference respectively from aspect of filtering labeled data and integrating CP information into model. Experimental results on the corpus of Freebase and New York Times (Freebase+NYT) show that the HA-BiGRU model outperforms baseline models, and the two co-occurrence probabilities based denoising strategies can improve robustness of HA-BiGRU model.

9:00AM *Multi-scale Stepwise Training Strategy of Convolutional Neural Networks for Diabetic Retinopathy Severity Assessment [#20096]*

Fangjun Li, Dongfeng Yuan, Mingqiang Zhang, Cong Liang, Xiaotian Zhou and Haixia Zhang, Shandong University, China

Diabetic retinopathy severity assessment is an important domain in which deep learning has benefited medical imaging analysis. In this regard, CNNs which perform well in ImageNet are incapable of extracting subtle lesion features from high-resolution retinal fundus images. So novel convolutional networks with higher input size were developed. But no prior work give deep investigation on the impact of image resolution in the context of DR severity assessment. In this paper, we first explore how the performance of diabetic retinopathy severity assessment task would change if higher-resolution input images were used. Next, we adopt the stepwise strategy of training convolutional networks with high input scales to avoid overfitting. Finally, rigorous analyses on the impact of image resolution are given, showing that as model expands with higher input image resolutions, the performance grows logarithmically while both time and space complexity increase exponentially. Our model obtains new state-of-the-art kappa score in the task of diabetic retinopathy severity assessment task on EyePACS dataset with convolutional networks whose input size is 896 * 896, and great progress in classification of mild diabetic retinopathy. There is great potential for generalizing this solution to other medical image analysis problems.

9:20AM *Spontaneous Facial Micro-Expression Recognition using 3D Spatiotemporal Convolutional Neural Networks [#20241]*

Sai Prasanna Teja Reddy, Surya Teja Karri, Shiv Ram Dubey and Snehasis Mukherjee, Indian Institute of Information Technology, Sri City, India

Facial expression recognition in videos is an active area of research in computer vision. However, fake facial expressions are difficult to be recognized even by humans. On the other hand, facial micro-expressions generally represent the actual emotion of a person, as it is a spontaneous reaction expressed through human face. Despite of a few attempts made for recognizing micro-expressions, still the problem is far from being a solved problem, which is depicted by the poor rate of accuracy shown by the state-of-the-art methods. A few CNN based approaches are found in the literature to recognize micro-facial expressions from still images. Whereas, a spontaneous micro-expression video contains multiple frames that have to be processed together to encode both spatial and temporal information. This paper proposes two 3D-CNN methods: MicroExpSTCNN and MicroExpFuseNet, for spontaneous facial micro-expression recognition by exploiting the spatiotemporal information in CNN framework. The MicroExpSTCNN considers the full spatial information, whereas the MicroExpFuseNet is based on the 3D- CNN feature fusion of the eyes and mouth regions. The experiments are performed over CAS(ME)^2 and SMIC micro-expression databases. The proposed MicroExpSTCNN model outperforms the state-of-the-art methods.

Applications and Data Mining

Thursday, July 18, 8:00AM-9:40AM, Room: Panorama II, Chair: Ao Feng

8:00AM *DICENet: Fine-Grained Recognition via Dilated Iterative Contextual Encoding [#20246]*

Abhishek Pal, Gautham Krishnan, Manav Moorthy, Narasimha Yadav, Adithya R Ganesh and Sree Sharmila, Sri Sivasubramaniya Nadar College of Engineering, India

Material Recognition is an intriguing problem in Computer Vision. While traditional approaches prefer an ensemble of networks to capture essential properties such as texture, more recent approaches leverage the power of Deep Learning to design end-to-end models. We do the same, and propose Dilated Iterative Contextual Encoding Network, a novel end-to-end framework for material recognition. As a result of gathering extensive knowledge on various characteristics of materials, our approach combines different components on the base network to address specific properties, which also helps in general recognition tasks. The traditional ResNet is replaced by a Dilated Residual Network to help capture fine-grained material information. Iterative deep aggregation helps capture and fuse global homogeneous material properties across multiple resolutions and scales. To enhance the discriminatory power of the learnt latent representation, we propose gramedial loss which is intuitively applied on a texture vector space. Spatial similarity loss is applied on strategic intermediate feature maps to effectively capture local non-homogeneous texture features from a global context, crucial for the primary classification task. Extensive experiments conducted on golden material datasets such as the FMD, MINC-2500, KTH-TIPS-2b, DTD and GTOS indicate improved performances over state of the art approaches on large datasets and two small datasets, while achieving compatible accuracies on the challenging FMD. Furthermore, our architecture also performed convincingly while categorizing general indoor and object classification datasets such as MIT-Indoor and CalTech-101.

8:20AM *Embeddings and Convolution, Is That the Best You can Do with Sentiment Features? [#19833]*

Ao Feng, Zhenghao Chen, Shuang Zhou and Xi Wu, Chengdu University of Information Technology, China

Rapid growth of digital media motivates research on machine-assisted text analysis. Sentiment analysis, among one of the prevalent applications, has drawn great attention. In addition to the traditional bag-of-words models, embedding methods have become de facto standard for text representation, and various convolutional, recurrent and recursive neural networks are dominating leaderboards. Despite the large number of deep learning models in publication, the performance benchmarks in sentiment analysis are approaching a limit. If language-specific syntactic and semantic knowledge is excluded, is there still room for significant improvements? Over a general neural network that is based on word embedding, 2D convolution and max-pooling, we conduct extensive experiments on its various components, including convolutional kernels, pooling methods, recurrent layers, and attention mechanism. Certain combinations show moderate improvements in classification accuracy which are comparable to more sophisticated networks, but no sign of major breakthrough is in sight. We also extend the scope with potential game changers, covering context-aware representations, linguistic information, and large scale knowledge transfer in natural languages. Reported metrics show their great value in breaking the current performance bottleneck.

8:40AM *3D Room Reconstruction from A Single Fisheye Image [#19993]*

Mingyang Li, Yi Zhou, Ming Meng, Yuehua Wang and Zhong Zhou, Beihang University, China; Bigview Technology Co. Ltd., China; Texas A&M University-Commerce, United States

We propose a rapid and accurate approach to recover the layout of a room automatically from a single fisheye image. It decomposes the fisheye image

to a set of perspective images and jointly extract line images from the fisheye image and perspective images for geometric information. The semantic information gained from semantic segmentation on a cylinder expansion of the fisheye image are then used for structure line determination. By considering distinct features contained in the perspective images, the invalid hypotheses are filtered effectively and the most accurate structure lines are selected to minimize computational cost. To evaluate the effectiveness of the proposed approach, we construct an annotated fisheye image dataset. Comprehensive experimental evaluation on the dataset illustrate that our proposed approach produces higher quality layout estimations than existing layout reconstruction approaches and being 6 times faster in the reconstruction time.

9:00AM *Incorporating Human Knowledge in Neural Relation Extraction with Reinforcement Learning [#19409]*

Bing Liu, Guilin Qi, Lu Pan, Shangfu Duan and Tianxing Wu, Southeast University, China; Baidu Inc., China; Nanyang Technological University, Singapore

Relation Extraction (RE) aims at extracting semantic relation of entities from text and it is a crucial task in natural language processing. Deep neural network (DNN) based models have achieved excellent performance in RE. However, they still have several problems remaining to be addressed: (1) humans can hardly take measures to amend the DNN-based RE systems because it is difficult to encode human intention to guide them to capture desired patterns. (2) DNN-based RE models may suffer from not having sufficient background information for making predictions. To handle these issues, we propose an RE framework based on reinforcement learning, which can enhance existing DNN-based RE models by incorporating human knowledge including soft rules and relation evidence. The introduction of soft rules enable human to impose an effect on the RE result and correct the RE system, while the relation evidence help supplement the background information without limiting its types and sources. The experimental results show that our approach can reinforce existing DNN-based RE models effectively and outperforms state-of-the-art RE methods.

9:20AM *Knowledge Adaptive Neural Network for Natural Language Inference [#19930]*

Zhang Qi, Yang Yan, Chen Chengcai, He Liang and Yu Zhou, Department of Computer Science and Technology, East China Normal University, China; Xiaoi Robot Technology Co., Ltd, China; Computer Science Department, University of California, Davis, United States

Natural language inference (NLI) has received widespread attention in recent years due to its contribution to various natural language processing tasks, such as question answering, abstract text summarization, and video caption. Most existing works focus on modeling the sentence interaction information, while the use of commonsense knowledge is not well studied for NLI. In this paper, we propose knowledge adaptive neural network (KANN) that adaptively incorporates commonsense knowledge at sentence encoding and inference stages. We first perform knowledge collection and representation to identify the relevant knowledge. Then we use a knowledge absorption gate to embed knowledge into neural network models. Experiments on two benchmark datasets, namely SNLI and MultiNLI for natural language inference, show the advantages of our proposed model. Furthermore, our model is comparable to if not better than the recent neural network based approaches on NLI.

Extreme Learning Machines (ELM) and Machine Learning

Thursday, July 18, 8:00AM-9:40AM, Room: Panorama III, Chair: WeiZhong Yan

8:00AM *Continuous Modeling of Power Plant Performance with Regularized Extreme Learning Machine [#19540]*

Rui Xu and WeiZhong Yan, GE Global Research, United States

Power plant modeling is critically important for power plant operation optimization and cost reduction. The inherently nonstationary characteristics of power plants raise a big challenge to the learning mechanisms and require the learning algorithms to adapt effectively and promptly to the continuously drifting environments. In our previous study, we proposed an online ensemble regression approach, with extreme learning machine (ELM) as the base model, to model power plant performance in a dynamic environment, which can autonomously update models to respond to environmental changes, either gradual or abrupt. However, one drawback we observed for the proposed approach is that the algorithm performance is not stable due to the randomness nature of ELMs. In this paper, we address this issue by applying regularized ELM as the base model within the online ensemble framework. The empirical results on three real power plant data sets demonstrate that the proposed modification can lead to more stable generalization performance of the algorithm. At the same time, the algorithm consistently achieves performance with mean average percentage error less than the required 1% threshold in real field operations.

8:20AM *Semi-Supervised Online Elastic Extreme Learning Machine with Forgetting Parameter to deal with concept drift in data streams [#20125]*

Carlos Silva and Renato Krohling, Federal University of Espirito Santo, Brazil

Concept drift is a common problem for online sequential algorithms that deal with data streams. Many supervised and unsupervised approaches to solve concept drifts were proposed recently, including some ELM-based algorithms. Due to its fast training, ELM-based algorithms can quickly adapt to dataset changes, detecting and preventing concept drift. SSOE-ELM is a semi-supervised online ELM-based algorithm with good accuracy and generalization ability, but as an online algorithm it is also affected by the concept drift problem. In this paper, a variation of SSOE-ELM algorithm with a semi-supervised concept drift detector and a forgetting parameter called SSOE-FP-ELM is proposed. This new approach is compared with standard SSOE-ELM and FP-ELM. Our experimental results show that SSOE-FP-ELM outperforms SSOE-ELM and FP-ELM in accuracy with two different concept drift types, without a considerable increase in training time.

8:40AM *A Hardware/Software Extreme Learning Machine Solution for Improved Ride Comfort in Automobiles [#20134]*

Oscar Mata-Carballeira, Ines del Campo, Victoria Martinez and Javier Echanobe, University of the Basque Country (UPV/EHU), Spain

Automotive ride comfort has become an important research topic in recent years due to the increasing level of automation in currently produced cars. These premises also apply to manned cars. In this work, a hybrid hardware/software extreme learning machine for improved ride comfort in automobiles is proposed. This system is based on a single-chip implementation able to provide real-time information about the level of ride comfort by classifying driving data into several comfort classes. To develop this system, unsupervised hierarchical clustering analysis (HCA) and supervised extreme learning machine (ELM) have been used jointly, to enhance the overall performance of the entire system, reaching classification success rates of up to 95%. This approach has been implemented on a Xilinx Zynq-7000 programmable system-on-chip. This chip is able to process data in real time and to identify the comfort class, achieving low latency marks and high operational frequencies due to its DSP-based implementation. These performance and accuracy marks, together with its low power consumption make this development suitable for novel practical implementations in current production cars.

9:00AM *Informative Instance Detection for Active Learning on Imbalanced Data [#19236]*

Xu Jian, Wang Xinyue, Cai Zixin, Yang Liu and Jing Liping, Beijing Jiaotong University, China; TianJin University, China

In imbalanced data classification, it is hard to learn the hidden pattern from the minority class due to its insufficient information. To solve this problem, a popular type of sampling methods is proposed based on Active Learning framework, but they still suffer from two key issues: how to keep the structure of the original data and avoid imbalance during the learning process. In this paper, we proposed a novel Active Learning framework (COAL) to select and generate informative instances. To keep the structure and enhance the diversity of the original data, we make use of Clustering-based uncertainty sampling to find informative instances. Meanwhile, to avoid imbalance in the active learning process, we make use of Oversampling method to balance the quantities between classes. Extensive experiments have been conducted by using real world datasets with a large range of imbalance ratio (from 2.78 to 66.67). The experimental results show that the proposed COAL outperforms state-of-the- art methods in terms of several well-known evaluation metrics.

9:20AM *Evolutionary Neural Architecture Search for Image Restoration [#19238]*

Gerard Jacques van Wyk and Anna Sergeevna Bosman, University of Pretoria, South Africa

Convolutional neural network (CNN) architectures have traditionally been explored by human experts in a manual search process that is time-consuming and ineffectively explores the massive space of potential solutions. Neural architecture search (NAS) methods automatically search the space of neural network hyperparameters in order to find optimal task-specific architectures. NAS methods have discovered CNN architectures that achieve state-of-the-art performance in image classification among other tasks, however the application of NAS to image-to-image regression problems such as image restoration is sparse. This paper proposes a NAS method that performs computationally efficient evolutionary search of a minimally constrained network architecture search space. The performance of architectures discovered by the proposed method is evaluated on a variety of image restoration tasks applied to the ImageNet64x64 dataset, and compared with human-engineered CNN architectures. The best neural architectures discovered using only 2 GPU-hours of evolutionary search exhibit comparable performance to the human-engineered baseline architecture.

S17: Biologically Inspired Computational Vision and S19: Ensemble Learning and Applications

Thursday, July 18, 8:00AM-9:40AM, Room: Panorama IV, Chair: Khan Iftekharuddin

8:00AM *3D Skeleton Estimation and Human Identity Recognition Using Lidar Full Motion Video [#20332]*

Alexander Glandon, Lasitha Vidyaratne, Nasrin Sadeghzadehyazdi, Nibir Dhar, Jide Familoni, Scott Acton and Khan Iftekharuddin, Old Dominion University, United States; University of Virginia, United States; Army NVESD, United States

This work proposes a novel computational modeling to estimate 3D dense skeleton and corresponding joint locations from Lidar (light detection and ranging) full motion video (FMV). Unlike motion capture (MoCap) video, where body mounted reflectors are used to capture 3D skeleton in a controlled research environment, the proposed model obtains full 3D dense skeleton in Lidar FMV. Our proposed method extracts 3D pose for subjects with walking motion from the 3D dense joints. The second contribution involves extraction of silhouette-based features and augmentation of the pose features with silhouette-based features generated over small windows of the video for human subject identification. We evaluate our model with a 10-person in-house Lidar FMV dataset and the proposed method offers 91.69% of cross-validated accuracy using a support vector machine (SVM) classification. For comparison, we implement transfer learning for another well-known deep learning-based human identification method, OpenPose, using the same Lidar FMV. The fully tuned OpenPose offers 85.00% cross-validated identification rate using the same dataset. The comparison suggests that the proposed computational modeling offers better human identification performance when compared to OpenPose transfer learning method using the 10-person Lidar FMV.

8:20AM *Adaptive Random Forests with Resampling for Imbalanced data Streams [#20476]*

Luis Eduardo Boiko Ferreira, Heitor Murilo Gomes, Albert Bifet and Luiz Eduardo Soares Oliveira, Federal University of Parana, Brazil; Telecom Paristech, France

The large volume of data generated by computer networks, smartphones, wearables and a wide range of sensors, which produce real-time data, are only useful if they can be efficiently processed so that individuals can make timely decisions based on them. In this context, machine learning techniques are widely used. While it performs better than humans in such tasks, every machine learning algorithm has a certain intrinsic bias, which means they assume that the data have specific characteristics, such as having a balanced distribution between classes. As many real-world applications present imbalanced traits in their data, this topic is gaining repercussion over time. In this work, we present the Adaptive Random Forest with Resampling (ARF_RE), which is a classifier designed to deal with imbalanced datasets. ARF_RE resample the instances based on the current class label distribution. We show through a set of extensive experiments on seven datasets that the proposed method can considerably improve the performance of the minority class(es) while avoiding degrading the performance in the majority class. On top of that, ARF_RE is more efficient regarding execution time in comparison to the standard ARF algorithm.

8:40AM *On Evaluating the Online Local Pool Generation Method for Imbalance Learning [#19443]*

Mariana A. Souza, George D. C. Cavalcanti, Rafael M. O. Cruz and Robert Sabourin, University of Quebec, Canada; Federal University of Pernambuco, Brazil; Stradigi AI, Canada

Imbalanced problems are characterized by a disproportion between the number of samples from the classes in a classification problem. This difference in amount of examples may lead to a bias toward the majority class, hindering the recognition of the minority class. Ensemble methods have been widely used for dealing with such problems. In this context, Dynamic Selection (DS) approaches, which perform the classification task on a local level, have been receiving some attention for their promising results. A DS-based approach proposed in a previous work, an online local pool generation method, generates on the fly locally accurate classifiers for labelling samples near the class borders. Though the local generation of the classifiers may reduce the impact of class imbalance on the performance of the technique, its suitability for imbalance learning was not yet evaluated. Thus, in this work we evaluate how well the online local pool generation method deals with imbalanced problems. We perform a comparative analysis with a baseline technique using three DS techniques over 64 imbalanced datasets and four performance measures. We also evaluate the use of a preprocessing and a balanced neighborhood definition steps on the online scheme to assess their impact on the performance. Moreover, we evaluate the online technique and its variants against seven state-of-the-art ensemble methods. Experimental results show that the approach of locally generating the classifiers is advantageous for imbalance learning, yielding state-of- the-art results. Furthermore, the addition of the noise removal and the balanced neighborhood definition steps improved the results, indicating the advantage of including such steps in DS-based techniques.

9:00AM *Vertical and Horizontal Partitioning in Data Stream Regression Ensembles [#19619]*

Jean Paul Barddal, PPGIa - Pontificia Universidade Catolica do Parana, Brazil

Data stream mining is an emerging topic in machine learning that targets the creation and update of predictive models over time as new data becomes available. Regarding existing works, classification is the most widely tackled task, which leaves regression nearly untouched. In this paper, the focus relies on ensemble learning for data stream regression, more specifically on vertical and horizontal data partitioning techniques. The goal is to determine whether and under which conditions partitioning can lessen the error rates of different types of learners in the data stream regression task. The proposed method combines vertical and horizontal partitioning, and it is compared with and against different types of learners and existing ensembles.

9:20AM *Evaluating Competence Measures for Dynamic Regressor Selection [#19604]*

Thiago J. M. Moura, George D. C. Cavalcanti and Luiz S. Oliveira, IFPB, Brazil; CIn - UFPE, Brazil; DInf - UFPR, Brazil

Dynamic regressor selection (DRS) systems work by selecting the most competent regressors from an ensemble to estimate the target value of a given test pattern. This competence is usually quantified using the performance of the regressors in local regions of the feature space around the test pattern. However, choosing the best measure to calculate the level of competence correctly is not straightforward. The literature of dynamic classifier selection presents a wide variety of competence measures, which cannot be used or adapted for DRS. In this paper, we review eight measures used with regression problems, and adapt them to test the performance of the DRS algorithms found in the literature. Such measures are extracted from a local region of the feature space around the test pattern, called region of competence, therefore competence measures. To better compare the competence measures, we perform a set of comprehensive experiments of 15 regression datasets. Three DRS systems were compared against individual regressor and static systems that use the Mean and the Median to combine the outputs of the regressors from the ensemble. The DRS systems were assessed varying the competence measures. Our results show that DRS systems outperform individual regressors and static systems but the choice of the competence measure is problem- dependent.

8: Other Applications

Thursday, July 18, 8:00AM-9:40AM, Room: Panorama V, Chair: Francesco Carlo Morabito

8:00AM *Analysis of Two Various Approaches for Attributes Classification Based on User-Submitted Photos [#19641]*

Wendy Damar Wisma Trisna Bayu, May Iffah Rizki, Lintang Matahari Hasani, Valian Fil Ahli, Ari Wibisono and Petrus Mursanto, Universitas Indonesia, Indonesia

There are some challenges in processing multilabel big data, namely the enormous size that may affect the computing time and the multilabel nature of the data which may further complicate the process. In a quest of exploring the right approaches to resolve such challenges, we experimented with two different big data classification approaches, which are the two- steps approach and the three-steps approach. The two-steps approach focuses on the classification of attributes of individual restaurant images as a basis for determining the attributes of a restaurant from calculating the score averages of each image labels. On the other hand, the three-steps approach focuses on the classification of restaurant attributes based on its photos' features average scores. Such approaches were tested in order to find out the different outcomes. The classifications were conducted on a dataset, which size reaches up to 13 gigabytes, consisting of 234,841 user- submitted restaurant photos from a crowdsourced restaurant reviews website. We found that the approaches produced different outcomes which have different applicability when those are intended to be implemented in a crowdsourced review site. Moreover, the two-steps approach has lower F-1 score, precision, and recall average score than three-steps approaches.

8:20AM *Synthetic Lung Nodule 3D Image Generation Using Autoencoders [#20009]*

Steve Kommrusch and Louis-Noel Pouchet, Colorado State University, United States

One of the challenges of using machine learning techniques with medical data is the frequent dearth of source image data on which to train. A representative example is automated lung cancer diagnosis, where nodule images need to be classified as suspicious or benign. In this work we propose an automatic synthetic lung nodule image generator. Our 3D shape generator is designed to augment the variety of 3D images. Our proposed system takes root in autoencoder techniques, and we provide extensive experimental characterization that demonstrates its ability to produce quality synthetic images.

8:40AM *Eye Gesture Based Communication for People with Motor Disabilities in Developing Nations [#19315]*

Sharan Pai and Anish Bhardwaj, IIIT Delhi, India

Current eye tracking systems for people with motor disabilities are expensive, restricted to only four eye movements, require constant re-calibration, assume literacy in English and offer no easy method to contact people during emergencies. Even though the above systems are flawed, they are crucial for the aforementioned individuals to communicate. Despite this, our surveys show that no such systems are in use for people in developing countries and there is a heavy dependence on simple yes or no answers. To overcome these obstacles we created E-ACE (Eye-based Alternative Communication Exchange): A smartphone-based system that allows the user to utilize nine eye movements rather than the traditional four. Our surveys show that many people with motor disabilities have difficulties in performing the conventional four eye movements. To counter this, E-ACE allows each user to create a personalized system which consists of their preferred four eye movements from a set of nine (four traditional, five new). To overcome the language barrier that many people face while using the current systems, E-ACE provides easy to understand pictures for a list of relevant English words and phrases which was curated from the doctors we surveyed. E- ACE also lets the user access an SOS Board that can call, message and alert the concerned people with just two eye movements. Thus, E-ACE offers a robust, portable and low-cost solution which solves the problems faced with current systems. Our Evaluations and User studies show that E-ACE has good user experience, is easily adopted by non-English speaking individuals, and is cost-effective, making it accessible to people in a developing nation.

9:00AM *Multi-Class Classification in Parkinson's Disease by Leveraging Internal Topological Structure of the Data and of the Label Space [#20094]*

Alex Frid, Larry Manevitz and Ohad Mosafi, Laboratory of Clinical Neurophysiology, Faculty of Medicine, Technion (IIT), Israel; Department of Computer Science Ariel University and University of Haifa, Israel; Department of Computer Science, University of Haifa, Israel

In recent work, attacks on automated classification of Parkinson's Disease have encountered difficulties, especially for cross-individual generalization. This is crucial since (i) Classifying the degree of Parkinson's disease is an important clinical necessity. (ii) The lack of such an automated system leaves current clinical methodology to use manual and subjective classification by a trained clinician. In earlier work, we have shown that, reliable classification as to the presence of the disease can be produced using a machine learning. However, this approach was unable to reliably classify the severity degree of the disease. In other work, a deep (convolutional) neural network was tried on the same data set (albeit without feature extraction), which again did not succeed on the multi-label case. In this work, we applied a Data Science approach to solve this problem by analyzing the topological structure of the label space and the internal topological structure of the data. Specifically we explored using (i) the linearity of the label-space to reduce the inherent noise in multi-class classifiers and (ii) to break the data into separate topological clusters (using unsupervised topological learning) and then applying separate classification for each cluster. While our interest was mainly directed to the Parkinson's classification problem, the methods seem relatively generic and should be applicable to many data sets. (As an example, we also applied this directly to a well-known baseline data set and obtained state of the art results.) On the Parkinson classification task, these methods obtained, on a 7 degree classification scale, results which are comparable to the best accuracy on simple two class classification.

9:20AM *Optimization of chemical processes applying surrogate models for phase equilibrium calculations [#19234]*

Corina Nentwich, Christopher Varela and Sebastian Engell, TU Dortmund University, Germany

The calculation of the thermo-physical phase equilibrium of a multicomponent mixture is employed in chemical process development to model the number and composition of phases in order to predict and to optimize the reaction and separation performance of a chemical process plant. Complex thermodynamic models as equations of state provide reliable predictions of phase equilibria over a broad operating range. But due to the need for iterative calculations, they are hardly applicable in optimization. This work shows how a combination of classification and regression can be used to replace these calculations in the optimization of the process of hydroformylation of 1-dodecene.

Thursday, July 18, 9:40AM-10:00AM

Special Lecture: Coffee Break

Thursday, July 18, 9:40AM-10:00AM, Room: Pre-function area Intercontinental

Thursday, July 18, 10:00AM-11:40AM

Plenary Poster Session: Poster Session 2

Thursday, July 18, 10:00AM-11:40AM, Room: Ballroom I + II +II, Chair: Manuel Roveri

P301 *Comparative study between Deep Face, Autoencoder and Traditional Machine Learning Techniques aiming at Biometric Facial Recognition [#20357]*

Jonnathann Finizola, Jonas Targino, Felipe Teodoro and Clodoaldo Lima, University of Sao Paulo, Brazil

Biometric technology is increasingly present in our daily lives whether in mobile devices or commercial sectors because it is an approach where there is great difficulty in being circumvented, unlike traditional models of security and identification. Biometrics is the means by which these technologies can identify individuals and uses physical or behavioral characteristics of the human being and the physical characteristics can be: the iris, face, palm, fingerprint, among others. The behavioral ones can be: way of walking and typing dynamics. With the emergence of Deep Learning, a number of problems that were once solved with traditional machine learning models, have come to better results with this approach, but in the face recognition environment there is still no evidence that Deep Learning can achieve better results than traditional models, with different extractors of characteristics, when applied in different databases facial data. Therefore, the objective of this work is to perform a comparative study between traditional models of machine learning and Deep Learning focusing on Convolutional Neural Networks and Autoencoders for facial recognition.

P302 *Estimating Betti Numbers using Deep Learning [#20363]*

Rahul Paul and Stephan Chalup, The University of Newcastle, Australia

This paper proposes an efficient computational approach for estimating the topology of manifold data as it may occur in applications. For two- or three-dimensional point cloud data, the computation of Betti numbers using persistent homology tools can already be computationally very expensive. We propose an alternative approach that employs deep learning to estimate Betti numbers of manifolds approximated by point clouds. A critical aspect in this new approach is the generation of suitable synthetic training data of scalable topological complexity. Once deep neural networks are trained on this data, inference can be computationally efficient and robust to noise. The pilot results of our study for two- and three-dimensional data support the hypothesis that deep convolutional neural networks can estimate Betti numbers of simulated data that has a topological complexity beyond immediate human visual comprehension. The approach could be generalised beyond estimating the numbers of holes, cavities and tunnels in low-dimensional manifolds to counting high-dimensional holes in high-dimensional data.

P303 *Neural Morphological Segmentation Model for Mongolian [#20397]*

Weihua Wang, Rashel Fam, Feilong Bao, Yves Lepage and Guanglai Gao, Inner Mongolia University, China; Waseda University, Japan

Morphological segmentation is useful for processing Mongolian. In this paper, we manually build a morphological segmentation data set for Mongolian. We then present a character-based encoder-decoder model with attention mechanism to perform the morphological segmentation task. We further investigate the influence of analogy features extracted from scratch and improve the performance of our model using multi languages setting. Experimental results show that our encoder-decoder model with attention mechanism provides a strong baseline for Mongolian morphological segmentation. The analogy features provide useful information to the model and improve the performance of the system. The use of multi languages data set shows the capability of our model to acquire knowledge through different languages and delivers the best result.

P304 *Motion Integration and Disambiguation by Spiking V1-MT-MSTl Feedforward-Feedback Interaction [#20399]*

Maximilian Paul Ruben Loehr, Daniel Schmid and Heiko Neumann, Ulm University, Germany

Motion detection registers items within restricted regions in the visual field. Early stages of cortical processing of motion advance this estimate by integrating spatio-temporal input responses in area V1 to build feature representations of direction and speed in area MT of primate cortex. The neural mechanisms underlying such processes are not yet fully understood. We propose a neural model of hierarchically organized areas V1, MT, and MSTl, with feedforward and feedback connections. Each area serves a distinct purpose and is formally represented by layers of model cortical columns composed of excitatory and inhibitory spiking neurons with conductance-based activation dynamics. Recurrent connections enhance activations by modulatory interaction and divisive normalization. MT population activities allow to estimate motion direction and speed which we show for various stimuli. The importance of the feedback connections for disambiguation is demonstrated in simulated lesion studies.

P305 *An End-to-End Location and Regression Tracker with Attention-based Fused Features [#20405]*

Qinyi Zhang, Shishuai Du and Huihua Yang, Beijing University Of Posts and Telecommunications, China

Visual object tracking, which continuously generating bounding box of a specific target initialized in the first frame in a video, is one of the most fundamental and challenging tasks in computer vision. In this paper, we derive a tracker that combines correlation filter and siamese network, which both achieve superior performance and complement each other, jointly improving the tracker's performance. However, this tracker still has some problems common to most state-of-the-art trackers, that is, discrimination of features is not strong enough, and the scale estimation method is limited. To solve these problems, we propose LR-AFNet, an end-to-end tracker for target location (via the siamese network) and bounding box regression, to better locate and estimate scales, and using attention-based multi features fusion method to extract more discriminative features. Firstly, the feature extraction backbone network is redesigned to improve the location accuracy. Secondly, in order to obtain more effective features, deep and shallow features are fused with an attention mechanism, which adaptively learn the fusion weights. Finally, the bounding box regression branch is added to the siamese similarity network, forming an end-to-end trained framework. Experiments proved that our proposed method achieves competitive performance on several benchmark datasets, compared with the state-of-the- art.

P306 *SE-GAN: A Swap Ensemble GAN Framework [#20411]*

Licheng Shen and Yan Yang, School of Information Science and Technology Southwest Jiaotong University, China

In recent years, Generative adversial network(GAN) becomes prevalent in the research and was applied in various fields. Though GAN attracting great attention undoubtedly, it was still blamed for the difficulty of training. Training a GAN is very likely to get into undesirable results such as gradient vanishing, gradient explosion, model collapse. In order to train GAN more efficiently, this study proposes the Swap Ensemble Generative Adversial Network(SE-GAN), a framework for training an ensemble GAN. Based on the concept of information sharing, SE-GAN allows serveral GANs to be trained in parallel: one of them is the leader, while the others are the workers. The leader pair swaps with each worker periodically to collect the information of all workers. At the same time, each worker acquires the information the leader learned. Under such implementation, the leader and workers would help each other. In this paper, a toy dataset was used to describe the concept and better show the effect of swap ensemble framework. And three real-world datasets were used to validate the effect. Results of the experiments shows this framework improves the performance and efficiency in training GAN.

P307 *A Novel Group-Aware Pruning Method for Few-shot Learning [#20434]*

Yin-Dong Zheng, Yun-Tao Ma, Ruo-Ze Liu and Tong Lu, National Key Lab for Novel Software Technology, Nanjing University, China

Few-shot learning, which focuses on solving machine learning problems in the setting of scarce data, has become a hot spot in the field of neural network and machine learning recently. Inspired by BlockDrop that performs pruning compression on the network using reinforcement learning, in this paper, we present a Group-Aware Pruning (GAP) method with new unifying pruning strategies and different training/testing ways to fit few-shot learning. The proposed GAP consists of three modules, that is, a pruning module, a strategy consensus module (SCM), and a classification module. The whole support set is first fed into the pruning module to get a pruning strategy for each image. Next, the strategies are fed into SCM to fuse into a group strategy for further pruning. When the classification module makes a prediction for a query image, it no longer needs re-output the pruning strategy for the query image since pruning strategies have been unified. Note that in SCM, strategy unification is proposed to achieve the group-aware strategy, which assures an existing deep network that has been pruned by the group-aware strategy work well for few-shot learning problems. Additionally, the testing will be speeded up due to the fact that only one classification model is required in the proposed GAP. Experiments on two benchmark datasets, namely, Omingnet and miniImageNet, and comparisons with the existing state-of-the-art methods show that the accuracy of the proposed GAP is 4.94% higher than the state-of-the-art methods on the 5-way 5-shot task, which shows the effectiveness of the proposed method.

P308 *K-Random Forests: a K-means style algorithm for Random Forest clustering [#19210]*

Manuele Bicego, Computer Science Department, University of Verona, Italy

In this paper we present a novel clustering approach based on Random Forests, a popular classification and regression technique whose usability in the clustering scenario has been investigated to a lesser extent. In the clustering context, the most used class of approaches is based on the exploitation of a single Random Forest to derive a proximity measure between points, to be used with any distance-based clustering technique. On the contrary, our scheme exploits {\itshape a set} of Random Forests, each one devoted to model one cluster, in a spirit similar to that of the mixture models approach. These Random Forests, which provide flexible cluster descriptors, are iteratively updated using a K-means-like clustering algorithm. The proposed scheme, which we call {\itshape K-Random Forests (K-RF)}, has been evaluated on five datasets: the obtained results suggest that it

represents a valid alternative to classic Random Forest clustering algorithms as well as to other established clustering approaches.

P309 *A Multivariate Fuzzy Kohonen Clustering Network [#19868]*

Rodrigo Cavalcanti, Bruno Pimentel, Carlos Almeida and Renata Souza, Universidade Federal de Pernambuco, Brazil; Universidade de Sao Paulo, Brazil; Universidade de Campina Grande, Brazil

Usually, in a fuzzy clustering, the memberships are the same for all the variables (features), i.e., the variables are considered equally important for the definition of the memberships. Fuzzy Kohonen Clustering network (FKCN) is a self-organizing fuzzy neural network that uses fuzzy membership values from the popular Fuzzy c-Means as learning rates. The replacement of the arbitrary learning rate by a fuzzy membership function can produce better clustering results. This paper introduces a new variant of the FKCN algorithm that finds a set of weights and a multivariate fuzzy partition minimizing an objective function. Here, the multivariate memberships allow to take account the intra- class and inter-class dispersion structures of the input data. Experiments with different configurations of synthetic data sets and applications with real data sets demonstrate the usefulness of this fuzzy clustering network model.

P310 *2 Learning Navigation via R-VIN on Road Graphs [#19544]*

Xiaojuan Wei, Jinglin Li, Quan Yuan, Xu Han and Fangchun Yang, Beijing University of Posts and Telecommunications, China

Guiding vehicles to their destination is an essential service. Nowadays navigation systems are mainly relying on the traffic conditions of road network, and other influence factors are not taken into account accurately, which is easy to lead to imbalance between the supply and demand of roads, resulting in congestion. In this paper, we introduce an online guiding approach via value iteration network on road graphs, R-VIN for short, which is an end-to-end planning model. In R-VIN, a largescale real GPS trajectories are mapped via map-matching based road topology, which enables R-VIN to catch the experienced driving knowledge. Then we propose a conversion method from irregular road graphs to regular grid images to formalize the learning model. For a global optimum, ConvLSTM is used to predict the future traffic situation to form "prediction reward" of R-VIN. Combining with "current reward", a double rewarded VIN is used to solve the plan-involved function. Lastly, we train and evaluate R-VIN on planning problem in road networks, showing that R-VIN can achieve segment-based autonomous navigation with high top-k accuracy and less commuting time.

P311 *MPSSD: Multi-Path Fusion Single Shot Detector [#19733]*

Shuyi Qu, Kaizhu Huang, Amir Hussain and Yannis Goulermas, Xi'an Jiaotong-Liverpool University, China; Edinburgh Napier University, United Kingdom; University of Liverpool, United Kingdom

Recent prevalent one stage detectors, such as single shot detector (SSD) and RetinaNet, are able to detect objects faster than two stage ones while maintaining comparable accuracy. To further boost the accuracy, many studies focus on enhancing the multi-scale feature pyramid. Most of these current proposals focus on strengthening features on one pyramid, ignoring the rich connection among different scale features. In contrast, we propose a novel multi-path design to fully utilize the localization and semantics information. First, we exploit the original SSD multi-scale features as our base pyramid. Then we fuse these features in different groups to generate multi-path feature pyramids. Finally, we combine these pyramids through a novel and effective aggregation module, to obtain the final informative pyramid for detection. Comparative experiments on benchmark PASCAL VOC and MS COCO datasets have shown that our proposed method outperforms many state- of-the-art detectors. As an illustrative example, for input image with size 512*512, we can achieve a mean Average Precision (mAP) of 81.8% on VOC2007 test and 33.1% on COCO test-dev2015.

P312 *Deep learning based domain knowledge integration for small datasets: Illustrative applications in materials informatics [#19941]*

Zijiang Yang, Reda Al-Bahrani, Andrew Reid, Stefanos Papanikolaou, Surya Kalidindi, Wei-keng Liao, Alok Choudhary and Ankit Agrawal, Northwestern University, United States; National Institute of Standards and Technology, United States; West Virginia University, United States; Georgia Institute of Technology, United States

Deep learning has shown its superiority to traditional machine learning methods in various fields, and in general, its success depends on the availability of large amounts of reliable data. However, in some scientific fields such as materials science, such big data is often expensive or even impossible to collect. Thus given relatively small datasets, most of data-driven methods are based on traditional machine learning methods, and it is challenging to apply deep learning for many tasks in these fields. In order to take the advantage of deep learning even for small datasets, a domain knowledge integration approach is proposed in this work. The efficacy of the proposed approach is tested on two materials science datasets with different types of inputs and outputs, for which domain knowledge-aware convolutional neural networks (CNNs) are developed and evaluated against traditional machine learning methods and standard CNN-based approaches. Experiment results demonstrate that integrating domain knowledge into deep learning can not only improve the model's performance for small datasets, but also make the prediction results more explainable based on domain knowledge.

P313 *FocalNet - Foveal Attention for Post-processing DNN Outputs [#19850]*

Burhan Ahmad Mudassar and Saibal Mukhopadhyay, Georgia Institute of Technology, United States

This paper presents FocalNet - an iterative information extraction algorithm that uses the concept of foveal attention to post-process the outputs of DNNs by performing variable sampling of the input/feature space. FocalNet is integrated into an existing task-driven deep learning model without modifying the weights of the network. Layers, at which to perform foveation are automatically selected using a data-driven approach. We apply FocalNet to the task of object detection using a state of the art convolutional detector, RFCN ResNet-101. On the PASCAL VOC 2007 dataset, we are able to achieve a mAP increase of 3.7\%. On the MS COCO 2017 validation dataset we achieve an increase in mAP by 0.3\%. Further, a higher increase in mAP is observed for a computationally efficient detector (1.7\% for Faster R-CNN with ResNet50). In additon to object detection, we show effectiveness of FocalNet to the problem of single object tracking with a 0.5\% increase in average IoU on the MOT17 dataset over a simple tracking by detection approach using DNNs.

P314 *Stochastic Variational Inference for Bayesian Sparse Gaussian Process Regression [#19464]*

Haibin Yu, Trong Nghia Hoang, Bryan Kian Hsiang Low and Patrick Jaillet, National University of Singapore, Singapore; MIT-IBM Watson AI Lab, United States; Massachusetts Institute of Technology, United States

This paper presents a novel variational inference framework for deriving a family of Bayesian sparse Gaussian process regression (SGPR) models whose approximations are variationally optimal with respect to the full-rank GPR model enriched with various corresponding correlation structures of the observation noises. Our variational Bayesian SGPR (VBSGPR) models jointly treat both the distributions of the inducing variables and hyperparameters as variational parameters, which enables the decomposability of the variational lower bound that in turn can be exploited for stochastic optimization. Such a stochastic optimization involves iteratively following the stochastic gradient of the variational lower bound to improve its

estimates of the optimal variational distributions of the inducing variables and hyperparameters (and hence the predictive distribution) of our VBSGPR models and is guaranteed to achieve asymptotic convergence to them. We show that the stochastic gradient is an unbiased estimator of the exact gradient and can be computed in constant time per iteration, hence achieving scalability to big data. We empirically evaluate the performance of our proposed framework on two real-world, massive datasets.

P315 *A Support Tensor Train Machine [#20155]*

Cong Chen, Kim Batselier, Ching-yun Ko and Ngai Wong, The University of Hong Kong, Hong Kong; Delft University of Technology, Netherlands

There has been growing interest in extending traditional vector-based machine learning techniques to their tensor forms. Support tensor machine (STM) and support Tucker machine (STuM) are two typical tensor generalization of the conventional support vector machine (SVM). However, the expressive power of STM is restrictive due to its rank-one tensor constraint, and STuM is not scalable because of the exponentially sized Tucker core tensor. To overcome these limitations, we introduce a novel and effective support tensor train machine (STTM) by employing a general and scalable tensor train as the parameter model. Experiments validate and confirm the superiority of the STTM over SVM, STM and STuM.

P316 *StepEncog: A Convolutional LSTM Autoencoder for Near-Perfect fMRI Encoding [#19397]*

Subba Reddy Oota, Vijay Rowtula, Manish Gupta and Raju S. Bapi, IIIT Hyderabad, India; IIIT Hyderabad / Microsoft, India; IIIT Hyderabad / University of Hyderabad, India

Learning a forward mapping that relates stimuli to the corresponding brain activation measured by functional magnetic resonance imaging (fMRI) is termed as estimating encoding models. Computational tractability usually forces current encoding as well as decoding solutions to typically consider only a small subset of voxels from the actual 3D volume of activation. Further, while reconstructing stimulus information from brain activation (brain decoding) has received wider attention, there have been only a few attempts at constructing encoding solutions in the extant neuro-imaging literature. In this paper, we present StepEncog, a convolutional LSTM autoencoder model trained on fMRI voxels. The model can predict the entire brain volume rather than a small subset of voxels, as presented in earlier research works. We argue that the resulting solution avoids the problem of devising encoding models based on a rule-based selection of informative voxels and the concomitant issue of wide spatial variability of such voxels across participants. The perturbation experiments indicate that the proposed deep encoder indeed learns to predict brain activations with high spatial accuracy. On challenging universal decoder imaging datasets, our model yielded encouraging results.

P317 *Multi-task Sentence Encoding Model for Semantic Retrieval in Question Answering Systems [#20437]*

Qiang Huang, Jianhui Bu, Weijian Xie, Shengwen Yang, Weijia Wu and Liping Liu, Baidu Inc., China

Question Answering (QA) systems are used to provide proper responses to users' questions automatically. Sentence matching is an essential task in the QA systems and is usually reformulated as a Paraphrase Identification (PI) problem. Given a question, the aim of the task is to find the most similar question from a QA knowledge base. In this paper, we propose a Multi-task Sentence Encoding Model (MSEM) for the PI problem, wherein a connected graph is employed to depict the relation between sentences, and a multi-task learning model is applied to address both the sentence matching and sentence intent classification problem. In addition, we implement a general semantic retrieval framework that combines our proposed model and the Approximate Nearest Neighbor (ANN) technology, which enables us to find the most similar question from all available candidates very quickly during online serving. The experiments show the superiority of our proposed method as compared with the existing sentence matching models.

P318 *Modular Multilayer Neural Networks Integrate Multisensory Information Near-optimally [#19845]*

Bat-Amgalan Bat-Erdene, He Wang and K. Y. Michael Wong, The Hong Kong University of Science and Technology, Hong Kong

Oftentimes the tasks that we regularly perform require our brain to integrate information from multiple sensory inputs. Experiments on human subjects have shown that the human brain combines multiple modalities in an optimal way that is predicted by the Bayesian inference. However, there have been few studies regarding whether deep neural networks can achieve this feat. Hence, we explore the capability of a modular multilayer neural network to integrate multiple sources of information. We designed a task in which two cameras, one left and one right, are monitoring a rotating chair. Our network comprises two modules, each processing a camera input, connected by an integrator layer. We find that the network successfully combines the information from the two cameras in a nearly Bayes-optimal manner, even when it is only trained with single-camera inputs. Further works can be done to study the capacity for integrating information from more than two inputs.

P319 *Melodious Micro-frissons: Detecting Music Genres From Skin Response [#19937]*

Jessica Sharmin Rahman, Tom Gedeon, Sabrina Caldwell, Richard Jones, Md Zakir Hossain and Xuanying Zhu, The Australian National University, Australia

The relationship between music and human physiological signals has been a topic of interest among researchers for many years. Understanding this relationship can not only lead to more enhanced music therapy methods, but it may also help in finding a cure to mental disorders and epileptic seizures that are triggered by certain music. In this paper, we investigate the effects of 3 different genres of music in participants' Electrodermal Activity (EDA). Signals were recorded from 24 participants while they listened to 12 music stimuli. Various feature selection methods were applied to a number of features which were extracted from the signals. A simple neural network using Genetic Algorithm (GA) feature selection can reach as high as 96.8% accuracy in classifying 3 different music genres. Classification based on participants' subjective rating of emotion reaches 98.3% accuracy with the Statistical Dependency (SD) / Minimal Redundancy Maximum Relevance (MRMR) feature selection technique. This shows that human emotion has a strong correlation with different types of music. In the future this system can be used to distinguish music based on their positive of negative effect on human mental health.

P320 *Enhanced Matching Network for Multi-turn Response Selection in Retrieval-Based Chatbots [#19710]*

Hui Deng, Xiang Xie and XueJun Zhang, Beijing Institute of Technology, China; Chinese Academy of Scienses, China

Semantic representation and dependency information are of great importance for matching a response with its multi-turn context. In this paper, we propose an enhanced matching network (EMN) to enhance the matching ability for the multi-turn response selection system in terms of both constructing semantic representation and extracting dependency information. First, the commonly used recurrent neural network (RNN) is replaced with gated convolutional neural network (GCNN) in the matching network to construct more expressive semantic representations of sentences. Second, local inference modeling and inference composition in the enhanced sequential inference model (ESIM) are utilized here to capture enhanced interactive information between the response and each utterance in the context. Finally, EMN is based on a similar multi-turn structure to sequential matching network (SMN) for extracting turns' dependency information in the chronological order. We furthermore propose a combined model (EMN-SMN) to integrate SMN into EMN for distilling more important dependency information from sentence pairs. Experiments are carried out on Ubuntu Corpus and Douban Conversation Corpus. The results show that EMN can

outperform the state-of-the-art methods and the combined model can further improve overall performance.

P321 *DeepHist: Towards a Deep Learning-based Computational History of Trends in the NIPS [#19862]*

Amna Dridi, Mohamed Medhat Gaber, R. Muhammad Atif Azad and Jagdev Bhogal, Birmingham City University, United Kingdom

Research in analysis of big scholarly data has increased in the recent past and it aims to understand research dynamics and forecast research trends. The ultimate objective in this research is to design and implement novel and scalable methods for extracting knowledge and computational history. While citations are highly used to identify emerging/rising research topics, they can take months or even years to stabilise enough to reveal research trends. Consequently, it is necessary to develop faster yet accurate methods for trend analysis and computational history that dig into content and semantics of an article. Therefore, this paper aims to conduct a fine-grained content analysis of scientific corpora from the domain of Machine Learning. This analysis uses DeepHist, a deep learning-based computational history approach; the approach relies on a dynamic word embedding. The scientific corpora come from 5991 publications from NIPS conference between 1987 and 2015 which are divided into six 5-year timespans. The analysis of these corpora generates visualisations produced by applying t-distributed stochastic neighbor embedding (t-SNE) for dimensionality reduction. The qualitative and quantitative study reported here reveals the evolution of the prominent Machine Learning keywords; this evolution supports the popularity of current research topics in the field. This support is evident given how well the popularity of the detected keywords correlates with the citation counts received by their corresponding papers: Spearman's positive correlation is 100%. With such a strong result, this work evidences the utility of deep learning techniques for determining the computational history of science.

P322 *Multi-label Classification Models for Detection of Phonetic Features in building Acoustic Models [#19387]*

Rupam Ojha and C Chandra Sekhar, Indian Institute of Technology Madras, India

Acoustic modeling in large vocabulary continuous speech recognition systems is commonly done by building the models for subword units such as phonemes, syllables or senones. In recent years, various end-to-end systems using acoustic models built at grapheme or phoneme level have also been explored. These systems either require a lot of data and/or heavily rely on the use of language models or pronunciation dictionary for good recognition performance. With the intention of reducing the dependence on data or external models, we have explored the usage of phonetic features in building acoustic models for speech recognition. The phonetic features describe a sound based on the speech production mechanism in humans. Multi-label classification models are built for detection of phonetic features in a given speech signal. The detected phonetic features are used along with the acoustic features as input to models for phoneme identification. The effectiveness of the proposed approach is demonstrated on TIMIT and Wall Street Journal corpora. Performance improvement over other phoneme recognition studies using the phonetic features is obtained.

P323 *Skeletonization Combined with Deep Neural Networks for Superpixel Temporal Propagation [#20272]*

Adam Fodor, Aron Fothi, Laszlo Kopacsi, Ellak Somfai and Andras Lorincz, Eotvos Lorand University, Hungary

Medial axis representation (a.k.a. shape skeleton) seems to be present in visual processing, but its relevance has remained unclear. Here, we show the potentials of the medial axis transformation in the temporal propagation of superpixels. We combine (i) state-of-the-art deep neural network 'sensors' for optical flow and for depth estimation and (ii) a superpixel algorithm with (iii) the medial axis transformation to obtain frame-to-frame propagation of visual objects. We study the precision of this deep learning facilitated superpixel temporal propagation. We discuss the advantages of the method compared to the temporal propagation of the superpixels themselves.

P324 *A Novel LSTM Approach for Asynchronous Multivariate Time Series Prediction [#19958]*

King Ma and Henry Leung, Department of Electrical and Computer Engineering, University of Calgary, Canada

Long-short-term-memory (LSTM) recurrent neural networks have difficulty in representing temporal and non-temporal inputs simultaneously, due to the sequential emphasis of the architecture. This limits the LSTM applicability in settings where multivariate data is difficult to align. In this paper, a modified hierarchical approach is proposed where a set of univariate LSTM's is trained for asynchronous temporal sequences. The resulting representation is jointly trained with an encoding of multivariate input for prediction. The approach is generalized to combine with non-sequential inputs. The proposed architecture is verified experimentally to improve prediction performance and convergence over conventional LSTM interpolation approaches on simulated Lorenz data and pipeline flow prediction.

P325 *RSLIME: An Efficient Feature Importance Analysis Approach for Industrial Recommendation Systems [#19708]*

Fan Zhu, Min Jiang, Yiming Qiu, Chenglong Sun and Min Wang, iQIYI Inc, China

iQIYI represents one of the largest scale and most sophisticated video recommendation systems in China. In this paper, we focus on the model architecture as well as feature interpretation of our short video recommendation system. We firstly describe the short video recommendation system of iQIYI at a high level, which follows a three- stage information retrieval dichotomy. A Recommendation System Boosted Local Interpretable Model-Agnostic Explanations (RSLIME) is then proposed for real-time feature interpretation and importance evaluation. Lastly, comprehensive online and offline experiments are conducted to demonstrate the effectiveness of RSLIME and prove its unique advantages in the task of feature importance analysis and feature selection on large scale industrial recommendation systems.

P326 *Deep Spiking Neural Network with Spike Count based Learning Rule [#19449]*

Jibin Wu, Yansong Chua, Malu Zhang, Qu Yang, Guoqi Li and Haizhou Li, National University of Singapore, Singapore; Institute for Infocomm Research, A*STAR, Singapore; Tsinghua University, China

Deep spiking neural networks (SNNs) support asynchronous event-driven computation, massive parallelism and demonstrate great potential to improve the energy efficiency and latency of its synchronous analog counterpart. Recently, the network conversion approach, with which pre-trained artificial neural networks are converted into their spiking counterparts, has set records for many machine learning and computer vision benchmarks. Despite the theoretical equivalence has been established between these rate-based models, the attempts for direct training on deep SNNs are still lacking. In this paper, we introduce a novel spike-based learning rule for such rate-coded deep SNNs, whereby the spike count of each neuron is used as a surrogate for gradient backpropagation. We evaluate the proposed learning rule by training deep spiking multi-layer perceptron (MLP) and spiking convolutional neural network (CNN) on the UCI machine learning and MNIST handwritten digit datasets. We show that the proposed learning rule achieves state-of-the-art accuracies on all benchmark datasets. The proposed learning rule allows introducing latency, spike rate, and hardware constraints into the SNN learning, which is superior to the indirect network conversion approach. Hence, it allows direct deployment to the target neuromorphic hardware and supports the efficient inference. Notably, a test accuracy of 98.40\% was achieved on the MNIST dataset in our experiments with only 10 simulation time steps, when the same latency constraint is imposed during training.

P327 *Improving Visual Road Condition Assessment by Extensive Experiments on the Extended GAPs Dataset [#20496]*

Ronny Stricker, Markus Eisenbach, Maximilian Sesselmann, Klaus Debes and Horst-Michael Gross, TU Ilmenau, Germany; LEHMANN + PARTNER GmbH, Germany

Aging public roads need frequent inspections in order to guarantee their permanent availability. In many countries, this includes the standardized visual assessment of millions of images. Due to the lack of sophisticated approaches, often, the evaluation is done manually and therefore requires excessive manual labor. GAPs is the most extensive publicly available dataset that provides standardized, high-quality images for training deep neural networks for pavement distress detection. We further enlarge this dataset and provide refined annotations. By conducting extensive experiments on the GAPs dataset, we improve the performance of automated visual road condition assessment. We evaluate the performance gain of several modern neural network architectures and advanced training techniques.

P328 *Hierarchical Dual Quaternion-Based Recurrent Neural Network as a Flexible Internal Body Model [#20474]*

Malte Schilling, Center of Excellence Cognitive Interaction Technology, Bielefeld University, Germany

Internal models of the body are assumed to serve a multitude of different function. Forward and inverse models are important concepts in motor control. Usually, it is distinguished between these function and there are different models for each individual function or there are even models that are specific to individual behaviors. Here, we present a concept of a core internal model of the body which can act both as an inverse and forward model. In addition, it provides a mechanism for integration of sensory data and can be in this way grounded in the continuous interaction with the environment. In this article, a hierarchical Mean of Multiple Computation neural network is presented that is based on an axis-angle representation of joint movements using dual quaternions. It is shown in detail how the network provides a solution for the forward kinematic problem applied for the case of a seven degrees of freedom robot manipulator. Furthermore, it is used in a complex scenario of a bimanual movement task. This demonstrates how the MMC approach can be easily scaled up from a representation of a single arm to a complex model of a complete body.

P329 *Additive Margin SincNet for Speaker Recognition [#20076]*

Joao Antonio Chagas Nunes, David Macedo and Cleber Zanchettin, Universidade Federal de Pernambuco, Brazil

Speaker Recognition is a challenging task with essential applications such as authentication, automation, and security. The SincNet is a new deep learning based model which has produced promising results to tackle the mentioned task. To train deep learning systems, the loss function is essential to the network performance. The Softmax loss function is a widely used function in deep learning methods, but it is not the best choice for all kind of problems. For distance-based problems, one new Softmax based loss function called Additive Margin Softmax (AM-Softmax) is proving to be a better choice than the traditional Softmax. The AM-Softmax introduces a margin of separation between the classes that forces the samples from the same class to be closer to each other and also maximizes the distance between classes. In this paper, we propose a new approach for speaker recognition systems called AM-SincNet, which is based on the SincNet but uses an improved AM-Softmax layer. The proposed method is evaluated in the TIMIT dataset and obtained an improvement of approximately 40\% in the Frame Error Rate (\%) compared to SincNet.

P330 *Recognition of patterns of optimal diel vertical migration of zooplankton using neural networks [#19332]*

Oleg Kuzenkov, Andrew Morozov and Galina Kuzenkova, Lobachevsky State University of Nizhni Novgorod, Russia; Shirshov Institute of Oceanolog, Russia

The problem of the recognition of evolutionarily stable strategy of diel vertical migrations of zooplankton is solved on the base of neural network using. The mathematical basis for solving the problem is the fitness maximum principle. The software is built in the form of two interconnected complexes. The first complex provides the numerical maximization of fitness function under environmental conditions. This allows us to find evolutionarily stable behavior by methods the calculus of variations. As a result, training samples are formed. The second complex provides the recognition of qualitative features of evolutionarily stable vertical migrations according to approximately given characteristics of the environment. It recognizes the presence or absence of detectable oscillations of zooplankton against the background of random noise. The second complex is created as the learning four-layouts neural network.

P331 *Dense-CAM: Visualize the Gender of Brains with MRI Images [#19352]*

Kai Gao, Hui Shen, Yadong Liu, Lingli Zeng and Dewen Hu, National University of Defense Technoloty, China

Studying the gender differences of brains is important to understand the brain cognitive mechanism. With the increasing amounts of brain imaging data, researchers start to use deep learning for brain image processing. However, the interpretability of deep models is the big gap between them. At present, deep model visualization is the main method to solve the interpretability problem of deep networks. Class Activation Map (CAM) is a commonly-used deep model visualization method. But the resolution of CAMs are restricted since it only uses the last layer for visualization. In this paper, we proposed a novel convolutional network named Dense-CAM by combining DenseNet and CAM and realized the visualization of the whole network so that it can generate more accurate and more robust deep model visualization. The network was tested with the gender classification problem using more than 6000 samples and achieved an accuracy of 92.93%. Brain regions with significant differences between men and women are found with the proposed method, which can be used for future brain imaging studies.

P332 *Using Deep Learning for Mobile Marketing User Conversion Prediction [#19327]*

Matos Luis Miguel, Cortez Paulo, Mendes Rui and Moreau Antoine, University of Minho, Portugal; OLAmobile, Portugal

Mobile performance marketing is a growing industry due to the massive adoption of smartphones and tablets. In this paper, we explore Deep Multilayer Perceptrons (MLP) to predict the Conversion Rate (CVR) of mobile users that are redirected to ad campaigns (i.e., if there will be a sale). We analyze recent real-world big data provided by a global mobile marketing company. Using a realistic rolling window validation, we conducted several experiments with different datasets (two sampling and two data traffic modes), in which we measure both the predictive binary classification performance and the computational effort. The modeling experiments include: two data preprocessing methods, the popular one-hot encoding and a proposed Percentage Categorical Pruning (PCP); and two MLP learning modes, offline (reset) and online (reuse). Overall, competitive classification results were achieved by the PCP transform and the two MLP learning modes, producing real-time predictions and comparing favorably against a Convolutional Neural Network and a Logistic Regression.

P333 *Angular Velocity Estimation of Image Motion Mimicking the Honeybee Tunnel Centring Behaviour [#19326]*

Huatian Wang, Qinbing Fu, Hongxin Wang, Jigen Peng, Paul Baxter, Cheng Hu and Shigang Yue, University of Lincoln, United Kingdom; Guangzhou University, China

Insects use visual information to estimate angular velocity of retinal image motion, which determines a variety of flight behaviours including speed regulation, tunnel centring and visual navigation. For angular velocity estimation, honeybees show large spatial-independence against visual stimuli, whereas the previous models have not fulfilled such an ability. To address this issue, we propose a bio-plausible model for estimating the image motion velocity based on behavioural experiments of the honeybee flying through patterned tunnels. The proposed model contains mainly three parts, the texture estimation layer for spatial information extraction, the delay-and-correlate layer for temporal information extraction and the decoding layer for angular velocity estimation. This model produces responses that are largely independent of the spatial frequency in grating experiments. And the model has been implemented in a virtual bee for tunnel centring simulations. The results coincide with both electro-physiological neuron spike and behavioural path recordings, which indicates our proposed method provides a better explanation of the honeybee's image motion detection mechanism guiding the tunnel centring behaviour.

P334 *Speech Emotion Recognition With Early Visual Cross-Modal Enhancement Using Spiking Neural Networks [#19775]*

Esma Mansouri-Benssassi and Juan Ye, University of St Andrews, Scotland

Speech emotion recognition (SER) is an important part of affective computing and signal processing research areas. A number of approaches, especially deep learning techniques, have achieved promising results on SER. However, there are still challenges in translating temporal and dynamic changes in emotions through speech. Spiking Neural Networks (SNN) have demonstrated as a promising approach in machine learning and pattern recognition tasks such as handwriting and facial expression recognition. In this paper, we investigate the use of SNNs for SER tasks and more importantly we propose a new cross-modal enhancement approach. This method is inspired by the auditory information processing in the brain where auditory information is preceded, enhanced and predicted by a visual processing in multisensory audio-visual processing. We have conducted experiments on two datasets to compare our approach with the state-of-the-art SER techniques in both uni-modal and multi-modal aspects. The results have demonstrated that SNNs can be an ideal candidate for modeling temporal relationships in speech features and our cross-modal approach can significantly improve the accuracy of SER.

P335 *Multi-Task Learning with Capsule Networks [#19215]*

Kai Lei, Qiuai Fu and Yuzhi Liang, Peking University, China

Multi-task learning is a machine learning approach learning multiple tasks jointly while exploiting commonalities and differences across tasks. A shared representation is learned by multi-task learning, and what is learned for each task can help other tasks be learned better. Most of existing multi-task learning methods adopt deep neural network as the classifier of each task. However, a deep neural network can exploit its strong curve-fitting capability to achieve high accuracy in training data even when the learned representation is not good enough. This is contradictory to the purpose of multi-task learning. In this paper, we propose a framework named multi-task capsule (MT-Capsule) which improves multi-task learning with capsule network. Capsule network is a new architecture which can intelligently model part-whole relationships to constitute viewpoint invariant knowledge and automatically extend the learned knowledge to different new scenarios. The experimental results on large real-world datasets show MT-Capsule can significantly outperform the state-of-the-art methods.

P336 *Coupled Dictionary Learning for Multi-label Embedding [#19469]*

Niu Sijia, Xu Qian, Zhu Pengfei, Hu Qinghua and Shi Hong, Tianjin University, China

With the booming of social networks, such as Facebook and Flickr, the candidate labels of an instance can be numerous. Hence, traditional multi-label learning algorithms are out of capability to handle a large quantity of labels for the unaffordable time complexity. To alleviate this problem, label space dimension reduction (LSDR) is proposed by transforming the original label space into a lower dimensional one. Inspired by the effectiveness of coupled dictionary learning (CDL) in dealing with cross-modal data, in this paper, we proposed a novel algorithm named Coupled Dictionary Learning for Multi-label Embedding (ML-CDL) to track the problem of LSDR. We novelly treat feature and label as coupled domains. Then CDL is utilized to generate the low-dimensional latent space that leverages the information between feature and label spaces. In particular, the sparse representation coefficients embody the properties of interpretability, discriminability and sparsity. Experimental results on benchmark datasets demonstrate the effectiveness of our algorithm.

P337 *Skip The Question You Don't Know: An Embedding Space Approach [#19359]*

Kaiyuan Chen and Jinghao Zhao, University of California, Los Angeles, United States

Deep neural network gives people power to generalize hidden patterns behind training data. However, due to limitations on available data collection methods, what neural networks learn should never be expected to deal with all the scenarios: predicting on samples that rarely appear in training set will have very low accuracy. Thus, we design an end-to-end neural network. It learns an inherent discriminative embedding on the training set to perform out-of- distribution(OOD) detection and classification at the same time: both OOD data points and points that resemble those with different labels can be visually observed in this embedding space. Based on this model, we also devise a training scheme that trains on only inliers. Experiments on various datasets and metrics validate that our method outperforms the state-of-art OOD detector.

P338 *Regularization and Iterative Initialization of Softmax for Fast Training of Convolutional Neural Networks [#19598]*

Qiang Rao, Bing Yu, Kun He and Bailan Feng, Huawei Technologies Co., Ltd., China

A softmax regularizer is proposed, a simple and elegant constraint on softmax weight distribution in the training process. Since the direct estimation of feature centers is neither memory efficient nor robust, the proposed regularizer utilizes the relations between feature centers and the classifier weights by adding constraints on the distances between softmax weight vectors. This apparently enlarges the distances between softmax weights to benefit the separation of different classes, and provides extra gradients for the optimization of softmax in order to speed up the training process. Furthermore, we argue that the massive amount of softmax parameters is the main cause that makes the network converge slowly, especial in the class classification tasks with large class number such as face recognition. Motivated by the analysis of the relations between deep features and softmax weights, a fast training process is presented, which splits the training into multiple stages and alternates training and initializing softmax weights for fast convergence when the class number is large. Since the softmax weights can be initialized with estimated deep feature centers, the scale of training data can be gradually increased along the stages. By this procedure, the total training computation cost can be reduced. To validate its effectiveness, our approach is applied on both face recognition and image classification tasks. It obtains comparable performance with the state-of-the-art methods while boasting a faster training process.

P339 *Efficient Deep Gaussian Process Models for Variable-Sized Inputs [#20261]*

Issam Laradji, Mark Schmidt, Vladimir Pavlovic and Minyoung Kim, UBC, Canada; Rutgers University, United States; Seoul Nat'l Univ. of Science & Technology, Korea (South)

Deep Gaussian processes (DGP) have appealing Bayesian properties, can handle variable-sized data, and learn deep features. Their limitation is that they do not scale well with the size of the data. Existing approaches address this using a deep random feature (DRF) expansion model, which makes inference tractable by approximating DGPs. However, DRF is not suitable for variable- sized input data such as trees, graphs, and sequences. We introduce the GP-DRF, a novel Bayesian model with an input layer of GPs, followed by DRF layers. The key advantage is that the combination of GP and DRF leads to a tractable model that can both handle a variable-sized input as well as learn deep long-range dependency structures of the data. We provide a novel efficient method to simultaneously infer the posterior of GP's latent vectors and infer the posterior of DRF's internal weights and random frequencies. Our experiments show that GP-DRF outperforms the standard GP model and DRF model across many datasets. Furthermore, they demonstrate that GP-DRF enables improved uncertainty quantification compared to GP and DRF alone, with respect to a Bhattacharyya distance assessment.

P340 *A Music Recommendation System Based on logistic regression and eXtreme Gradient Boosting [#19514]*

Haoye Tian, Haini Cai, Junhao Wen, Shun Li and Yingqiao Li, School of Big Data and Software Engineering, Chongqing University, Chongqing, China

With the rapid growth of music industry data, it is difficult for people to find their favorite songs in the music library. Therefore, people urgently need an efficient music recommendation system to help them retrieve music. Traditional collaborative filtering algorithms are applied to the field of music recommendation. However, collaborative filtering does not handle data sparse problems very well when new items are introduced. To solve this problem, some people use the logistic regression method as a classifier to predict the user's music preferences to recommend songs. Logistic regression is a linear model that does not handle complex non-linear data features. In this paper, we propose a hybrid LX recommendation algorithm by integrating logistic regression and eXtreme Gradient Boosting(xgboost). A series of experiments are conducted on a real music dataset to evaluate the effectiveness of our proposed LX model. Our results show that the error and AUC of our LX model are better than other methods.

P341 *Brain Dynamics Encoding from Visual Input during Free Viewing of Natural Videos [#19366]*

Zhen Liang, Hiroshi Higashi, Shigeyuki Oba and Shin Ishii, Kyoto University, Japan

Neuroscientific studies have revealed that alpha oscillations in electroencephalography (EEG) signals play an active role in the cognitive aspects of human daily life. In this paper, we developed an artificial encoding model to mimic the dynamic brain visual processing system in terms of alpha oscillations. We analyzed the induced alpha power during free viewing of natural videos and extracted relevant features in a participant-independent way. Meanwhile, the visual characteristics that would trigger involuntary attention in the early stage of visual processing were extracted from videos and further employed to estimate the alpha fluctuations at every 0.5 s. We compared the encoding performances across 19 electrode channel locations and selected an optimal encoding model with a support vector regression. The results demonstrated a promising brain encoding model in EEG signals, that could further contribute to the development of brain-computer interface and visual design.

P342 *Deep Fusion: An Attention Guided Factorized Bilinear Pooling for Audio-video Emotion Recognition [#19842]*

Yuanyuan Zhang, Zi-Rui Wang and Jun Du, University of Science and Technology of China, China

Automatic emotion recognition (AER) is a challenging task due to the abstract concept and multiple expressions of emotion. Although there is no consensus on a definition, human emotional states usually can be apperceived by auditory and visual systems. Inspired by this cognitive process in human beings, it's natural to simultaneously utilize audio and visual information in AER. However, most traditional fusion approaches only build a linear paradigm, such as feature concatenation and multi-system fusion, which hardly captures complex association between audio and video. In this paper, we introduce factorized bilinear pooling (FBP) to deeply integrate the features of audio and video. Specifically, the features are selected through the embedded attention mechanism from respective modalities to obtain the emotion-related regions. The whole pipeline can be completed in a neural network. Validated on the AFEW database of the audio-video sub-challenge in EmotiW2018, the proposed approach achieves an accuracy of 62.48%, outperforming the state-of-the-art result.

P343 *Your Eyes Say You're Lying: An Eye Movement Pattern Analysis for Face Familiarity and Deceptive Cognition [#19623]*

Jiaxu Zuo, Tom Gedeon and Zhenyue Qin, Australian National University, Australia

Eye movement patterns reflect human latent internal cognitive activities. We aim to discover eye movement patterns during face recognition under different cognitions of information concealing. These cognitions include the degrees of face familiarity and deception or not, namely telling the truth when observing familiar and unfamiliar faces, and deceiving in front of familiar faces. We apply Hidden Markov models with Gaussian emission to generalize regions and trajectories of eye fixation points under the above three conditions. Our results show that both eye movement patterns and eye gaze regions become significantly different during deception compared with truth-telling. We show the feasibility of detecting deception and further cognitive activity classification using eye movement patterns.

P344 *Unsupervised Learning of Eye Gaze Representation from the Web [#20230]*

Neeru Dubey, Shreya Ghosh and Abhinav Dhall, Indian Institute of Technology Ropar, India

Automatic eye gaze estimation has interested researchers for a while now. In this paper, we propose an unsupervised learning based method for estimating the eye gaze region. To train the proposed network "Ize-Net" in self-supervised manner, we collect a large 'in the wild' dataset containing 1,54,251 images from the web. For the images in the database, we divide the gaze into three regions based on an automatic technique based on pupil-centers localization and then use a feature- based technique to determine the gaze region. The performance is evaluated on the Tablet Gaze and CAVE datasets by fine-tuning results of Ize-Net for the task of eye gaze estimation. The feature representation learned is also used to train traditional machine learning algorithms for eye gaze estimation. The results demonstrate that the proposed method learns a rich data representation, which can be efficiently finetuned for any eye gaze estimation dataset.

P345 *Video Super Resolution with Estimation of Motion Information by Using Higher Resolution Images Obtained by Single Image Super Resolution [#19300]*

Jonathan Mojoo, Motaz Sabri and Takio Kurita, Hiroshima University, Dept. of Information Engineering, Japan

Video Super resolution algorithms usually utilize the motion information of each pixel in consecutive frames to interpolate pixel values in a higher resolution and reconstruct frames of a higher resolution video from a lower resolution input video. Recently, architectures based on deep neural

networks have gained popularity and can generate higher resolution videos with better visual quality. Also, deep neural networks make single image super resolution possible. In single image super resolution, a higher resolution image is constructed from a given input image by learning the transformation from lower resolution images to the higher resolution images. In this paper we propose to apply the single image super resolution algorithm to each frame in the original video and then use the resulting frames to estimate the movements of each pixel. Since the spatial resolution of the estimated motion of each pixel affects the visual quality of the higher resolution video, it is expected that the proposed approach can improve the visual quality of video super resolution.

P346 *Aspect-level Sentiment Classification with Reinforcement Learning [#19726]*

Tingting Wang, Jie Zhou, Qinmin Vivian Hu and Liang He, East China Normal University, China; Ryerson University, Canada

Aspect-level sentiment classification aims to predict the sentiment polarity of a given aspect in a sentence. However, most of the existing methods focus on the information of the entire sentence rather than a segment that describes the aspect, making it difficult to identify the mapping between an aspect and a segment. Moreover, these methods are prone to the noise in the sentence. To alleviate this problem, we propose a novel approach that models the specific segments for aspect-level sentiment classification in a reinforcement learning framework. Our approach consists of two parts: an aspect segment extraction (ASE) model and an aspect sentiment classification (ASC) model. Specifically, the ASE model extracts the corresponding segment with reinforcement learning and feeds the extracted segment into the ASC model. Then, the ASC model makes the segment-level prediction and provides rewards to the ASE model. The experimental results indicate that our proposed approach can extract the segment towards the aspect effectively, and thus obtains competitive performance. Furthermore, we provide an intuitive understanding of why our ASE model is more effective for aspect-level sentiment classification via case studies.

P347 *DOAD: An Online Dredging Operation Anomaly Detection Method based on AIS Data [#19478]*

Bin Cheng, Shiyou Qian, Jian Cao, Guangtao Xue, Jiadi Yu, Yanmin Zhu and Minglu Li, Shanghai Jiao Tong University, China

Dredging is the removal of sediment from the bottom of lakes, rivers, harbors, and other water bodies. It is a routine necessity in waterways around the world because the natural process of sand and silt washing downstream results in the sediment gradually filling channels and harbors. However, during the dredging operation, some dredgers do not transport the sediment to the designated area as expected but throw it near the waterway meaning sediment may return to the waterway in a short period. This paper proposes an online dredging operation anomaly detection (DOAD) method to detect this kind of irregular behavior during the dredging operation based on automatic identification system (AIS) data. First, we establish a feature system to extract behavior features from AIS data. Furthermore, we jointly utilize t-distributed stochastic neighbor embedding (t-SNE) with neural networks and a Gaussian mixture model (GMM) to train a detection model in a semi-supervised way. Through the trained model, irregular behaviors can be efficiently detected in real time during the dredging operation. The effectiveness of DOAD is evaluated according to a series of experiments. To the best of our knowledge of the published literature, this work is the first to introduce the application of AIS data to detect irregular behaviors during dredging operations.

P348 *MDLDA: A New Multi-Dimension Topic Approach [#19617]*

Juncheng Ding and Wei Jin, University of North Texas, United States

Topic models have been prevalent because topics capture the hidden topics of documents and words are the reflection of topics appearing in documents. However, current topic models fail to provide explicit representation of the relationship between topics. We address this problem by proposing a new topic model, Multi-Dimension Latent Dirichlet Allocation (MDLDA). Our algorithm models the topics together with their mutual relationship. The model describes the relationship between topics as a mixture of words as topics in traditional topic models. The topics extracted by our model can better represent the documents in multiple classic tasks such as document classification. We conduct a case study to explain the model. Quantitative experiments on different datasets show improvements regarding modeling ability and performance in tasks as clustering and classification.

P349 *Analysing and Inferring of Intimacy Based on fNIRS Signals and Peripheral Physiological Signals [#19757]*

Chao Li, Qian Zhang, Ziping Zhao, Li Gu, Nicholas Cummins and Bjorn Schuller, Tianjin Normal University, China; University of Augsburg, Germany; Imperial College London, United Kingdom

Intimacy refers to a relatively long-lasting affinity relationship between individuals, which involves complex neuronal activities and physiological changes in the body. Recent advancements in the field of neuroimaging have demonstrated that functional near- infrared spectroscopy (fNIRS) has excellent potential for intimate relationship analysis. Signals such as fNIRS and physiological signals are increasingly utilised in this regard due to their consistency and complementarity. In this paper, first, we apply fNIRS and physiological database collected from 26 subjects when viewing lover, friend and stranger pictures to analyse and infer the intimacy. Then, the time domain information from both the fNIRS and physiological signals are utilised to exploit the representation of intimacy by General Linear Model (GLM) and Complex Brain Network Analysis (CBNA) methods. Based on these two methods, the intimacy can be analysed with different brain activation patterns. Finally, different machine learning techniques are utilised to predict the intimate relationship. The results demonstrate that multi-modal features are more efficient for intimacy research. Moreover, the average classification accuracy of ensemble learning is 98.72% whereas for KNN it is 91.03%.

P350 *Extreme Dimensionality Reduction for Network Attack Visualization with Autoencoders [#19240]*

Daniel C. Ferreira, Felix Iglesias Vazquez and Tanja Zseby, TU Wien, Austria

The visualization of network traffic flows is an open problem that affects the control and administration of communication networks. Feature vectors used for representing traffic commonly have from tens to hundreds of dimensions and hardly tolerate visual conceptualizations. In this work we use neural networks to obtain extremely low-dimensional data representations that are meaningful from an attack-detection perspective. We focus on a simple Autoencoder architecture, as well as an extension that benefits from pre-knowledge, and evaluate their performances by comparing them with reductions based on Principal Component Analysis and Linear Discriminant Analysis. Experiments are conducted with a modern Intrusion Detection dataset that collects legitimate traffic mixed with a wide variety of attack classes. Results show that feature spaces can be strongly reduced up to two dimensions with tolerable classification degradation while providing a clear visualization of the data. Visualizing traffic flows in two-dimensional spaces is extremely useful to understand what is happening in networks, also to enhance and refocus classification, trigger refined analysis, and aid the security experts' decision-making. We additionally developed a tool prototype that covers such functions, therefore supporting the optimization of network traffic attack detectors in both design and application phases.

P351 *Learning Topological Representation for Networks via Hierarchical Sampling [#19727]*

Guoji Fu, Chengbin Hou and Xin Yao, Southern University of Science and Technology, China

The topological information is essential for studying the relationship between nodes in a network. Recently, Network Representation Learning (NRL), which projects a network into a low-dimensional vector space, has been shown their advantages in analyzing large-scale networks. However, most existing NRL methods are designed to preserve the local topology of a network and they fail to capture the global topology. To tackle this issue, we

propose a new NRL framework, named HSRL, to help existing NRL methods capture both local and global topological information of a network. Specifically, HSRL recursively compresses an input network into a series of smaller networks using a community-awareness compressing strategy. Then, an existing NRL method is used to learn node embeddings for each compressed network. Finally, the node embeddings of the input network are obtained by concatenating the node embeddings resulting from all compressed networks. Empirical studies of link prediction on five real-world datasets demonstrate the advantages of HSRL over state-of-the-art methods.

P352 *Application Inference using Machine Learning based Side Channel Analysis [#19947]*

Nikhil Chawla, Arvind Singh, Monodeep Kar and Saibal Mukhopadhyay, Georgia Institute of Technology, United States; Intel Corporation, United States

The proliferation of ubiquitous computing requires energy-efficient as well as secure operation of modern processors. Side channel attacks are becoming a critical threat to security and privacy of devices embedded in modern computing infrastructures. Unintended information leakage via physical signatures such as power consumption, electromagnetic emission (EM) and execution time have emerged as a key security consideration for SoCs. Also, information published on purpose at user privilege level accessible through software interfaces results in software only attacks. In this paper, we used a supervised learning based approach for inferring applications executing on android platform based on features extracted from EM side-channel emissions and software exposed dynamic voltage frequency scaling (DVFS) states. We highlight the importance of machine learning based approach in utilizing these multi-dimensional features on a complex SoC, against profiling-based approaches. We also show that learning the instantaneous frequency states polled from on-board frequency driver (cpufreq) is adequate to identify a known application and flag potentially malicious unknown application. The experimental results on benchmarking applications running on ARMv8 processor in Snapdragon 820 board demonstrates early detection of these apps in 700 ms, and atleast 85% accuracy in detecting unknown applications. Overall, the highlight is to utilize a low-complexity path to application inference attacks through learning instantaneous frequency states pattern of CPU core.

P353 *A Hybrid Character Representation for Chinese Event Detection [#19768]*

Xiangyu Xi, Tong Zhang, Wei Ye, Jinglei Zhang, Rui Xie and Shikun Zhang, National Engineering Research Center for Software Engineering, Peking University, China

For the Chinese language, event triggers in a sentence may appear inside or across words after word segmentation. Thus recent works on Chinese event detection often formulate the task as a character-wise sequence labeling problem instead of a word-wise one. Due to a limited amount of corpus, however, it is more difficult in practice to train character-wise models to capture the inner structure of event triggers and the semantics of sentence-level context compared with word-wise ones. In this paper, we propose to improve character-wise models by incorporating word information and language model representation into Chinese character representation. More specifically, the former consists of the position of the character inside a word and the word's embedding, which can aid structural pattern learning; the latter is obtained by BERT, which contains long-distance semantic information. We construct a sequence tagging model equipped with the hybrid representation and evaluate our model on ACE 2005 Chinese corpus. Experiment results show that both word information and language model representation are effective enhancements, and our model gains an increase of 4.5 (6.5%) and 6.1 (9.4%) in F1-score in event trigger identification task and classification task respectively over the state-of-the-art method.

P354 *Skin lesion segmentation using deep learning for images acquired from smartphones [#20107]*

Gabriel G. De Angelo, Andre G. C. Pacheco and Renato A. Krohling, Federal University of Espirito Santo, Brazil

Skin lesion is an abnormal growth of skin cells that may become a skin cancer, which is a major health issue around the world and its incidence has been increasing throughout the years. Early detection is critical to increase the survival probability and, for economically emerging countries, it is a real problem, since there is a lack of specialists and medical tools. Smartphones are a low-cost device that may help to tackle this problem. However, acquiring a satisfactory amount of images taken using this tool is a hard task. In this sense, in partnership with a group of dermatologists, we developed an application to collect skin lesion images using smartphones' camera and create a new clinical dataset. In this work, we present a methodology using deep learning, color space combination and conditional random fields to segment the created dataset. We present a carefully investigation regarding the color spaces and the post-processing that allow us to raise some important remarks about the skin lesion ground truth images that strongly affect the final segmentation.

P355 *Classification and Regression Analysis of Lung Tumors from Multi-level Gene Expression Data [#20033]*

Pratheeba Jeyananthan and Mahesan Niranjan, PhD Student, United Kingdom; Supervisor, United Kingdom

We study classification and regression problems in lung tumours where high throughput gene expression is measured at multiple levels: epi-genetics, trancription and protein. We uncover the correlates of smoking and gender-specificity in lung tumors. Different genes are indicative of smoking levels, gender and survival rates at these different levels. We also carry out an integrative anaysis, by feature selection from the pool of all three levels of features. Our results show that the epigenetic information in DNA methylation is a better marker for smoking status than gene expression either at the transcript or protein levels. Further, surprisingly, integrative anlysis using multi-level gene expression offers no significant advantage over the individual levels in the classification and survival prediction problems considered.

P356 *Common Fate Based Episodic Segmentation by Combining Supervoxels with Deep Neural Networks [#20273]*

Laszlo Kopacsi, Aron Fothi, Adam Fodor, Ellak Somfai and Andras Lorincz, Eotvos Lorand University, Hungary

We estimated the contribution of different factors in segmentation tasks by means of deep neural networks. Results indicated that texture and optical flow have similar power, but they seem not to add up. In turn, we decided to study the 'Common Fate Principle' of the 100 years gestaltism suggesting that elements that move together belong together. We developed a simple, fast, and efficient episodic segmentation method that -- to some extent -- resembles the 'how system' of the visual processing: we dropped every piece of information except motion, and started from pure optical flow estimations on 2D videos. For the sake of segmentation, we used a parallel and fast hierarchical supervoxel algorithm. We studied (i) grid topology in space and time, (ii) 2D grid in space and topology dictated by the optical flow in time, and (iii) added deep network based depth estimation from 2D images. We measure performances on episodic foreground-background segmentation task of the Davis benchmark videos. Results are competitive to state-of-the-art segmentation techniques.

P357 *Spatial Event Prediction via Multivariate Time Series Analysis of Neighboring Social Units using Deep Neural Networks [#19403]*

Bonaventure Chidube Molokwu and Ziad Kobti, School of Computer Science, University of Windsor, Windsor, Ontario, Canada N9B-3P4, Canada

Event prediction in social network structures is a crucial research problem in social network analysis. This impels understanding the intrinsic relationship patterns preserving a given social network system, based on the study of several structural properties computed on the constituent social units, with respect to space and time. In this regard, tackling problems of this nature is considered NP-Complete. Consequently, this paper proposes an original and unique approach which involves making event predictions about a target social unit, y, based on the intrinsic patterns of relationship learnt from one or more neighboring social units. Our methodology is based on Deep Learning (DL) architectures, and is developed using deep-layer stacks of Multilayer Perceptron (MLP) appended with an adjustment-bias (a_b) vector at the output layer in a bid to improve the accuracy and precision of predictions made with respect to the target unit (or node). Also, we trained and tested our technique on a real world social clique comprising 5 connected cities; thereafter, we performed a comparative analysis of our approach against 9 other models drawn from the fields of Deep Learning, Machine Learning, and Statistics.

P358 *Risk Prediction for Imbalanced Data in Cyber Security : A Siamese Network-based Deep Learning Classification Framework [#19908]*

Degang Sun, Zhengrong Wu, Yan Wang, Qiujian Lv and Bo Hu, University of Chinese Academy of Sciences, China

Risk prediction plays an important role in network security which can be used to predict riskiest parts and then proactive measures can be adopted to avoid potential damage. Most existing literature model risk prediction problems as binary classification problems by using machine learning methods. However, these traditional machine learning models have poor performance -- tending to misclassify the risky ones into the category of risk-free -- on risk prediction task when the datasets are imbalanced or small in size. In this paper, we propose a Siamese Network Classification Framework (SNCF) that can map the Siamese network to a classification based on the similarity to alleviate imbalance for risk prediction. Experimental results on imbalanced data in risk prediction verify that the deep learning-based classification architecture SNCF has better efficiency when compared with other algorithms.

P359 *PROMISE: A Taxi Recommender System Based on Inter-regional Passenger Mobility [#19151]*

Xiaojun Li, Yu-E Sun, Qian Liu, Zhiwei Shen, Benjian Song, Yang Du and He Huang, School of Rail Transportation, Soochow University, China; School of Computer Science and Technology,University of Science and Technology of China, China; School of Computer Science and Technology, Soochow University, China

Taxi recommender systems have remarkably benefited the taxi business by providing a sequence of pick-up points to reduce the passenger waiting time or the taxi cruising time. In reality, taxi drivers may have their preferred destination regions to avoid traffic jams or to execute arranged pickup orders. However, no prior work managed to maximize the profit of drivers and satisfy the requirements for destination regions at the same time. To tackle this challenge, we propose a PROfit Maximization recommendatIon SystEm (PROMISE) based on the inter-regional probability. In this paper, we first divide the city map into sub-regions and compute the mobility probability of passengers among pick-up points and sub-regions, namely inter-regional mobile probability. Then, we propose an efficient Driving Route Suggestion (DRS) algorithm based on inter-regional probability, which can maximize the profit of taxi drivers with designated destination regions. The experimental results based on the taxi traces collected from Shanghai, China, validate the effectiveness of the proposed recommender system.

P360 *Ideal Neighbourhood Mask for Speech Enhancement Using Deep Neural Networks [#19725]*

Christian Arcos, Marley Vellasco and Abraham Alcaim, Pontifical Catholic University of Rio de Janeiro, Brazil

Degradation of speech signal due to adverse conditions is the major challenge for automatic speech recognition (ASR) systems. This paper introduces a novel approach to estimate an Ideal Neighbourhood Mask (INM) for speech segregation based on deep neural networks estimator. The method described here is based on the local binary patterns (LBP) technique often used in digital image processing. Ideal Neighbourhood Mask will indicate which time-frequency (T-F) units of the noisy speech are cancelled. The performance assessment of the proposed application in conjunction with the traditional mask techniques, i.e., Ideal Binary Mask (IBM) and Ideal Ratio Mask (IRM), are carried out under various environments regarding the objective speech quality measures. The recognition experiments including results in the AURORA IV framework indicate that the proposed scheme, when applied in adverse environments yield significantly better performance than the conventional techniques.

P361 *Knowledge graph-based entity importance learning for multi-stream regression on Australian fuel price forecasting [#19589]*

Dennis Chow, Anjin Liu, Guangquan Zhang and Jie Lu, FEIT, UTS, Australia; CAI,FEIT,UTS, Australia

A knowledge graph (KG) represents a collection of interlinked descriptions of entities. It has become a key focus for organising and utilising this type of data for applications. Many graph embedding techniques have been proposed to simplify the manipulation while preserving the inherent structure of the KG. However, scant attention has been given to the investigation of the importance of the entities (the nodes of KGs). In this paper, we propose a novel entities importance learning framework that investigates how to weight the entities and use them as a prior knowledge for solving multi-stream regression problems. The framework consists of KG feature extraction, multi-stream correlation analysis, and entity importance learning. To evaluate the proposed method, we implemented the framework based on Wikidata and applied it to Australian retail fuel price forecasting. The experiment results indicate that the proposed method reduces prediction error, which supports the weighted knowledge graph information as a means for improving machine learning model accuracy.

P362 *An Initial Study on the Relationship Between Meta Features of Dataset and the Initialization of NNRW [#19297]*

Weipeng Cao, Muhammed J. A. Patwary, Pengfei Yang, Xizhao Wang and Zhong Ming, Shenzhen University, China; University of Chinese Academy of Sciences, China

The initialization of neural networks with random weights (NNRW) has a significant impact on model performance. However, there is no suitable way to solve this problem so far. In this paper, the relationship between meta features of a dataset and the initialization of NNRW is studied. Specifically, we construct seven regression datasets with known attributes' distributions, then initialize NNRW with different distributions and trained them based on the datasets to get the corresponding models respectively. The relationship between the attributes' distributions of the datasets and the initialization of NNRW is analyzed by the performance of the models. Several interesting phenomena are observed: firstly, initializing NNRW with the Gaussian distribution can help the model to have a faster convergence rate than ones with the Gamma and Uniform distribution. Secondly, if one or more attributes in a dataset that follow the Gamma distribution, using Gamma distribution to initialize NNRW may result in a slower convergence rate and easy overfitting. Thirdly, initializing NNRW with a specific distribution with smaller variances can always achieve faster convergence rate and better generalization

performance than the one with larger variances. The above experimental results are not sensitive to the activation function and the type of NNRW. Some theoretical analyses about the above observations are also given in the study.

P363 *Multi-Objective Ensemble Model for Short-Term Price Forecasting in Corn Price Time Series [#19074]*

Matheus Henrique Dal Molin Ribeiro, Victor Henrique Alves Ribeiro, Gilberto Reynoso-Meza and Leandro dos Santos Coelho, Federal Technological University of Parana and Pontifical Catholic University of Parana, Brazil; Pontifical Catholic University of Parana, Brazil; Federal University of Parana and Pontifical Catholic University of Parana, Brazil

Short-term forecasting plays an important role in the economic area. Several studies have been carried, where models with good forecast capacity, focusing on accuracy or stability, were built. Modeling only one of these characteristics without the other can lead to a model with lower generalization capacity. To deal with such situation, this study proposes an ensemble model (EM) to forecast one, two and three months ahead the 60 kg corn bag prices received by producers in the state of Parana (Brazil). Additionally, feature extraction by means of principal component analysis is employed. The EM is built using machine learning models as base (weak) learners (BL) combined by weighted sum. The adopted BL are: Extremely randomized trees, partial least squares, k-nearest neighbors, neural network, bagging and multivariate adaptive regression splines. The weights are chosen through multi-objective optimization, where bias and variance are minimized. The multi-objective differential evolution with spherical pruning algorithm is adopted, while physical programming is used in order to obtain the preferred set of weights. The model built is appointed as multi-objective ensemble model (MOEM). The performance of the model is evaluated using mean absolute percentage error, mean squared error and root mean squared error. Additionally, the Diebold-Mariano test is used to evaluate the reduction on forecasting errors. In general lines, the results show that forecasting using MOEM with two, three or four BL is more stable and accurate than forecasting with single BL. Therefore, this approach is recommended to make short-term forecast of corn prices, which leads to a more assertive decision making.

P364 *Proactive Minimization of Convolutional Networks [#20176]*

Bendeguz Jenei, Gabor Berend and Laszlo Varga, University of Szeged, Institute of Informatics, Hungary

Optimizing the performance of convolutional neural networks (CNNs) in real applications (i.e., production programs) is crucial. The desired network can perform a task while having minimal evaluation time and resource requirement. The evaluation time of a network strongly depends on the number of layers and the number of convolutional kernels in each layers. Therefore, by minimizing the network while keeping its accuracy high is a frequent task. In this paper, we present variations of a method for the automatic minimization of convolutional networks. Our method minimizes the neural network by omitting convolutional kernels during the training process, while also keeping the quality of the results high.

P365 *Text Attention and Focal Negative Loss for Scene Text Detection [#19875]*

Randong Huang and Bo Xu, Institute of Automation, Chinese Academy of Sciences, Beijing, China, China

This paper proposes a novel attention mechanism and a fancy loss function for scene text detectors. Specifically, the attention mechanism can effectively identify the text regions by learning an attention mask automatically. The fine-grained attention mask is directly incorporated into the convolutional feature maps of a neural network to produce graininess-aware feature maps, which essentially obstruct the background inference and especially emphasize the text regions. Therefore, our graininess-aware feature maps concentrate on text regions, in especial those of exceedingly small size. Additionally, to address the extreme text-background class imbalance during training, we also propose a newfangled loss function, named Focal Negative Loss (FNL). The proposed loss function is able to down-weight the loss assigned to easy negative samples. Consequently, the proposed FNL can make training focused on hard negative samples. To evaluate the effectiveness of our text attention module and FNL, we integrate them into the efficient and accurate scene text detector (EAST). The comprehensive experimental results demonstrate that our text attention module and FNL can increase the performance of EAST by F-score of 3.98% on ICDAR2015 dataset and 1.87% on MSRA-TD500 dataset.

P366 *Unsupervised Meta-Learning for Clustering Algorithm Recommendation [#19885]*

Bruno Pimentel and Andre Carvalho, Instituto de Ciencias Matematicas e de Computacao (ICMC-USP), Brazil

In this work, the goal is to use clustering algorithms as recommender in a meta-learning system and, thus, to propose an unsupervised meta-learning approach. Meta-learning has been successfully used for recommendation of Machine Learning algorithms in several Data Mining tasks. Meta-learning can rank algorithms according to their adequacy for a new dataset and use this ranking to recommend algorithms. The recommendations are usually made by predictive meta-models induced by supervised Machine Learning techniques, therefore needing a target attribute. In many situations, the target attribute is not available or has a high computational cost. In these situations, the use of unsupervised meta-models (as clustering algorithms) could provide important insights from Machine Learning experiments, like the interpretation of the partitions found by these clustering algorithms. Here, clustering algorithms are used as unsupervised meta-models. Experimental results show that the proposed approach achieved better clustering quality.

P367 *Strong-Background Restrained Cross Entropy Loss for Scene Text Detection [#19894]*

Randong Huang and Bo Xu, Institute of Automation, Chinese Academy of Sciences, Beijing, China, China

In this paper, we investigate the issue of class imbalance in scene text detection. Class Balanced Cross Entropy (CBCE) loss is often adopted for addressing this imbalance problem. We find that CBCE excessively restrains the backward gradients of background. Negative samples own extremely small weights which are offered by CBCE during training of text detectors. These tiny weight values lead to insufficient learning of background. As a result, the CBCE-based text detection methods only can achieve sub-optimal performance. We propose a novel loss function, Strong-Background Restrained Cross Entropy (SBRCE), to deal with the disadvantage in CBCE. Specifically, SBRCE effectively down-weights the loss assigned to the strong background which means well-classified negative samples. Our SBRCE can make training focused on all positive samples and weak background(i.e., hard-classified negative samples). Moreover, it can prevent the enormous amount of strong background from overwhelming text detectors during training. Experimental results show that the proposed SBRCE can improve the performance of the efficient and accurate scene text detector (EAST) by F-score of 3.3% on ICDAR2015 dataset and 1.12% on MSRA-TD500 dataset, without sacrificing the training and testing speed of EAST.

P368 *Heteroclinic Orbits and Chaos in A Ring of Three Unidirectionally Coupled Nonmonotonic Neurons [#20012]*

Horikawa Yo and Fujimoto Ken'ichi, Faculty of Engineering, Kagawa University, Japan

Bifurcations and chaos in a ring of three unidirectionally coupled nonmonotonic neurons are examined. The Hopf bifurcations of steady states generated through saddle-node bifurcations occur, which are accompanied with homoclinic bifurcations. Heteroclinic orbits connecting these steady states are generated and Shil'nikov's chaos emerges. It is indicated that the second bump in the nonmonotonic output function of neurons is intrinsic for the emergence of homoclinic and heteroclinic orbits as well as chaos.

P369 *Exploring Writing Pattern with Pop Culture Ingredients for Social User Modeling [#20014]*
Chiyu Cai, Linjing Li, Daniel Zeng and Hongyuan Ma, Institute of Automation, Chinese Academy of Sciences, China; CNCERT/CC, China

Social networks have significantly altered the behavior patterns of netizens all around the world. Therefore, accurate and expressive model of social users is increasingly demanded as it pose great value in a variety of scenarios, such as e- commerce, cyber security, and entertainment to name a few. In this paper, we propose the Pop Culture Attention Writing Model (PAWM) to explore the writing patterns of social users by explicitly capturing the influence of Internet pop culture ingredients with an attention mechanism. The writing pattern representations are learned by a memory network through storing and updating historical latent patterns. We then develop the Deep Social User Model via jointly modeling basic properties of social users, temporal contents, and the learned writing patterns based on PAWM. This paper is the first trial, to the best of our knowledge, which captures Internet pop culture information and applies deep neural network to model user writing pattern. A series of experiments conducted on social bot detection and social user identification demonstrate and validate the effectiveness of the proposed models.

P370 *DeepSqueezeNet-CRF: A Lightweight Deep Model for Semantic Image Segmentation [#20019]*
Danyu Lai, Yique Deng and Long Chen, Sun Yat-sen University, China

Deep convolutional neural networks (DCNNs) has shown the powerful capability in image semantic segmentation, along with the huge model sizes and massive parameters. Our key insight is to build a new neural network for image semantic segmentation with fewer parameters and smaller model size under the premise of maintaining state-of-the-art performance. In this paper, we propose a lightweight deep model named DeepSqueezeNet, which introduces the SqueezeNet to the fullyconvolutional networks, replacing parts of convolutional layers by Fire Module layers. Our DeepSqueezeNet contains fewer parameters, smaller model size and higher accuracy than the original fully-convolutional networks. Further more, we propose DeepSqueezeNet-CRF by combining the responses at the final DCNN layer with a fully-connected Conditional Random Field (CRF) to improve localization accuracy. Both quantitative and qualitative results on PASAL VOC 2012 demonstrate the priority and feasibility of our methods, comparing with several other state-of-the-art methods. Especially, our methods can reduce the number of network parameters, model size and training time of the original fully-connected network, with even better performance.

P371 *A GAN Model With Self-attention Mechanism To Generate Multi-instruments Symbolic Music [#20066]*
Faqian Guan, Chunyan Yu and Suqiong Yang, Fuzhou University, China

GAN has recently been proved to be able to generate symbolic music in the form of piano-rolls. However, those existing GAN-based multi-track music generation methods are always unstable. Moreover, due to defects in the temporal features extraction, the generated multi-track music does not sound natural enough. Therefore, we propose a new GAN model with self-attention mechanism, DMB-GAN, which can extract more temporal features of music to generate multi-instruments music stably. First of all, to generate more consistent and natural single-track music, we introduce self-attention mechanism to enable GAN-based music generation model to extract not only spatial features but also temporal features. Secondly, to generate multi-instruments music with harmonic structure among all tracks, we construct a dual generative adversarial architecture with multi-branches, each branch for one track. Finally, to improve generated quality of multi-instruments symbolic music, we introduce switchable normalization to stabilize network training. The experimental results show that DMB- GAN can stably generate coherent, natural multi-instruments music with good quality.

P372 *ADPR: An Attention-based Deep Learning Point-of-Interest Recommendation Framework [#20072]*
Junjie Yin, Yun Li, Zheng Liu, Jian Xu, Bin Xia and Qianmu Li, Nanjing University of Posts and Telecommunications, China; Nanjing University of Science and Technology, China

With the development of location-based social networks (LBSNs), Point-of-Interest (POI) recommendation has attracted lots of attention. Most of the existing studies focus on recommending POIs to users based on their recent check-ins. However, the recent check-ins may contain some daily check-ins that users are not really interested in. If a model treats the recent check-ins equally, it is non-trivial to capture the actual preference of users. To address the issue of mining the actual preferences of users in the POI recommendation, we propose an attention-based deep learning POI recommendation framework (ADPR), which consists of a latent representation method and an attention-based deep convolutional neural network. To learn the embedding of users and POIs, we propose a latent representation method, which incorporates the geographical influence and the categories of POIs to capture the relationships between POIs better. Further, we propose an attention-based deep convolutional neural network, which employs the attention mechanism to filter the important information in the recent check-ins, to recommend POIs to users based on the latent representations of users and the recent check-ins. We conduct experiments on a real-world LBSN dataset to evaluate our framework, and the experimental results show the effectiveness of our framework.

P373 *Closer to Optimal Angle-Constrained Path Planning [#20124]*
Changwu Zhang, Hengzhu Liu and Yuchen Tang, National University of Defense Technology, China; The University of Hong Kong, China

Planning on grids and planning via sampling are the two classical mainstreams of path planning for intelligent agents, whose respective representatives are A* and RRT, including their variants, Theta* and RRT*. However, in the nonholonomic path planning, such us being under angle constraints, Theta* and Lazy Theta* may fail to generate a feasible path because the line-of-sight check (LoS-Check) will modify the original orientation of a state, which makes the planning process incomplete (cannot visit all possible states). Then, we propose a more delayed evaluation algorithm called Late LoS-Check A* (LLA*) to relax the angle constraints. Due to the nature of random sampling, RRT* is asymptotically optimal but still not optimal, then we propose LoS-Check RRT* (LoS-RRT*). In order to solve the problems caused by improper settings of the planning resolution, we propose the LoS-Slider (LoSS) smoothing method. Through experimental comparison, it can be found that angle-constrained versions of LLA* and LoS-RRT* can both generate the near-optimal paths. Meanwhile, the experiment result shows that LLA* performs better than Theta* and Lazy Theta* under angle constraints. The planned path will be even closer to the optimal (shortest) solution after the smoothing of LoSS algorithm.

P374 *Composing Multi-Instrumental Music with Recurrent Neural Networks [#20153]*
David Samuel and Martin Pilat, Charles University, Faculty of Mathematics and Physics, Czech Republic

We propose a generative model for artificial composition of both classical and popular music with the goal of producing music as well as humans do. The problem is that music is based on a highly sophisticated hierarchical structure and it is hard to measure its quality automatically. Contrary to other's work, we try to generate a symbolic representation of music with multiple different instruments playing simultaneously to cover a broader musical space. We train three modules based on LSTM networks to generate the music; a lot of effort is put into reducing the high complexity of multi-instrumental music representation by a thorough musical analysis. We believe that the proposed preprocessing techniques and symbolic representation constitute a useful resource for future research in this field.

P375 *Self-Attention based Network For Medical Query Expansion [#20157]*

Su Chen, Qinmin Vivian Hu, Yang Song, Yun He, Huaying Wu and Liang He, East China Normal University, China; Ryerson University, Canada; Texas A&M University, United States

The aim of clinical decision support implementing electronic health records is to satisfy the physicians' information needs. We are motivated to propose a self-attention based network on query expansion. Considering the difficulty and cost of medical text annotation and inspired by the idea of migration learning, we choose the Semantic Textual Similarity dataset for model training. Different from the previous work, the proposed approach is not only considering the score of a single term as an expansion term, but also taking the score of term combination into account. Our model utilizes Convolutional Neural Networks (CNN) to obtain sentence representation and self-attention mechanism for entity representation. With selfattention, it is able to estimate the weight of each entity to learn better representation for all entities. We conduct the experiments on three standard datasets of Text REtrieval Conference Clinical Decision Support Track, where the approach has a promising overall performance over the strong baselines.

P376 *Static Crowd Scene Analysis via Deep Network with Multi-branch Dilated Convolution Blocks [#20158]*

Haoran Liu, Aiwen Jiang, Qiaosi Yi, Xiaolin Deng, Jianyi Wan and Mingwen Wang, Jiangxi Normal University, China

In this paper, we have proposed a static crowd scene analysis network via multi-branch dilated convolution block, called MDBNet. It focuses on a joint task of estimating crowd count and high-quality density map from static single image. The proposed MDBNet follows one-stage object detection framework, and consists of two parts: pre-trained convolutional layers as the front end for high-level feature extraction and cascaded multi-branch dilated convolution block as the back end for context information aggregation on different ranges. Pixel-wise objectness probabilities are predicted and regressed to generate density map. The proposed MDBNet is an easy training model with strong learning ability. We have tested it on two public datasets (ShanghaiTech dataset and the UFC CC 50 dataset). On almost all evaluation criterions, the proposed method has achieved superior performance. Especially on structure quali- ty criterions, including our newly introduced spatial adjusted mutual information measurement, the MDBNet reports a new state-of-the-art performance. The source code will be distributed depending on publication of our work.

P377 *Hybrid K-Means and Improved Self-Adaptive Particle Swarm Optimization for Data Clustering [#20172]*

Luciano Pacifico and Teresa Ludermir, UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO, Brazil; UNIVERSIDADE FEDERAL DE PERNAMBUCO, Brazil

Data Clustering has become an important mechanism for data exploration and understanding. K-Means algorithm is currently one of the most popular clustering techniques, due to its simplicity and scalability. However, K-Means performance is highly influenced by the choice of initial cluster centers, which may lead to suboptimal solutions. In this paper, a novel hybrid partitional clustering algorithm is proposed, named IDKPSOC-k-means, based on an improved self-adaptive Particle Swarm Optimization (PSO) and K-Means, which uses a crossover operator to improve PSO capability to escape from local minima points from the problem space. To evaluate the performance of the proposed approach, experiments have been performed on sixteen benchmark data sets obtained from UCI Machine Learning Repository. The experimental evaluation, conducted by the use of Friedman hypothesis tests in relation to four clustering metrics, has shown the effectivity of the proposed model in relation to the comparison algorithms.

P378 *Improving Retrieval-Based Question Answering with Deep Inference Models [#20175]*

George Sebastian Pirtoaca, Traian Rebedea and Stefan Ruseti, University Politehnica of Bucharest, Romania

Question answering is one of the most important and difficult applications at the border of information retrieval and natural language processing, especially when we talk about complex questions which require some form of inference to determine the correct answer. In this paper, we present a two-step method that combines information retrieval techniques optimized for question answering with deep learning models for natural language inference in order to tackle the multiple-choice question answering problem. In the first stage, each question- answer pair is fed into an information retrieval engine to find relevant candidate contexts that serve as the underlying knowledge for the inference models. In the second stage, deep learning architectures are used to predict if a candidate answer can be inferred from the context extracted in the first stage. We deploy multiple deep learning architectures pre-trained on different datasets in order to capture semantic features and to enlarge the scope of the questions we can answer correctly. As it will be described, each dataset used for training the inference models has particular characteristics that can be exploited. In the end, all these solvers are combined in an ensemble model to predict the correct answer. This proposed two-step model outperforms the best retrieval-based solver by over 3% in absolute accuracy. Moreover, the model can answer both simple, factoid questions and more complex questions that require reasoning or inference.

P379 *Leveraging Recursive Processing for Neural-Symbolic Affect-Target Associations [#20179]*

Alexander Sutherland, Sven Magg and Stefan Wermter, University of Hamburg, Germany

Explaining the outcome of deep learning decisions based on affect is challenging but necessary if we expect social companion robots to interact with users on an emotional level. In this paper, we present a commonsense approach that utilizes an interpretable hybrid neural-symbolic system to associate extracted targets, noun chunks determined to be associated with the expressed emotion, with affective labels from a natural language expression. We leverage a pre-trained neural network that is well adapted to tree and sub-tree processing, the Dependency Tree-LSTM, to learn the affect labels of dynamic targets, determined through symbolic rules, in natural language. We find that making use of the unique properties of the recursive network provides higher accuracy and interpretability when compared to other unstructured and sequential methods for determining target-affect associations in an aspect-based sentiment analysis task.

P380 *An ensemble strategy for Haplotype Inference based on the internal variability of algorithms [#20265]*

Rogerio Rosa, Lucas Cambuim and Edna Barros, Center for Strategic Technologies of Brazilian Northeast, Brazil; Pernambuco Federal University, Brazil

In this paper, we present an ensemble strategy for haplotype inference problem. The proposed approach generates an ensemble solution from several haplotype matrices yielded by a non-deterministic algorithm. We performed extensive experiments and statistical performance evaluation. Besides the inference accuracy based on Switch Error, our analysis controls the execution time as well. The results show that the proposed method: (1) generates more accurate solutions compared to the existing strategies, (2) improves the precision of haplotyping techniques, such as fastPHASE, Beagle, and Mach, and (3) the Beagle based ensemble produced solutions with quality comparable to the more accurate but more computing intensive method: fastPHASE.

P381 *Hierarchical Intention Enhanced Network for Automatic Dialogue Coherence Assessment [#20353]*
Yunxiao Zhou, Man Lan and Wenting Wang, East China Normal University, China; Alibaba Group, China

Dialogue coherence across multiple turns is still an open challenge. The entity grid model is arguably the most popular approach for coherence modeling. However, it heavily relies on the distribution of entities across adjacent sentences but ignores the emotional context embedded in non-entity text and fails to model long dependencies between speech intentions. These limitations become even more severe when applied to dialogue domain since sentences in dialogue are short, informal and colloquial, thereby, less entities could be extracted and less coherence information could be expressed in these grids. To address the limitations of entity gird methods and incorporate the structure knowledge of dialogue, we propose a new neural network architecture, Hierarchical Intention Enhance Network, to integrate semantic context and speech intention in both utterance and dialogue levels to hierarchically model the global coherence without any entity grids. Our proposed model outperforms the state-of-the-art entity-grid based coherence model on text discrimination task by 17.13% increase in accuracy, confirming the effectiveness of our hierarchical modeling in dialogue context and the crucial importance of intention information in dialogue coherence assessment.

P382 *Learning Distributed Coordinated Policy in Catching Game with Multi-Agent Reinforcement Learning [#19070]*
Xiangyu Liu and Ying Tan, Peking University, China; Peking Univeristy, China

Although learning-based methods such as reinforcement learning have been applied to multi-agent systems design successfully, it is still difficult to learn efficient coordinated policies for agents in partially observed environment settings. Centralized learners contain much more information, but add more complexity, while independent learners suffer from partial observation. To address these problems, we propose a directed multi-agent actor-critic algorithm to directly learn the coordinated policy from experience. The directed critic model can obtain all information including global information and actions, which provides effective learning signals for distributed learning actors. We take Multi-Agent Catching Game as the test scenario, where the task is to coordinate multiple moving paddles to catch balls dropping from the top of the screen. We perform several experimental evaluations and show that our method leads to superior results in learning performance, coordination effect and scalability, compared with both centralized and independent learning approach.

P383 *Biploar fuzzy rough cognitive network [#20525]*
Hua Zheng, School of Information Science, Beijing Language and Culture University, China

Fuzzy Rough Cognitive Network, which is constructed on the theory of fuzzy rough set and fuzzy cognitive maps, have shown good performance in the decision-making. As we know, traditional fuzzy rough set theory concerns the satisfaction degree between objects in universe. Nevertheless, in reality, decision making problems involving the relation which describe the dissatisfaction degree between objects to counterproperty are quite common. This means that there are some hidden information may not be fully utilized. For this reason, we employ a novel mixture decision model, named Bipolar Fuzzy Rough Cognitive Networks(BiFRCN) in this paper. BiFRCN is based on bipolar fuzzy rough set theory and fuzzy cognitive maps by making full use of bipolar fuzzy rough granular information according to three-way decision model. Due to BiFRCN not only focus on satisfaction degree, but also dissatisfaction degree, simulation experiments show that BiFRCN is good and robust in functionality and performance compare with traditional methods.

S25: Artificial Intelligence in Health and Medicine: from Theory to Applications and S27: Deep Neural image and text processing

Thursday, July 18, 10:00AM-11:40AM, Room: Duna Salon I, Chair: Wei Chang Yeh

10:00AM *Benchmarking Multi-task Learning in Predictive Models for Drug Discovery [#20136]*
Philippa Grace McCabe, Sandra Ortega-Martorell and Ivan Olier, Liverpool John Moores University, United Kingdom

Being able to predict the activity of chemical compounds against a drug target (e.g. a protein) in the preliminary stages of drug development is critical. In drug discovery, this is known as Quantitative Structure Activity Relationships (QSARs). Datasets for QSARs are often ill-posed for traditional machine learning to provide meaningful insights (e.g. very high dimensionality). Here, we propose a multi-task learning (MTL) approach to enrich the original QSAR datasets with the hope of improving overall QSAR performance. The proposed approach, henceforth named MTL-AT, increases the size of the useable data by the use of an assistant task: a supplementary dataset formed by compounds automatically extracted from other possibly related tasks. The main novelty in our MTL-AT approach is the addition of control for data leakage. We tested MTL- AT in two drug discovery scenarios: 1) using 100 unrelated QSAR datasets, and 2) using 20 QSAR datasets that are related to the same protein family. Results were compared against equivalent single-task approach (STL). MTL-AT outperformed STL in 45 tasks of scenario 1, and in 12 tasks of scenario 2. The best overall method appears to be MTL-AT on both scenarios, with the small datasets yielded the best performance improvement from using multi-task learning. These results show that implementing multi-task learning with QSAR data has promise, but more investigation is required to test its applicability to certain features in datasets to make it suitable for widespread use in the drug discovery area. To the best of our knowledge, this is the first study that benchmarks the use of MTL on a large number of small datasets, which represents a more realistic scenario in drug development.

10:20AM *An Application of Convolutional Neural Networks for the Early Detection of Late-onset Neonatal Sepsis [#19944]*
Yifei Hu, Vincent Lee and Kenneth Tan, Monash University, Australia; Monash Children's Hospital, Australia

Preterm newborns are vulnerable and easily infected due to the immature immune system. Late-onset neonatal sepsis occurring 48 hours after birth is a widespread disease among preterm newborns leading to high mortality and morbidity rates. The diagnosis is primarily based on biochemistry test, and the prescribed treatment is to use antibiotics. Risk-averse clinicians, often applied overdose to reduce the mortality. A non-invasive method on monitoring vital sign signals deterioration to predict late-onset neonatal sepsis is proposed in this paper. First, we set up collectors within the local networks in Neonatal Intensive Care Units (NICUs) where bedside monitoring machines locate to capture the necessary data. Then they were transformed into images according to specific rules. Finally, a convolutional neural network was built to predict the onset of sepsis. Pilot experiments conducted on data we have collected demonstrated the feasibility of this deep learning model. This method could be incorporated into the current clinical workflow as a decision support system and provide useful information for clinicians.

10:40AM *Deep Capsule Network based Automatic Batch Code Identification Pipeline for a Real-life Industrial Application [#20212]*

Chandan Kumar Singh, Vivek Kumar Gangwar, Harsh Vardhan Singh, Karan Narain, Anima Majumder and Swagat Kumar, Tata Consultancy Services-Research, India

Automatic recognition of text, such as a batch code printed on a box placed on a moving conveyor belt, is still a challenging problem. This paper proposes an end-to-end character recognition technique while addressing the major challenges encountered in a real environment, such as motion blur in the acquired images, slanted or oriented characters, creased batch codes due to wear and tear of boxes, variations in label formats, and variations in printing styles. The major contribution of this work lies in development of three sequential modules: text localization using Connectionist Text Proposal Network(CTPN), character detection and character recognition using a modified version of the capsule network (CapsNet). In contrast to CapsNet, where only a standard single convolution is used, the proposed method uses a series of feature blocks, making it a deep CapsNet which is later proven to generate more comprehensive and better separable feature vectors over its counterpart. The feature generation module is further enhanced by setting a smaller kernel size than CapsNet. The proposed system is validated on a real-world box / packet dataset generated in a retail manufacturing industry. The proposed recognition network architecture is also validated on a standard public dataset (ICDAR 2013). The comparative results are presented with statistical analysis in the experimental results section.

11:00AM *A TOI based CNN with Location Regression for Insurance Contract Analysis [#19259]*

Kai Zhang, Lin Sun and Fule Ji, Zhejiang University City College, China

Contract analysis with AI techniques can significantly ease the work for humans. This paper shows a problem of Element Tagging on Insurance Policy (ETIP). We present a novel Text-Of-Interest (TOI) convolutional neural network for the ETIP solution. We introduce a TOI pooling layer to replace traditional pooling layer for processing the nested phrasal or clausal elements in insurance policies. The advantage of TOI pooling layer is that the nested elements from one sentence could share computation and context in the forward and backward passes. The computation of backpropagation through TOI pooling is also demonstrated in the paper. In addition, a location regressor is trained to improve the precision of element localization, called TOI-CNN+LR. In the detection, we devise a novel non-maximum suppression method with the fusion of length and score metrics, called LS-NMS. A large Chinese insurance contract dataset was collected to test the performance of the proposed method. An extensive set of experiments is performed to investigate how TOI-CNN+LR can work effectively in insurance elements tagging and outperforms other state-of- the-art nested NER models.

11:20AM *Transformation-gated LSTM: efficient capture of short-term mutation dependencies for multivariate time series prediction tasks [#19607]*

Jun Hu and Wendong Zheng, College of Computer Science and Electronic Engineering Hunan University, China

Most multivariate time series data have very complex long-term and short-term dependencies that change over time. Currently, some recurrent neural network (RNN) variants for sequence tasks enhance the learning ability of long-term dependence on time series data. However, there lack of RNN network for capturing short-term mutation information for multivariate time series. In the present work, we proposed a transformation-gated LSTM (TG-LSTM) to enhance the ability of capturing short- term mutation information. First, the transformation gate introduced a hyperbolic tangent function to the memory cell state of the previous time step and the input gate information of the current time step without losing the memory cell state information. Then, the function value range of the partial derivative corresponding to the transformation gate during the backpropagation fully reflected the gradient change, thereby obtaining a better error gradient flow. We further extended to multi-layer TG-LSTM network and compared its stability and robustness with all baseline models. The multi-layer TG-LSTM network was superior to all baseline models in terms of prediction accuracy and performance stability on two different multivariate time series tasks.

S29: Biologically Inspired Learning for Cognitive Robotics and S02: Low Power Hardware for Spiking Neural Networks

Thursday, July 18, 10:00AM-11:40AM, Room: Duna Salon II, Chair: Chris Yakopcic

10:00AM *Effect of pruning on catastrophic forgetting in Growing Dual Memory Networks [#19745]*

Wei Shiung Liew, Chu Kiong Loo, Vadym Gryshchuk, Cornelius Weber and Stefan Wermter, University of Malaya, Malaysia; University of Hamburg, Germany

Grow-when-required networks such as the Growing Dual-Memory (GDM) networks possess a dynamic network structure, expanding to accommodate new neurons in response to learning novel concepts. Over time, it may be necessary to prune extraneous neurons and/or neural connections to meet performance or resource limitations. GDM networks utilize an age-based pruning strategy, whereby older neurons and neural connections that have not been activated recently are removed. Catastrophic forgetting occurs when knowledge learned by the networks in previous learning iterations are lost due to being overwritten by newer learning iterations, or to the pruning process. In this work, we investigate catastrophic forgetting in GDM networks in response to different pruning strategies. The age-based pruning method was shown to significantly sparsify the GDM network topology while improving the network's ability to recall newly-acquired concepts with a slight decrease in performance with older knowledge. A significance-based pruning method was tested as a replacement for the age-based pruning, but was not as effective at pruning even though it performed better at recalling older knowledge.

10:20AM *Heartbeat Detection Based on Pulse Neuron Model for Heart Rate Variability Analysis [#20508]*

Takenori Obo, Daiki Takaguchi, Daisuke Katagami, Junji Sone, Takahito Tomoto, Yuta Ogai and Yoshihisa Udagawa, Tokyo Polytechnic University, Japan

In this study, we aim to develop an educational environment and platform based on color science, merging information technology, robotics and artificial intelligence. In the educational environment, the robot requires functional capability for human-like communication. However, the robot may have difficulty in grasping the small cues to represent feelings and emotional changes of human by itself. This paper presents a measurement system with pneumatic pressure sensor for heart rate variability analysis. We proposed a fuzzy spiking neural network to extract heartbeat signal from the measured data. We furthermore conducted experiments to investigate the effects on student's psychological states while solving some conundrums by using the measurement system.

10:40AM *Action Acquisition Method for Constructing Cognitive Development System Through Instructed Learning [#19923]*

Ryosuke Tanaka, Jinseok Woo and Naoyuki Kubota, Tokyo Metropolitan University, Japan

At the beginning of the research on artificial intelligence, intelligence was thought to be almost synonymous with thought, but as the research of psychology and robot progresses, it began to be considered that the essence of intelligence is to have physicality. In research on cognitive developmental robotics, various methodologies related to cognitive developmental learning by communication based on physicality are proposed, and even in human cognitive development, cognitive ability, communication skill, and physical exercise ability interrelate with each other is important. In this paper, for instructed learning between humans and humanoid robots, we focus on the analysis of nonverbal information and construct a cognitive development the system with the objective of recognizing the target motion, extracting action segments, and acquiring actions. We propose a method to recognize the motion of human beings in real time by using Spiking Neural Network, cluster direction vectors using Multi Layer Growing Neural Gas, and obtain motion. In addition, in order to confirm the effectiveness of the proposed method, actions are acquired using data actually measured, and the acquired action is confirmed.

11:00AM *A Spiking Neural Network with a Global Self-Controller for Unsupervised Learning Based on Spike-Timing-Dependent Plasticity Using Flash Memory Synaptic Devices [#19979]*

Won-Mook Kang, Chul-Heung Kim, Soochang Lee, Sung Yun Woo, Jong-Ho Bae, Byung-Gook Park and Jong-Ho Lee, Seoul National University, Korea (South)

Neuromorphic engineering aims to implement a brain-inspired computing architecture as an alternative paradigm to the von Neumann processor. In this work, we propose a hardware-based spiking neural network (SNN) architecture for unsupervised learning with spike-timing-dependent plasticity (STDP) synapse array using flash memory synaptic array. This novel architecture includes a global self-controller to make each neuron in a single neuron layer operate systematically, which can be also an excellent benefit in terms of area required for system configuration. Therefore, the proposal of this architecture configuration is significant in terms of suggesting a methodology for extending a single neuron into a network. We perform circuit simulation using HSPICE to verify systematic operations of multiple neuron system such as feed-forward and -back pulses generation, a refractory period, a lateral inhibition, and a homeostasis. Various operation in the proposed architecture are designed based on MATLAB simulation result of 28 x 28 MNIST handwritten digit learning and recognition in the SNN having an array of thin-film transistor (TFT)- type NOR flash memory synaptic devices. The results of circuit simulation reflect the specifications required for the STDP operation using the long-term potentiation (LTP) and long-term depression (LTD) characteristics of the proposed synaptic device. The pulse scheme required for STDP in this paper is shown to be suitable for unsupervised learning with flash memory synaptic device.

11:20AM *High Speed Cognitive Domain Ontologies for Asset Allocation Using Loihi Spiking Neurons [#19994]*

Chris Yakopcic, Nayim Rahman, Tanvir Atahary, Tarek Taha, Alex Beigh and Scott Douglass, University of Dayton, United States; University of Dayton Research Institute, United States; Human Effectiveness Directorate, Air Force Research Laboratory, United States

Cognitive agents are typically utilized in autonomous systems for automated decision making. These systems interact at real time with their environment and are generally heavily power constrained. Thus, there is a strong need for a real time agent running on a low power platform. The agent examined is the Cognitively Enhanced Complex Event Processing (CECEP) architecture. This is an autonomous decision support tool that reasons like humans and enables enhanced agent-based decision-making. It has applications in a large variety of domains including autonomous systems, operations research, intelligence analysis, and data mining. One of the key components of CECEP is the mining of knowledge from a repository described as a Cognitive Domain Ontology (CDO). One problem that is often tasked to CDOs is asset allocation. Given the number of possible solutions in this allocation problem, determining the optimal solution via CDO can be very time consuming. In this work we show that a grid of isolated spiking neurons is capable of generating solutions to this problem very quickly, although some degree of approximation is required to achieve the speedup. However, the approximate spiking approach presented in this work was able to complete all allocation simulations with greater than 99.9% accuracy. To show the feasibility of low power implementation, this algorithm was executed using the Intel Loihi manycore neuromorphic processor. Given the vast increase in speed (greater than 1000 times in larger allocation problems), as well as the reduction in computational requirements, the presented algorithm is ideal for moving asset allocation to low power, portable, embedded hardware.

2b: Unsupervised learning and clustering, (including PCA, and ICA)

Thursday, July 18, 10:00AM-11:40AM, Room: Duna Salon III, Chair: Samet Akcay

10:00AM *A Novel Clustering Algorithm based on Directional Propagation of Cluster Labels [#19152]*

Na Xiao, Kenli Li, Xu Zhou and Keqin Li, Hunan University, China; State University of New York, United States

Clustering is an important topic in the field of machine learning. There are abundant algorithms which are proposed for clustering, and they are mainly based on two basic physical metrics, distance and density. However, it is difficult to reflect the orientation relationship between data by distance and density alone, while the orientation relationship can be easily expressed by direction. Inspired by this, we regard direction as the main basic physical metric, and propose a clustering algorithm based on directional propagation of cluster labels, namely DBC (Direction Based Clustering). DBC doesn't need to make specific assumptions about the density of data, but uses the orientation relationship between data to help clustering. Similar to the density-based clustering algorithms DBSCAN (Density Based Spatial Clustering Of Applications with Noise) and DPC (Density Peak Clustering), DBC can also find clusters of arbitrary shapes. It requires two parameters, but its two parameters have a fixed range of empirical values. Moreover, it is less affected by uneven density distribution than DBSCAN and DPC. We compare DBC with four well-known clustering algorithms, including DPC, DBSCAN, AP (Affinity Propagation) and K-means++. Experiments on artificial data sets show that DBC performs better than DBSCAN and DPC on data sets with uneven density distribution, and can effectively identify clusters with overlapping regions. Experiments on real-world data sets show that DBC has advantages over DPC, AP and K-means++ in clustering effect.

10:20AM *Automatic detection of the support points in relational clustering [#19480]*

Parisa Rastin, Younes Bennani and Rosanna Verde, UP13, Sorbonne Paris Cite, France; Universit della Campania Luigi Vanvitelli, Italy

The task of clustering is at the same time challenging and very important in Artificial Intelligence. One of the most popular family of clustering algorithms is the prototype-based approach. Prototype-based algorithms compute a representation of the clusters in the form of a set of prototypes, usually vectors approximating each cluster's barycenter. However, the objects in a data set are not necessarily vectors, especially in real-world applications. These non-vectorial data sets are often represented by the dissimilarities,

distances, or relations between all pairs of objects. They are usually referred as relational data sets. For this kind of data, the algorithms must be adapted to different measures of distance. There are a few state-of-the-art algorithms adapted to relational data sets through the use of barycentric coordinates formalism, in which the objects of a relational data sets are embedded in a space defined by the distances between a subset of the objects, called support points. In this paper, we propose an approach that is able to automatically select the optimal set of support points. We also extend the method to relational data streams, in order to detect variations in the intrinsic dimensionality of the representation space over time. We have compared experimentally the quality of the proposed algorithms on real and artificial data sets. We show that the automatic selection of support points allows an optimal quality in a minimal computation time.

10:40AM *Learning with Coherence Patterns in Multivariate Time-series Data via Dynamic Mode Decomposition [#19278]*

Takehito Bito, Masashi Hiraoka and Yoshinobu Kawahara, Osaka University, Japan; Osaka University / RIKEN, Japan; Kyushu University / RIKEN, Japan

Understanding complex dynamics in the real world is a fundamental problem in various engineering and scientific fields. Dynamic mode decomposition (DMD) has attracted attention recently as a prominent way to obtain global modal descriptions of nonlinear dynamical processes from data without requiring explicit prior knowledge about the underlying systems. In this paper, we propose a novel learning method for multivariate time-series data involving complex dynamics using coherence patterns among attributes extracted by DMD. To this end, we develop kernels defined with Grassmann subspaces spanned by dynamic modes which are calculated by DMD and represent coherence patters among attributes with respect to the estimated modal dynamics. To incorporate information in labels attached to a set of time-series sequences, we employ a supervised embedding step in the DMD procedure. We illustrate and investigate the empirical performance of the proposed method using real-world data.

11:00AM *Unifying Unsupervised Domain Adaptation and Zero-Shot Visual Recognition [#19887]*

Qian Wang, Penghui Bu and Toby Breckon, Durham University, United Kingdom; Xi'an Jiaotong University, China

Unsupervised domain adaptation aims to transfer knowledge from a source domain to a target domain so that the target domain data can be recognized without any explicit labelling information for this domain. One limitation of the problem setting is that testing data (despite no labels) from the target domain is needed during training, which prevents the trained model being directly applied to classify newly arrived test instances. We formulate a new cross-domain classification problem arising from real-world scenarios where labelled data are available for a subset of classes (known classes) in the target domain, and we expect to recognize new samples belonging to any class (known and unseen classes) once the model is learned. This is a generalized zero-shot learning problem where the side information comes from the source domain in the form of labelled samples instead of class-level semantic representations commonly used in traditional zero-shot learning. We present a unified domain adaptation framework for both unsupervised and zero-shot learning conditions. Our approach learns a joint subspace from source and target domains so that the projections of both data in the subspace can be domain invariant and easily separable. We use the supervised locality preserving projection (SLPP) as the enabling technique and conduct experiments under both unsupervised and zero-shot learning conditions, achieving state-of-the-art results on three domain adaptation benchmark datasets: Office-Caltech, Office31 and Office-Home.

11:20AM *Skip-GANomaly: Skip Connected and Adversarially Trained Encoder-Decoder Anomaly Detection [#20178]*

Samet Akcay, Amir Atapour-Abarghouei and Toby Breckon, Durham University, United Kingdom

Despite its ill-definition, anomaly detection is a research endeavour of great interest within machine learning and visual scene understanding. Most commonly, anomaly detection is considered as the detection of outliers within a given distribution based on some measure of normality. The most significant challenge in real-world anomaly detection problems is that available data is highly imbalanced towards normality and contains at most a sub-set of all possible anomalous samples -hence limiting the use of well-established supervised learning methods. By contrast, we introduce an unsupervised anomaly detection model, trained only on the normal (non-anomalous, plentiful) samples in order to learn the normality distribution of the domain, and hence detect abnormality based on deviation from this model. Our proposed approach employs an encoder- decoder convolutional neural network with skip connections to thoroughly capture the multi-scale distribution of the normal data distribution in image space. Furthermore, utilizing an adversarial training scheme for this chosen architecture provides superior reconstruction both within image space and a higher-dimensional latent vector space encoding. Minimizing the reconstruction error metric within both the image and hidden vector spaces during training aids the model to learn the distribution of normality as required. Higher reconstruction metrics during subsequent test and deployment are thus indicative of a deviation from this normal distribution, hence indicative of an anomaly. Experimentation over established anomaly detection benchmarks and challenging real-world datasets, within the context of X-ray security screening, shows the unique promise of such a proposed approach.

S07: Advanced Machine Learning Methods for Big Graph Analytics

Thursday, July 18, 10:00AM-11:40AM, Room: Panorama I, Chair: Guodong Long

10:00AM *ICNet: Incorporating Indicator Words and Contexts to Identify Functional Description Information [#19939]*

Qu Liu, Zhenyu Zhang, Yanzeng Li, Tingwen Liu, Diying Li and Jinqiao Shi, Institute of Information Engineering, Chinese Academy of Sciences., China; DiDi Chuxing., China; Beijing University of Posts and Telecommunications., China

Functional description information refers to the texts that describe the functionality or performance characteristics of a certain object. This type of information is of great potential value for the field of intelligence discovery. Thus automatically and accurately identifying this information from large amounts of texts on the web is very important. In this paper we reduce the functional description problem to a binary classification task deciding whether the input sentence is a functional description sentence or not. However, there exist lots of comment texts in the web data, which are semantically very similar to description texts, making our task quite difficult. Also, existing methods only provide general sentence representation models, which can't lead to targeted ways to solve our problem. Therefore, to address the problem, we not only exploit contexts, like many other previous work did, but also introduce indicator word information to learn rich representations. And in order to incorporate them both, we propose two models, namely ICNet(multi-tasks) and ICNet(ensemble). ICNet(multi-tasks) exploits them jointly in a integrated process of learning representations, while ICNet(ensemble) exploits them by two respective but concatenated sub-models. Experimental results on the collected real-world dataset indicate that both ICNet(multi-tasks) and ICNet(ensemble) achieve higher F1 scores compared with FaxtText, CNN, RNN, LSTM and Bi-LSTM, QuickThought models on this task.

10:20AM *Smooth Deep Network Embedding [#19989]*

Mengyu Zheng, Chuan Zhou, Jia Wu and Li Guo, Institute of Information Engineering, Chinese Academy of Sciences, China; Department of Computing, Faculty of Science and Engineering, Macquarie University, Australia

Network embedding is an efficient method to learn low-dimensional representations of vertexes in networks since the network structure can be captured and preserved through this process. Unlike shallow models, deep neural network framework is able to capture the highly non-linear network structure. Therefore, it can achieve much better performance in comparison of traditional network embedding methods. However, few attentions have been paid to the smoothness of such models, in contrast to numerous research works for image and text fields. Methods without smoothness are not robust enough, which means that slight changes on network may lead dramatic changes on the embedding results. Hence, how to find a smooth deep framework is still an open yet important problem. To this end, in this paper, we propose a Smooth Deep Network Embedding method, namely SmNE, which generates stable and reliable embedding results. Empirically, we conduct experiments on real-world networks. The results show that compared to the state- of-the-art methods, our proposed method can achieve significant gains in several applications.

10:40AM *Evolutionary Community Detection in Dynamic Social Networks [#20102]*

Fanzhen Liu, Jia Wu, Chuan Zhou and Jian Yang, Department of Computing, Macquarie University, Australia; Institute of Information Engineering, Chinese Academy of Sciences, China

Evolutionary clustering is a way of detecting the evolving patterns of communities in dynamic social networks. In principle, the detection process seeks to simultaneously maximize clustering accuracy at the current time step and minimize the clustering drift between two successive time steps. Several evolutionary clustering methods have been developed in an attempt to find the best trade-off between clustering accuracy and temporal smoothness, but the classic genetic operators in these methods do not make the best of the inter- and intra-connected relationships between nodes, which limits their effectiveness. To overcome this problem, we propose a novel migration operator to work in tandem with classic genetic operators to improve the discovery of evolving community structures. The operator is implemented within an existing genetic algorithm which relies on a genome representation under a decomposition framework that formulates evolutionary community detection as a multiobjective optimization problem. Moreover, we present a new method of calculating modularity directly from a genome matrix as the objective for measuring the snapshot quality, which results in a wider search space for finding the optimal solution. Experimental results over several synthetic networks and one real-world dynamic social network suggest that our method is superior to two other state-of-the-art methods in terms of both accuracy and smoothness in discovering evolving community structures in dynamic social networks.

11:00AM *RASE: Relationship Aware Social Embedding [#19714]*

Aravind Sankar, Adit Krishnan, Zongjian He and Carl Yang, University of Illinois, Urbana-Champaign, United States

This paper studies the problem of learning latent representations or embeddings for users in social networks, by leveraging relationship semantics associated with each link. User embeddings are low-dimensional vector-space representations designed to preserve structural proximity indicated by the pairwise relationships. In social networks, the closeness (or proximity) between pairs of users is very different w.r.t. multiple social relationships and thus cannot be represented accurately using a single embedding space. Furthermore, social networks pose a unique challenge of relationship label sparsity that precludes the application of knowledge-graph embedding techniques. In this paper, we associate each observed link with multiple relationship types through relationship weights and learn projection matrices for each relationship type to model the social distance (or proximity) between users specific to each relationship. We propose a novel two-step mutual enhancement framework to iteratively (a) learn user embeddings preserving relationship-specific proximity, and (b) link-relationship weights capturing the role of each link in multiple relationship types. The first step learns user embeddings optimizing relationship-specific proximity, while fixing the relationship weights (or roles) for each link. In the second step, the user embeddings and corresponding projection matrices are assumed to be fixed, while the link-relationship weights are learned. We demonstrate that the relationship-aware user embeddings learned through this mutual enhancement framework, are more effective in representing the users and outperform representative baseline techniques in multilabel classification and relationship prediction tasks.

11:20AM *Meta-Learning for User Cold-Start Recommendation [#19471]*

Homanga Bharadhwaj, IIT Kanpur, India

Recent studies in recommender systems emphasize the importance of dealing with the cold-start problem i.e. the modeling of new users or items in the recommendation system. Meta-learning approaches have gained popularity recently in the Machine Learning (ML) community for learning representations useful for a wide-range of tasks. Inspired by the generalizable modeling prowess of Model-Agnostic Meta Learning, we design a recommendation framework that is trained to be reasonably good enough for a wide range of users. During testing, to adapt to a specific user, the model parameters are updated by a few gradient steps. We evaluate our approach on three different benchmark datasets, from Movielens, Netflix, and MyFitnessPal. Through detailed simulation studies, we show that this framework handles the user cold-start model much better than state-of-the art benchmark recommender systems. We also show that the proposed approach performs well on the task of general recommendations to non cold-start users and effectively takes care of routine and eclectic preference trends of users.

Deep Learning and Algorithms

Thursday, July 18, 10:00AM-11:40AM, Room: Panorama II, Chair: Thomas Trappenberg

10:00AM *A Deep Learning Based Approach to Skin Lesion Border Extraction With a Novel Edge Detector in Dermoscopy Images [#19358]*

Abder-Rahman Ali, Jingpeng Li, Sally Jane O'Shea, Guang Yang, Thomas Trappenberg and Xujiong Ye, University of Stirling, United Kingdom; Mater Private Hospital, Ireland; Imperial College London, United Kingdom; Dalhousie University, Canada; University of Lincoln, United Kingdom

Lesion border detection is considered a crucial step in diagnosing skin cancer. However, performing such a task automatically is challenging due to the low contrast between the surrounding skin and lesion, ambiguous lesion borders, and the presence of artifacts such as hair. In this paper we propose a two-stage approach for skin lesion border detection: (i) segmenting the skin lesion dermoscopy image using U-Net, and (ii) extracting the edges from the segmented image using a novel approach we call FuzzEdge. The proposed approach is compared with another published skin lesion border detection approach, and the results show that our approach performs better in detecting the main borders of the lesion and is more robust to artifacts that might be present in the image. The approach is also compared with the manual border drawings of a dermatologist, resulting in an average Dice similarity of 87.7%.

10:20AM *Query recommendation based on user behavior and query semantics [#19353]*

Jialu Xu, Feiyue Ye, Hang Yu and Bo Wang, Shanghai University, China; University of Technology Sydney, Australia

Search engines contain many user search behaviours. Based on the idea of group wisdom, analysis of the search behaviours in the search log to recommend a query has attracted significant attention from researchers. It can improve the accuracy of search recommendation. Query flow graph is a query recommendation method based on search behaviours. In a query flow graph, nodes representing queries are connected if they have the same search intention. In the process of query recommendation, user intention is associated both with current queries and with the clicked URL. The URL can serve as important information to locate user search intentions. However, a query flow graph only considers query information in the search log to determine the relation between the queries. Accordingly, a novel method based on the improved query flow graph is proposed, which expands the query flow graph by adding the clicked URL information and semantic information. The clicked URL information and semantic information can make the next query closer to user search intentions. Empirical experiments are performed in accordance with AOL log, and the results confirm the effectiveness of our approach in suggesting queries. The results demonstrate that the performance of our query recommendation algorithm is superior to those of other algorithms.

10:40AM *Predicting Household Water Consumption Events: Towards a Personalised Recommender System to Encourage Water-conscious Behaviour [#20078]*

Md Shamsur Rahim, Khoi Anh Nguyen, Rodney Anthony Stewart, Damien Giurco and Michael Blumenstein, Centre for Artificial Intelligence, School of Software, University of Technology Sydney, Australia; School of Engineering and Built Environment, Griffith University, Australia; Institute for Sustainable Futures, University of Technology Sydney, Australia

Recommender systems assist customers to make decisions; however, the modest adoption of digital technology in the water industry means no such system exists for household water users. Such a system for the water industry would suggest to consumers the most effective ways to conserve water based on their historical data from smart water meters. The advantage for water utilities in metropolitan areas is in managing demand, such as low pressure during peak hours or water shortages during drought. For customers, effective recommendations could save them money. This paper presents a novel vision of a recommender system prototype and discusses the benefits both for the consumers and the water utility companies. The success of this type of system would depend on the ability to anticipate the time of the next major water use so as to make useful, timely recommendations. Hence, the prototype is based on a long short-term memory (LSTM) neural network that predicts significant water consumption events (i.e., showers, baths, irrigation, etc.) for 83 households. The preliminary results show that LSTM is a useful method of prediction with an average root mean square error (RMSE) of 0.403. The analysis also provides indications of the scope of further research required for developing a commercially successful recommender system.

11:00AM *SAI: A Sensible Artificial Intelligence that plays Go [#19394]*

Francesco Morandin, Gianluca Amato, Rosa Gini, Carlo Metta, Maurizio Parton and Gian-Carlo Pascutto, Universita` di Parma, Italy; Universita` di Chieti-Pescara, Italy; Agenzia regionale di sanita` della Toscana, Italy; Universita` di Firenze, Italy; Mozilla Corporation, Belgium

We propose a multiple-komi modification of the AlphaGo Zero/Leela Zero paradigm. The winrate as a function of the komi is modeled with a two-parameters sigmoid function, hence the winrate for all komi values is obtained, at the price of predicting just one more variable. A second novel feature is that training is based on self-play games that occasionaly branch -- with changed komi-- when the position is uneven. With this setting, reinforcement learning is shown to work on 7x7 Go, obtaining very strong playing agents. As a useful byproduct, the sigmoid parameters given by the network allow to estimate the score difference on the board, and to evaluate how much the game is decided. Finally, we introduce a family of agents which target winning moves with a higher score difference.

11:20AM *The Emergent-Context Emergent-Input Framework for Temporal Processing [#20406]*

Xiang Wu and Juyang Weng, Nanjing University of Science and Technology, China; Michigan State University, United States

Many temporal processing tasks face great challenges. The handcrafted features and structure used in these temporal processing methods result in the brittle systems. To avoid handcrafted designs, we analyze and propose the Emergent-Context Emergent-Input (ECEI) framework in this work. The Developmental Network (DN-1) can be considered as the first ECEI framework that used emergent motor vectors as temporal context (state) for top- down attention. This method suggests that the developmental root of temporal states is at the (open) motor end and always expressible by the motor end, not originally internal and hidden. Indeed, it has enabled the control of a Turing machine to incrementally emerge inside the network. However, DN-1 used a handcrafted partition of motor neurons as concept regions, which does not allow context to automatically emerge. The same is true for hidden (internal) areas other than the motor area. We analyze why this limitation is fundamentally damaging. To overcome this limitation, Developmental Network 2 (DN-2) is proposed with some new biology inspired mechanisms. DN-2 is an ECEI framework which can directly use any naturally emerged motor patterns, allowing a freedom of thought that has been largely overlooked. The phoneme recognition experiments are used to investigate DN-2's performance with two different settings about the motor area: handcrafted motor concept zones and emergent motor area. Based on the experimental results, we compared and analyzed such two mechanisms. This work is a step toward our goal of empowering DN-2 to not only free thoughts, but such free-thoughts are also necessary for handling increasingly complicated temporal processing tasks.

Neural Network Models

Thursday, July 18, 10:00AM-11:40AM, Room: Panorama III, Chair: Ata Kaban

10:00AM *Compressive Learning of Multi-layer Perceptrons: An Error Analysis [#20494]*

Ata Kaban, University of Birmingham, United Kingdom

We consider the class of 2-layer feed-forward neural networks with sigmoidal activations -- one of the oldest black-box learning machines -- and ask the question: Under what conditions can it successfully learn from a random linear projection of the data? Part of this question has been previously attempted in the literature: A high probability bound has been given on the absolute difference between the outputs of the network on the sample before and after random projection -- provided that the target dimension is at least $\Omega(M^2 (\log MN))$, where M is the size of the hidden layer, and N is the number of training points. By contrast, in this paper we prove that a lower target dimension independent of both N and M suffices, not only to guarantee low distortion of the outputs but also to ensure good generalisation for learning the network on randomly projected data. We do not require a sparse representation of the data, instead our target dimension bound depends on the regularity of the problem expressed as norms of the weights. These are uncovered in our analysis by the use of random projection, which fulfils a regularisation role on the input layer weights.

10:20AM *Relearning procedure to adapt pollutant prediction neural model: Choice of relearning algorithm [#19144]*

Philippe Thomas, Marie-Christine Suhner and William Derigent, University of Lorraine CRAN, France

Predict the indoor air quality becomes a global public health issue. That's why Airbox lab company develops a smart connected object able to measure different physical parameters including concentration of pollutants (volatile organic compounds, carbon dioxide and fine particles). This smart object must embed prediction capacities in order to avoid the exceedance of an air quality threshold. This task is performed by neural network models. However, when some events occur (change of people's behaviors, change of place of the smart connected object as example), the embedded neural models become less accurate. So a relearning step is needed in order to refit the models. This relearning must be performed by the smart connected object, and therefore, it must use the less computing time as possible. To do that, this paper propose to combine a control chart in order to limit the frequency of relearning, and to compare three learning algorithms (backpropagation, Levenberg-Marquardt, neural network with random weights) in order to choose the more adapted to this situation.

10:40AM *Accelerating Deep Unsupervised Domain Adaptation with Transfer Channel Pruning [#19085]*

Chaohui Yu, Jindong Wang, Yiqiang Chen and Zijing Wu, University of Chinese Academy of Sciences, China; Columbia University, United States

Deep unsupervised domain adaptation (UDA) has recently received increasing attention from researchers. However, existing methods are computationally intensive due to the computation cost of Convolutional Neural Networks (CNN) adopted by most work. To date, there is no effective network compression method for accelerating these models. In this paper, we propose a unified Transfer Channel Pruning (TCP) approach for accelerating UDA models. TCP is capable of compressing the deep UDA model by pruning less important channels while simultaneously learning transferable features by reducing the cross-domain distribution divergence. Therefore, it reduces the impact of negative transfer and maintains competitive performance on the target task. To the best of our knowledge, TCP is the first approach that aims at accelerating deep UDA models. TCP is validated on two benchmark datasets - Office- 31 and ImageCLEF-DA with two common backbone networks - VGG16 and ResNet50. Experimental results demonstrate that TCP achieves comparable or better classification accuracy than other comparison methods while significantly reducing the computational cost. To be more specific, in VGG16, we get even higher accuracy after pruning 26% floating point operations (FLOPs); in ResNet50, we also get higher accuracy on half of the tasks after pruning 12% FLOPs. We hope that TCP will open a new door for future research on accelerating transfer learning models.

11:00AM *Attention-driven Multi-sensor Selection [#19120]*

Stefan Braun, Daniel Neil, Jithendar Anumula, Enea Ceolini and Shih-Chii Liu, Institute of Neuroinformatics, Zurich, Switzerland

Recent encoder-decoder models for sequence-to-sequence mapping show that integrating both temporal and spatial attention mechanisms into neural networks considerably improve network performance. The use of attention for sensor selection in multi-sensor setups and the benefit of such an attention mechanism is less studied. This work reports on a sensor transformation attention network (STAN) that embeds a sensory attention mechanism to dynamically weigh and combine individual input sensors based on their task-relevant information. We demonstrate the correlation of the attentional signal to changing noise levels of each sensor on the audio-visual GRID dataset and synthetic noise; and on CHiME-4, a multi-microphone real-world noisy dataset. In addition, we demonstrate that the STAN model is able to deal with sensor removal and addition without retraining, and is invariant to channel order. Compared to a two-sensor model that weighs both sensors equally, the equivalent STAN model has a relative parameter increase of only 0.09%, but reduces the relative character error rate (CER) by up to 19.1% on the CHiME-4 dataset. The attentional signal helps to identify a lower SNR sensor with up to 94.2% accuracy.

11:20AM *DGFFM: Generalized Field-aware Factorization Machine based on DenseNet [#19720]*

Qing-Long Zhang, Lu Rao and Yubin Yang, State Key Laboratory for Novel Software Technology at Nanjing University, China

In this paper, we build a generalized field-aware factorization machine (GFFM) based on FFM, which stores different feature embeddings in multiple files separately instead of one single file. By making use of the corresponding location relationship between the field index and feature index, GFFM can significantly reduce the computation time. Also, features in GFFM are subdivided into dynamic ones and static ones. When modeling, the temporal variation of dynamic features such as fashion trends and user behavior preferences are considered to establish an accurate dynamic model based on the time window. We further propose DGFFM which uses the wide \& deep framework to jointly train GFFM and DenseNet, aiming to combine the benefits of traditional machine learning methods including their faster learning speed on low- rank features and ability to extract high dimensional features. DGFFM add extra first-order identity features into the FM/FFM related models for the first time by adopting the multi-storage pattern designed in this paper. Experimental results on four real-world datasets demonstrate that GFFM can obtain higher accuracy and computational efficiency compared with the state-of-the-art methods used in recommendation and CTR tasks, and DGFFM can achieve an even higher prediction accuracy than recent DNN models.

S16: Explainable Machine Learning

Thursday, July 18, 10:00AM-11:40AM, Room: Panorama IV, Chair: Davide Bacciu

10:00AM *Scalable implementation of measuring distances in a Riemannian manifold based on the Fisher Information metric [#19892]*

Raul V. Casana-Eslava, Jose D. Martin-Guerrero, Sandra Ortega-Martorell, Paulo J. Lisboa and Ian H. Ian, Liverpool John Moores University, United Kingdom; Universitat de Valencia, Spain

This paper focuses on the scalability of the Fisher Information manifold by applying techniques of distributed computing. The main objective is to investigate methodologies to improve two bottlenecks associated with the measurement of distances in a Riemannian manifold formed by the Fisher Information metric. The first bottleneck is the quadratic increase in the number of pairwise distances. The second is the computation of global distances, approximated through a fully connected network of the observed pairwise distances, where the challenge is the computation of the all sources shortest path (ASSP). The scalable implementation for the pairwise distances is performed in Spark. The scalable global distance computations are addressed by applying the Dijkstra algorithm (SSSP) sequentially on two proposed networks based on prototypes that approximate the manifold. The proposed solutions are compared with a single-machine implementation in Matlab with experiments showing the first bottleneck solution is faster in Spark, but the distributed solutions for the second bottleneck is slower. Therefore, our conclusion is that the most appropriate method is a hybrid approach, where in terms of runtime and scalability a hybrid approach performs best; running the distributed method and the single-machine approach to solve bottleneck one then two, respectively.

10:20AM *How to produce complementary explanations using an Ensemble Model [#20304]*

Wilson Silva, Kelwin Fernandes and Jaime S. Cardoso, INESC TEC, Portugal; NILG.AI, Portugal

In order to increase the adoption of machine learning models in areas like medicine and finance, it is necessary to have correct and diverse explanations for the decisions that the models provide, to satisfy the curiosity of decision-makers and the needs of the regulators. In this paper, we introduced a method, based in a previously presented framework, to explain the decisions of an Ensemble Model. Moreover, we instantiate the proposed approach to an ensemble composed of a Scorecard, a Random Forest, and a Deep Neural Network, to produce accurate decisions along with correct and diverse explanations. Our methods are tested on two biomedical datasets and one financial dataset. The proposed ensemble leads to an improvement in the quality of the decisions, and in the correctness of the explanations, when compared to its constituents alone. Qualitatively, it produces diverse explanations that make sense and convince the experts.

10:40AM *On The Stability of Interpretable Models [#19575]*

Riccardo Guidotti and Salvatore Ruggieri, ISTI-CNR, Italy; University of Pisa, Italy

Interpretable classification models are built with the purpose of providing a comprehensible description of the decision logic to an external oversight agent. When considered in isolation, a decision tree, a set of classification rules, or a linear model, are widely recognized as human-interpretable. However, such models are generated as part of a larger analytical process. Bias in data collection and preparation, or in model's construction may severely affect the accountability of the design process. We conduct an experimental study of the stability of interpretable models with respect to feature selection, instance selection, and model selection. Our conclusions should raise awareness and attention of the scientific community on the need of a stability impact assessment of interpretable models.

11:00AM *Contrastive Relevance Propagation for Interpreting Predictions by a Single-Shot Object Detector [#19595]*

Hideomi Tsunakawa, Yoshitaka Kameya, Hanju Lee, Yosuke Shinya and Naoki Mitsumoto, Meijo University, Japan; DENSO CORPORATION, Japan

Object detection is a widely-used computer vision task, in which we identify a bounding box around each object in an image and classify the object into one of pre- defined classes. Single Shot MultiBox Detector (SSD) is a real-time object detector based on a single convolutional neural network. SSD is popular and known for high speed and accuracy, but its black-box nature is not ignorable when it is applied to critical systems. In this paper, we propose Contrastive Relevance Propagation (CRP), an extension of Layer-wise Relevance Propagation (LRP) tailored for SSD. CRP can consistently deal with SSD's heterogeneous output, i.e. confidences for object classes and location offsets, and create a heatmap that highlights a crucial part of the input, which is not available with a standard use of LRP. By experiments with the Pascal VOC 2012 dataset, we confirmed the quality of heatmaps created by CRP, and with such heatmaps, we conducted a simple analysis on prediction errors made by SSD.

11:20AM *Explainable Classifier Supporting Decision-making for Breast Cancer Diagnosis from Histopathological Images [#19794]*

Patrik Sabol, Peter Sincak, Kana Ogawa and Pitoyo Hartono, Technical University of Kosice, Slovakia; Chukyo University, Japan

This paper presents an application of semantically explainable classifier, Cumulative Fuzzy Class Membership Criterion (CFCMC), in medical domain, specifically for breast cancer detection from histopathological images. This classifier, in contrast with commonly used classifiers for image classification, is able to provide additional information about its classification results in human-friendly form. In this paper, we proposed a means for presenting the additional semantical informations that is potentionally useful for decision-making in medical domain. First, the classifier provides semantic explanation, regarding the possibility of misclassification of the test sample. Alongside with semantics, it provides visualization of similar and non-similar samples of different class. The classification performance of CFCMC is compared against three commonly used classifiers for image classification, Convolutional Neural Network (CNN), Stacked Auto-encoder (SAE) and Deep Multi-layered Perceptron. The experimental result shows that the CFCMC is not necessarily the best classifier. However, the ability to provide semantic and visual explanation of classification result allows the classifier to be applied as a supporting tool for pathologists in diagnostic of breast cancer.

S32: Deep Reinforcement Learning for Games

Thursday, July 18, 10:00AM-11:40AM, Room: Panorama V, Chair: Yuanheng Zhu

10:00AM *End-to-end Learning Method for Self-Driving Cars with Trajectory Recovery Using a Path-following Function [#19741]*

Tadashi Onishi, Toshiyuki Motoyoshi, Yuki Suga, Hiroki Mori and Tetsuya Ogata, Waseda University, Japan

We propose an end-to-end learning method for autonomous driving systems in this article. End-to-end model for autonomous driving systems has recently been based on neural networks, which are popular for their good recognition ability. A common problem is how to return a car to the driving lane when the car goes off the track. In our research, we collect recovery data based on the distance from a desired track (the nearest waypoint link) during a road test with a simulator. To train the recovery behavior, instead of collecting human driving data, we use a path-following module (which means the car automatically drives on a pre-decided route using the car's current position). Our proposed method is divided into three phases. In phase 1, we collect data only using a path-following module during 100 laps of driving. In phase 2, we generate driving behavior using a neural driving module trained by the

data collected in phase 1. This includes switching between the accelerator, brake and steering based on a threshold. We collect further data on the recovery behavior using the path-following module during 100 laps of driving. In phase 3, we generate driving behavior using the neural driving module trained by the data collected in phases 1 and 2. To assess the proposed method, we compared the average distance from the nearest waypoint link and the average distance traveled per lap for datasets with no recovery, for datasets with random recovery, and for datasets for the proposed method with recovery. A model based on the proposed method drove well and paid more attention to the road rather than the sky and other unrelated objects automatically for both untrained and trained courses and weather.

10:20AM *Modified State Observer Based Two-Way ETNAC Design For Uncertain Linear Systems [#20379]*

Abdul Ghafoor and Sivasubramanya N Balakrishnan, Missouri University of Sciences and Technology, Rolla, Missouri., United States

In this study, a neural network (NN) based two-way event-triggered controller (ETC) design is presented for uncertain linear systems. Dynamic triggering conditions are developed for both state and control data transmission. These triggering conditions are based on real performance parameters including estimation/tracking errors which make control execution more relevant instead of the extended time sampling as seen in most ETC literature. The other unique feature of this ETNAC scheme is the online uncertainty approximation even during inter-event time which makes control robust and efficient. A Modified State Observer (MSO) used in this study to replicate the system to estimate the system uncertainty. In this way it not only saves data communication over network but also reduces computational efforts. System stability and event-triggering conditions are established using Lyapunov analysis. A benchmark numerical example is used to illustrate the effectiveness of the proposed method.

10:40AM *Optimal Pedestrian Evacuation in Building with Consecutive Differential Dynamic Programming [#19916]*

Yuanheng Zhu, Haibo He, Dongbin Zhao and Zhongsheng Hou, Institute of Automation, Chinese Academy of Sciences, China; University of Rhode Island, United States; Qingdao University, China

Fast and efficient evacuation of pedestrians from an enclosed area is a difficult but crucial issue in modern society. In this paper, the optimization of evacuation from a building is studied. A graph is adopted to describe the building layout with nodes representing areas and edges representing connections. The dynamics of the evacuation process in the graph is formulated by a nonlinear discrete-time model at a macroscopic level. To find the optimal evacuation plan, a consecutive differential dynamic programming is developed. It inherits the differential dynamic programming property that solves the value and optimal policy locally. Additionally, it consecutively executes actions for multiple steps in the trajectory, which is beneficial to reduce computational burden and lower optimization difficulty. Simulations on a four-storey building layout demonstrates our method is efficient and suitable for on-site evacuation plan making.

11:00AM *Formation Control with Collision Avoidance through Deep Reinforcement Learning [#19932]*

Zezhi Sui, Zhiqiang Pu, Jianqiang Yi and Tianyi Xiong, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China

Generating collision free, time efficient paths for followers is a challenging problem in formation control with collision avoidance. Specifically, the followers have to consider both formation maintenance and collision avoidance at the same time. Recent works have shown the potentialities of deep reinforcement learning (DRL) to learn collision avoidance policies. However, only the collision factor was considered in the previous works. In this paper, we extend the learning-based policy to the area of formation control by learning a comprehensive task. In particular, a two-stage training scheme is adopted including imitation learning and reinforcement learning. A fusion reward function is proposed to lead the training. Besides, a formation-oriented network architecture is presented for environment perception and long short-term memory (LSTM) is applied to perceive the information of an arbitrary number of obstacles. Various simulations are carried out and the results show the proposed algorithm is able to anticipate the dynamic information of the environment and outperforms traditional methods.

11:20AM *Strategy Selection in Complex Game Environments Based on Transfer Reinforcement Learning [#20395]*

Hongwei Ge, Mingde Zhao, Kai Zhang and Liang Sun, Dalian University of Technology, China; McGill University, Canada

Boosting the learning process in the new task by making use of previously obtained knowledge has been a challenging task in many fields of industrial engineering and scientific. In this paper, we propose a transfer reinforcement learning model with knowledge Inheritance and decision-making Assistance (trIA). In the stage of knowledge inheritance, trIA adopts a model that employs a simultaneous multi-task and multi-instance learning strategy to compress acquired experts knowledge from distinct task into a global multi-task agent. In the stage of decision-making assistance, trIA adopts a dualcolumn progressive neural network framework to fully utilize the previous knowledge in the global multi-task agent and the acquired knowledge in the new task. The experimental results on the Atari domain demonstrate that the proposed knowledge inheritance model can performed at nearly the same level as the experts on the distinct source task environments. The results also demonstrate that the decision-making assistance model can transfer knowledge from the source tasks to the target tasks effectively. Moreover, the comparative results with the state-of-the- art algorithms validate the effectiveness of the proposed trIA for strategy selection in complex game environments.

Thursday, July 18, 11:50AM-1:30PM

Plenary Poster Session: Poster Session 3

Thursday, July 18, 11:50AM-1:30PM, Room: Ballroom I + II +II, Chair: Khan M. Iftekharuddin

P501 *A Novel Two-Factor Attention Encoder-Decoder Network through Combining Temporal and Prior Knowledge for Weather Forecasting [#20141]*

Minglei Yuan, Xiaozhong Ji, Tong Lu, Pengfei Chen and Hualu Zhang, Nanjing University, China; Nari Group Corporation, China

This paper proposes a novel two-factor attention based encoder-decoder model (TwoFactorEncoderDecoder) for multivariate weather prediction. The proposed model learns attention weights from two factors, namely, temporal information and prior knowledge inferred information. Here, temporal information contains change patterns hidden in observed time series data, while prior knowledge inferred information gives various types of meteorological observations in weather forecasting. Attention weights of the two factors are used to select the intermediate outputs of the encoder, and then combine the selected result with information inferred by prior knowledge for weather forecasting by a more effective way. In addition, this paper proposes a loss function for multivariate prediction. Compared with Mean Square Error (MSE) loss function, the proposed loss function can fit small

variances more accurately in performing multivariate prediction. Compared with the attention model that only uses temporal information or the prior knowledge inferred information, the proposed TwoFactorEncoderDecoder model has encouraging improvements in prediction accuracy on the public weather forecasting dataset, namely, the MAPE of t2m is increased by 5.42%, the MAPE of rh2m is increased by 2.92%, and the MAPE of w2m is increased by 1.67%, which shows the effect of the two-factor attention mechanism. Source code for the complete system will be available at https://github.com/YuanMLer/TFAEncoderDecoder.

P502 *Synaptic Learning of Long-Term Cognitive Networks with Inputs [#20482]*

Richar Sosa, Alejandro Alfonso, Gonzalo Napoles, Rafael Bello, Koen Vanhoof and Ann Nowe, Artificial Intelligence Lab, Vrije Universiteit Brussel(VUB), Belgium; Universidad Central de Las Villas (UCLV), Cuba; Faculty of Business Economics, Hasselt University (UHasselt), Belgium

In contrast with the extense variety of machine learning algorithms, to fully automate the reasoning process, only a few can take advantage of the expert knowledge. Fuzzy Cognitive Maps (FCMs) are neural networks that can naturally integrate this kind of knowledge in the inference process. Nevertheless, FCMs have serious drawbacks difficult to overcome from the absence of an intrinsically learning algorithm or limited prediction horizon of the activation space of the neurons. Recently, some variants of the FCMs like Short-Term Cognitive Networks (STCN) and Long Term Cognitive Networks (LTCN) have been proposed to solve these problems. In this paper, we propose a new neural network model as a variant of LTCNs called Long-Term Cognitive Networks with Inputs (LTCNIs). A new kind of input neuron which is not present in the traditional FCMs approach or the derived algorithms STCNs and LTCNs is introduced, in order to model inputs like energy or mass in physical systems. The performance of the method is discussed through the modeling of a passive circuit problem. As a second contribution, a new flexible reasoning strategy, which preserves the expert knowledge through synaptic learning is presented. A synaptic learning based on a gradient descent method is implemented limited by a set of restrictions that preserves the model semantics.

P503 *A temporal encoding method based on expansion representation [#19470]*

Yan Dai, Mengwen Yuan, Huajin Tang and Rui Yan, College of Computer Science,Sichuan University, China

Temporal encoding of visual stimulus based on the spiking neural networks is important and challenging. Inspired by the representation of neuron population in the sensory pathway, we propose a new hierarchical encoding method, which amplifies the difference among spike trains through the synaptic projection matrix and leaves the top neurons keeping active by a winner-take-all scheme. Finally, the neurons are assigned to positive and negative subthreshold membrane oscillation to fire new spike trains. In the optical character recognition task, we compare the readout performance of the proposed encoding method and the phase encoding method. When the image is severely destroyed, for example, at 25% noise, our method still maintains 88% recognition accuracy, which is 30% higher than the accuracy of phase encoding. The robustness of readout neurons to perform recognition tasks is enhanced through the proposed encoding method. We performed the benchmark two class classification experiment on Caltech 101 dataset, and the accuracy can reach 99.4%, which is 17% higher than the phase encoding. Compared with the other spiking deep neural network, our method also shows great improvement and strong scalability.

P504 *Cellular Computational Network for Distributed Power Flow Inferencing in Electric Distribution Systems [#20374]*

Hasala Dharmawardena and Ganesh K. Venayagamoorthy, Clemson University, United States

modern power system is undergoing a rapid transformation from centralized generation to distributed generation. The distributed generation consists of a large number of small distributed energy resource (DER) based generators with stochastic power output connected at different geographic locations throughout the feeder. Therefore, the operation of the future distribution grid requires efficient, scalable, and robust techniques for system analysis and control. This study presents a data driven approach to solve the electric power flow problem based on the framework known as the Cellular Computational Network (CCN). It is scalable since diakoptics resolves the system to computational cells, which are then connected together to form the full system. The results for the IEEE 4 Bus system shows that this approach can generate accurate power flow solutions.

P505 *From Content Text Encoding Perspective: A Hybrid Deep Matrix Factorization Approach for Recommender System [#19654]*

Jianing Zhou, Junhao Wen, Shun Li and Wei Zhou, School of Big Data & Software Engineering, Chongqing University, China

Recommender systems usually make personalized recommendation for users by analyzing interaction ratings between users and items. In plenty of application domains, some implicit feedback, content text, and other additional auxiliary information are widely used for improving the performance of recommendation. Memory-based collaborative filtering approaches that are widely used like matrix factorization (MF) predicts a personalized ranking for an individual user by leveraging latent factor models to handle common problems in recommender systems. However, most of previous MF methods adopt only explicit ratings to make personalized recommendation, ignoring the importance of implicit feedback on both users and items in-formation. Latent factor model can be easily extended with content text by mapping text information into factors, therefore we propose an approach by using neural networks to acquire the latent factors. Meanwhile, recurrent neural networks (RNNs) are applied to convert the textual data to an auxiliary factor feature to promote the representation of items. The experimental results prove that our model is effective on two benchmark datasets, outperforming some state-of-art approaches.

P506 *Spatio-temporal Active Learning for Visual Tracking [#19498]*

Chenfeng Liu, Pengfei Zhu and Qinghua Hu, Tianjin University, China

The success of state-of-the-art deep learning based trackers is fuelled by the large-scale datasets. However, not all training data has a gain on model performance, and some data can even degrade the performance of the model. Therefore, we aim to train state-of-the-art trackers using a part of labeled frame with high information and less data. To this end, we propose a novel framework called STAL(spatio-temporal active learning strategy), which is integrated with an efficient deep tracker and a spatio- temporal active learning strategy. Specifically, we first mine the most informative frames to boost the deep tracker based on the corresponding spatio-temporal response scores of the target. Then, the hard and easy frames are labeled from annotation and machine auto- annotation, respectively. Hard and easy sample pairs are generated from selected frames. To alleviate the impact of sample pairs with large loss, a self-paced fully convolutional Siamese network is proposed by introducing a norm negative regularization. The STAL framework will converge well and output promising tracking performance on several publicly available datasets.

P507 *CARL: Aggregated Search with Context-Aware Module Embedding Learning [#20343]*

Xinting Huang, Jianzhong Qi, Yu Sun, Rui Zhang, Hai-Tao Zheng and Xiaojie Wang, The University of Melbourne, Australia; Twitter Inc., United States; Tsinghua University, China

Aggregated search aims to construct search result pages (SERPs) from blue-links and heterogeneous modules (such as news, images, and videos). Existing studies have largely ignored the correlations between blue-links and heterogeneous modules when selecting the heterogeneous modules to be presented. We observe that the top ranked blue-links, which we refer to as the context, can provide important information about query intent. For example, informative terms like "streamed" and "recorded" in the context imply that a video module may better satisfy the query. To model and utilize the context information for aggregated search, we propose a model with context attention and representation learning (CARL). Our model applies a recurrent neural network with attention mechanism to encode the context, and incorporate the encoded context information into module embeddings. The context-aware module embeddings together with the ranking policy are jointly optimized under the Markov decision process (MDP) formulation. To achieve a more effective joint learning, we further propose an optimization function with self-supervision loss to provide auxiliary supervision signals. Experimental results based on two public datasets demonstrate the superiority of CARL over multiple baseline approaches, and confirm the effectiveness of the proposed optimization function in boosting the joint learning process.

P508 *Continuous Gesture Recognition through Selective Temporal Fusion [#19974]*

Pradyumna Narayana, Ross Beveridge and Bruce Draper, Colorado State University, United States

Gesture recognition is an important task with the potential to revolutionize human/computer interfaces (HCI). Gestures, however, are dynamic. While a few gestures may be static poses, most gestures are complex temporal sequences of motions. For most HCI applications, gestures must be recognized in real-time in streaming data. Therefore, most recognition systems analyze each frame as it comes in, fusing data across time to detect gestures. This paper presents results of the first systematic study of temporal fusion techniques for streaming gesture recognition. These results show that the choice of the best fusion strategy depends on whether the input is global (i.e. full-frame) or a spatially focused window, and on whether the input is unprocessed RGB or depth depth versus a flow field. This conclusion is then used to extend a state-of-the-art architecture for isolated gesture recognition, FOANet, to continuous gesture recognition. The result is a system that established a new state-of-the-art for recognition performance on the ChaLearn ConGD data set, with a mean Jaccard Index of 0.77 compared to the previous best result of 0.61. This paper also establishes a baseline of performance for the newer, continuous version of the NVIDIA dataset.

P509 *AuxBlocks: Defense Adversarial Examples via Auxiliary Blocks [#20403]*

Yueyao Yu, Pengfei Yu and Wenye Li, The Chinese University of Hong Kong, Shenzhen, China

Deep learning models are vulnerable to adversarial examples, which poses an indisputable threat to their applications. However, recent studies have observed that gradient-masking defenses are self-deceiving methods if an attacker can realize this defense. In this paper, we propose a new defense method based on appending information. We introduce the Aux Block model to produce extra outputs as a self-ensemble algorithm and analytically investigate the robustness mechanism of Aux Block. We have empirically studied the efficiency of our method against adversarial examples in two types of white-box attacks, and found that even in the full white-box attack where an adversary can craft malicious examples from defense models, our method has a more robust performance of about 54.6% precision on Cifar10 dataset and 38.7% precision on Mini-Imagenet dataset. Another advantage of our method is that it is able to maintain the prediction accuracy of the classification model on clean images, and thereby exhibits its high potential in practical applications.

P510 *TA-STAN: A Deep Spatial-Temporal Attention Learning Framework for Regional Traffic Accident Risk Prediction [#19880]*

Lei Zhu, Tianrui Li and Shengdong Du, Southwest Jiaotong University, China

Accurate and effective prediction of future traffic accident risks is critical to reducing the number of traffic accidents, which is also of great help to personal safe travel. Different from the previous traffic accident prediction research, we divide the traffic flow into multiple traffic volumes according to the vehicle type, since the road traffic flow is roughly regarded as a single influencing factor. In order to better model the dynamic impact of different traffic flow data and traffic accident data in the local region and global regions for future traffic accident risks prediction, we design a deep learning framework to predict regional Traffic Accident risk that includes a Spatial-Temporal Attention Network (named TA-STAN), which incorporates a spatial-temporal attention mechanism including local spatial attention, global spatial attention and temporal attention. We also integrate many external environmental factors to further improve the accuracy of the model. We evaluate the proposed TA-STAN model on the real traffic accident dataset in New York City. The experimental results show that TA-STAN outperforms the 6 baseline models in 3 evaluation metrics. More importantly, by visualizing the weight of attention, we can reasonably interpret the actual meaning of attention weights, which plays a crucial role in our model.

P511 *Simulating Brain Signals: Creating Synthetic EEG Data via Neural-Based Generative Models for Improved SSVEP Classification [#20251]*

Nik Khadijah Nik Aznan, Amir Atapour-Abarghouei, Stephen Bonner, Jason Connolly, Noura Al Moubayed and Toby Breckon, Durham University, United Kingdom

Despite significant recent progress in the area of Brain-Computer Interface (BCI), there are numerous shortcomings associated with collecting Electroencephalography(EEG) signals in real-world environments. These include, but are not limited to, subject and session data variance, long and arduous calibration processes and predictive generalisation issues across different subjects or sessions. This implies that many downstream applications, including Steady State Visual Evoked Potential(SSVEP) based classification systems, can suffer from a shortage of reliable data. Generating meaningful and realistic synthetic data can therefore be of significant value in circumventing this problem. We explore the use of modern neural-based generative models trained on a limited quantity of EEG data collected from different subjects to generate supplementary synthetic EEG signal vectors, subsequently utilised to train an SSVEP classifier. Extensive experimental analysis demonstrates the efficacy of our generated data, leading to improvements across a variety of evaluations, with the crucial task of cross-subject generalisation improving by over 35% with the use of such synthetic data.

P512 *SFSegNet: Parse Freehand Sketches using Deep Fully Convolutional Networks [#19360]*

Junkun Jiang, Ruomei Wang, Shujin Lin and Fei Wang, School of Data and Computer Science, Sun Yat-Sen University, China; School of Communication and Design, Sun Yat-Sen University, China

Parsing sketches via semantic segmentation is attractive but challenging, because (i) free-hand drawings are abstract with large variances in depicting objects due to different drawing styles and skills; (ii) distorting lines drawn on the touchpad make sketches more difficult to be recognized; (iii) the high-performance image segmentation via deep learning technologies needs enormous annotated sketch datasets during the training stage. In this paper, we propose a Sketch-target deep FCN Segmentation Network(SFSegNet) for automatic free-hand sketch segmentation, labeling each sketch in a single object with multiple parts. SFSegNet has an end-to-end network process

between the input sketches and the segmentation results, composed of 2 parts: (i) a modified deep Fully Convolutional Network(FCN) using a reweighting strategy to ignore background pixels and classify which part each pixel belongs to; (ii) affine transform encoders that attempt to canonicalize the shaking strokes. We train our network with the dataset that consists of 10,000 annotated sketches, to find an extensively applicable model to segment stokes semantically in one ground truth. Extensive experiments are carried out and segmentation results show that our method outperforms other state-of-the-art networks.

P513 *Absolute Human Pose Estimation with Depth Prediction Network [#19559]*

Marton Veges and Andras Lorincz, Eotvos Lorand University, Hungary

The common approach to 3D human pose estimation is predicting the body joint coordinates relative to the hip. This works well for a single person but is insufficient in the case of multiple interacting people. Methods predicting absolute coordinates first estimate a root-relative pose then calculate the translation via a secondary optimization task. We propose a neural network that predicts joints in a camera centered coordinate system instead of a root-relative one. Unlike previous methods, our network works in a single step without any post-processing. Our network beats previous methods on the MuPoTS-3D dataset and achieves state-of-the-art results.

P514 *DR-NET: A Stacked Convolutional Classifier Framework for Detection of Diabetic Retinopathy [#20457]*

Sathiya Narayan Chakravarthy, Himanshu Singhal and Narasimha Yadav R.P., SSN College of Engineering, India

Diabetic Retinopathy is the main cause of vision impairment among people suffering from diabetes and often leads to blindness. It has no early warning signs. Hence, it is of utmost importance to detect it as early as possible in order to provide adequate treatment. Most of the current research in this field focuses on a manual process of feature extraction such as annotation of lesions and optic disk segmentation so as to detect the presence of DR. In this paper, a framework DR-NET using stacked convolutional neural networks for diabetic retinopathy detection from digital fundus images is proposed. A network consisting of convolutional layers with different filters stacked in parallel, the output of which is concatenated and global max pooling performed on it, is developed. This architecture helps extract intricate features during the classification task along with minimizing the learnable parameters and reduces overfitting, thus, improving the overall performance of the model. Various preprocessing methods were applied to further improve accuracy. Visualization techniques were also used to gain insights into the learning of the model. The experimental results were performed on about 12,000 images which were an ensemble of various online datasets, yielded an accuracy of 81% and a kappa score of 0.6.

P515 *Convolutional Neural Network based Eye Recognition from Distantly Acquired Face Images for Human Identification [#19551]*

Kazi Shah Nawaz Ripon, Lasker Ershad Ali, Nazmul Siddique and Jinwen Ma, Norwegian University of Science and Technology, Norway; Khulna University, Bangladesh; University of Ulster, United Kingdom; Peking University, China

Eye image recognition from a face image acquired at a distant is a promising physical biometric technique to use for human identification. This contemporary field of research depends on image preprocessing, feature extraction, and reliable classification techniques. In this work, we separate eye images from an image of the entire face of a subject and then extract features from these eye images utilizing a convolutional neural network (CNN) model. In general, CNN models convolve images in different layers to extract effective features and then use the softmax function to produce a probability output in the final layer. In our approach, we use CNN features and a kernel extreme learning machine (KELM) classifier instead of softmax to modify the original CNN model. The modified CNN-KELM model has been verified using the publicly available CASIA.v4 distance image database. The experimental results demonstrate that our proposed approach obtains a satisfactory recognition result when compared with several current state-of-the-art human identification approaches.

P516 *Competitive Online Generalised Linear Regression with Multidimensional Outputs [#19874]*

Raisa Dzhamtyrova and Yuri Kalnishkan, Royal Holloway, University of London, United Kingdom

We apply online prediction with expert advice to construct a universal algorithm for multi-class classification problem. Our experts are generalised linear regression models with multidimensional outputs, i.e. neural networks with multiple output nodes but no hidden nodes. We allow the final layer transfer function to be a softmax function with the linear activations to all the output neurons. We build an online algorithm competitive with all the experts of the relevant models of this type and derive the upper bound on the cumulative loss of the algorithm. We carry on experiments on three data sets and compare the cumulative losses of our algorithm and single neuron with multiple output nodes.

P517 *GMM-based Undersampling and Its Application for Credit Card Fraud Detection [#19370]*

Fengjun Zhang, Guanjun Liu, Zhenchuan Li, Chungang Yan and Changjun Jiang, Tongji University, China

The class imbalance problem exists in many real-world applications such as fraud detection, medical diagnosis and spam filtering, and seriously influences the performance of learning algorithms. Randomly undersampling is a famous method to solve the problem. However, it cannot well extract the samples nearby the cross-edge of majority and minority classes due to its randomness, while these samples are very important for a classifier since they influence the classification performance. In this paper, we propose a novel Gaussian Mixture Undersampling (GMUS for short). GMUS mainly contains three steps. Firstly, a Gaussian Mixture Model (GMM) is applied to fit the majority samples. Secondly, considering the probability density function (PDF) of predicted minority samples on the well-fitted GMM, the maximum of PDF is selected as the cross-edge of two classes. Finally, we undersample the majority samples near the cross-edge. We do experiments on 16 public datasets and the results demonstrate that GMUS can sample more informative instances and thus improve the performance of classifiers compared with the state-of-the-art undersampling methods. We also apply GMUS to the credit card fraud detection and obtain a good performance.

P518 *Efficient and Robust Convolutional Neural Networks via Channel Prioritization and Path Ensemble [#19404]*

Chun-Min Chang, Chia-Ching Lin and Kuan-Ta Chen, Institute of Information Science, Academia Sinica, Taiwan

With the growing recognition of both efficiency and security issues in machine learning models, we propose a novel convolutional neural networks (CNNs) training algorithm, called channel prioritization and path ensemble (CPPE), to not only allow dynamically trade-offs between different resource and performance requirements but also enable secure inference without any extra computational cost or memory overhead. Our approach not only prioritizes channels to prune the network in a structured way and ensemble multiple inference paths over different utilization conditions. We demonstrated the effectiveness of channel prioritization by the experiment of the VGG-16 network on various benchmark datasets. The experimental results show that, on the CIFAR-10 dataset, a 10x parameters reduction and a 4x FLOPs reduction can be achieved, with only a 0.2% accuracy drop. Furthermore, we allow CNNs to dynamically trade-offs between resource demand and accuracy with only 4% degradation in accuracy in exchange for 16x FLOPs reduction. By ensembling multiple inference paths, our model can improve robustness against various adversarial attacks without any additional computational cost and memory overhead. Finally, our method is simple and easily applied to any convolutioanl neural networks.

P519 *Deep Generative State-Space Modeling of FMRI Images for Psychiatric Disorder Diagnosis [#20028]*
Koki Kusano, Tetsuo Tashiro, Takashi Matsubara and Kuniaki Uehara, Kobe University, Japan

An early and accurate diagnosis of psychiatric disorders is critical for patients' quality of life and deep understanding of the disorders. For this reason, many studies have proposed machine learning-based diagnostic procedures for functional magnetic resonance imaging (fMRI) data. Especially, these procedures often employed temporal models due to the time- varying nature of the brain activities and probabilistic generative models for understanding the underlying mechanism of the disorders. For leveraging the recent advantage of deep learning, we proposed a state-space model of fMRI images based on deep learning. The proposed deep state-space model is more flexible than conventional models and less likely to suffer from overfitting than a straightforward deep learning-based classifier. The proposed model estimates the subjects' conditions more accurately than existing diagnostic procedures. Also, the proposed model potentially identifies brain regions related to the disorders.

P520 *Exploring Spatiotemporal Functional Connectivity Dynamics of the Human Brain using Convolutional and Recursive Neural Networks [#19362]*
Zachary Harper and Charles Welzig, Medical College of Wisconsin, United States; Tufts Medical Center, United States

Many cognitive pathologies can be characterized by abnormalities in functional connectivity. Deep artificial neural networks may provide an innovative tool, capable of modeling highly complex spatiotemporal connectivity dynamics. We combine convolutional and recursive artificial neural network architectures to differentiate functional connectivity patterns associated with high and low working memory load conditions. Source-level Pearson correlation coefficient- estimated parcellated connectomes, detected using magnetoencephalography, are processed as machine learning input. Initial findings show patterns of increasing artificial neural network accuracy linked to task evoked functional network coherence dynamics. Planned artificial neural network architecture optimizations may improve feature detection and projection of learned features onto input space for translational applications.

P521 *An Analysis on the Learning Rules of the Skip-Gram Model [#20415]*
Canlin Zhang, Xiuwen Liu and Daniel Bis, Florida State University, United States

In this paper, we derive the learning rules for the skip-gram model and establish their close relationship to competitive learning. In addition, we provide the global optimal solution constraints for the skip-gram model and validate them by experimental results.

P522 *Micro-states based dynamic brain connectivity in understanding the commonality and differences in gender-specific emotion processing [#19407]*
Rakib Al-Fahad and Mohammed Yeasin, The University of Memphis, United States

In this paper, we present a data-driven micro-states based complex network analysis on cortical surface data to understand the connectives in modeling elicited emotion. In particular, we focus on processing arousal in identifying the differences and similarities between males and females. Microstates are transient, patterned, quasi-stable states of a time series (e.g., Electroencephalography recording) that allows visualizing dynamic coupling and possibly time-varying levels of correlated or mutually informed activity between brain regions. To obtain cortical surface data from EEG recording, we use source localization method from Brainstorm. We adopted t-distributed stochastic neighbor embedding (t-SNE), for better visualization and finding the optimal number of clusters using different algorithms (e.g., Gaussian Mixture Models(GMM), K-means, etc). Centroids of these clusters are considered as micro-states. We used Hidden Markov Model (HMM) to compute the transition probabilities among micro-states. Subsequently, p-values on graph theoretic measures (e.g., modularity, small-worldness, etc) computed from micro-sites were used to determine the significantly distinguishable, highly segregated and densely integrated network of brain connectivity. Empirical analysis using DEAP dataset reveals that males and females have mostly complimentary micro-states with some commonalities. Males are more likely to stay in specific stable state and females are more likely to stay in transient states. Both groups utilize highly segregated and densely integrated network structure among brain regions in processing arousal.

P523 *Predicting Group Cohesiveness in Images [#19501]*
Shreya Ghosh, Abhinav Dhall, Nicu Sebe and Tom Gedeon, Indian Institute of Technology Ropar, India; University of Trento, Italy; Australian National University, Australia

The cohesiveness of a group is an essential indicator of the emotional state, structure and success of a group of people. We study the factors that influence the perception of group-level cohesion and propose methods for estimating the human-perceived cohesion on the group cohesiveness scale. In order to identify the visual cues (attributes) for cohesion, we conducted a user survey. Image analysis is performed at a group-level via a multi-task convolutional neural network. For analyzing the contribution of facial expressions of the group members for predicting the Group Cohesion Score (GCS), a capsule network is explored. We add GCS to the Group Affect database and propose the 'GAF- Cohesion database'. The proposed model performs well on the database and is able to achieve near human-level performance in predicting a group's cohesion score. It is interesting to note that group cohesion as an attribute, when jointly trained for group-level emotion prediction, helps in increasing the performance for the later task. This suggests that group-level emotion and cohesion are correlated.

P524 *Evaluating Incomplete DCOP Algorithms On Large-Scale Problems [#19110]*
Allan Leite and Fabricio Enembreck, Pontifical Catholic University of Parana (PUCPR), Brazil

The distributed constraint optimization problem (DCOP) has emerged as one of the most promising coordination techniques in multi-agent systems (MAS). However, because DCOP is known to be NP-hard, the existing DCOP techniques are often unsuitable for large-scale applications, which require scalable algorithms to deal with severely limited computing and communication. Moreover, the selection of DCOP algorithm is a challenging and critical task for obtaining a desirable performance on certain MAS domains. In this paper, we present a performance analysis of incomplete DCOP algorithms on large-scale DCOPs. We experimentally evaluate the state-of-the-art incomplete algorithms on two types of problems involving hundreds of variables with different network topologies and densities. Such performance analysis can help to mitigate the challenges of selection of algorithm for a number of realistic large-scale, complex MAS applications.

P525 *CSSD: Cascade Single Shot Face Detector [#19310]*
Shuainan Wang, Tong Xu, Wei Li and Haifeng Sun, Beijing University of Posts and Telecommunications, China

Face detection has achieved great success with the development of convolution neural network. However, it remains a challenging problem to detect small and blurred faces in unconstrained environment. This paper presents a novel cascade single-shot face detector, named Cascade Single Shot Face Detector (CSSD), which introduces novel cascade classification and regression network in an anchor-based face detector to reject false positives and improve location accuracy. We have contributed in the following three aspects: 1) proposing a feature enchanted and scale-invariable face detection architecture to process faces with different scales; 2) regressing bounding boxes of faces in two steps with a cascade method; 3) filtering negative anchors online after anchor refinement and rebalancing

puzzle negative anchors and positive anchors with rate of 3:1. As a consequence, our method achieves state-of-the-art detection performance on FDDB and WIDER FACE dataset.

P526 *Missing Entity Synergistic Completion across Multiple Isomeric Online Knowledge Libraries [#20409]*

Bowen Dong, Jiawei Zhang, Chenwei Zhang, Yang Yang and Philip S. Yu, University of Illinois at Chicago, United States; Florida State University, United States; Beihang University, China

Online knowledge libraries refer to the online data warehouses that systematically organize and categorize the knowledge-based information about different kinds of concepts and entities. In the era of big data, the setup of online knowledge libraries is an extremely challenging and laborious task, in terms of efforts, time and expense required in the completion of knowledge entities. Especially nowadays, a large number of new knowledge entities, like movies, are keeping on being produced and coming out at a continuously accelerating speed, which renders the knowledge library setup and completion problem more difficult to resolve manually. In this paper, we will take the online movie knowledge libraries as an example, and study the multiple aligned Isomeric Online Knowledge Libraries Completion problem. We aim at identifying the missing entities for multiple knowledge libraries synergistically and ranking them for editing based on certain ranking criteria. To solve the problem, a thorough investigation of two isomeric online knowledge libraries, Douban and IMDB, have been carried out in this paper. Based on analyses results, a novel deep online knowledge library completion framework is introduced to solve the problem. By projecting the entities from multiple isomeric knowledge libraries to a shared feature space, our model solves the problem via three steps: (1) entity feature space unification via embedding, (2) knowledge library fusion based missing entity identification, and (3) missing entity ranking. Extensive experiments done on the real-world online knowledge library dataset have demonstrated the effectiveness of our model in addressing the problem.

P527 *Real-time Accurate Object Counting for Smart Farms [#19730]*

Hao Shang, Rui Li, Xu He, Jilong Wang and Xinhui Peng, Hunan University, China

Automatic methods that count the number of living bodies is extremely valuable for modern farms. Recently, there are an increasing number of studies conducted on imagery based automatic object counting. However, these studies suffer from extreme overlap of objects, scene perspective, the size of instances and etc. Moreover, real-time performance needs to be met in many application scenarios. In this study, we propose a video imagery based pipeline for real-time accurate object counting on smart farms. Firstly, we introduce the concept - overlap degree of a frame (ODF). Next, an end-to-end frame filtering algorithm based on ODF is established. Then, we propose an object detection and counting algorithm based on Faster region-based convolutional neural networks (Faster RCNN). Finally, the counting results are fused to accurately obtain the number of detected objects. The experimental results demonstrate that our proposed method can automatically and accurately count the living bodies on farms in real-time. In our experiments, the average speed of detection is 40 frames per second, and the total difference is 1.69%. Therefore, our method is considered to be a suitable tool to conduct automatic object counting on smart farms.

P528 *Sports Motion Recognition based on Foot Trajectory State Sequence Mapping [#20127]*

Lingjia Huang, Hao Ma, Weichao Yan, Wuda Liu, Haoyang Liu and Zaiyue Yang, Southern University of Science and Technology, China; Noitom Ltd, China; Beijing Sport University, China

Quantitative motion analysis to evaluate the performance of athletes has been actively studied recently. Although various methods based on wearable inertial sensors have been developed for simple and repetitive movements recognition, the understanding of continuous complex movements of in-field sports is still challenging. In this paper, we propose a new motion segmentation and recognition method based on foot swing trajectory state to achieve robust and efficient recognition of motion of interest (MOI) in the lower limbs from continuous and complex movements. In order to segment complex movements in the lower limbs, a series of foot motion states are defined based on foot-ground contact status and foot trajectory during swing. The lower body motion state sequence combining the states of both feet is matched to a prior knowledge of MOI cycle sequences obtained in advance, so as to obtain a motion type candidate set. In this case, the continuous movement is segmented based on the prescreened motion types to realize adaptive time window for feature extraction. Finally, according to the prescreened motion type candidate set, corresponding trained neural network binary classifiers are used to make the classification based on the calculated kinematic features. The proposed method is verified through experiments of football movements consisting of walking, dribbling and stepover. As the result, the motion type recognition accuracy is 95%.

P529 *On Dissimilarity Representation and Transfer Learning for Offline Handwritten Signature Verification [#19342]*

Victor L. F Souza, Adriano L. I. Oliveira, Rafael M. O. Cruz and Robert Sabourin, Centro de Informatica - Universidade Federal de Pernambuco, Brazil; Stradigi AI, Canada; Ecole de Technologie Superieure - Universite du Quebec, Canada

When compared to Writer-Dependent (WD) Handwritten Signature Verification, in which a model is trained for each individual writer, the Writer-Independent (WI) approach offers greater scalability, since only a single model is trained for all users from a dissimilarity space generated by the dichotomy transformation. However, many samples from the dissimilarity space are redundant and have little influence during the training of the verification model. This work investigates whether prototype selection (PS) preprocessing can be used in the space resulting from the dichotomy transformation without degrading the performance of the classifier. Furthermore, an investigation is also performed to examine the use of a WI classifier in a transfer learning scenario, i.e., where the classifier is trained in one dataset, and is used to verify signatures in other datasets. The experiments reported herein show that the use of prototype selection in the dissimilarity space allows a reduction in the complexity of the classifier without degrading its generalization performance. In addition, the results show that the WI classifier is scalable enough to be used in a transfer learning approach, with a resulting performance comparable to that of a classifier trained and tested in the same dataset. An analysis of the results obtained based on the instance hardness (IH) measure and dendrogram diagrams is performed in order to better understand the behavior of the resulting dichotomy transformation.

P530 *Adaptive Neural Network Time-varying Formation Tracking Control for Multi-agent Systems via Minimal Learning Parameter Approach [#19935]*

Tianyi Xiong, Zhiqiang Pu, Jianqiang Yi and Zezhi Sui, School of Artificial Intelligence, University of Chinese Academy of Sciences; Institute of Automation, Chinese Academy of Sciences, China

This paper investigates the time-varying formation tracking control problem for multi- agent systems with consideration of model uncertainties. For each dimension of an agent, a radial basis function neural network (RBFNN) is first adopted to approximate the model uncertainties online. Taking the square of the norm of the neural network weight vector as a newly developed adaptive parameter, a novel RBFNN-based adaptive control law with minimal learning parameter (MLP) approach is then constructed to tackle the time-varying formation tracking problem. The uniformly ultimately boundedness (UUB) of formation tracking errors is guaranteed through Lyapunov analysis. Compared with other traditional RBFNN-based formation tracking control laws for multi-agent systems, very few parameters need to be updated online in our proposed one, which can greatly lessen the computational burden.

Finally, comparative simulation results demonstrate the effectiveness and superiority of the proposed adaptive control law.

P531 *Celebrities-ReID: A Benchmark for Clothes Variation in Long-Term Person Re-Identification [#19581]*

Yan Huang, Qiang Wu, Jingsong Xu and Yi Zhong, University of Technology, Sydney, Australia

This paper considers person re-identification (re-ID) in the case of long-time gap (i.e., long-term re-ID) that concentrates on the challenge of clothes variation of each person. We introduce a new dataset, named Celebrities-reID to handle that challenge. Compared with current datasets, the proposed Celebrities- reID dataset is featured in two aspects. First, it contains 590 persons with 10,842 images, and each person does not wear the same clothing twice, making it the largest clothes variation person re-ID dataset to date. Second, a comprehensive evaluation using state of the arts is carried out to verify the feasibility and new challenge exposed by this dataset. In addition, we propose a benchmark approach to the dataset where a two-step fine-tuning strategy on human body parts is introduced to tackle the challenge of clothes variation. In experiments, we evaluate the feasibility and quality of the proposed Celebrities-reID dataset. The experimental results demonstrate that the proposed benchmark approach is not only able to best tackle clothes variation shown in our dataset but also achieves competitive performance on a widely used person re-ID dataset Market1501, which further proves the reliability of the proposed benchmark approach.

P532 *GCGAN: Generative Adversarial Nets with Graph CNN for Network-Scale Traffic Prediction [#19230]*

Yuxuan Zhang, Senzhang Wang, Bing Chen and Jiannong Cao, Nanjing University of Aeronautics and Astronautics, China; Nanjing University of Aeronautics and Astronautics & The Hong Kong Polytechnic University, China; The Hong Kong Polytechnic University, China

Traffic prediction is practically important to facilitate many real applications in urban areas such as relieving traffic congestion. Traditional traffic prediction models are mostly statistic based methods, and they cannot effectively capture the nonlinear, stochastic and time-varying characteristics of the urban transportation systems. Another limitation of these methods is that they are not capable to predict the traffic conditions of all the road segments in a large transportation network of a city as a whole. Therefore, in recent years, deep neural network based methods for forecasting the road network-scale traffic have been emphasized greatly. However, most existing deep neural network methods suffer from the blurry prediction issue and do not perform well on the task of multi-step traffic prediction. In this paper, We propose a network-scale deep traffic prediction model called GCGAN by combining adversarial training and graph CNN. Specifically, we propose a Generative Adversarial Net based prediction framework to address the blurry prediction issue by introducing the adversarial training loss. To predict the traffic conditions in multiple future time intervals simultaneously, we design a sequence to sequence model as the generator of GCGAN. To fully capture the spatial correlations among the road segments of a transportation network, we propose to apply a graph convolution network in both generator and discriminator of GCGAN for feature learning. We evaluate our proposal over a large real traffic dataset in the arterial road network of downtown Chicago. The results show that GCGAN significantly outperforms both traditional statistic based methods and recent state-of-the-art deep learning methods.

P533 *Nonlinear Transformation for Multiple Auxiliary Information in Music Recommendation [#20258]*

Junwei Zhang, Min Gao, Junliang Yu, Xinyi Wang, Yuqi Song and Qingyu Xiong, Chongqing University, China; The University of Queensland, Australia; Chingqing University, China

Online music recommender systems are becoming increasingly prevalent because of the popularity of digital music and music recommendation generally caters to users by discovering songs that match their preferences. However, these systems have to face a challenge: how to recommend new songs in a situation where prior knowledge is scarce. Some researches take auxiliary information into consideration in new recommendation approaches to deal with this problem. Nevertheless, they rarely pay attention to complex relationships among different feature spaces when they map those information to a latent space. To this end, this paper proposes an approach that uses non-linear transformation to integrate different auxiliary information into the songs' latent representations. Unlike other studies which directly map auxiliary information to the feature space, the proposed music recommendation model (NeuTrans) maps different information features to low-dimensional vector representation by non-linear neural networks. Specifically, the NeuTrans separately employs matrix factorization and attribute network embedding to extract auxiliary information (historical interaction, network structure and attributes of songs). The feature space of different information is obtained by nonlinearly mapping the feature space of the songs. Experimental analysis on two real-world datasets shows that our framework outperforms the state-of-the-art approaches on Top-N music recommendation.

P534 *Deep Learning-Based Strategy For Macromolecules Classification with Imbalanced Data from Cellular Electron Cryotomography [#19400]*

Ziqian Luo, Xiangrui Zeng, Zhipeng Bao and Min Xu, Beijing University of Posts and Telecommunications, China; Carnegie Mellon University, United States; Tsinghua University, China

Deep learning model trained by imbalanced data may not work satisfactorily since it could be determined by major classes and thus may ignore the classes with small amount of data. In this paper, we apply deep learning based imbalanced data classification for the first time to cellular macromolecular complexes captured by Cryo-electron tomography (Cryo-ET). We adopt a range of strategies to cope with imbalanced data, including data sampling, bagging, boosting, Genetic Programming based method and. Particularly, inspired from Inception 3D network, we propose a multi-path CNN model combining focal loss and mixup on the Cryo-ET dataset to expand the dataset, where each path had its best performance corresponding to each type of data and let the network learn the combinations of the paths to improve the classification performance. In addition, extensive experiments have been conducted to show our proposed method is flexible enough to cope with different number of classes by adjusting the number of paths in our multi-path model. To our knowledge, this work is the first application of deep learning methods of dealing with imbalanced data to the internal tissue classification of cell macromolecular complexes, which opened up a new path for cell classification in the field of computational biology.

P535 *VN-GAN: Identity-preserved Variation Normalizing GAN for Gait Recognition [#19476]*

Peng Zhang, Qiang Wu and Jingsong Xu, University of Technology Sydney, Australia

Gait is recognized as a unique biometric characteristic to identify a walking person remotely across surveillance networks. However, the performance of gait recognition severely suffers challenges from view angle diversity. To address the problem, an identity-preserved Variation Normalizing Generative Adversarial Network (VN-GAN) is proposed for learning purely identity-related representations. It adopts a coarse-to-fine manner which firstly generates initial coarse images by normalizing view to an identical one and then refines the coarse images by injecting identity-related information. In

specific, Siamese structure with discriminators for both camera view angles and human identities is utilized to achieve variation normalization and identity preservation of two stages, respectively. In addition to discriminators, reconstruction loss and identity- preserving loss are integrated, which forces the generated images to be the same in view and to be discriminative in identity. This ensures to generate identity-related images in an identical view of good visual effect for gait recognition. Extensive experiments on benchmark datasets demonstrate that the proposed VN- GAN can generate visually interpretable results and achieve promising performance for gait recognition.

P536 *On the Linear Separability of Random Points in the d-dimensional Spherical Layer and in the d-dimensional Cube [#19253]*

Sergey Sidorov and Nikolai Zolotykh, Lobachevsky State University of Nizhni Novgorod, Russia

A.N. Gorban and I.Y. Tyukin (Neural Networks 94, 255--259 (2017)) propose a method for correcting errors of artificial intelligence systems by separating erroneous cases with the Fisher linear discriminant. It turned out that if the dimension is large this approach works well even for an exponential (of the dimension) number of samples. In this paper, we specify the limits of applicability of this approach by estimating the number of points that are linearly separable with a probability close to 1 in two particular cases: when the points drawn randomly, independently and uniformly from a d-dimensional spherical layer and from the d-dimensional cube. Our bounds for these two cases improve some bounds obtained by A.N. Gorban and I.Y. Tyukin.

P537 *Deep Convolutional Neural Networks for Text Localisation in Figures From Biomedical Literature [#20388]*

Ibrahim Almakky, Vasile Palade and Ariel Ruiz-Garcia, Coventry University, United Kingdom

Text contained within figures is an important source of information in biomedical literature. Despite this, end-to-end text extraction from biomedical figures remains a challenging task. This paper presents a novel approach to address the founding block of this task, text detection, not only from biomedical figures but also from images in general. Particularly, the paper proposes an approach that simplifies the text detection problem into a reconstruction problem using a deep convolutional neural network. Designed to overcome the specific challenges of text detection from biomedical figures, our proposed model reports promising results on the DETEXT dataset.

P538 *Urban Area Vehicle Re-Identification With Self-Attention Stair Feature Fusion and Temporal Bayesian Re-Ranking [#19325]*

Chenghuan Liu, Du Huynh and Mark Reynolds, University of Western Australia, Australia

Vehicle re-identification (Re-ID) plays a key role in many smart traffic management systems. Re-identifying a vehicle can be very challenging because the differences in visual appearances between pairs of vehicles are sometimes extremely subtle if they have the same colour and even the same model. Given an image of a vehicle, most existing techniques adopt a global feature representation where details may be ignored. In this paper, we propose a Self- Attention Stair Feature Fusion model to learn the discriminative features for vehicle Re-ID. The model is designed to extract multi-level features in order to capture as much small details as possible. We also propose a Temporal Bayesian Re-Ranking method to exploit the spatial-temporal information in the vehicles' travel patterns. Our algorithm has been tested against state-of-the-art techniques on popular benchmarks. The results show that our algorithm outperforms other state-of-the-art techniques by a large margin.

P539 *Combining convolutional side-outputs for road image segmentation [#20252]*

Felipe Reis, Raquel Almeida, Ewa Kijak, Simon Malinowski, Silvio Jamil F. Guimaraes and Zenilton Patrocinio Jr., Pontifical Catholic University of Minas Gerais, Brazil; Univ Rennes, Inria, CNRS, IRISA, France

Image segmentation consists in creating partitions within an image into meaningful areas and objects. It can be used in scene understanding and recognition, in fields like biology, medicine, robotics, satellite imaging, amongst others. In this work we take advantage of the learned model in a deep architecture, by extracting side-outputs at different layers of the network for the task of image segmentation. We study the impact of the amount of side- outputs and evaluate strategies to combine them. A post-processing filtering based on mathematical morphology idempotent functions is also used in order to remove some undesirable noises. Experiments were performed on the publicly available KITTI Road Dataset for image segmentation. Our comparison shows that the use of multiples side outputs can increase the overall performance of the network, making it easier to train and more stable when compared with a single output in the end of the network. Also, for a small number of training epochs (500), we achieved a competitive performance when compared to the best algorithm in KITTI Evaluation Server.

P540 *Exploiting Action-Value Uncertainty to Drive Exploration in Reinforcement Learning [#19466]*

Carlo D'Eramo, Andrea Cini and Marcello Restelli, Politecnico di Milano, Italy

Most of the research in Reinforcement Learning (RL) focuses on balancing exploration and exploitation. Indeed, the reasons for the success or failure of an RL algorithm often deal with the choice between the execution of exploratory actions and the exploitation of actions that are known to be good. In the context of Multi-Armed Bandits (MABs), many algorithms have addressed this dilemma. In particular, Thompson Sampling (TS) is a solution that, besides having good theoretical properties, usually works very well in practice. Unfortunately, the success of TS in MAB problems has not been replicated in RL, where it has shown to scale very poorly w.r.t. the dimensionality of the problem. Nevertheless, the application of TS in RL, instead of more myopic strategies such as epsilon-greedy, remains a promising solution. This paper addresses such issue proposing several algorithms to use TS in RL and deep RL in a feasible way. We present these algorithms explaining the intuitions and theoretical considerations behind them and discussing their advantages and drawbacks. Furthermore, we provide an empirical evaluation on an increasingly complex set of RL problems, showing the benefit of TS w.r.t. other sampling strategies available in classical and more recent RL literature.

P541 *Curse of Dimensionality in Adversarial Examples [#19975]*

Nandish Chattopadhyay, Anupam Chattopadhyay, Sourav Sen Gupta and Michael Kasper, Nanyang Technological University & Fraunhofer Singapore, Singapore; Nanyang Technological University, Singapore; Fraunhofer Singapore, Singapore

While machine learning and deep neural networks in particular, have undergone massive progress in the past years, this ubiquitous paradigm faces a relatively newly discovered challenge, adversarial attacks. An adversary can leverage a plethora of attacking algorithms to severely reduce the performance of existing models, therefore threatening the use of AI in many safety-critical applications. Several attempts have been made to try and understand the root cause behind the generation of adversarial examples. In this paper, we try to relate the geometry of the high-dimensional space in which the model operates and optimizes, and the properties and problems therein, to such adversarial attacks. We present the mathematical background, the intuition behind the existence of adversarial examples and substantiate them with empirical results from our experiments.

P542 *Improve L2-normalized Softmax with Exponential Moving Average [#19582]*

Xuefei Zhe, Le Ou-Yang and Hong Yan, City University of Hong Kong, Hong Kong; Shenzhen University, China

In this paper, we propose an effective training method to improve the performance of L2-normalized softmax for convolutional neural networks. Recent studies of deep learning show that by L2-normalizing the input features of softmax, the accuracy of CNN can be increased. Several works proposed novel loss functions based on the L2-normalized softmax. A common property shared by these modified normalized softmax models is an extra set of parameters introduced as the class center. Although the physical meaning of this parameter is clear, few attentions have been paid to how to learn these class centers, which limits the performance improvements brought by L2-normalized softmax. In this paper, we address the problem of learning the class centers in the L2-normalized softmax. By treating the CNN training process as a time series, we propose a novel learning algorithm that combines the general used gradient descent with the exponential moving average. Extensive experiments show that our model not only achieves better performance but also has a higher tolerance to the imbalance data.

P543 *A Character-Enhanced Chinese Word Embedding Model [#20429]*

Gang Yang, Hongzhe Xu, Tianhao He and Zaishang Cai, Xi'an Jiaotong University, China

Distributed word representation has demonstrated its advantages in many natural language processing tasks. Such as named entity recognition, entity relation extraction, and text classification. Traditional one-hot word representation represents a word as a high-dimensional and sparse vector. Instead, distributed word representation represents a word as a low-dimensional and dense vector, which are more suitable as inputs of deep neural networks. Furthermore, distributed word representation can express the semantic relatedness and syntactic regularities between different words. Word embedding is a kind of distributed word representation technology, which is very popular and useful in many natural language processing tasks. Recently, more and more researches have focused on learning word embeddings with internal morphological knowledge in words, such as character, subwords, and other kinds of morphological information. For example, Chinese characters contain rich semantic information related to words they compose. Thus, characters can help improving the representation of words. In this paper, we present a character- enhanced Chinese word embeddings model (CCWE). In the model, we train character and word embeddings simultaneously in two parallel tasks. The framework of our model is based-on Skip-Gram. We evaluate CCWE on word similarity, analogical reasoning, text classification, and named entity recognition. The results demonstrate that our model can learn both better Chinese word and character embeddings than other baseline models.

P544 *A Shortcut-Stacked Document Encoder for Extractive Text Summarization [#19289]*

Peng Yan, Linjing Li and Daniel Zeng, The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences, China; The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China

While doing summarization, human needs to understand the whole document, rather than separately understanding each sentence in the document. However, inter- sentence features within one document are not adequately modeled by previous neural network-based models that almost use only one layer recurrent neural network as document encoder. To learn high quality context-aware representation, we propose a shortcut-stacked document encoder for extractive summarization. We use multiple stacked bidirectional long short-term memory (LSTM) layers and add shortcut connections between LSTM layers to increase representation capacity. The shortcut-stacked document encoder is built on a temporal convolutional neural network-based sentence encoder to capture the hierarchical structure of the document. Then sentence representations encoded by document encoder are fed to a sentence selection classifier for summary extraction. Experiments on the well- known CNN/Daily Mail dataset show that the proposed model outperforms several recently proposed strong baselines, including both extractive and abstractive neural network-based models. Furthermore, the ablation analysis and position analysis also demonstrate the effectiveness of the proposed shortcut-stacked document encoder.

P545 *Towards a Smarter Fault Tolerant Indoor Localization System Through Recurrent Neural Networks [#19526]*

Eduardo Carvalho, Bruno Ferreira, Geraldo P. R. Filho, Pedro H. Gomes, Gustavo M. Freitas, Patricia A. Vargas, Jo Ueyama and Gustavo Pessin, SENAI Innovation Institute for Mineral Technologies, Brazil; University of Brasilia, Brazil; University of Southern California, United States; Federal University of Minas Gerais, Brazil; Heriot-Watt University, United Kingdom; University of Sao Paulo, Brazil; Instituto Tecnologico Vale, Brazil

This paper proposes a fault-tolerant indoor localization system that employs Recurrent Neural Networks (RNNs) for the localization task. A decision module is designed to detect failures and this is responsible for the allocation of RNNs that are suitable for each situation. As well as the fault-tolerant system, several architectures and models for RNNs are exploited in the system: Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM) and Simple RNN. The system uses as inputs a collection of Wi-Fi Received Signal Strength Indication (RSSI) signals, and the RNN classifies the position of an agent on the basis of this collection. A fault-tolerant mechanism has been designed to handle two types of failures: (i) momentary failure, and (ii) permanent failure. The results show that the RNNs are suitable for tackling the problem and that the whole system is reliable when employed for a series of failures.

P546 *Cropout: A General Mechanism for Reducing Overfitting on Convolutional Neural Networks [#19487]*

Wenbo Hou, Wenhai Wang, Ruo-Ze Liu and Tong Lu, Nanjing University, China

Recently, a lot of Convolutional Neural Networks(CNNs) have been proposed for computer vision applications. However, how to improve the generalization ability of them remains challenging. In this paper, we propose a novel mechanism, namely, Cropout, to further reduce overfitting on Convolutional Neural Networks. The proposed Cropout is able to enlarge the diversity of the feature-map produced by convolutional layer, and further improve the generalization ability of deep CNNs. It mainly consists of three operations: grouping, cropping, and concatenating. Specifically, we first divide the feature-map produced by convolutional layer into different groups, and each group is considered as one transformation path. Next, each transformation path is assigned with a random crop transformation. Finally, all the transformation paths are concatenated into a new feature-map for further training. Extensive experiments on two benchmark datasets CIFAR-10/100 validate the effectiveness and generality of Cropout. Specially, the ResNeXt-29, 8*64d (with 34M parameters) with our proposed Cropout achieve the error rate of 3.38% and 16.89% on CIFAR-10/100 dataset, and surpass the standard ResNeXt-29, 16*64d (with 68M parameters) which is twice larger in model size. In addition, the proposed Cropout is able to be applied to different modern deep networks (e.g.ResNeXt, ResNet and DenseNet) to further boost the performance on image classification tasks.

P547 *Exploiting Machine Learning Models to Avoid Texting While Driving [#19431]*

Renato Torres, Orlando Ohashi, Gabriel Garcia, Filipe Rocha, Hector Azpurua and Gustavo Pessin, Federal University of Para (UFPA), Brazil; Federal Rural University of Amazonia (UFRA), Brazil; Federal University of Ouro Preto, Brazil; Instituto Tecnologico Vale, Brazil

Text while driving is a worldwide phenomenon which is acknowledged as a meaningful problem for road safety. Text while driving is one of the most dangerous types of distraction because it involves visual, cognitive and physical distraction. Due to the advancement of technology, in the same way that it increases the possibilities of drivers' distractions, also increases the number of proposals that aim to avoid this behavior. In this context, we propose an intelligent system to identify text while driving. We deploy four machine learning models: Decision Tree; Support Vector Machine; Random Forest and Gradient Boosting. We compare their performances and significance for the domain of this work. The evaluation of the models were carried out with a real word collected dataset. Regarding the accuracy of the models, Random Forest and Gradient Boosting presented performance superior to 0.93. Recall 0.94 and Precision 0.94 also presented good results. In this work, these two metrics are highlighted because they are related, respectively, to driver's safety and passenger's convenience. In addition, we also carried out an analysis of the social and economic impacts that the proposed model can cause.

P548 *Character-Aware Convolutional Recurrent Networks with Self-Attention for Emotion Detection on Twitter [#20061]*

Jiangping Huang, Chunli Xiang, Shuwei Yuan, Desen Yuan and Xiaorui Huang, School of Software Engineering, Chongqing University of Posts and Telecommunications, China; School of Cyber Science and Engineering, Wuhan University, China; School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, China; International College, Chongqing University of Posts and Telecommunications, China

Despite myriad efforts in the literature designing neural representation system for emotion detection, very few works consider constructing effective model for apperceiving various emotion intensity on social media because of the informal expression and lack of context. In this paper, we proposed a character-aware convolutional recurrent networks with self-attention for emotion detection on user-generated content. The proposed model contains three parts: the character-level convolutional layer is designed to learn word representation based on character n-grams for capturing subword information and independent on pre- trained word embedding. The recurrent neural networks learn the sequential context information used both forward and backward recurrent neural network. And the self-attention module is used to extract different emotion aspects of the sentence into multiple vector representations. The attention module performs on top of recurrent networks which enables attention to be used in special domain or task when there are no extra inputs. We evaluate the proposed model on two public emotion datasets including both emotion intensity detection and emotion classification. We compare our model with the state-of-the-art methods on these datasets and the experimental results demonstrate that the proposed model outperforms several baselines on most emotion types detection and indicates the effectiveness of the designed model. In addition, the training of the proposed model for these tasks relies exclusively on initialized character vector which can be used for morphologically rich languages with long-tailed frequency distributions or domains with dynamic vocabuaries.

P549 *A Riemannian Primal-dual Algorithm Based on Proximal Operator and its Application in Metric Learning [#19644]*

Shijun Wang, Baocheng Zhu, Lintao Ma and Yuan Qi, Ant Financial Services Group, United States; Ant Financial Services Group, China

In this paper, we consider optimizing a smooth, convex, lower semicontinuous function in Riemannian space with constraints. To solve the problem, we first convert it to a dual problem and then propose a general primal-dual algorithm to optimize the primal and dual variables iteratively. In each optimization iteration, we employ a proximal operator to search optimal solution in the primal space. We prove convergence of the proposed algorithm and show its non-asymptotic convergence rate. By utilizing the proposed primal-dual optimization technique, we propose a novel metric learning algorithm which learns an optimal feature transformation matrix in the Riemannian space of positive definite matrices. Preliminary experimental results on an optimal fund selection problem in fund of funds (FOF) management for quantitative investment showed its efficacy.

P550 *Hierarchical Recurrent Attention Networks for Context-Aware Education Chatbots [#19345]*

Jean-Baptiste Aujogue and Alex Aussem, Computer Science Department, University of Lyon 1, France; LIRIS UMR CNRS 5205, University of Lyon 1, France

We propose a hierarchical network architecture for context-aware dialogue systems, that chooses which parts of the past conversation to focus on through a two-layer attention mechanism. The model can encode the parts of the historical dialog that are relevant to the current question to reason about the required response. We first assess the performance of our model on the Dialog bAbI task that involves a restaurant reservation system, where the goal is to book a table at a restaurant. We then train our model on a new hand-crafted dialogue data set, consisting of 7500 dialogues, to inform prospective students about the Data Science master program at University of Lyon.

P551 *Fashion Outfit Composition Combining Sequential Learning and Deep Aesthetic Network [#20498]*

Zhen Wang and Hongyan Quan, School of Computer Science and Software Engineering, East China Normal University, China

A proper outfit should consist of different categories of items that are visually compatible and share a similar style. Besides, personal aesthetic preference is also an important criterion when creating an overall outfit. Nevertheless, only few studies deal with aesthetic information in previous work on outfit composition. In this paper, we exploit both sequential learning and deep aesthetic network to train an end-to-end model for composing aesthetic outfits automatically. In detail, we firstly introduce a bidirectional long short-term memory (Bi-LSTM) model to discover the concept of compatibility among fashion items in an outfit. Then, an aesthetic-based model is proposed to parallel supervise the Bi-LSTM model so that the trained model can capture the aesthetic features from outfits. Meanwhile, we leverage visual-semantic descriptions to guarantee the uniqueness of each fashion item in an outfit. Moreover, to evaluate the effectiveness and practicability of the proposed model, we also design two representative tasks. One is evaluating the aesthetic scores of existing outfits. The other is composing aesthetic outfits conditioned on the given fashion item. Considering that aesthetic is highly correlated with personal preference, we additionally conduct the experiment that we can create proper outfits according to personal aesthetic preferences. Extensive experiments indicate that our method can generate aesthetic outfits that meet personal preferences and outperform the state-of-the-art.

P552 *Hierarchical Multi-Task Learning for Healthy Drink Classification [#19223]*

Homin Park, Homanga Bharadhwaj and Brian Y. Lim, National University of Singapore, Singapore; Indian Institute of Technology Kanpur, India

Recent advances in deep convolutional neural networks have enabled convenient diet tracking exploiting photos captured with smartphone cameras. However, most of the current diet tracking apps focus on recognizing solid foods while omitting drinks despite their negative impacts on our health when consumed without moderation. After an extensive analysis of drink images, we found that such an absence is due to the following challenges that conventional convolutional neural networks trained under the single-task learning framework cannot easily handle. First, drinks are amorphous. Second, visual cues of the drinks are often occluded and distorted by their container properties. Third, ingredients are inconspicuous because they often blend into the drink. In this work, we present a healthy drink classifier trained under a hierarchical multi-task learning framework composed of a shared residual network with hierarchically shared convolutional layers between similar tasks and task-specific fully-connected layers. The proposed structure includes two main tasks, namely sugar level classification and alcoholic drink recognition, and six auxiliary tasks, such as classification and recognition of drink name, drink type, branding logo, container transparency, container shape, and container material. We also curated a drink dataset, Drink101, composed of 101 different drinks including 11,445 images overall. Our experimental results demonstrate improved classification precision compared to single-task learning and baseline multi-task learning approaches.

P553 *Deep Learning and One-class SVM based Anomalous Crowd Detection [#19570]*

Meng Yang, Sutharshan Rajasegarar, Sarah M. Erfani and Christopher Leckie, The University of Melbourne, Australia; Deakin University, Australia

Anomalous event detection in videos is an important and challenging task. This paper proposes a deep representation approach to the problem, which extracts and represents features in an unsupervised way. This algorithm can detect anomalous activity like standing statically and loitering among a crowd of people. Our proposed framework is a two-channel scheme by using feature channels extracted from the appearance and foreground of the original video. Two hybrid deep learning architectures SDAE-DBN-PSVM (a four-layer Stacked Denoising Auto-encoder with three-layer Deep Belief Nets and Plane-based one class SVM) are implemented for these two channels to learn the high-level feature representation automatically and produce two anomaly scores. Finally, a fusion scheme is proposed for combining anomaly scores and detecting anomalous events. Experimental results on a large real-world dataset (MCG) and two benchmark datasets (UCSD and Subway) demonstrate the effectiveness of this approach. Furthermore, quantitative analyses of the effects of the amount of training data and the illumination conditions of the video on the accuracy of anomaly detection are presented.

P554 *Pose estimator and tracker using temporal flow maps for limbs [#19414]*

Jihye Hwang, Jieun Lee, Sungheon Park and Nojun Kwak, Seoul National University, Korea (South); Ajou University, Korea (South)

For human pose estimation in videos, it is significant how to use temporal information between frames. In this paper, we propose temporal flow maps for limbs (TML) and a multi-stride method to estimate and track human poses. The proposed temporal flow maps are unit vectors describing the limbs' movements. We constructed a network to learn both spatial information and temporal information end-to- end. Spatial information such as joint heatmaps and part affinity fields is regressed in the spatial network part, and the TML is regressed in the temporal network part. We also propose a data augmentation method to learn various types of TML better. The proposed multi-stride method expands the data by randomly selecting two frames within a defined range. We demonstrate that the proposed method efficiently estimates and tracks human poses on the PoseTrack 2017 and 2018 datasets.

P555 *Fusion of Multiple Representations Extracted from a Single Sensor's Data for Activity Recognition Using CNNs [#20080]*

Farzan Majeed Noori, Enrique Garcia-Ceja, Md Zia Uddin, Michael Riegler and Jim Torresen, University of Oslo, Norway

With the emerging ubiquitous sensing field, it has become possible to build assistive technologies for persons during their daily life activities to provide personalized feedback and services. For instance, it is possible to detect an individual's behavioral information (e.g. physical activity, location, and mood) by using sensors embedded in smartwatches and smartphones. To detect human's daily life activities, accelerometers have been widely used in wearable devices. In the current research, usually a single data representation is used, i.e., either image or feature vector representations. In this paper, a novel method is proposed to address two key aspects for the future development of robust deep learning methods for Human Activity Recognition (HAR): (1) multiple representations of a single sensor's data and (2) fusion of these multiple representations. The presented method utilizes Deep Convolutional Neural Networks (CNNs) and was evaluated using a publicly available HAR dataset. The proposed method showed promising performance, with the best result reaching an overall accuracy of 0.97, which outperforms current conventional approaches.

P556 *Dual-stream Self-Attentive Random Forest for False Information Detection [#19965]*

Manqing Dong, Lina Yao, Xianzhi Wang, Boualem Benatallah, Xiang Zhang and Quan Z. Sheng, University of New South Wales, Australia; University of Technology Sydney, Australia; Macquarie University, Australia

The prevalence of online social media facilitates massive knowledge acquisition and sharing throughout the Web. Meanwhile, it inevitably poses the risk of generating and disseminating false information by both benign and malicious users. Despite there has been considerable research on false information detection from both the opinion-based and fact-based perspectives, they mostly focus on tailored solutions for a particular domain and carry out limited work on leveraging multi-faceted clues such as textual cues, behavioral trails, and relational connection. We propose a novel dual-stream attentive random forest that is capable of selecting clues of discriminative information from individuals, collective information (e.g., texts), and correlations of entities (e.g., social interactions) adaptively. In particular, we use an interpretive attention model for learning textual contents. The model treats the important and unimportant content differently when constructing the textual representation and employs a multilayer perceptron to capture the hidden complex relationships among features of side information. We further propose a unified framework for leveraging the above clues, where we use attentive forests to provide probabilistic distribution as predictions over the two learned representations, which are then leveraged to make a better estimation. We conduct extensive experiments on three real-world benchmark datasets for fake news and fake review detection. The results show our approach outperforms multiple baselines in the accuracy of detecting false information.

P557 *TA-BLSTM: Tag Attention-based Bidirectional Long Short-Term Memory for Service Recommendation in Mashup Creation [#20294]*

Min Shi, Yufei Tang and Jianxun Liu, Florida Atlantic University, United States; Hunan University of Science and Technology, China

The service-oriented architecture makes it possible for developers to create value-added Mashup applications by composing multiple available Web services. Due to the overwhelming number of Web services online, it is often hard and time-consuming for developers to find their desired ones from the

entire service repository. In the past, various approaches aim at recommending Web services for automatic Mashup creation have been proposed, i.e., TFIDF, collaborative filtering and topic model-based methods, which rely on the original service descriptions given by service providers. However, most traditional methods fail to capture the function-related features of services since words contained in service descriptions usually correspond to different intent aspects (e.g., functional and non-functional related). To tackle this problem, we propose a tag attention-based recurrent neural networks model for Web service recommendation. The model consists of two Siamese bidirectional Long Short-Term Memory (LSTM) networks, which jointly learn two embeddings representing the functional features of Web services and the functional requirements of Mashups. In addition, by considering the tags of services as functional context information, the model can learn to assign attention scores to different words in service descriptions according to their intent importance, thus words used to reveal the functional properties of Web service will be given special attention. We compare our approach with the state-of-the-art methods (e.g., RTM, Word2vec, etc.) on a real-world dataset crawled from ProgrammableWeb, and the experimental results demonstrate the effectiveness of the proposed model.

P558 *An Efficient Framework by Topic Model for Multi-label Text Classification [#19809]*

Sun Wei, Ran Xiangying, Luo Xiangyang and Wang Chongjun, Department of Computer Science and Technology National Key Laboratory for Novel Software Technology at Nanjing University, China

Most existing multi-label text classification (MLTC) approaches only exploit label correlations from label pairwises or label chains. However, in the real world, features of instances have much importance for classification. In this paper, we propose a simple but efficient framework for MLTC called Hybrid Latent Dirichlet Allocation Multi-Label (HLDAML). To be specific, the topics of text features (i.e., a concrete description of documents) and the topics of label sets (i.e., a summarization of documents) can be obtained from training data by topic model before building models for multi-label classification. After that, hybrid topics can be used in existing approaches to improve the performance of MLTC. Experiments on several benchmark datasets demonstrate that the proposed framework is general and effective when taking text features and label sets into consideration simultaneously. It is also worth mentioning that we construct a new multi-label dataset called Parkinson about diagnosing parkinson disease by Traditional Chinese Medicine.

P559 *Deep learning price momentum in US equities [#19216]*

Stephen Choi and Tyler Renelle, LORA Technologies, Hong Kong

Research into deep learning techniques for stock price trend identification is limited. This can partly be attributed to the aversion of technical analysis within the academic community. One popular investment strategy that has been accepted by both academics and professionals, based purely on historical prices, is price momentum. However, the recent performance of this strategy has been disappointing. In this paper, we construct a new framework integrating state-of-the-art deep learning and machine learning methods to identify price trends of US equities: a "deep learning price momentum" portfolio. We first replicate the conventional price momentum calculations and compare the results with the market benchmarks and standard implementations of deep learning. We examine the issues of applying standard deep learning techniques to a limited noisy data set. Then we propose a new modular approach, built on deep learning clustering methods and recurrent neural networks that shows significant improvement on conventional price momentum while addressing the deficiencies of conventional deep learning methods. While the best-performing conventional price momentum portfolio yields 12.88% annual return and -0.49% market neutral annual returns for the 15-year period (2003 - 2017), our model improves these to 15.44% and +1.93% respectively with a significantly enhanced Sharpe ratio.

P560 *Quantitative Trading on Stock Market Based on Deep Reinforcement Learning [#19821]*

Jia Wu, Chen Wang, Lidong Xiong and Hongyong Sun, University of Electronic Science and Technology of China, China; Quantitative Trading on Stock Market Based on Deep Reinforcement Learning, China

With the development of computer science technology and artificial intelligence, quantitative trading attracts more investors due to its efficiency and stable performance. In this paper, we explore the potential of deep reinforcement learning in quantitative trading. A LSTM-based agent is proposed to learn the temporal pattern in data and automatically trades according to the current market condition and the historical data. The input to the agent is the raw financial data and the output of the agent is decision of trading. The goal of the agent is to maximize the ultimate profit. Besides, to reduce the influence of noise in the market and to improve the performance of the agent, we use several technical indicators as an extra input. The proposed system has been back-tested on the stock market. The results demonstrate that our method performs well in most conditions.

P561 *Compensating Supervision Incompleteness with Prior Knowledge in Semantic Image Interpretation [#19302]*

Ivan Donadello and Luciano Serafini, Fondazione Bruno Kessler, Italy

Semantic Image Interpretation is the task of extracting a structured semantic description from images. This requires the detection of visual relationships: triples (subject, relation, object) describing a semantic relation between a subject and an object. A pure supervised approach to visual relationship detection requires a complete and balanced training set for all the possible combinations of (subject, relation, object). However, such training sets are not available and would require a prohibitive human effort. This implies the ability of predicting triples which do not appear in the training set. This problem is called zero-shot learning. State-of-the-art approaches to zero-shot learning exploit similarities among relationships in the training set or external linguistic knowledge. In this paper, we perform zero-shot learning by using Logic Tensor Networks, a novel Statistical Relational Learning framework that exploits both the similarities with other seen relationships and background knowledge, expressed with logical constraints between subjects, relations and objects. The experiments on the Visual Relationship Dataset show that the use of logical constraints outperforms the current methods. This implies that background knowledge can be used to alleviate the incompleteness of training sets.

P562 *Deep Cyclic Group Networks [#19658]*

Zhe-Cheng Fan, Tak-Shing Chan, Yi-Hsuan Yang and Jyh-Shing Jang, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan; Research Center for Information Technology Innovation, Academia Sinica, Taiwan

We propose a new network architecture called deep cyclic group network (DCGN) that uses the cyclic group algebra for convolutional vector-neuron learning. The input to DCGN is a three-way tensor, where the mode-3 dimension corresponds to the dimensionality of the input data, e.g., three for RGB images. To handle vector-valued inputs, we replace scalar multiplication with circular convolution for the feedforward and backpropagation processes. As a result, every feature map and kernel map is a three-way tensor with the same mode-3 dimension as the input data. This way, DCGN may capture more of the relations among different data dimensions, especially for regression tasks where the target output has the same dimensionality as the input data. Moreover, DCGN can deal with input data of arbitrary dimensions, a property that existing architectures such as deep complex networks and deep quaternion networks (DQN) lack. Experiments show that DCGN indeed performs better than convolutional neural networks and DQN for two regression tasks, namely color image inpainting and multispectral image denoising.

P563 *Spatial and Channel Restraint for Complementary Feature Learning [#19277]*

Donghui Liu, Wei Fang and Ziwei Wang, Beijing University of Posts and Telecommunications, China; Information Science Academy, China Electronics Technology Group Corporation, China

Convolutional neural networks (CNNs) have been widely used in the field of computer vision in recent years, and have greatly improved the effect. At present, most of the CNNs extract the features from the network, obtain the probability distribution of the categories through the fully connected layer, then obtain and optimize the softmax loss functions. In this process, the networks gradually put the focus on discriminative features that are helpful for classification, while other features of the object may be ignored by the network. In order to help the networks, in addition to focusing on discriminative features of the object, also focus on other features of the object in training, we proposed spatial restraint module and channel restraint module. Our methods impose constraints on the discriminative features which have high response during the training process, forcing the network to mine useful complementary features. Our methods could improve the generalization performance of the model. Compared with baselines, our CNNs can achieve performance boost in both image classification task and weakly supervised object localization task.

P564 *Dynamic Fusion of Convolutional Features based on Spatial and Temporal Attention for Visual Tracking [#19324]*

Dongcheng Zhao and Yi Zeng, Institute of Automation, Chinese Academy of Sciences, China

Convolutional neural networks (CNN) based trackers have been widely employed in visual object tracking due to their powerful representations. Features from different CNN layers encode different information. Deeper layers contain more semantic information, while the resolution is too coarse to localize the target. Shallower layers carry more detail information but are less robust for appearance variations. In this paper, we propose an algorithm which incorporates the Spatial and Temporal attention to take full advantage of the Hierarchical Convolutional Features for Tracking (STHCFT). We firstly learn correlation filters on each convolutional layer. Based on the spatial attention inspired by the paraventricular thalamus (PVT) in the brain, we choose the most important layer to build the base response, and the others to be the auxiliary responses. In addition, we make full use of the temporal attention to determine the weights of the auxiliary responses. Finally, the target is located by the maximum value of the fused responses. Extensive experimental results on the benchmark OTB-2013 and OTB-2015 have shown the proposed algorithm performs favorably against several state-of-the-art trackers.

P565 *Testing the Robustness of Manifold Learning on Examples of Thinned-Out Data [#20087]*

Fayeem Aziz and Stephan Chalup, School of Electrical Engineering and Computing, The University of Newcastle, Australia

Manifold learning can only be successful if enough data is available. If the data is too sparse, the geometrical and topological structure of the manifold extracted from the data cannot be recognised and the manifold collapses. In this paper we used data from a simulated two-dimensional double pendulum and tested how well several manifold learning methods could extract the expected manifold, a two-dimensional torus. The experiments were repeated while the data was downsampled in several ways to test the robustness of the different manifold learning methods. We also developed a neural network-based deep autoencoder for manifold learning and demonstrated that it performed in most of our test cases similarly or better than traditional methods such as principal component analysis and isomap.

P566 *Parallel Convolution Algorithm Using Implicit Matrix Multiplication on Multi-Core CPUs [#20120]*

Qinglin Wang, Songzhu Mei, Jie Liu and Chunye Gong, National University of Defense Technology, China

Convolution neural networks (CNNs) have been extensively used in machine learning applications. The most time-consuming part of CNNs are convolution operations. A common approach to implementing convolution operations is to recast them as general matrix multiplication, known as the im2col+GEMM approach. There are two main drawbacks of this approach. One is that large additional memory space is required. The other is the packing on the input elements of convolution operations are not memory-efficient enough. In this paper, we present a new parallel convolution algorithm using implicit matrix multiplication on multi-core CPUs. In comparison with Im2col+GEMM, our new algorithm can reduce the memory footprints and improve the packing efficiency. The experiment results on two ARV8-based multi-core CPUs demonstrate that our new algorithm gives much better performance and scalability than the im2col+GEMM method in most cases.

P567 *COMC: A Framework for Online Cross-domain Multistream Classification [#20367]*

Hemeng Tao, Zhuoyi Wang, Yifan Li, Mahmoud Zamani and Latifur Khan, The University of Texas at Dallas, United States

With the tremendous increase of the online data, training a single classifier may suffer because of the large variety of data domains. One solution could be to learn separate classifiers for each domain. However, this would arise a huge cost to gather annotated training data for a large number of domains and ignore similarity shared across domains. Hence, it leads to our problem setting: can labeled data from a related source domain help predict the unlabeled data in the target domain? In this paper, we consider two independent simultaneous data streams, which are referred to as the source and target streams. The target stream continuously generates data instances from one domain where the label is unknown, while the source stream continuously generates labeled data instances from another domain. Most likely, the two data streams would have different but related feature spaces and different data distributions. Moreover, these streams may have asynchronous concept drifts between them. Our problem setting, which is called Cross-domain Multistream Classification, is to predict the class labels of data instances in the target stream using a classifier trained on the labeled source stream. In this paper, we propose an efficient solution for cross-domain multistream classification by integrating change detection into online data stream adaptation. The class labels of data instances in the target stream are predicted using the sufficient amount of label information in the related source stream. And the concept drifts along the two independent streams are continuously being addressed at the same time. Experimental results on real- world data sets indicate significantly improved performance over baseline methods.

P568 *Improving Fast Adaptive Stacking of Ensembles [#19983]*

Laura Maria Palomino Marino, Juan Isidro Gonzalez Hidalgo, Roberto Souto Maior de Barros and Germano Crispim Vasconcelos, Universidade Federal de Pernambuco-UFPE, Brazil

The treatment of large data streams in the presence of concept drifts is one of the main challenges in the fields of machine learning and data mining. This article presents two ensemble algorithms designed to quickly adapt to concept drifts, both abrupt and gradual. Fast Stacking of Ensembles boosting the Best (FASEB) and Fast Ensemble boosting the Best with Combined Weighting Voting (FASEBwv) are adaptations of Fast Adaptive Stacking of Ensembles (FASE), designed to improve their execution time without causing considerable losses in the algorithm's accuracy. In order to achieve a more efficient model, adjustments were made in the update strategy and voting procedure of the ensemble. To evaluate the proposals against the state of the art, we used Naive Bayes (NB) and Hoeffding Tree (HT) as learners to compare the performance of the algorithms on artificial and real-world data

sets. The experiments showed that FASEB and FASEBwv are more efficient than FASE in many scenarios, often also achieving better results in accuracy.

P569 *Deep Reinforcement Learning for Chatbots Using Clustered Actions and Human-Likeness Rewards [#20122]*

Heriberto Cuayahuitl, Donghyeon Lee, Seonghan Ryu, Sungja Choi, Inchul Hwang and Kim Jihie, University of Lincoln, United Kingdom; Samsung Research, Korea (South)

Training chatbots using the reinforcement learning paradigm is challenging due to high-dimensional states, infinite action spaces and the difficulty in specifying the reward function. We address such problems using clustered actions instead of infinite actions, and a simple but promising reward function based on human- likeness scores derived from human-human dialogue data. We train Deep Reinforcement Learning (DRL) agents using chitchat data in raw text---without any manual annotations. Experimental results using different splits of training data report the following. First, that our agents learn reasonable policies in the environments they get familiarised with, but their performance drops substantially when they are exposed to a test set of unseen dialogues. Second, that the choice of sentence embedding size between 100 and 300 dimensions is not significantly different on test data. Third, that our proposed human-likeness rewards are reasonable for training chatbots as long as they use lengthy dialogue histories of $\geq$10 sentences.

P570 *Pyramid Attention Dense Network for Image Super-Resolution [#19383]*

Si-Bao Chen, Chao Hu, Bin Luo, Chris Ding and Shi-Lei Huang, Anhui University, China; University of Texas at Arlington, United States; PKU-HKUST Shenzhen Hong Kong Institution, China

Recent deep convolution neural networks has made remarkable progress in single images super-resolution area. They achieved very high Peak Signal to Noise Ratio (PSNR) and structural similarity (SSIM), by improved learning of high-frequency details to enhance visual perception. However, current models usually ignore relations between adjacent pixels. In this work, we propose a network that incorporate gradients of adjacent pixels in addition to per-pixel loss and perceptual loss. In addition, we utilize multi-stage network learning to progressively generate high resolution images, by incorporating a new inter-stage feedback in the Laplacian pyramid network structure. Furthermore, we adopted recently proposed attention mechanism and dense block structure. The proposed Pyramid Attention Dense model for image super-resolution achieved state-of-the-art performance in experiments on four benchmark datasets.

P571 *SpaMHMM: Sparse Mixture of Hidden Markov Models for Graph Connected Entities [#19017]*

Diogo Pernes and Jaime S. Cardoso, INESC TEC; University of Porto, Portugal

We propose a framework to model the distribution of sequential data coming from a set of entities connected in a graph with a known topology. The method is based on a mixture of shared hidden Markov models (HMMs), which are jointly trained in order to exploit the knowledge of the graph structure and in such a way that the obtained mixtures tend to be sparse. Experiments in different application domains demonstrate the effectiveness and versatility of the method.

P572 *Deep Structured Cross-Modal Anomaly Detection [#19481]*

Yuening Li, Ninghao Liu, Jundong Li, Mengnan Du and Xia Hu, Texas A&M University, United States; Arizona State University, United States

Anomaly detection is a fundamental problem in data mining field, with many data-driven applications. A vast majority of existing anomaly detection methods predominately focused on data collected from a single source. In real-world applications, instances often have multiple types of features, such as images (ID photos, finger prints) and texts (bank transaction histories, user online social media posts), resulting in the so-called multi-modal data. In this paper, we focus on identifying anomalies whose patterns are disparate across different modalities, i.e., cross-modal anomalies. Some of the data instances within a multi-modal context are often not anomalous when they are viewed separately in each individual modality, but contains inconsistent patterns when multiple sources are jointly considered. The existence of multi-modal data in many real-world scenarios brings both opportunities and challenges to the canonical task of anomaly detection. On the one hand, in multi-modal data, information of different modalities may complement each other in improving the detection performance. On the other hand, complicated distributions across different modalities call for a principled framework to characterize their inherent and complex correlations, which is often difficult to capture with conventional linear models. To this end, we propose a novel deep structured anomaly detection framework to identify the cross-modal anomalies embedded in the data. Experiments on real-world datasets demonstrate the effectiveness of the proposed framework comparing with the state-of-the-art.

P573 *Cystoid Fluid Color Map Generation in Optical Coherence Tomography Images Using a Densely Connected Convolutional Neural Network [#19427]*

Placido Vidal, Joaquim de Moura, Jorge Novo and Marcos Ortega, Universidade da Coruna, Spain

Optical Coherence Tomography (OCT) is a medical imaging modality that is currently the focus of many advancements in the field of ophthalmology. It is widely used to diagnose relevant diseases like Diabetic Macular Edema (DME) or Age-related Macular Degeneration (AMD), both among the principal causes of blindness. These diseases have in common the presence of pathological cystoid fluid accumulations inside the retinal layers that tear its tissues, hindering the correct vision of the patient. In the last years, several works proposed a variety of methodologies to obtain a precise segmentation of these fluid regions. However, many cystoid patterns present several difficulties that harden significantly the process. In particular, some of these cystoid bodies present diffuse limits, others are deformed by shadows, appear mixed with other tissues and other complex situations. To overcome these drawbacks, a regional analysis has been proven to be reliable in these problematic regions. In this work, we propose the use of the DenseNet architecture to perform this regional analysis instead of the classical machine learning approaches, and use it to represent the pathological identifications with an intuitive color map. We trained, validated and tested the DenseNet neural network with a dataset composed of 3247 samples labeled by an expert. They were extracted from 156 images taken with two of the principal OCT devices of the domain. Then, this network was used to generate the color map representations of the cystoid areas in the OCT images. Our proposal achieved robust results in these regions, with a satisfactory 97.48% +- 0.7611 mean test accuracy as well as a mean AUC of 0.9961 +- 0.0029.

P574 *FKIMNet: A Finger Dorsal Image Matching Network Comparing Component (Major, Minor and Nail) Matching with Holistic (Finger Dorsal) Matching [#20441]*

Daksh Thapar, Gaurav Jaswal and Aditya Nigam, Indian Institute of Technology Mandi, India

Current finger knuckle image recognition systems, often require users to place fingers' major or minor joints flatly towards the capturing sensor. To extend these systems for user non-intrusive application scenarios, such as consumer electronics, forensic, defence etc, we suggest matching the full dorsal fingers, rather than the major/ minor region of interest (ROI) alone. In particular, this paper makes a comprehensive study on the comparisons between full finger and fusion of finger ROI's for finger knuckle image recognition. These experiments suggest that using full-finger, provides a more elegant solution. Addressing the finger matching problem, we propose a CNN (convolutional neural network) which creates a 128-D feature embedding of an image. It is trained via. triplet loss function, which enforces the L2 distance between the embeddings of the same subject to be approaching zero, whereas the distance between any 2 embeddings of

different subjects to be at least a margin. For precise training of the network, we use dynamic adaptive margin, data augmentation, and hard negative mining. In distinguished experiments, the individual performance of finger, as well as weighted sum score level fusion of major knuckle, minor knuckle, and nail modalities have been computed, justifying our assumption to consider full finger as biometrics instead of its counterparts. The proposed method is evaluated using two publicly available finger knuckle image datasets i.e., PolyU FKP dataset and PolyU Contactless FKI Datasets.

P575 *A Unified Approach on Active Learning Dual Supervision [#20117]*

Adrian Chriswanto, Hsing-Kuo Pao and Yuh-Jye Lee, National Taiwan University of Science and Technology, Taiwan; National Chiao Tung University, Taiwan

Active Learning (AL) is a machine learning framework that aims to efficiently select a limited labeled data to construct an effective model given huge amount of unlabeled data on the side. Most studies in AL focus on how to select the unlabeled data to be labeled by a human oracle in order to maximize the performance gain of the model with as little labeling effort as possible. In this work, we focus not only on how to select appropriate data instances but also how to select informative features, more specifically, categorical features to be labeled by the oracle in a unified manner. The unification means that we select the best choice of item to label where the item can be either a feature or an instance on each iteration given a unified scoring function to make the decision. The method that we propose is by synthesizing new instances that represent a set of features. By utilizing synthesized instances, we can treat this set of features as if they are regular instances. Therefore they could be compared on an equal ground when the model tries to select which instance to be labeled by the oracle. We demonstrate the effectiveness of the proposed method through a few data sets that consist of categorical features where the feature labeling makes more sense to labeling oracles. The experiment results show that adopting the unified approach creates clear benefit to model construction, especially in the early stage where we can efficiently obtain an effective model through only a few iterations, compared to the one using only instance labeling for model construction.

P576 *Mixture of Pre-processing Experts Model for Noise Robust Deep Learning on Resource Constrained Platforms [#19977]*

Taesik Na, Minah Lee, Burhan A. Mudassar, Priyabrata Saha, Jong Hwan Ko and Saibal Mukhopadhyay, Georgia Institute of Technology, United States

Deep learning on an edge device requires energy efficient operation due to ever diminishing power budget. Intentional low quality data during the data acquisition for longer battery life, and natural noise from the low cost sensor degrade the quality of target output which hinders adoption of deep learning on an edge device. To overcome these problems, we propose simple yet efficient mixture of pre- processing experts (MoPE) model to handle various image distortions including low resolution and noisy images. We also propose to use adversarially trained auto encoder as a pre-processing expert for the noisy images. We evaluate our proposed method for various machine learning tasks including object detection on MS-COCO 2014 dataset, multiple object tracking problem on MOT-Challenge dataset, and human activity classification on UCF 101 dataset. Experimental results show that the proposed method achieves better detection, tracking and activity classification accuracies under noise without sacrificing accuracies for the clean images. The overheads of our proposed MoPE are 0.67\% and 0.17\% in terms of memory and computation compared to the baseline object detection network.

P577 *A Convolutional Neural Network with Two-Channel Input for Image Super-Resolution [#20354]*

Purbaditya Bhattacharya and Udo Zoelzer, Helmut Schmidt University, Germany

In this work a convolutional neural network with two input channels is proposed for image super-resolution. Initially, the original image is downsampled with bicubic and nearest neighbour interpolation methods and the low-resolution image pair is used as the input to our network. Additionally, the input channels are randomly swapped as an augmentation method during training. The proposed network is a combination of a feedforward architecture and a residual network architecture, to provide a smooth image estimate. We show that the additional input image obtained with nearest neighbour interpolation improves image super-resolution performance with our network, since it acts as an image specific prior information and constrains the solution space for the model to learn. We train the network end-to-end for multiple upscale factors, evaluate the method on standard test datasets, and show that the proposed approach can produce state-of-the-art results.

P578 *Improving the realism of synthetic images through a combination of adversarial and perceptual losses [#20355]*

Charith Atapattu and Banafsheh Rekabdar, Southern Illinois University, United States

In recent years, deep learning methods are becoming more widely used; however, large quantities of labeled training data are required for most models. Labeling large datasets is tedious, expensive, and time consuming. Generating large labeled synthetic datasets, on the other hand, is easier and less expensive since annotations are available. But there is usually a large gap between the distribution of the synthetic and real data. In this paper, we propose a novel method based on Generative Adversarial Networks (GANs) to improve the realism of the synthetic images while preserving the annotation information. In our work, the inputs of the GANs are synthetic images instead of random vectors. Furthermore, we describe how a perceptual loss can be utilized while introducing the basic features and techniques from adversarial networks for obtaining better results. We evaluate our approach for appearance- based gaze direction classification on the MPIIGaze dataset. The results show that our generated refined images are more realistic and better preserve the annotation information than the refined images generated by the state-of-the-art methods.

P579 *Active visual object exploration and recognition with an unmanned aerial vehicle [#19613]*

Uriel Martinez-Hernandez, Victor Cedeno-Campos and Adrian Rubio-Solis, University of Bath, United Kingdom; University of Sheffield, United Kingdom

In this paper, an active control method for visual object exploration and recognition with an unmanned aerial vehicle is presented. This work uses a convolutional neural network for visual object recognition, where input images are obtained with an unmanned aerial vehicle from multiple objects. The object recognition task is an iterative process actively controlled by a saliency map module, which extracts interesting object regions for exploration. The active control allows the unmanned aerial vehicle to autonomously explore better object regions to improve the recognition accuracy. The iterative exploration task stops when the probability from the convolutional neural network exceeds a decision threshold. The active control is validated with offline and real-time experiments for visual exploration and recognition of five objects. Furthermore, passive exploration is also tested for performance comparison. Experiments show that the unmanned aerial vehicle is capable to autonomously explore interesting object regions. Results also show an improvement in recognition accuracy from 88.14% to 95.66% for passive and active exploration, respectively. Overall, this work offers a framework to allow robots to autonomously decide where to move and look next, to improve the performance during a visual object exploration and recognition task.

P580 *Keyphrase Guided Beam Search for Neural Abstractive Text Summarization [#19103]*
Xuewen Chen, Jinlong Li and Haihan Wang, University of Science and Technology of China, China

As a recently proposed way of text summarization, abstractive text summarization features the use of new phrases to obtain a condensed version of the source text. Most approaches nowadays in this category are under the sequence-to-sequence framework, which is the bedrock of many text generation tasks. However, these approaches usually fail to get a sound representation of the source text, and they are liable to produce summaries with a low semantic relevance. In this work, we devise a novel structure of convolutional recurrent neural network-based encoder to get a better latent representation of the source text. Further, we propose to exploit keyphrases to guide the summary selection in a modified beam search process, thus contributing to a closer semantic relevance between the source text and the generated summary. With this new approach, we outperform the state-of-the-art systems on several datasets at no extra cost of training, particularly on the benchmark CNN/Daily Mail dataset where the source texts are lengthy.

P581 *Deep Representation Learning for Code Smells Detection using Variational Auto-Encoder [#20433]*
Mouna Hadj-Kacem and Nadia Bouassida, Miracl Laboratory, Sfax University, Tunisia

Detecting code smells is an important research problem in the software maintenance. It assists the subsequent steps of the refactoring process so as to improve the quality of the software system. However, most of existing approaches have been limited to the use of structural information. There have been few researches to detect code smells using semantic information although its proven effectiveness in many software engineering problems. In addition, they do not capture entirely the semantic embedded in the source code. This paper attempts to fill this gap by proposing a semantic-based approach that detects bad smells which are scattered at different levels of granularity in the source code. To this end, we use an Abstract Syntax Tree with a Variational Auto-Encoder in the detection of three code smells. The code smells are Blob, Feature Envy and Long Method. We have performed our experimental evaluation on nine open-source projects and the results have achieved a considerable overall accuracy. To further evaluate the performance of our approach, we compare our results with a state-of-the-art method on the same publicly available dataset.

S34: Mind, Brain, and Cognitive Algorithms and Other Cross-Disciplinary Topics

Thursday, July 18, 11:50AM-1:30PM, Room: Duna Salon I, Chair: Robert Kozma

11:50AM *Interpretation of Mesoscopic Neurodynamics by Simulating Conversion Between Pulses and Waves [#20511]*
Joshua J.J. Davis and Robert Kozma, Embassy of Peace, Whitianga & U Auckland, New Zealand; U Memphis, TN, United States

Cognition and brain dynamics manifests a delicate balance between processes at various temporal and spatial scales. The conversion between microscopic neural pulses and waves of mesoscopic activity of neural masses is a crucial research subject. In this work we analyze the hierarchy of neural structures and dynamics, with an emphasis on pulse-wave-pulse conversion. Our models describe pulse-to- wave conversion, as well as the feedback of action potentials on neurons in the recovery mode. We study the behavior of neural populations in the background state of activity, as well as under perturbations due to external stimuli. Simulations results are employed for the interpretation of electrocorticogram data obtained with rabbits trained using classical conditioning paradigm.

12:10PM *Nonmodular Architectures of Cognitive Systems based on Active Inference [#20216]*
Manuel Baltieri and Christopher Laurie Buckley, EASY group, Sussex Neuroscience - Department of Informatics - University of Sussex, United Kingdom

In psychology and neuroscience it is common to describe cognitive systems as input/output devices where perceptual and motor functions are implemented in a purely feedforward, open-loop fashion. On this view, perception and action are often seen as encapsulated modules with limited interaction between them. While embodied and enactive approaches to cognitive science have challenged the idealisation of the brain as an input/output device, we argue that even the more recent attempts to model systems using closed-loop architectures still heavily rely on a strong separation between motor and perceptual functions. Previously, we have suggested that the mainstream notion of modularity strongly resonates with the separation principle of control theory. In this work we present a minimal model of a sensorimotor loop implementing an architecture based on the separation principle. We link this to popular formulations of perception and action in the cognitive sciences, and show its limitations when, for instance, external forces are not modelled by an agent. These forces can be seen as variables that an agent cannot directly control, i.e., a perturbation from the environment or an interference caused by other agents. As an alternative approach inspired by embodied cognitive science, we then propose a nonmodular architecture based on active inference. We demonstrate the robustness of this architecture to unknown external inputs and show that the mechanism with which this is achieved in linear models is equivalent to integral control.

12:30PM *Exploring Deep Models for Comprehension of Deictic Gesture-Word Combinations in Cognitive Robotics [#19677]*
Gabriella Pizzuto and Angelo Cangelosi, University of Manchester, United Kingdom

In the early stages of infant development, gestures and speech are integrated during language acquisition. Such a natural combination is therefore a desirable, yet challenging, goal for fluid human-robot interaction. To achieve this, we propose a multimodal deep learning architecture, for comprehension of complementary gesture-word combinations, implemented on an iCub humanoid robot. This enables human-assisted language learning, with interactions like pointing at a cup and labelling it with a vocal utterance. We evaluate various depths of the Mask Regional Convolutional Neural Network (for object and wrist detection) and the Residual Network (for gesture classification). Validation is carried out with two deictic gestures across ten real-world objects on frames recorded directly from the iCub's cameras. Results further strengthen the potential of gesture-word combinations for robot language acquisition.

12:50PM *A comparison of machine learning algorithms as surrogate model for net present value prediction from wells arrangement data [#19818]*
Joao Bertini, Mei Funcia, Antonio Santos and Denis Schiozer, University of Campinas, Brazil

Net Present Value (NPV) measures whether an investment will be profitable within a given period of time. In oil production planning, it consists of an important indicator to evaluate different production strategies. The NPV estimate is calculated on the basis of the production data which are generally obtained by means of numerical simulations, which consider the strategy details and the physical reservoir model. However, the simulator demands high computational resource, which may take hours or days of processing time to evaluate a single strategy, depending on the size of the reservoir model. To speed up this process a simpler model, referred to as a surrogate model, can be used to approximate the simulator output. In this work, we hypothesize that it is possible to predict the NPV using only wells arrangement data as predictors. Moreover, we present a comparison among six machine learning algorithms used as a surrogate model: Linear

Regression, K-Nearest Neighbor, Multi- Layer Perceptron, Kernel Ridge Regression, Support Vector Regression, and Gradient Tree Boosting. Results confirm it is viable to predict NPV from wells arrangement data, in special with kernel-based methods.

1:10PM *Autoencoder-Based Articulatory-to-Acoustic Mapping for Ultrasound Silent Speech Interfaces [#20143]*

Gabor Gosztolya, Adam Pinter, Laszlo Toth, Tamas Grosz, Alexandra Marko and Tamas Gabor Csapo, MTA-SZTE Research Group on Artificial Intelligence, Hungary; University of Szeged, Hungary; Eotvos Lorand University, Hungary; Budapest University of Technology and Economics, Hungary

When using ultrasound video as input, Deep Neural Network-based Silent Speech Interfaces usually rely on the whole image to estimate the spectral parameters required for the speech synthesis step. Although this approach is quite straightforward, and it permits the synthesis of understandable speech, it has several disadvantages as well. Besides the inability to capture the relations between close regions (i.e. pixels) of the image, this pixel-by-pixel representation of the image is also quite uneconomical. It is easy to see that a significant part of the image is irrelevant for the spectral parameter estimation task as the information stored by the neighbouring pixels is redundant, and the neural network is quite large due to the large number of input features. To resolve these issues, in this study we train an autoencoder neural network on the ultrasound image; the estimation of the spectral speech parameters is done by a second DNN, using the activations of the bottleneck layer of the autoencoder network as features. In our experiments, the proposed method proved to be more efficient than the standard approach: the measured normalized mean squared error scores were lower, while the correlation values were higher in each case. Based on the result of a listening test, the synthesized utterances also sounded more natural to native speakers. A further advantage of our proposed approach is that, due to the (relatively) small size of the bottleneck layer, we can utilize several consecutive ultrasound images during estimation without a significant increase in the network size, while significantly increasing the accuracy of parameter estimation.

8c: Bioinformatics and Other Applications

Thursday, July 18, 11:50AM-1:30PM, Room: Duna Salon II, Chair: Heung-Il Suk

11:50AM *Representation-dimensionality Trade-off in Biological Sequence-based Inference [#20023]*

Bahman Asadi and Niranjan Mahesan, University of Southampton, United Kingdom

Statistical inference from the analysis of biological sequences is widely used in the prediction of structure and biochemical functions of newly found macromolecules. For the application of machine learning methodologies such as kernel methods and artificial neural networks for such inference, variable length sequence data is often embedded in a finite dimensional real-valued space. The corresponding embedding dimensions are often high, leading to technical difficulties of the curse of dimensionality. We demonstrate a trade-off between fidelity of representation and the resulting dimensionality of the embedding space. Clustering chemically similar amino acids, thereby reducing the alphabet size, reduces the accuracy in their variation, but achieves a reduction in the corresponding feature space. We show this trade-off in three different problems of statistical inference, namely, protein-protein interaction, remote homology and secondary structure prediction. We show that in the reduced space performance often improves similar to what is seen in ``diminishing returns'' type reward-effort curves. We find alphabet reduction schemes taken from the literature perform significantly better than arbitrary random clustering of the alphabets. Statistical feature selection from the full 20 amino acid representation is not competitive with any of these. Dimensionality of representation has animportant role when mapping sequence data onto fixed dimensions of an Euclideanspace. This work shows that dimensionality reduction based on compressing the amino acid alphabet improves inference performance in two widely studied problems and degrades gracefully in the third.

12:10PM *Stochastic Imputation and Uncertainty-Aware Attention to EHR for Mortality Prediction [#20430]*

Eunji Jun, Ahmad Wisnu Mulyadi and Heung-Il Suk, Department of Brain and Cognitive Engineering, Korea University, Korea (South)

Electronic Health Records (EHR) has become a vital source of a patient data but characterized by a variety of missing values. Using the variational inference of Bayesian framework, Variational autoencoders (VAEs), a deep generative model, has been shown to be efficient and accurate to capture the latent structure of vast mount of complex high-dimensional data. Recently, it has been used for missing data imputation. In this paper, we propose a general framework that incorporates effective missing data imputation using VAEs and multivariate time series prediction. We utilize the uncertainty obtained from the generative network of the VAE and employ uncertainty-aware attention in imputing the missing values. We illustrate the performance of our architecture on real-world clinical dataset (MIMIC-III) for in-hospital mortality prediction task. Our results showed higher performance than other competing methods in mortality prediction task.

12:30PM *GADGET: Using Gated GRU for Biomedical Event Trigger Detection [#19202]*

Zeng Cheng, Zhang Yi, Lu Heng-Yang and Wang Chong-Jun, National Key Laboratory for Novel Software Technology, Nanjing University, China

Biomedical event extraction plays an important role in the field of biomedical text mining, and the event trigger detection is the first step in the pipeline process of event extraction. Event trigger can clearly indicates the occurrence of related events. There have been many machine learning based methods applied to this area already. However, most previous work have omitted two crucial points: (1) Class Difference: They simply regard non-trigger as same level class label. (2) Information Isolation: Most methods only utilize token level information. In this paper, we propose a novel model based on gate mechanism, which identifies trigger and non- trigger words in the first stage. At the same time, we also introduce additional fusion layer in order to incorporate sentence level information for event trigger detection. Experimental results on the Multi Level Event Extraction (MLEE) corpus achieve superior performance than other state-of-the-art models. We have also performed ablation study to show the effectiveness of proposed model components.

12:50PM *Study of Short-Term Personalized Glucose Predictive Models on Type-1 Diabetic Children [#19145]*

Maxime De Bois, Mounim A. El Yacoubi and Mehdi Ammi, CNRS-LIMSI, France; Telecom SudParis, France; Universite Paris 8, France

Research in diabetes, especially when it comes to building data-driven models to forecast future glucose values, is hindered by the sensitive nature of the data. Because researchers do not share the same data between studies, progress is hard to assess. This paper aims at comparing the most promising algorithms in the field, namely Feedforward Neural Networks (FFNN), Long Short-Term Memory (LSTM) Recurrent Neural Networks, Extreme Learning Machines (ELM), Support Vector Regression (SVR) and Gaussian Processes (GP). They are personalized and trained on a population of 10 virtual children from the Type 1 Diabetes Metabolic Simulator software to predict future glucose values at a prediction horizon of 30 minutes. The performances of the models are evaluated using the Root

Mean Squared Error (RMSE) and the Continuous Glucose-Error Grid Analysis (CG-EGA). While most of the models end up having low RMSE, the GP model with a Dot-Product kernel (GP-DP), a novel usage in the context of glucose prediction, has the lowest. Despite having good RMSE values, we show that the models do not necessarily exhibit a good clinical acceptability, measured by the CG-EGA. Only the LSTM, SVR and GP-DP models have overall acceptable results, each of them performing best in one of the glycemia regions.

1:10PM *Bidirectional Associative Memory for Multimodal Fusion : a Depression Evaluation Case Study [#20299]*

Stephane Cholet, Helene Paugam-Moisy and Sebastien Regis, Universite des Antilles, Guadeloupe

In this study, we step aside from traditional fusion strategies and propose to use a bidirectional associative memory to combine abstract representations of several modalities. The innovative contribution of our strategy is the fact that the fusion is performed neither at the input data level nor at the output score level. The model achieves outperforming results on depression evaluation from the videos of the AVEC2014 challenge corpus.

8e: Data analysis and pattern recognition and Other Applications

Thursday, July 18, 11:50AM-1:30PM, Room: Duna Salon III, Chair: Rama Murthy

11:50AM *Si-GCN: Structure-induced Graph Convolution Network for Skeleton-based Action Recognition [#19285]*

Rong Liu, Chunyan Xu, Tong Zhang, Wenting Zhao, Zhen Cui and Jian Yang, Nanjing University of Science and Technology, Nanjing, China

In recent years, the graph-convolution networks have been used to solve the problem of skeleton-based action recognition. Previous works often adopted a structure-fixed graph to model the physical joints of human skeleton, but cannot well consider these interactions of different human parts (e.g., the right arm and the left leg) to some extent. To deal with this problem, we propose a novel structure-induced graph convolution network (Si-GCN) framework to boost the performance of the skeleton- based action recognition task. Given a video sequence of human skeletons, the Si-GCN can produce the sample-wise category in an end-to-end way. Specifically, according to the natural divisions of human body, we define a collection of intra-part graphs for each input human skeleton (i.e., each part-graph denotes a specific part/global of human skeleton), and then formulate an interactive graph to model the relationships of different part-graphs. The Si-GCN framework, which will then perform the spectral graph convolutions on these constructed intra/inter-part graphs, can not only capture the internal modalities of each human part/subgraph, but also consider the interactions/relationships between different human parts. A temporal convolution follows to model the temporal and spatial dynamics of the skeleton in combination with the characteristics of time and space. Comprehensive evaluations on two public datasets (including NTU RGB+D and HDM05) well demonstrate the superiority of our proposed Si-GCN when compared with existing skeleton-based action recognition approaches.

12:10PM *VT-GAN: View Transformation GAN for Gait Recognition Across Views [#19549]*

Peng Zhang, Qiang Wu and Jingsong Xu, University of Technology Sydney, Australia

Recognizing gaits without human cooperation is of importance in surveillance and forensics because of the benefits that gait is unique and collected remotely. However, change of camera view angle severely degrades the performance of gait recognition. To address the problem, previous methods usually learn mappings for each pair of views which incurs abundant independently built models. In this paper, we proposed a View Transformation Generative Adversarial Networks (VT-GAN) to achieve view transformation of gaits across two arbitrary views using only one uniform model. In specific, we generated gaits in target view conditioned on input images from any views and the corresponding target view indicator. In addition to the classical discriminator in GAN which makes the generated images look realistic, a view classifier is imposed. This controls the consistency of generated images and conditioned target view indicator and ensures to generate gaits in the specified target view. On the other hand, retaining identity information while performing view transformation is another challenge. To solve the issue, an identity distilling module with triplet loss is integrated, which constrains the generated images inheriting identity information from inputs and yields discriminative feature embeddings. The proposed VT- GAN generates visually promising gaits and achieves promising performances for cross-view gait recognition, which exhibits great effectiveness of the proposed VT-GAN.

12:30PM *An Inferable Representation Learning for Fraud Review Detection with Cold-start Problem [#19434]*

Qian Li, Qiang Wu, Chengzhang Zhu, Jian Zhang and Wentao Zhao, University of Technology Sydney, Australia; National University of Defense Technology, China

Fraud review significantly damages the business reputation and also customers' trust to certain products. It has become a serious problem existing on the current social media. Various efforts have been put in to tackle such problems. However, in the case of cold-start where a review is posted by a new user who just pops up on the social media, common fraud detection methods may fail because most of them are heavily depended on the information about the user's historical behavior and its social relation to other users, yet such information is lacking in the cold-start case. This paper presents a novel Joint-bEhavior-and-Social-relaTion-infERable (JESTER) embedding method to leverage the user reviewing behavior and social relations for cold-start fraud review detection. JESTER embeds the deep characteristics of existing user behavior and social relations of users and items in an inferable user-item-review-rating representation space where the representation of a new user can be efficiently inferred by a closed-form solution and reflects the user's most probable behavior and social relations. Thus, a cold-start fraud review can be effectively detected accordingly. Our experiments show JESTER (i) performs significantly better in detecting fraud reviews on four real-life social media data sets, and (ii) effectively infers new user representation in the cold-start problem, compared to three state-of-the-art and two baseline competitors.

12:50PM *Dynamic Bus Arrival Time Prediction exploiting Non-linear Correlations [#19142]*

Avinash Achar, Rohith Regikumar and B Anil Kumar, Tata Consultancy Services, India; Nanyang Technological University, Singapore

Public transport buses exhibit uncertainties in their arrival/travel times, due to several factors such as signals, dwell times at bus stops, seasonal variations, fluctuating travel demands etc. Further, factors like excess vehicles, acute lack of lane discipline and diversity in modes of transport additionally plague the traffic in the developing world. Owing to these diverse factors, the bus arrival time prediction (BATP) continues to be a challenging problem especially in the developing world. The current work proposes a data-driven method to perform bus arrival time prediction in real-time. Unlike existing approaches, the proposed method explicitly learns both the spatial and temporal correlations/patterns of traffic in a novel and general fashion. In particular, it first detects the unknown order of spatial dependence and then learns general nonlinear and non-stationary correlations using supervised

learning, for this detected order. The real time prediction problem is now posed as an inference problem on an associated non-linear dynamical system (NLDS) model. We propose to use an Extended Kalman Filter (EKF) to solve this inference problem in an approximate but efficient manner. We demonstrate the effectiveness of our proposed algorithm on real field data from challenging traffic conditions in the developing world. Our experiments demonstrate that the proposed EKF outperforms diverse existing state-of-art data driven approaches proposed for the same problem.

1:10PM *Non-Traditional Input Encoding Schemes for Spiking Neuromorphic Systems [#19330]*

Catherine Schuman, James Plank, Grant Bruer and Jeremy Anantharaj, Oak Ridge National Laboratory, United States; University of Tennessee, United States

A key challenge for utilizing spiking neural networks or spiking neuromorphic systems for most applications is translating numerical data into spikes that are appropriate to apply as input to a spiking neural network. In this work, we present several approaches for encoding numerical values as spikes, including binning, spike-count encoding, and charge-injection encoding, and we show how these approaches can be combined hierarchically to form more complex encoding schemes. We demonstrate how these different encoding approaches perform on four different applications, running on four different neuromorphic systems that are based on spiking neural networks. We show that the input encoding method can have a significant effect on application performance and that the best input encoding method is application- specific.

Deep Learning and Neural Network Models

Thursday, July 18, 11:50AM-1:30PM, Room: Panorama I, Chair: Chi-Jen Lu

11:50AM *Nested Variance Estimating VAE/GAN for Face Generation [#19165]*

Hong-You Chen and Chi-Jen Lu, Academia Sinica, Taiwan

We study the task of conditional image generation and provide a general framework for combining autoencoders (AEs) with generative adversarial networks (GANs). Our framework provides a principled way to avoid well-known problems such as mode collapse and training instability. We use two AEs, a big parent- AE and a small child-AE, to play different roles. Our parent-AE is trained to minimize only one single objective: the reconstruction loss, which makes its training process stable and efficient. It is then fixed and used for several purposes, including initializing the generator and providing powerful features for a simple discriminator of GAN. It also plays the role of reducing the initial harder task of image generation to a simpler one: sampling from its latent distribution, which is given to child-AE to accomplish. Child-AE can be trained very efficiently due to its small size, and we only need to modify it if necessary for different applications, which makes our framework very flexible. Our experiments show that our model is capable of generating high quality novel images with controllable attributes.

12:10PM *Generate Desired Images from Trained Generative Adversarial Networks [#19141]*

Ming Li, Rui Xi, Beier Chen, Mengshu Hou, Daibo Liu and Lei Guo, University of Electronic Science and Technology of China, China; Ohio State University, Columbus, United States

The emerging of Generative Adversarial Networks (GANs) gives rise to a significant improvement in image generation. However, a controllable way of synthesizing images with specific characteristics still is a challenging issue. Many existing methods are not efficient enough that require additional information and pre-designed attributes, and are with much more human intervention. In this paper, we propose GAGAN, an extension method to the Generative Adversarial Network, that is the first work to generate specific images from a trained GAN model. To control the characteristics of images, a DNA pool of the trained GAN model is introduced and evolved by a genetic algorithm (GA). Then, with the DNA pool, GAGAN can generate the corresponding latent vector (DNA) of target images. Furthermore, GAGAN can synthesize images containing a single specific characteristic or multiple specific attributes (including AND and OR relation). Moreover, several fitness evaluation strategies are also proposed to make GAGAN flexible to control the target characteristics. Experiments on CelebA and MNIST are conducted, and results show that the proposed method is feasible and effective in specific image generation problem.

12:30PM *Multiple-Instance Learning through Optimum-Path Forest [#19104]*

Luis Claudio Sugi Afonso, Danilo Colombo, Clayton Reginaldo Pereira, Kelton Augusto Pontara Costa and Joao Paulo Papa, Federal University of Sao Carlos - UFSCar, Brazil; Petroleo Brasileiro - Petrobras, Brazil; Sao Paulo State University - UNESP, Brazil

Multiple-instance (MI) learning aims at modeling problems that are better described by several instances of a given sample instead of individual descriptions often employed by standard machine learning approaches. In binary- driven MI problems, the entire bag is considered positive if one (at least) sample is labeled as positive. On the other hand, a bag is considered negative if it contains all samples labeled as negative as well. In this paper, we introduced the Optimum-Path Forest (OPF) classifier to the context of multiple-instance learning paradigm, and we evaluated it in different scenarios that range from molecule description, text categorization, and anomaly detection in well-drilling report classification. The experimental results showed that two different OPF classifiers are very much suitable to handle problems in the multiple-instance learning paradigm.

12:50PM *Long-Term Prediction of Small Time-Series Data Using Generalized Distillation [#19154]*

Shogo Hayashi, Akira Tanimoto and Hisashi Kashima, Kyoto University, Japan; NEC, Japan

The recent increase of ``big data" in our society has led to major impacts of machine learning and data mining technologies in various fields ranging from marketing to science. On the other hand, there still exist areas where only small-sized data are available for various reasons, for example, high data acquisition costs or the rarity of targets events. Machine learning tasks using such small data are usually difficult because of the lack of information available for training accurate prediction models. In particular, for long-term time-series prediction, the data size tends to be small because of the unavailability of the data between input and output times in training. Such limitations on the size of time-series data further make long-term prediction tasks quite difficult; in addition, the difficulty that the far future is more uncertain than the near future. In this paper, we propose a novel method for long-term prediction of small time-series data designed in the framework of generalized distillation. The key idea of the proposed method is to utilize the middle-time data between the input and output times as ``privileged information," which is available only in the training phase and not in the test phase. We demonstrate the effectiveness of the proposed method on both synthetic data and real-world data. The experimental results show the proposed method performs well, particularly when the task is difficult and has high input dimensions.

1:10PM *Not All Adversarial Examples Require a Complex Defense: Identifying Over-optimized Adversarial Examples with IQR-based Logit Thresholding [#19374]*

Utku Ozbulak, Arnout Van Messem and Wesley De Neve, Ghent University, Belgium

Detecting adversarial examples currently stands as one of the biggest challenges in the field of deep learning. Adversarial attacks, which produce adversarial examples, increase the prediction likelihood of a target class for a particular data point. During this process, the adversarial example can be further optimized, even when it has already been wrongly classified with 100% confidence, thus making the adversarial example even more difficult to detect. For this kind of adversarial examples, which we refer to as over-optimized adversarial examples, we discovered that the logits of the model provide solid clues on whether the data point at hand is adversarial or genuine. In this context, we first discuss the masking effect of the softmax function for the prediction made and explain why the logits of the model are more useful in detecting over-optimized adversarial examples. To identify this type of adversarial examples in practice, we propose a non-parametric and computationally efficient method which relies on interquartile range, with this method becoming more effective as the image resolution increases. We support our observations throughout the paper with detailed experiments for different datasets (MNIST, CIFAR-10, and ImageNet) and several architectures.

Machine Learning

Thursday, July 18, 11:50AM-1:30PM, Room: Panorama II, Chair: Eric Bax

11:50AM *Optimizing Weight Value Quantization for CNN Inference [#19192]*

Wakana Nogami, Tsutomu Ikegami, Shin-ichi O'uchi, Ryosei Takano and Tomohiro Kudoh, The University of Tokyo, Japan; National Institute of Advanced Industrial science and Technology, Japan

The size and complexity of CNN models are increasing. Use of a lower bit width numerical representation has been studied extensively to reduce the required resources. There are some prior studies that use moderate lower bit width with well-known numerical representations such as fixed point or logarithmic representation. It is not apparent, however, whether those representations are optimal for maintaining high accuracy. In this paper, we investigated the numerical quantization and introduced a novel Variable Bin-size Quantization (VBQ) representation in which quantization bin boundaries are optimized to obtain maximum accuracy for each CNN model. A genetic algorithm was employed to optimize the bin boundaries of VBQ. Additionally, since the appropriate bit width to obtain sufficient accuracy cannot be determined in advance, we attempted to use the parameters obtained by a training process using higher precision representation and used quantization in inference only. This reduced the required large computational resource cost for training. During the process of tuning VBQ boundaries using a genetic algorithm, we discovered that the optimal distribution of bins can be approximated by a formula with two parameters. We then used simulated annealing for finding the optimal parameters of the formula for AlexNet and VGG16. As a result, AlexNet and VGG16 with our 4-bit quantization achieved top-5 accuracy at 74.8\% and 86.3\% respectively, which were comparable to 76.3\% and 88.1\% obtained by FP32. Thus, VBQ combined with the approximate equation and the simulated annealing scheme can achieve similar levels of accuracy with less resources and reduced computational cost compared to other current approaches.

12:10PM *Coral Classification Using DenseNet and Cross-modality Transfer Learning [#19118]*

Lian Xu, Mohammmed Bennamoun, Farid Boussaid, Senjian An and Ferdous Sohel, The University of Western Australia, Australia; Curtin University, Australia; Murdoch University, Australia

Coral classification is a challenging task due to the complex morphology and ambiguous boundaries of corals. This paper investigates the benefits of Densely connected convolutional network (DenseNet) and multi-modal image translation techniques in boosting image classification performance by synthesizing missing fluorescence information. To this end, an image-conditional Generative Adversarial Network (GAN) based image translator is trained to model the relationship between reflectance and fluorescence images. Through this image translator, fluorescence images can be generated from the available reflectance images to provide complementary information. During the classification phase, reflectance and translated fluorescence images are combined to obtain more discriminative representations and produce improved classification performance. We present results on the EFC and MLC datasets and report state-of-the-art coral classification performance.

12:30PM *A Multiple Local Model Learning for Nonlinear and Time-Varying Microwave Heating Process [#19061]*

Tong Liu, Shan Liang, Sheng Chen and Chris J. Harris, School of Automation Chongqing University, China; School of Electronics and Computer Science University of Southampton, United Kingdom

This paper proposes a multiple local model learning approach for nonlinear and nonstationary microwave heating process (MHP). The proposed local learning framework performs model adaption at two levels: (1) adaptation of the local linear model set, which adaptively partitions the process's data into multiple process states, each fitted with a local linear model; (2) online adaptation of model prediction, which selects a subset of candidate local linear models and linearly combines them to produce the model prediction. Adaptive process state partition and fitting a new local linear model to the newly emerging process state is based on statistical hypothesis testing, and the optimal combining coefficients of the selected subset linear models are obtained by minimizing the mean square error with the constraint that the sum of these coefficients is unity. A case study involving a real-world industrial MHP is used to demonstrate the superior performance of the proposed multiple local model learning approach, in terms of online modeling accuracy and computational efficiency.

12:50PM *Using a Recurrent Kernel Learning Machine for Small-Sample Image Classification [#19071]*

Mihael Cudic and Jose Principe, University of Florida, United States

Many machine learning algorithms, like Convolutional Neural Networks (CNNs), have excelled in image processing tasks; however, they have many practical limitations. For one, these systems require large datasets that accurately represent the sample distribution in order to optimize performance. Secondly, they have difficulty transferring previously learned knowledge when evaluating data from slightly different sample distributions. To overcome these drawbacks, we propose a recurrent kernel-based approach for image processing using the Kernel Adaptive Autoregressive Moving Average algorithm (KAARMA). KAARMA minimizes the amount of training data required by using the Reproducing Kernel Hilbert Space to build inference into the system. The recurrent nature of KAARMA additionally allows the system to better learn the spatial correlations in the images through one-shot or near one-shot learning. We demonstrate KAARMA's superiority for small-sample image classification using the JAFFE Face Dataset and the UCI hand written digit dataset.

1:10PM *Ensemble Validation: Selectivity has a Price, but Variety is Free [#19018]*

Eric Bax and Farshad Kooti, Verizon, United States; Facebook, United States

Suppose some classifiers are selected from a set of hypothesis classifiers to form an equally-weighted ensemble that selects a member classifier at random for each input example. Then the ensemble has an error bound consisting of the average error bound for the member classifiers, a term for selectivity that varies from zero (if all hypothesis classifiers are selected) to a standard uniform error bound (if only a single classifier is selected), and small constants. There is no penalty for using a richer hypothesis set if the same fraction of the hypothesis classifiers are selected for the ensemble.

Applications

Thursday, July 18, 11:50AM-1:30PM, Room: Panorama III, Chair: Yan Yang

11:50AM *Selective Expression For Event Coreference Resolution on Twitter [#19175]*

Chao Wenhan, Wei Ping, Luo Zhunchen, Liu Xiao and Sui Guobin, Beihang University, China; PLA Academy of Military Science, China; Beijing Institute of Technology, China

With the growth in popularity and size of social media, there is an urgent need for systems that for systems that can recognize the coreference relation between two event mentions in texts from s in texts from social media. In existing event coreference resolution research, a rich set of li rich set of linguistic features derived from pre-existing NLP tools and various knowledge bas knowledge bases is often required. This kind of methods restricts domain scalability and leads to the leads to the propagation of errors. In this paper, we present a novel selective expression expression approach based on event trigger to explore the coreferential relationship in high- volu high-volume Twitter texts. Firstly, we exploit a bidirectional Long Short Term Memory (Bi-LSTM) to extract the sentence level and mention level features. Then, to selectively express the essential parts of generated features, we apply a gate on sent on sentence level features. Next, to integrate the time information of event mention mention pairs, we design an auxiliary feature based on triggers and time attributes of the t the two event mentions. Finally, all these features are concatenated and fed into a classifier to predict the binary coreference relationship between the event ment mention pair. To evaluate our method, we publish a new dataset EventCoreOnTweet (ECT)1 tha that annotates the coreferential relationship between event mentions and event tr trigger of each event mention. The experimental results demonstrate that our approach a achieves significant performance in the ECT dataset.

12:10PM *An LSTM based Encoder-Decoder Model for Multi-Step Traffic Flow Prediction [#19005]*

Shengdong Du, Tianrui Li, Yan Yang, Xun Gong and Shi-Jinn Horng, School of Information Science and Technology, Southwest Jiaotong University, China; Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taiwan

Traffic flow prediction has been regarded as a key research problem in the intelligent transportation system. In this paper, we propose an encoder-decoder model with temporal attention mechanism for multi-step forward traffic flow prediction task, which uses LSTM as the encoder and decoder to learn the long dependencies features and nonlinear characteristics of multivariate traffic flow related time series data, and also introduces a temporal attention mechanism for more accurately traffic flow prediction. Through the real traffic flow dataset experiments, it has shown that the proposed model has better prediction ability than classic shallow learning and baseline deep learning models. And the predicted traffic flow value can be well matched with the ground truth value not only under short step forward prediction condition but also under longer step forward prediction condition, which validates that the proposed model is a good option for dealing with the real-time and forward-looking problems of traffic flow prediction task.

12:30PM *SkiDNet: Skip Image Denoising Network for X-Rays [#20277]*

Swaraj Kumar, Sandipan Dutta, Shaurya Chaturvedi and Mps Bhatia, Netaji University of Technology, India

Medical imaging has evolved to become an essential tool for screening and diagnosing diseases, but they have certain limitations just like every other technology. X-rays, which is one of the most common radiological examinations, is not immune to imperfections. In this paper, we aim to tackle one such imperfection in X-rays, which is the presence of undesirable noises which causes aberrations in the output projections. This makes diagnosis and analysis difficult since such noises shroud the intricate details that these images contain. Distinctive denoising algorithms have been proposed in the past for a spectrum of vision datasets, but a very few of them are for X-rays. We introduce a new denoising network called SkiDNet, a deep learning approach using an encoder-decoder architecture with skip connections of varying length. The network has been trained on the NIH Chest X-Ray Dataset. With the unique properties injected by different types of connections, SkiDNet is able to surpass the performance of existing models. Furthermore, adopting a different approach to weight initialization and batch normalization makes the network more robust. Denoised X-rays obtained from the network were objectively evaluated using different metrics namely mean squared error, peak signal-to- noise ratio, and the structural similarity index.

12:50PM *A Multi-model Ensemble Method Using CNN and Maximum Correntropy Criterion for Basal Cell Carcinoma and Seborrheic Keratoses Classification [#19196]*

Leida Guo, Shaoyi Du, Yuting Chi, Wenting Cui, Panpan Song, Jihua Zhu, Songmei Geng and Meifeng Xu, School of Software Engineering, Xi'an Jiaotong University, China; Institute of Artificial Intelligence and Robotics, School of Electronic and Information Engineering, Xi'an Jiaotong University, China; The Second Affiliated Hospital of Xi'an Jiaotong University, China

Basal cell carcinoma is very similar to the clinical traits of seborrheic keratosis, which is still a difficult problem in medical image analysis. To accurately classify it, this paper proposes a multi-model ensemble method based on the maximum correntropy criterion(MCC) and convolutional neural network(CNN). First of all, it is well known that the CNN single models like ResNet, Xception, DensNet, etc. have a good effect in the classification, but the accuracy is still limited, so the multi-model ensemble method is presented to improve the accuracy. Secondly, the traditional multi-model ensemble methods, such as voting and linear regression, can improve the accuracy of the model, but it means that the weight computation of each model do not consider the noise, and could not obtain good results. Therefore, we propose the MCC for model ensemble, which overcomes the noise in the data and effectively improves the classification accuracy. Finally, our proposed multi-model ensemble algorithm based on the MCC achieved an accuracy of 97.07% in the basal cell carcinoma and seborrheic keratosis classification experiments, surpassing the CNN single model and traditional multi-model ensemble method.

1:10PM *Hierarchical Classification Feature Extraction for Moving Target Detection Using Radar Echo [#19054]*

Chunhua Zhou, Huiting Xia, Jiejun Yin, Liang Gao and Yaqi Liu, 1. Shanghai Radio Equipment Research Institute 2. Shanghai Engineering Research Center of Target Identification and Environment Perception, China

Firstly, this paper presents a novel hierarchical classification feature extraction technology of radar echo for moving target detection using radar echo. Secondly, a progress is made to acquire the envelope and Doppler feature in the MATLAB simulation analysis. Finally, an efficient combination of the empirical mode decomposition (EMD) with the all-phase FFT (apFFT) to implement the hierarchical classification feature extraction of the moving target. The computation used by this technology is very simple and efficient and needs much less memory resource and should be perfectly adequate for computation in real time.

S33: Transferable neural models for language understanding; Applications

Thursday, July 18, 11:50AM-1:30PM, Room: Panorama IV, Chair: Zhiwei Lin

11:50AM *A Transformer-Based Variational Autoencoder for Sentence Generation [#19705]*

Danyang Liu and Gongshen Liu, Shanghai Jiao Tong University, China

The variational autoencoder(VAE) has been proved to be a most efficient generative model, but its applications in natural language tasks have not been fully developed. A novel variational autoencoder for natural texts generation is presented in this paper. Compared to the previously introduced variational autoencoder for natural text where both the encoder and decoder are RNN-based, we propose a new transformer- based architecture and augment the decoder with an LSTM language model layer to fully exploit information of latent variables. We also propose some methods to deal with problems during training time, such as KL divergency collapsing and model degradation. In the experiment, we use random sampling and linear interpolation to test our model. Results show that the generated sentences by our approach are more meaningful and the semantics are more coherent in the latent space.

12:10PM *Gated Task Interaction Framework for Multi-task Sequence Tagging [#19497]*

Isaac Kojo Essel Ampomah, Sally McClean, Zhiwei Lin and Glenn Hawe, Ulster University, United Kingdom

Recent studies have shown that neural models can achieve high performance on several sequence labelling/tagging problems without the explicit use of linguistic features such as part-of-speech (POS) tags. These models are trained only using the character-level and the word embedding vec vectors as inputs. Others have shown that linguistic features can improve the performance of ne neural models on tasks such as chunking and named entity recognition (NER). However, the change in performance depends on the degree of semantic r relatedness between the linguistic features and the target task; in some i instances, linguistic features can have a negative impact on performance. This paper presents an approach to jointly learn these linguistic features along with the target sequence labelling tasks with a new multi-task learning ((MTL) framework called Gated Tasks Interaction (GTI) network for solving m multiple sequence tagging tasks. The GTI network exploits the relations between the multiple tasks via neural gate modules. These gate modules control the flow of information between the different tasks. Experiments on benchmark datasets for chunking and NER show that our framework outperforms other competitive baselines trained with and without external training resources.

12:30PM *Emergent Multilingual Language Acquisition using Developmental Networks [#20377]*

Juan Castro-Garcia and Juyang Weng, Michigan State University, United States

There has been work on language acquisition but such prior work was based on symbolic representations and non-incremental learning. Neural Networks are meant for incremental learning but their performance has been weak. This situation was mainly due to a "lack of logic" in neural networks. By language acquisition here we mean incremental learning from lifetime experience. Since developmental networks (DN) has clearly understandable emergent ""logic" in terms of finite automata and Turing machines, this is the first work on language acquisition based on a clearly understandable emergent Turing machine. We show how symbolic words are represented by patterns to simulate naturally grounded and emergent inputs. The context as states/actions are also represented by patterns to simulate naturally grounded and emergent inputs. Our work demonstrates that symbolic state features can be fully automated by emergent input-context pattern pairs. This is a step toward fully automated acquisition of language by a grounded robot, but we are not there yet.

12:50PM *Across-Sensor Feature Learning for Energy-Efficient Activity Recognition on Mobile Devices [#19879]*

Yuriy Gavrilin and Adil Khan, Innopolis University, Russia

In this paper we propose across-sensor representation learning framework for improving power-accuracy trade-off in multi-sensor human activity recognition (HAR). The goal of the study is to achieve the level of performance comparable to one of multi-sensor HAR systems by using fewer or even single sensor. Such performance is achieved by learning relations between these sensors at training time and utilizing them at test time. These relations are learned by supervised deep models which use multi-sensor data during training only. The absence of need for having multiple sensors during test time allows turning these sensors off and replacing them with a single sensor coupled with learned across-sensor relations. These across-sensor relations make up for the information lost from the turned-off sensors. Using fewer sensors reduces energy consumption of HAR systems deployed on a smartphone. Moreover, it allows building HAR systems for situations when collection of multi-sensor data is possible only during training. This work presents preliminary results achieved with the proposed approach on the SHL dataset. Obtained results show an improvement of up to 14% in classification accuracy of single-sensor HAR.

S32: Deep Reinforcement Learning for Games

Thursday, July 18, 11:50AM-1:30PM, Room: Panorama V, Chair: Sagar Verma

11:50AM *Mixing Update Q-value for Deep Reinforcement Learning [#20036]*

Zhunan Li and Xinwen Hou, Institute of Automation, Chinese Academy of Sciences, China

The value-based reinforcement learning methods are known to overestimate action values such as deep Q-learning, which could lead to suboptimal policies. This problem also persists in an actor-critic algorithm. In this paper, we propose a novel mechanism to minimize its effects on both the critic and the actor. Our mechanism builds on Double Q-learning, by mixing update action value based on the minimum and maximum between a pair of critics to limit the overestimation. We then propose a specific adaptation to the Twin Delayed Deep Deterministic policy gradient algorithm (TD3) and show that the resulting algorithm not only reduces the observed overestimations, as hypothesized, but that this also leads to much better performance on several tasks.

12:10PM *MAPEL: Multi-Agent Pursuer-Evader Learning using Situation Report [#20184]*

Sagar Verma, Richa Verma and P.B. Sujit, CVN, CentraleSupelec, Universite Paris-Saclay, France; TCS Innovation Lab, India, India; IIIT Delhi, India, India

In this paper, we consider a territory guarding game involving pursuers, evaders and a target in an environment that contains obstacles. The goal of the evaders is to capture the target, while that of the pursuers is to capture the evaders before they reach the target. All the agents have limited sensing range and can only detect each other when they are in their observation space. We focus on the challenge of effective cooperation between agents of a team. Finding exact solutions for such multi-agent systems is difficult because of the inherent complexity. We present Multi-Agent Pursuer-Evader Learning (MAPEL), a class of algorithms that use spatio-temporal graph representation to learn structured cooperation. The key concept is that the learning takes place in a decentralized manner and agents use situation report updates to learn about the whole environment from each others' partial observations. We use Recurrent Neural Networks (RNNs) to parameterize the spatio-temporal graph. An agent in MAPEL only updates all the other agents if an opponent or the target is inside its observation space by using situation report. We present two methods for cooperation via situation report update: a) Peer-to-Peer Situation Report (P2PSR) and b) Ring Situation Report (RSR). We present a detailed analysis of how these two cooperation methods perform when the number of agents in the game are increased. We provide empirical results to show how agents cooperate under these two methods.

12:30PM *RevCuT Tree Search Method in Complex Single-player Game with Continuous Search Space [#19807]*

Hongming Zhang, Fangjuan Cheng, Bo Xu, Feng Chen, Jiachen Liu and Wei Wu, Institute of Automation, Chinese Academy of Sciences, China; Xi'an Jiaotong University, China; China Ship Development and Design Center, China

Monte-Carlo Tree Search (MCTS) has achieved great success in combinatorial game, which has the characteristics of finite action state space, deterministic state transition and sparse reward. AlphaGo Zero combined MCTS and deep neural networks defeated the world champion Lee Sedol in the Go game, proving the advantages of tree search in combinatorial game with enormous search space. However, when the search space is continuous and even with chance factors, tree search methods like UCT failed. Because each state will be visited repeatedly with probability zero and the information in tree will never be used, that is to say UCT algorithm degrades to Monte Carlo rollouts. Meanwhile, the previous exploration experiences cannot be used to correct the next tree search process, and makes a huge increase in the demand for computing resources. To solve this kind of problem, this paper proposes a step-by-step Reverse Curriculum Learning with Truncated Tree Search method (RevCuT Tree Search). In order to retain the previous exploration experiences, we use the deep neural network to learn the state-action values at explored states and then guide the next tree search process. Besides, taking the computing resources into consideration, we establish a truncated search tree focusing on continuous state space rather than the whole trajectory. This method can effectively reduce the number of explorations and achieve the effect beyond the human level in our well designed single-player game with continuous state space and probabilistic state transition.

12:50PM *Data-to-Text Generation with Attention Recurrent Unit [#19731]*

Hechong Wang, Wei Zhang, Yuesheng Zhu and Zhiqiang Bai, Peking University, China

Recurrent Neural Networks (RNNs) have shown promising results in many text generation tasks with their ability in modeling complex data distribution. However, text generation model in their encoder or decoder RNNs still can not use the context efficiently. In this paper, we propose a novel Attention Recurrent Unit (ARU) to generate short descriptive texts conditioned on database records. Different from conventional approaches Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), ARU allows the context information from the encoder to be aligned first inside the unit, which can improve the ability of content selection and surface realization for the model. And we also design a method called DoubleAtten to enhance the attention distribution when computing the generation probabilities. On the recently released ROTOWIRE dataset, extensive experimental results demonstrate that the ARU and DoubleAtten can efficiently improve the model performance for data-to-text generation task.

1:10PM *Attentive Dual Embedding for Understanding Medical Concept in Electronic Health Record [#20253]*

Xueping Peng, Guodong Long, Shirui Pan, Jing Jiang and Zhendong Niu, University of Technology Sydney, Australia; Monash University, Australia; Beijing Institute of Technology, China

Electronic health records contain a wealth of information on a patients healthcare over many visits, such as diagnoses, treatments, drugs administered, and so on. The untapped potential of these data in healthcare analytics is vast. However, given that much of medical information is a cause and effect science, new embedding methods are required to ensure the learning representations reflect the comprehensive interplays between medical concepts and their relationships over time. Unlike one-hot encoding, a distributed representation should preserve these complex interactions as high-quality inputs for machine learning-based healthcare analytics tasks. Therefore, we propose a novel attentive dual embedding method called MC2Vec. MC2Vec captures the proximity relationships between medical concepts through a two-step optimization framework that recursively refines the embedding for superior output. The framework comprises a Skip-gram model to generate the initial embedding and an attentive CBOW model to fine-tune the embedding with temporal information gleaned from sequences of patient visits. Experiments with two public datasets demonstrate that MC2Vecs produces embeddings of higher quality than five state-of-the-art methods.

Thursday, July 18, 1:30PM-2:30PM

Special Lecture: Lunch

Thursday, July 18, 1:30PM-2:30PM, Room: Various locations in the area

Thursday, July 18, 2:30PM-6:30PM

Workshop: Advances in Learning from/with Multiple Learners (ALML) Learn more
Thursday, July 18, 2:30PM-6:30PM, Room: Sofitel Bellevue 1, Chair: Paris 13 University, Razvan Andonie, Central Washington, Parisa Rastin, Paris 13 University, Nicoleta Rogovschi, University Paris Descartes, Basarab Matei, Paris 13 University, Gu{\'e}na{\"e}l Cabanes, Paris 13 University Nistor Grozavu

Workshop: Computational Sport Science: Human Motion Modelling and Analysis
Thursday, July 18, 2:30PM-6:30PM, Room: Sofitel Bellevue 2, Chair: Auckland University of Technology, New Zealand Dr. Boris Ba{\v{c}}i{\'c}

Workshop: Causality and Dynamics in Brain Networks
Thursday, July 18, 2:30PM-6:30PM, Room: Sofitel Bellevue 3, Chair: Wigner Research Centre for Physics, Zolt{\'a}n Somogyv{\'a}ri, Wigner Research Centre for Physics, Vaibhav Diwadkar, Wayne State University, L{\'a}szl{\'o} N{\'e}gyessy, Wigner Research Centre for Physics Andr{\'a}s Telcs

Friday, July 19, 9:00AM-1:00PM

Workshop: Advances in Learning from/with Multiple Learners (ALML)
Friday, July 19, 9:00AM-1:00PM, Room: Sofitel Bellevue 1, Chair: Paris 13 University, Razvan Andonie, Central Washington, Parisa Rastin, Paris 13 University, Nicoleta Rogovschi, University Paris Descartes, Basarab Matei, Paris 13 University, Gu{\'e}na{\"e}l Cabanes, Paris 13 University Nistor Grozavu

Workshop: Ethical AI Challenges
Friday, July 19, 9:00AM-1:00PM, Room: Sofitel Bellevue 2, Chair: Rebecca Raper, Matthias Rolf, Chrisina Jayne, Oxford Brookes University, UK Nigel Crook

Workshop: Causality and Dynamics in Brain Networks
Friday, July 19, 9:00AM-1:00PM, Room: Sofitel Bellevue 3, Chair: Wigner Research Centre for Physics, Zolt{\'a}n Somogyv{\'a}ri, Wigner Research Centre for Physics, Vaibhav Diwadkar, Wayne State University, L{\'a}szl{\'o} N{\'e}gyessy, Wigner Research Centre for Physics Andr{\'a}s Telcs

AUTHOR INDEX

Page numbers of papers where a person is the first author are shown in bold. The italic page numbers point to sessions of which the person is a chair.

Please click HERE and hit F9 to update the field.