

11th International Workshop on Multiple Classifier Systems  
2nd International Workshop on Partially Supervised Learning

# MCS-PSL 2013 Conference Program

May, 13-17, Nanjing, China



11th International Workshop on Multiple Classifier Systems

## MCS 2013

organized by



UNIVERSITY OF  
SURREY



LAMDA Group  
National Key Laboratory for  
Novel Software Technology  
Nanjing University (China)

Center for Vision, Speech  
and Signal Processing  
University of Surrey (UK)

Dept. of Electrical and  
Electronic Engineering  
University of Cagliari (Italy)

sponsored by

endorsed by



National Science  
Foundation of China



IEEE Computer Society  
Nanjing Chapter



The International Association  
for Pattern Recognition

2nd International Workshop on Partially Supervised Learning

## PSL 2013

organized by



LAMDA Group  
National Key Laboratory for  
Novel Software Technology  
Nanjing University (China)



University of Ulm  
(Germany)

sponsored by



National Science  
Foundation of China



IEEE Computer Society  
Nanjing Chapter

# TABLE OF CONTENTS

Program Overview	3
Sessions & Abstracts	5
May 14 (PSL)	5
May 15 (MCS Day 1)	11
May 16 (MCS Day 2)	15
May 17 (MCS Day 3)	24
Conference Organization	31

Participant Guidelines	32
Nanjing Mini-Map	32
New Era Hotel Surrounding Map	33
Venue Surrounding Map	34
A Simple Metro Map	35
Shuttle Bus Schedule	36
Excursion	37
Living in Nanjing	39
Food Adventure	40
Useful Sentences in Chinese	41

# Program Overview

## PSL'13 Program

### May 13

09:00 - 18:00     *Registration (in New Era Hotel / CS Building in Xianlin campus)*  
18:00             *Reception (in New Era Hotel)*

### May 14

07:20 - 07:30     *Bus pickup (at New Era Hotel)*  
08:30 - 08:40     **Opening**  
08:40 - 09:40     **Invited Talk: Dale Schuurmans**  
09:40 - 09:55     *Coffee break*  
09:55 - 11:10     **Morning Session**  
11:10 - 12:40     *Lab visit & Lunch*  
12:40 - 14:20     **Afternoon Session 1**  
14:20 - 14:35     *Coffee break*  
14:35 - 15:50     **Afternoon Session 2**  
15:50 - 18:00     *Bus transfer & campus tour*  
18:00 - 20:00     *MCS-PSL joint reception (in New Era Hotel)*

### May 15

12:30 - 18:30     *Excursion (bus dispatch from CS Building in Xianlin campus)*

- \* PSL participants are encouraged and free to join the MCS morning session in May 15; otherwise, please come to the tour bus boarding place directly.
- \* All sessions will be in Room 233 of the CS Building (Xianlin campus)

# MCS'13 Program

## May 14

- 09:00 - 18:00     *Registration (in New Era Hotel / CS Building in Xianlin campus)*  
18:00 - 20:00     *MCS-PSL joint reception (in New Era Hotel)*

## May 15

- 07:20 - 07:30     *Bus pickup (at New Era Hotel)*  
08:30 - 08:40     **Opening**  
08:40 - 09:40     **Invited Talk: Bin Yu**  
09:40 - 09:55     *Coffee break*  
09:55 - 12:00     **Session 1: Diversity and Selective Ensemble**  
12:00 - 12:30     *Lunch*  
12:30 - 18:30     *Excursion*  
18:30 - 21:00     *Banquet*

## May 16

- 07:20 - 07:30     *Bus pickup (at New Era Hotel)*  
08:30 - 09:30     **Invited Talk: Marcello Pelillo**  
09:30 - 09:45     *Coffee break*  
09:45 - 11:50     **Session 2: Class-Imbalance Learning**  
11:50 - 12:50     *Lunch*  
12:50 - 14:55     **Session 3: Feature Processing**  
14:55 - 15:10     *Coffee break*  
15:10 - 17:40     **Session 4: Classifier Ensemble**  
17:40 - 18:30     *Lab visit & campus tour*  
18:30 - 20:00     *Dinner (in Xianlin campus)*

## May 17

- 07:20 - 07:30     *Bus pickup (at New Era Hotel)*  
08:30 - 09:30     **Invited Talk: Xin Yao**  
09:30 - 09:45     *Coffee break*  
09:45 - 11:50     **Session 5: Multi-Instance and Part-Based Learning**  
11:50 - 12:50     *Lunch*  
12:50 - 14:05     **Session 6: Clustering Ensemble**  
14:05 - 14:20     *Coffee break*  
14:20 - 16:25     **Session 7: Applications**  
16:25 - 16:30     **Closing**

\* All sessions will be in Room 233 of the CS Building (Xianlin campus)

## May 14 (PSL)

08:40 - 09:40 **Invited Talk**

Session chair: Friedhelm Schwenker (University of Ulm, Germany)

### *Learning with Output Kernels and Latent Kernels*

Dale Schuurmans

Department of Computing Science

University of Alberta, Canada



#### Abstract:

Kernels provide a finite representation of abstract feature mappings, which is an invaluable modeling tool in computational data analysis. Although kernel methods are frequently applied to input representations, particularly when learning functions, much less work has considered applying kernels to output representations or representations of latent structure. In this talk, I will discuss some new applications of kernel methods to both output and latent representations. I will first discuss how losses beyond squared error can be exactly and efficiently accommodated within output kernels. Then I will discuss how kernel methods can be extended to partially supervised learning problems, such as latent representation discovery. Such kernel based extensions are able to achieve state of the art results in robust regression, multi-label classification, and hidden-layer network training.

#### Short Biography:

Dale Schuurmans is a Professor of Computing Science and Canada Research Chair in Machine Learning at the University of Alberta. He received his PhD in Computer Science from the University of Toronto, and has been employed at the National Research Council Canada, University of Pennsylvania, NEC Research Institute and the University of Waterloo. He is an Associate Editor of JAIR and AIJ, and currently serves on the IMLS and NIPS Foundation boards. He has previously served as a Program Co-chair for NIPS-2008 and ICML-2004, and as an Associate Editor for IEEE TPAMI, JMLR and MLJ. His research interests include machine learning, optimization, probability models, and search. He is author of more than 130 refereed publications in these areas and has received paper awards at IJCAI, AAIL, ICML, IEEE ICAL and IEEE ADPRL.

## **Anomaly Detection using Convex Hulls on A Pairwise Estimated Parameter Space**

*Gabriel B. P. Costa and Moacir P. Ponti Jr.  
(Universidade de São Paulo, Brazil)*

Anomaly detection is the problem of identifying objects appearing to be inconsistent with the remainder of that set of data. Detecting such samples is useful on various applications such as fault detection, fraud detection and diagnostic systems. Partially supervised methods for anomaly detection are interesting because they only need data labeled as one of the classes (normal or abnormal). In this paper, we propose a partially supervised framework for anomaly detection based on convex hulls in a parameter space, assuming a given probability distribution. It can be considered a framework since it support any distribution or combination of distribution, and allow variations in the training step, modelling only normal samples, only anomalies, or both. We investigated an algorithm based on this framework, assuming the normal distribution for the normal (not anomalous) data and compared the results with statistical algorithms and the Naive Bayes classifier. The proposed method performed well and showed results comparable or better than the competing methods. Furthermore, the proposed method can handle any probability distribution or mixture of distributions, allowing the user to choose a parameter space that better models the problem of finding anomalies.

## **Adaptive Graph Constrained NMF for Semi-Supervised Learning**

*Qian Li, Liping Jing and Jian Yu  
(Beijing Jiaotong University, China)*

Recently, graph-based semi-supervised learning (GB-SSL) has received a lot of attentions in pattern recognition, computer vision and information retrieval. The key parts of GB-SSL are designing loss function and constructing graph. In this paper, we proposed a new semi-supervised learning method where the loss function is modeled via graph constrained non-negative matrix factorization (GCNMF). The model can effectively cooperate the precious label information and the local consistency among samples including labeled and unlabeled data. Meanwhile, an adaptive graph construction method is presented so that the selected neighbors of one sample are as similar as possible, which makes the local consistency be correctly preserved in the graph. The experimental results on real world data sets including object image, face and handwritten digit have shown the superiority of our proposed method.



## **A New Weighted Sparse Representation Based on MSLBP and Its Application to Face Recognition**

*He-Feng Yin and Xiao-Jun Wu*  
(Jiangnan University, China)

Face recognition via sparse representation-based classification has received more and more attention in recent years. This approach has achieved state-of-the-art results, which outperforms traditional methods, especially when face image pixels are corrupted or occluded. In this paper, we propose a new weighted sparse representation method called WSRC-MSLBP which utilizes the multi-scale LBP (MSLBP) feature to measure similarity between face images, and to form the weight matrix. The proposed WSRC-MSLBP method not only represents the test sample as a sparse linear combination of all the training samples, but also makes use of locality of local binary pattern. Experimental results on publicly available databases show that the proposed WSRC-MSLBP method is more effective than sparse representation-based classification algorithm and the original weighted sparse representation method.

**12:40 - 14:20 Afternoon Session 1**

Session chair: Ming Li (Nanjing University, China)

## **Conscientiousness Measurement from Weibo's Public Information**

*Dong Nie, Tingshao Zhu and Lin Li*  
(University of Chinese Academy of Sciences, China)

We apply a graph-based semi-supervised learning algorithm to identify the conscientiousness of Weibo users. Given a set of Weibo users' public information (e.g., number of followers) and a few labeled Weibo users, the task is to predict conscientiousness assessment for numeric unlabeled Weibo users. Singular value decomposition (SVD) technique is taken for feature reduction, and K nearest neighbor (KNN) method is used to recover a sparse graph. The local and global consistency algorithm is followed to deal with our data. Experiments demonstrate the advantage of semi-supervised learning over standard supervised learning when limited labeled data are available.

## **A Semi-Supervised Dictionary Learning Approach for Sparse Emotion Recognition**

*Markus Kächele and Friedhelm Schwenker*  
(University of Ulm, Germany)

This work presents a technique for the classification of emotions based on biophysiological data. A dictionary learning approach is used to generate sparse representations of blood volume pulse signals, which are then used as input for classification algorithms. Unlabeled data, i.e. data that is not relevant for the classification task, is used to enrich the dictionary learning stage. Superior representation abilities of the underlying structure of the data are demonstrated by the learnt dictionaries. As a result, classification rates are improved. Experimental validation in the form of

different classification experiments is presented. The results are presented with a discussion about the benefits of the approach and the existing limitations.

## **Neighborhood Co-regularized Multi-view Spectral Clustering for Microbiome Data**

*Evgeni Tsivtsivadze (The Netherlands Organization for Applied Scientific Research, The Netherlands)*

*Hanneke Borgdorff (University of Amsterdam and Amsterdam Institute for Global Health and Development, The Netherlands)*

*Janneke van de Wijert (University of Amsterdam and Amsterdam Institute for Global Health and Development, The Netherlands)*

*Frank Schuren (The Netherlands Organization for Applied Scientific Research, The Netherlands)*

*Rita Verhelst (Ghent University, Belgium)*

*Tom Heskes (Radboud University, The Netherlands)*

In many unsupervised learning problems data can be available in different representations, often referred to as views. By leveraging information from multiple views we can obtain clustering that is more robust and accurate compared to the one obtained via the individual views. We propose a novel algorithm that is based on neighborhood co-regularization of the clustering hypotheses and that searches for the solution which is consistent across different views. In our empirical evaluation on publicly available datasets, the proposed method outperforms several state-of-the-art clustering algorithms. Furthermore, application of our method to recently collected biomedical data leads to new insights, critical for future research on determinants of the vaginal microbiome and the vaginal microbiome as a risk factor for the transmission of HIV. These insights could have an influence on the interpretation of clinical presentation of women with bacterial vaginosis and treatment decisions.

## **A Robust Image Watermarking Scheme Based on BWT and ICA**

*Tao Wang (Anhui University, Key Lab of Industrial Image Processing & Analysis of Anhui Province, China)*

*Jin Tang (Anhui University, Key Lab of Industrial Image Processing & Analysis of Anhui Province, China)*

*Bin Luo (Anhui University, Key Lab of Industrial Image Processing & Analysis of Anhui Province, China)*

*Cheng Zhang (Anhui University, China)*

A robust image watermarking scheme combined with the human visual characteristics is proposed. Berkeley wavelet transform (BWT) which is used in watermarking embedding procedure simulates physiology characteristics of the mammalian primary visual cortex (V1). Independent Component Analysis (ICA) which is blind separation technology will be adapted to the watermarking extracting procedure. By combining the advantages of BWT and ICA, a robust image watermarking scheme is proposed and a simulation of the scheme is designed. Experimental results demon-

strate that the proposed watermarking technique combines the imperceptibility, robustness, real-time and high capacity of digital watermarking algorithms.

**14:35 - 15:50 Afternoon Session 2**

**Session chair: Moacir P. Ponti Jr. (Universidade de São Paulo, Brazil)**

## **Kernel Parameter Optimization in Stretched Kernel-Based Fuzzy Clustering**

*Chunhong Lu, Zhaomin Zhu and Xiaofeng Gu  
(Jiangnan University, China)*

Although the kernel-based fuzzy c-means (KFCM) algorithm utilizing a kernel-based distance measure between patterns and cluster prototypes outperforms the standard fuzzy c-means clustering for some complex distributed data, it is quite sensitive to selected kernel parameters. In this paper, we propose the stretched kernel-based fuzzy clustering method with optimized kernel parameter. The kernel parameters are updated in accordance with the gradient method to further optimize the objective function during each iteration process. To solve the local minima problem of the objective function, a function stretching technique is applied to detect the global minimum. Experiments on both synthetic and real-world datasets show that the stretched KFCM algorithm with optimized kernel parameters has better performance than other algorithms.

## **Meta-Learning of Exploration/Exploitation Parameters on a Robot with Limited Resources**

*Michel Tokic (University of Ulm, Germany; University of Applied Sciences  
Ravensburg-Weingarten, Germany)  
Friedhelm Schwenker (University of Ulm, Germany)  
Guenther Palm (University of Ulm, Germany)*

Controlling exploration and exploitation is one of the main challenges when developing autonomous learning agents. In general, exploratory actions lead to an increase in knowledge about the long-term utility of actions (long-term optimization), but often may cause the income of negative reward due to randomly selected bad actions. However, exploiting knowledge (short-term optimization) may also lead to sub-optimal action selections if the utility of an optimal action is underestimated. As a consequence, the dilemma between exploration and exploitation arises.

In this paper we consider the problem of adapting the amount of exploration and exploitation in model-free reinforcement learning. We apply the VDBE-Softmax approach to real-world reward data obtained from a little crawling robot. The results are compared to the action-selection policies  $\epsilon$ -Greedy and Softmax, where the latter one is also known to be a plausible model for adapting exploration and exploitation in the human brain. In artificial experiments, however, setting the exploratory parameter of Softmax and  $\epsilon$ -Greedy is non-trivial, for which reason we investigate meta-learning of this parameter. We investigate a reward model sampled from real-

world interactions of a crawling robot aiming at learning to move forward. Since the real hardware robot has only 2kB of memory due to an ATmega32 microcontroller, the problem of learning under limited resources arises, which is solved by discretizing the continuous state-space model.

## **Self-Practice Imitation Learning from Weak Policy**

*Qing Da, Yang Yu and Zhi-Hua Zhou*  
(*Nanjing University, China*)

Imitation learning is an effective strategy to reinforcement learning, which avoids the delayed reward problem by learning from mentor-demonstrated trajectories. A limitation for imitation learning is that collecting sufficient qualified demonstrations is quite expensive. In this work, we study how an agent can automatically improve its performance from a weak policy, by automatically acquiring more demonstrations for learning. We propose the LEWE framework to sample tasks for the weak policy to execute, and then learn from the successful trajectories to achieve an improvement. As the sampling strategy is the key to the efficiency of LEWE, we further propose to incorporate active learning for the sampling strategy for LEWE. Experiments in a spatial positioning task show that LEWE with active learning can effectively and efficiently improve the weak policy and achieves a better performance than the comparing sampling approaches.

# May 15 (MCS Day 1)

08:40 - 09:40 **Invited Talk**

Session chair: Zhi-Hua Zhou (Nanjing University, China)

## *Stability and Large Scale Statistical Inference*

Bin Yu

Department of Statistics

Department of Electrical Engineering & Computer Science

University of California, Berkeley, USA



### Abstract:

Information technology has enabled collection of large amounts of and high-dimensional data across fields in science, engineering, social science, commerce, and beyond. Reproducibility is imperative for any meaningful discovery from these data. Statistical machine learning analyses are often used to bring the discoveries and, as a minimal manifestation of reproducibility, conclusions from statistical analyses have to be stable or robust to "reasonable" perturbations to data and to the model used. Examples of data perturbation schemes include Jackknife, bootstrap, and cross-validation while robust statistics aims at studying perturbations to models. In this talk, the indispensable requirement of stability is advocated for large scale statistical inference. A movie-reconstruction from fMRI data motivates the need for stability, and a new estimation stability revision on CV for Lasso is proposed to lead to a much simpler and more reliable model. Finally, a novel stability analytical argument is seen to drive new results that shed light on the intriguing interactions between sample to sample variability and heavier tail error distribution (e.g. double-exponential) in high-dim robust statistics.

### Short Biography:

Bin Yu is Chancellor's Professor in the Departments of Statistics and of Electrical Engineering & Computer Science at UC Berkeley. She has published over 100 scientific papers in premier journals in Statistics, EECS, remote sensing and neuroscience, in a wide range of research areas including empirical process theory, information theory (MDL), MCMC methods, signal processing, machine learning, high dimensional data inference (boosting and Lasso and sparse modeling in general), and interdisciplinary data problems. She has served on many editorial boards for journals such as Annals of Statistics, Journal of American Statistical Association, and Journal of Machine Learning Research.

She was a 2006 Guggenheim Fellow, co-recipient of the Best Paper Award of IEEE Signal Processing Society in 2006, and the 2012 Tukey Memorial Lecturer of the Bernoulli Society (selected every four years). She is a Fellow of AAAS, IEEE, IMS (Institute of Mathematical Statistics) and ASA (American Statistical Association).

She is currently President-Elect of IMS (Institute of Mathematical Statistics). She is serving on the Scientific Advisory Board of IPAM (Institute of Pure and Applied Mathematics) and on the Board of Mathematical Sciences and Applications of NAS. She was co-chair of the National Scientific Committee of SAMSI (Statistical and Applied Mathematical Sciences Institute), and on the Board of Governors of IEEE-IT Society.

### **Diversity in Classifier Ensembles: Fertile Concept or Dead End?**

*Luca Didaci, Giorgio Fumera, and Fabio Roli*

*(University of Cagliari, Italy)*

Diversity is deemed a crucial concept in the field of multiple classifier systems, although no exact definition has been found so far. Existing diversity measures exhibit some issues, both from the theoretical viewpoint, and from the practical viewpoint of ensemble construction. We propose to address some of these issues through the derivation of decompositions of classification error, analogue to the well-known bias-variance-covariance and ambiguity decompositions of regression error. We then discuss whether the resulting decompositions can provide a more clear definition of diversity, and whether they can be exploited more effectively for the practical purpose of ensemble construction.

### **Can Diversity amongst Learners Improve Online Object Tracking?**

*Georg Nebehay (Austrian Institute of Technology, Austria)*

*Walter Chibamu (University of Birmingham, U.K.)*

*Peter R. Lewis (University of Birmingham, U.K.)*

*Arjun Chandra (University of Oslo, Norway)*

*Roman Pflugfelder (Austrian Institute of Technology, Austria)*

*Xin Yao (University of Birmingham, U.K.)*

We present a novel analysis of the state of the art in object tracking with respect to diversity found in its main component, an ensemble classifier that is updated in an online manner. We employ established measures for diversity and performance from the rich literature on ensemble classification and online learning, and present a detailed evaluation of diversity and performance on benchmark sequences in order to gain an insight into how the tracking performance can be improved.

### **Selective Ensemble of Classifier Chains**

*Nan Li, and Zhi-Hua Zhou*

*(Nanjing University, China)*

In multi-label learning, the relationship among labels is well accepted to be important, and various methods have been proposed to exploit label relationships. Amongst them, ensemble of classifier chains (ECC) which builds multiple chaining classifiers by random label orders has drawn much attention. However, the ensembles generated by ECC are often unnecessarily large, leading to extra high computational and storage cost. To tackle this issue, in this paper, we propose selective ensemble of classifier chains (SECC) which tries to select a subset of classifier chains to composite the ensemble whilst keeping or improving the performance. More precisely, we focus on the performance measure F1-score, and formulate this problem as

a convex optimization problem which can be efficiently solved by the stochastic gradient descend method. Experiments show that, compared with ECC, SECC is able to obtain much smaller ensembles while achieving better or at least comparable performance.

## **ECOC Matrix Pruning Using Accuracy Information**

*Cemre Zor, Terry Windeatt, and Josef Kittler*  
(University of Surrey, U.K.)

The target of ensemble pruning is to increase efficiency by reducing the ensemble size of a multi classifier system and thus computational and storage costs, without sacrificing and preferably enhancing the generalization performance. However, most state-of-the-art ensemble pruning methods are based on unweighted or weighted voting ensembles; and their extensions to the Error Correcting Output Coding (ECOC) framework is not strongly evident or successful. In this study, a novel strategy for pruning ECOC ensembles which is based on a novel accuracy measure is presented. The measure is defined by establishing the link between the accuracies of the two-class base classifiers in the context of the main multiclass problem. The results show that the method outperforms the ECOC extensions of the state-of-the-art pruning methods in the majority of cases and that it is even possible to improve the generalization performance by only using 30% of the initial ensemble size in certain scenarios.

## **Selective Clustering Ensemble Based on Covariance**

*Xuyao Lu, Yan Yang, and Hongjun Wang*  
(Southwest Jiaotong University, China)

Clustering Ensemble effectively improves clustering accuracy, stability and robustness, which is most resulted from the diversity of the base clustering results. It is a key point to measure the diversity of clustering results. This paper proposes a method to measure diversity of base clustering results and a covariance-based selective clustering ensemble algorithm. Experiments on 20 UCI data sets show that this algorithm effectively improves the clustering performance.



08:30 - 09:30 **Invited Talk**

Session chair: Fabio Roli (University of Cagliari, Italy)

## *Similarity-Based Pattern Recognition: A Game-Theoretic Perspective*

Marcello Pelillo

Department of Environmental Sciences, Informatics and Statistics

Università Ca' Foscari Venezia, Italy



### Abstract:

The classical approach to deal with non-geometric (dis)similarities is "embedding", which refers to any procedure that takes a set of (dis)similarities as input and produces a vectorial representation of the data as output, such that the proximities are either locally or globally preserved. This is an (approximate or ideal) isometric mapping which finds a set of vectors in an instance-specific Euclidean space that are capable of describing the data satisfactorily. Embedding approaches are all based on the assumption that the non-geometricity of similarity information can be eliminated or some-how approximated away. When this is not the case, i.e., when there is significant information content in the non-geometricity of the data, however, alternative approaches are needed.

In this talk, I will maintain that game theory offers an elegant and powerful conceptual framework which serves well this purpose. The development of game theory in the early 1940's by von Neumann was a reaction against the then dominant view that problems in economic theory can be formulated using standard methods from optimization theory. Indeed, most real-world economic problems typically involve conflicting interactions among decision-making agents that cannot be adequately captured by a single (global) objective function, thereby requiring a different, more sophisticated treatment. Accordingly, the main point made by game theorists is to shift the emphasis from optimality criteria to equilibrium conditions, namely to the search of a balance among multiple interacting forces. Interestingly, the development of evolutionary game theory in the late 1970's by Maynard Smith offered a dynamical systems perspective to game theory, an element which was totally missing in the traditional formulation, and provided powerful tools to deal with the equilibrium selection problem.

As it provides an abstract theoretically-founded framework to elegantly model complex scenarios, game theory has found a variety of applications not only in economics and, more generally, in the social sciences but also in different fields of engineering and information technologies. In the talk, I will describe recent attempts aimed at formulating (or interpreting) several pattern recognition and machine learning problems from a game-theoretic perspective. Indeed, many problems within these fields can naturally be formulated at an abstract level in terms of a game where (pure) strategies correspond to class labels and the payoff function is expressed in terms of competition between the hypotheses of class membership. In particular, I will focus on data clustering, semi-supervised learning, structural matching, and contextual pattern recognition. Further, I will discuss potential applications of game theory within the context of multiple classifier systems.

#### Short Biography:

Marcello Pelillo joined the faculty of the University of Bari, Italy, as an Assistant professor of computer science in 1991. Since 1995, he has been with the University of Venice, Italy, where he is currently a Full Professor of Computer Science.

He leads the Computer Vision and Pattern Recognition Group and has served from 2004 to 2010 as the Chair of the board of studies of the Computer Science School. He held visiting research positions at Yale University, the University College London, McGill University, the University of Vienna, York University (UK), and the National ICT Australia (NICTA). He has published more than 130 technical papers in refereed journals, handbooks, and conference proceedings in the areas of computer vision, pattern recognition and neural computation.

He serves (or has served) on the editorial board for the journals IEEE Transactions on Pattern Analysis and Machine Intelligence and Pattern Recognition, and is regularly on the program committees of the major international conferences and workshops of his fields. In 1997, he co-established a new series of international conferences devoted to energy minimization methods in computer vision and pattern recognition (EMMCVPR) which has now reached the ninth edition. He is (or has been) scientific coordinator of several research projects, including SIMBAD, an EU-FP7 project devoted to similarity-based pattern analysis and recognition. Prof. Pelillo is a Fellow of the IEEE and of the IAPR.

### **Random Oracle Ensembles for Imbalanced Data**

*Juan J. Rodríguez, José-Francisco Díez-Pastor, and César García-Osorio  
(University of Burgos, Spain)*

In the Random Oracle ensemble method, each base classifier is a mini-ensemble of two classifiers and a randomly generated oracle that selects one of the two classifiers. The performance of this method have been previously studied, but not for imbalanced data sets. This work studies its performance for this kind of data. As the Random Oracle ensemble method can be combined with any other ensemble method, this work considers its combination with four ensemble methods: Bagging, SMOTE-Boost, SMOTEBagging and RUSBoost. The last three methods combine classical, not specific for imbalance, ensemble methods (i.e., Bagging, Boosting), with pre-processing approaches designed for imbalance (i.e., random undersampling, SMOTE). The results show that Random Oracles improves all these methods.

### **Ensembles of Optimum-Path Forest Classifiers Using Input Data Manipulation and Undersampling**

*Moacir P. Ponti Jr., and Isadora Rossi  
(University of São Paulo, Brazil)*

The combination of multiple classifiers was proven to be useful in many applications to improve the classification task and stabilize results. In this paper we used the Optimum-Path Forest (OPF) classifier to investigate input data manipulation techniques in order to use less data from the training set without hampering the classification accuracy. The data undersampling can be useful to speed-up the classification task, and could be specially useful with large datasets. The results indicate that the OPF-based ensemble methods allow a significant reduction on the size of the training set, while maintaining or slightly improving accuracy. We provide intuition for a case of failure and report the results of synthetic and real datasets.

### **Similarity Weighted Ensembles for Relocating Models of Rare Events**

*Claire D'Este, and Ashfaqur Rahman  
(Intelligent Sensing and Systems Laboratory, Australia)*

Spatially distributed regions may have different influences that affect the underlying physical processes and make it inappropriate to directly relocate learned models. We may also be aiming to detect rare events for which we have examples in some regions, but not others. A novel method is presented for combining classifiers trained on regions with known sensor data and predicting rare events in new regions, specifically the closure of shellfish farms. The proposed similarity weighted ensemble method demonstrates an average 10 fold improvement in accuracy over One Class classification and 3 fold improvement over rules hand-crafted by an expert.

## **Adaptive Ensemble Selection for Face Re-identification under Class Imbalance**

*Paulo Radtke (Université du Québec, Canada)*

*Eric Granger (Université du Québec, Canada)*

*Robert Sabourin (Université du Québec, Canada)*

*Dmitry Gorodnichy (Canada Border Services Agency, Canada)*

Systems for face re-identification over a network of video surveillance cameras are designed with a limited amount of reference data, and may operate under complex environments. Furthermore, target individuals provide a small proportion of the facial captures for design and during operations, and these proportions may change over time according to operational conditions. Given a diversified pool of base classifiers and a desired false positive rate (fpr), the Skew-Sensitive Boolean Combination (SSBC) technique allows to adapt the selection of ensembles based on changes to levels of class imbalance, as estimated from the input video stream. Initially, a set of BCs for the base classifiers is produced in the ROC space, where each BC curve corresponds to reference data with a different level of imbalance. Then, during operations, class imbalance is periodically estimated using the Hellinger distance between the data distribution of inputs and that of imbalance levels, and used to approximate the most accurate BC of classifiers among operational points of these curves viewed in the precision-recall space. Simulation results on real-world video surveillance data indicate that, compared to traditional approaches, FR systems based on SSBC allow to select BCs that provide a higher level of precision for target individuals, and a significantly smaller difference between desired and actual fpr. Performance of this adaptive approach is also comparable to full recalculation of BCs (for a specific level of imbalance), but for a considerably lower complexity. Using face tracking, a high level of discrimination between target and non-target individuals may be achieved by accumulating SSBC predictions for faces captured corresponding to a same track in video footage.

## **Cascaded Reduction and Growing of Result Sets for Combining Object Detectors**

*Uwe Knauer (Humboldt University, Germany)*

*Udo Seiffert (Fraunhofer IFF, Germany)*

In this paper cascaded reduction and growing of result sets is introduced as a principle for combining the results of different object detectors. First, different candidate operating points are selected for each object detection algorithm. This procedure is based on the analysis of precision and recall of the individual methods. Selecting an appropriate operating point prior to fusion is important because it regulates the cardinal number of the result set. As diversity and correlation between object detectors also depend on the elements of the result sets, this and the application of set operations allow to create a final set of detected objects by including missing and excluding false detections. The approach allows both diverse and correlated detectors to contribute to the performance of the combined detector. The performance of the proposed algorithm is compared to other combining algorithms. It outperforms or com-

petes with existing state of the art combiners for several datasets. Additionally, the results provide a significant improvement in the interpretability of the combining rules. As a unique feature of the proposed algorithm, the found operating points can be used to reconfigure the object detection algorithms to adapt their individual results to the needs of the combination procedure allowing a reduction in runtime.

**12:50 - 14:55 Session 3: Feature Processing**

**Session chair: Marco Loog (Delft University of Technology, The Netherlands)**

### **Dimensionality Reduction Using Stacked Kernel Discriminant Analysis for Multi-label Classification**

*Muhammad Atif Tahir (Al-Imam Mohammad Ibn Saud Islamic University, KSA)*

*Ahmed Bouridane (Northumbria University, U.K.)*

*Josef Kittler (University of Surrey, U.K.)*

Multi-label classification in which each instance may belong to more than one class is a challenging research problem. Recently, a considerable amount of research has been concerned with the development of “good” multi-label learning methods. Despite the extensive research effort, many scientific challenges posed by e.g. curse-of-dimensionality and correlation among labels remain to be addressed. In this paper, we propose a new approach to multi-label classification which combines stacked Kernel Discriminant Analysis using Spectral Regression (SR-KDA) with state-of-the-art instance-based multi-label (ML) learning method. The proposed system is validated on two multi-label databases. The results indicate significant performance gains when compared with the state-of-the art multi-label methods for multi-label classification.

### **Stable L2-Regularized Ensemble Feature Weighting**

*Yun Li (Nanjing University of Posts and Telecommunications, China)*

*Shasha Huang (Nanjing University of Posts and Telecommunications, China)*

*Songcan Chen (Nanjing University of Aeronautics and Astronautics, China)*

*Jennie Si (Arizona State University, USA)*

When selecting features for knowledge discovery applications, stability is a highly desired property. By stability of feature selection, here it means that the feature selection outcomes vary only insignificantly if the respective data change slightly. Several stable feature selection methods have been proposed, but only with empirical evaluation of the stability. In this paper, we aim at providing a try to give an analysis for the stability of our ensemble feature weighting algorithm. As an example, a feature weighting method based on L2-regularized logistic loss and its ensembles using linear aggregation is introduced. Moreover, the detailed analysis for uniform stability and rotation invariance of the ensemble feature weighting method is presented. Additionally, some experiments were conducted using real-world microarray data sets.

Results show that the proposed ensemble feature weighting methods preserved stability property while performing satisfactory classification. In most cases, at least one of them actually provided better or similar tradeoff between stability and classification when compared with other methods designed for boosting the stability.

## **Ensemble of Feature Chains for Anomaly Detection**

*Lena Tenenboim-Chekina, Lior Rokach, and Bracha Shapira  
(Ben-Gurion University of the Negev, Israel)*

Along with recent technological advances more and more new threats and advanced cyber-attacks appear unexpectedly. Developing methods which allow for identification and defense against such unknown threats is of great importance. In this paper we propose new ensemble method (which improves over the known cross-feature analysis, CFA, technique) allowing solving anomaly detection problem in semi-supervised settings using well established supervised learning algorithms. Theoretical correctness of the proposed method is demonstrated. Empirical evaluation results on Android malware datasets demonstrate effectiveness of the proposed approach and its superiority against the original CFA detection method.

## **A New Feature Fusion Approach Based on LBP and Sparse Representation and Its Application to Face Recognition**

*He-Feng Yin and Xiao-Jun Wu  
(Jiangnan University, China)*

In this paper, we propose a new feature fusion approach based on local binary pattern (LBP) and sparse representation (SR). Firstly, local features are extracted by LBP and global features are sparse coefficients which are obtained via decomposing samples based on the over-complete dictionary. Then the global and local features are fused in a serial fashion. Afterwards PCA is used to reduce the dimensionality of the fused vector. Finally, SVM is employed as a classifier on the reduced feature space for classification. Experimental results obtained on publicly available databases show that the proposed feature fusion method is more effective than other methods like LBP+PCA, Gabor+PCA and Gabor+SR in terms of recognition accuracy.

## **Feature Level Multiple Model Fusion Using Multilinear Subspace Analysis with Incomplete Training Set and Its Application to Face Image Analysis**

*Zhen-Hua Feng (Jiangnan University, China)  
Josef Kittler (University of Surrey, U.K.)  
William Christmas (University of Surrey, U.K.)  
Xiao-Jun Wu (Jiangnan University, China)*

In practical applications of pattern recognition and computer vision, the performance of many approaches can be improved by using multiple models. In this paper, we develop a common theoretical framework for multiple model fusion at the feature level using multilinear subspace analysis (also known as tensor algebra). One disad-

vantage of the multilinear approach is that it is hard to obtain enough training observations for tensor decomposition algorithms. To overcome this difficulty, we adopted the M2SA algorithm to reconstruct the missing entries of the incomplete training tensor. Furthermore, we apply the proposed framework to the problem of face image analysis using Active Appearance Model (AAM) to validate its performance. Evaluations of AAM using the proposed framework are conducted on Multi-PIE face database with promising results.

15:10 - 17:40 **Session 4: Classifier Ensemble**

Session chair: Xiao-Jun Wu (Jiangnan University, China)

## **Towards a Framework for Designing Full Model Selection and Optimization Systems**

*Quan Sun, Bernhard Pfahringer, and Michael Mayo*  
(The University of Waikato, New Zealand)

People from a variety of industrial domains are beginning to realise that appropriate use of machine learning techniques for their data mining projects could bring great benefits. End-users now have to face the new problem of how to choose a combination of data processing tools and algorithms for a given dataset. This problem is usually termed the Full Model Selection (FMS) problem. Extended from our previous work [10], in this paper, we introduce a framework for designing FMS algorithms. Under this framework, we propose a novel algorithm combining both genetic algorithms (GA) and particle swarm optimization (PSO) named GPS (which stands for GA-PSO-FMS), in which a GA is used for searching the optimal structure for a data mining solution, and PSO is used for searching optimal parameters for a particular structure instance. Given a classification dataset, GPS outputs a FMS solution as a directed acyclic graph consisting of diverse data mining operators that are available to the problem. Experimental results demonstrate the benefit of the algorithm. We also present, with detailed analysis, two model-tree-based variants for speeding up the GPS algorithm.

## **Coding Theory Tools for Improving Recognition Performance in ECOC Systems**

*Claudio Marrocco, Paolo Simeone, and Francesco Tortorella*  
(Università degli Studi di Cassino, Italia)

Error-correcting output coding (ECOC) is nowadays an established technique to build polychotomous classification systems by aggregating highly efficient dichotomizers. This approach has exhibited good classification performance and generalization capabilities in many practical applications. In this field much work has been devoted to study new solutions both for the coding and the decoding phase, but little attention has been paid to the algebraic tools typically employed in the Coding Theory, which could provide an ECOC design approach based on robust theoretical foundations. In this paper we propose an ECOC classification system based on Low

Density Parity Check (LDPC) Codes, a well known technique in Coding Theory. Such framework is particularly suitable to define an ECOC system that employs dichotomizers provided of a reject option. The experiments on some public data sets have demonstrated that, in this way, the ECOC system can reach good recognition rates when a suitable reject level is imposed to the dichotomizers.

## **A Directed Inference Approach towards Multi-class Multi-model Fusion**

*Tianbao Yang, Lei Wu, and Piero P. Bonissone*

*(GE Global Research Center, USA)*

In this paper, we propose a directed inference approach for multi-class multi-model fusion. Different from traditional approaches that learn a model in training stage and apply the model to new data points in testing stage, directed inference approach constructs (one) general direction of inference in training stage, and constructs an individual (ad-hoc) rule for each given test point in testing stage. In the present work, we propose a framework for applying the directed inference approach to multiple model fusion problems that consists of three components: (i) learning of individual models on the training samples, (ii) nearest neighbour search for constructing individual rules of bias correction, and (iii) learning of an optimal combination weights of individual models for model fusion. For inference on a test sample, the prediction scores of individual models are first corrected with bias estimated from the nearest training data points, and then the corrected scores are combined using the learned optimal weights. We conduct extensive experiments and demonstrate the effectiveness of the proposed approach towards multi-class multiple model fusion.

## **Randomized Bayesian Network Classifiers**

*Qing Wang (Fudan University, China)*

*Ping Li (Anhui University of Technology, China)*

In this paper, we propose Randomized Bayesian Network Classifiers (RBNC). It borrows the idea of ensemble learning by constructing a collection of semi-naive Bayesian network classifiers and then combines their predictions as the final output. Specifically, the structure learning of each component Bayesian network classifier is performed by just randomly choosing the parent of each attribute in addition to class attribute, and parameter learning is performed by using maximum likelihood method. RBNC retains many of naive Bayes' desirable property, such as scaling linearly with respect to both the number of instances and attributes, needing a single pass through the training data and robust to noise, etc. On the 60 widely used benchmark UCI datasets, RBNC outperforms state-of-the-art Bayesian classifiers.



# Self-Organizing Neural Grove and Its Application to Incremental Learning

*Hiroataka Inoue*

*(Kure National College of Technology, Japan)*

Recently, multiple classifier systems have been used for practical applications to improve classification accuracy. Self-generating neural networks (SGNN) are one of the most suitable base-classifiers for multiple classifier systems because of their simple settings and fast learning ability. However, the computation cost of the multiple classifier system based on SGNN increases in proportion to the numbers of SGNN. In this paper, we propose a novel pruning method for efficient classification and we call this model a self-organizing neural grove (SONG). Experiments have been conducted to compare the SONG with bagging and the SONG with boosting, the multiple classifier system based on C4.5, and support vector machine (SVM). The results show that the SONG can improve its classification accuracy as well as reducing the computation cost. Additionally, we investigate SONG's incremental learning performance.

## A Novel Pattern Rejection Criterion Based on Multiple Classifiers

*Wei-Na Wang (Concordia University, Canada)*

*Xu-Yao Zhang (Chinese Academy of Sciences, China)*

*Ching Y. Suen (Concordia University, Canada)*

Aiming at improving the reliability of a recognition system, this paper presents a novel SVM-based rejection measurement (SVMM) and voting based combination methods of multiple classifier system (MCS) for pattern rejection. Compared with the previous heuristic designed criteria, SVMM is more straight-forward and can make use of much more information from the training data. The voting based combination methods for rejection is a preliminary attempt to adopt MCS for rejection. Comparison of SVMM with other well-known rejection criteria proves that it achieves the highest performance. Two different methods (structural modification and dataset re-sampling) are used to build MCSs. The basic classifier is the convolution neural network (CNN) which has achieved promising performances in numerous applications. Rejection based on MCS is then evaluated on MNIST and CENPARMI digit databases. Specifically, different rejection criteria (FRM, FTRM and SVMM) are individually combined with MCS for pattern rejection. Experimental results indicate that these combinations improve the rejection performance consistently and MCS built by dataset re-sampling works better than that with structural modification in rejection.

# May 17 (MCS Day 3)

08:30 - 09:30 **Invited Talk**

Session chair: Josef Kittler (University of Surrey, U.K.)

## *From Evolutionary Computation to Ensemble Learning*

Xin Yao

School of Computing Science

University of Birmingham, U.K.



### Abstract:

Designing a monolithic system for a large and complex learning task is hard. Divide-and-conquer is a common strategy in tackling such large and complex problems. Artificial speciation and niching have long been used in evolutionary computation as one way towards automatic divide-and-conquer. It turns out that several key ideas in speciated evolutionary algorithms are closely linked to ensemble learning in general, especially diversity in ensembles. This talk reviews selected work on this topic and illustrates the link between evolutionary computation and ensemble learning using the example of negative correlation learning. Then the importance of diversity is demonstrated using examples from online learning and class imbalance learning. As a practical solution towards accurate and diverse ensembles, multi-objective ensemble learning is advocated. Finally, a recent application of multi-objective learning to software effort estimation will be described.

### Short Biography:

Xin Yao is a professor of computer science from the University of Birmingham, UK. He was also a Distinguished Visiting Professor of the University of Science and Technology of China (USTC), P. R. China, and a visiting professor of three other universities. He is a Fellow of the IEEE, Editor-in-Chief (2003-08) of IEEE Transactions on Evolutionary Computation, an associate editor or an editorial board member of 11 other journals. He was the winner of 2001 IEEE Donald G. Fink prize paper award and several other best paper awards. His research interests include evolutionary computation, neural network ensembles, and their applications. He has more than 300 refereed publications in those areas. He is currently the Director of the Centre of Excellence for Research in Computational Intelligence and Applications (CER-CIA), which is focused on applied research and knowledge transfer to industry.

## **The Link between Multiple-Instance Learning and Learning from Only Positive and Unlabelled Examples**

*Yan Li, David M.J. Tax, Robert P.W. Duin, and Marco Loog*  
(Delft University of Technology, The Netherlands)

This paper establishes a link between two supervised learning frameworks, namely multiple-instance learning (MIL) and learning from only positive and unlabelled examples (LOPU). MIL represents an object as a bag of instances. It is studied under the assumption that its instances are drawn from a mixture distribution of the concept and the non-concept. Based on this assumption, the classification of bags can be formulated as a classifier combining problem and the Bayes classifier for instances is shown to be closely related to the classification in LOPU. This relationship provides a possibility to adopt methods from LOPU to MIL or vice versa. In particular, we examine a parameter estimator in LOPU being applied to MIL. Experiments demonstrate the effectiveness of the instance classifier and the parameter estimator.

## **Combining Instance Information to Classify Bags**

*Veronika Cheplygina, David M.J. Tax, and Marco Loog*  
(Delft University of Technology, The Netherlands)

Multiple Instance Learning is concerned with learning from sets (bags) of feature vectors (instances), where the bags are labeled, but the instances are not. One of the ways to classify bags is using a (dis)similarity space, where each bag is represented by its dissimilarities to certain prototypes, such as bags or instances from the training set. The instance-based representation preserves the most information, but is very high-dimensional, whereas the bag-based representation has lower dimensionality, but risks throwing away important information. We show a connection between these representations and propose an alternative representation based on combining classifiers, which can potentially combine the advantages of the other methods. The performances of the ensemble classifiers are disappointing, but require further investigation. The bag-based representation preserves sufficient information to classify bags correctly and produces the best results on several datasets.

## **Transfer Learning with Part-Based Ensembles**

*Shiliang Sun, Zhijie Xu, and Mo Yang*  
(East China Normal University, China)

Transfer learning is one of the most important directions in current machine learning research. In this paper, we propose a new learning framework called Multi-source part-based Transfer Learning (Ms-pbTL), which is one kind of parameter transfer with multiple related source tasks. Dissimilar to many traditional works, we consider how to transfer information from one task to another in the form of integrating transferred information between parts. We regard all the complex tasks as a col-

lection of several constituent parts respectively. It means that transfer learning between two complex tasks can be accomplished by sub-transfer learning between their parts. Then, after completing the above information transfer between the source and target tasks, we integrate the models of all the parts in the target task into a whole. Experiments on some real data sets with support vector machines (SVMs) validate the effectiveness of our proposed learning frameworks.

## **Improving Simple Collaborative Filtering Models Using Ensemble Methods**

*Ariel Bar (Ben-Gurion University of the Negev, Beer-Sheva, Israel)*

*Lior Rokach (Ben-Gurion University of the Negev, Beer-Sheva, Israel)*

*Guy Shani (Ben-Gurion University of the Negev, Beer-Sheva, Israel)*

*Bracha Shapira (Ben-Gurion University of the Negev, Beer-Sheva, Israel)*

*Alon Schclar (Academic College of Tel Aviv-Yaffo, Israel)*

In this paper we examine the effect of applying ensemble learning to the performance of collaborative filtering methods. We present several systematic approaches for generating an ensemble of collaborative filtering models based on a single collaborative filtering algorithm (single-model or homogeneous ensemble). We present an adaptation of several popular ensemble techniques in machine learning for the collaborative filtering domain, including bagging, boosting, fusion and randomness injection. We evaluate the proposed approach on several types of collaborative filtering base models: k-NN, matrix factorization and a neighborhood matrix factorization model. Empirical evaluation shows a prediction improvement compared to all base CF algorithms. In particular, we show that the performance of an ensemble of simple (weak) CF models such as k-NN is competitive compared with a single strong CF model (such as matrix factorization) while requiring an order of magnitude less computational cost.

## **Single Classifier Based Multiple Classifications**

*Albert Hung-Ren Ko and Robert Sabourin*

*(University of Quebec, Canada)*

In this paper, a Single Classifier-based Multiple Classification Scheme (SMCS) is proposed as an alternative multiple classification scheme. The SMCS uses only a single classifier to generate multiple classifications for a given test data point. Because of the presence of multiple classifications, classification combination schemes, such as majority voting, can be applied, and so the mechanism may improve the recognition rate in a manner similar to that of Multiple Classifier Systems (MCS). The experimental results confirm the validity of the proposed SMCS as applicable to many classification systems.

**Soft-Voting Clustering Ensemble**

*Haishen Wang, Yan Yang, Hongjun Wang, and Dahai Chen*  
(Southwest Jiaotong University, China)

Clustering ensemble is a framework for combining multiple based clustering results of a set of objects without accessing the original feature of the objects. The majority voting method is widely used in clustering ensemble because of its simplicity, robustness and stability. In general, the existing voting methods only accept hard clustering results as input. In this paper we propose a new algorithm, Soft-Voting Clustering Ensemble (SVCE), which has better flexibility and generalization. The theory of SVCE is illustrated and the algorithm of SVCE is stated in detail firstly. Then 15 UCI datasets are used for the experiment and the results show that the proposed method has a better performance than state of the art ensemble methods in most cases, such as Majority Voting, Weighted Majority Voting, CSPA, MCLA, HGPA.

**Semi-Supervised Clustering Ensemble for Web Video Categorization**

*Amjad Mahmood, Tianrui Li, Yan Yang, Hongjun Wang, and*  
*Mehtab Afzal*  
(Southwest Jiaotong University, China)

Recently, web video categorization has been an ever interesting research with the popularity of web videos. Clustering ensemble has become a good alternative for categorization. Semi-supervised clustering ensemble has shown a better performance since it may incorporate the known prior knowledge, e.g., pairwise constraints. In this paper, we propose a Semi-supervised Cluster-based Similarity Partitioning Algorithm (SS-CSPA) to categorize the videos containing textual data provided by their up-loaders. The feature of this algorithm is the introduction of an unsupervised learning, consensus between clustering and additional support of pairwise constraints to formulate semi-supervised clustering ensemble paradigm. Experimental results on the real-world web videos show that the proposed algorithm outperforms existing methods for categorization of web videos.

**A New Perspective of Support Vector Clustering with Boundary Patterns**

*Yuan Ping, Huina Li, Yong Zhang, and Zhili Zhang*  
(Xuchang University, China)

To overcome the pricey computation required by redundant kernel function matrix and poor label performance, in a novel perspective, we present support vector clustering with boundary patterns (BPSVC for abbreviation) for efficiency. For the first phase, the conventional method of estimating the support vector function with the whole data is altered by only essential boundary patterns. Thence, BPSVC only need to solve a much simpler optimization problem. For the second phase of cluster labeling, both convex decomposition and cone cluster labeling method are employed by

an ensemble labeling strategies for further improvements on accuracy and efficiency. Both theoretical analysis and experimental results show its superiorities in comparison of the state-of-the-art methods, especially for large-scale data analysis.

14:20 - 16:25 **Session 7: Applications**

Session chair: Eric Granger (Université du Québec, Canada)

## **Multi-View Multi-Class Classification for Identification of Pathogenic Bacterial Strains**

*Evgeni Tsvitsivadze (The Netherlands Organization for Applied Scientific Research, The Netherlands)*

*Tom Heskes (Radboud University, The Netherlands)*

*Armand Paauw (The Netherlands Organization for Applied Scientific Research, The Netherlands)*

In various learning problems data can be available in different representations, often referred to as views. We propose multi-class classification method that is particularly suitable for multi-view learning setting. The algorithm uses co-regularization and error-correcting techniques to leverage information from multiple views and in our empirical evaluation notably outperforms several state-of-the-art classification methods on publicly available datasets. Furthermore, we apply the proposed algorithm for identification of the pathogenic bacterial strains from the recently collected biomedical dataset. Our algorithm gives a low classification error rate of 5%, allows rapid identification of the pathogenic microorganisms, and can aid effective response to an infectious disease outbreak.

## **MRF-Based Multiple Classifier System for Hyperspectral Remote Sensing Image Classification**

*Junshi Xia, Peijun Du, and Xiyan He*

*Junshi Xia (Grenoble Institute of Technology, France)*

*Peijun Du (Nanjing University, China)*

*Xiyan He (Grenoble Institute of Technology, France)*

Hyperspectral remote sensing image (HSI) classification is a challenging problem because of its large amounts of spectral channels. Meanwhile, labeled samples for supervised classifier is very limited. The above two reasons often lead to unstable classification result and poor generalization capacity. Recent research has demonstrated the potential of multiple classifier system (MCS) for producing more accurate classification result. In addition, another vital aspect of HSI classification is spatial contents. Markov random field (MRF), which takes the spatial dependence among neighborhood pixels based on the intensity field from observed data into consideration, is always adopted as an effective way to integrate the spatial information. In this paper, we proposed an effective framework for classifying HSI image, called MRF-based MCS, which are based on the aforementioned two powerful algorithms. The proposed model is validated by multinomial logistic regression (MLR) classifier. Ex-

perimental results with hyperspectral images collected by the NASA Jet Propulsion Laboratory's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) demonstrate that MRF-based MCS is a promising strategy in the context of hyperspectral image classification.

### **Kalman Filter Based Classifier Fusion for Affective State Recognition**

*Michael Glodek, Stephan Reuter, Martin Schels, Klaus Dietmayer, and  
Friedhelm Schwenker  
(University of Ulm, Germany)*

The combination of classifier decisions is a common approach to improve classification performance. However, non-stationary fusion of decisions is still a research topic which draws only marginal attention, although more and more classifier systems are deployed in real-time applications. Within this work, we study Kalman filters [4] as a combiner for temporally ordered classifier decisions. The Kalman filter is a linear dynamical system based on a Markov model. It is capable of combining a variable number of measurements (decisions), and can also deal with sensor failures in a unified framework. The Kalman filter is analyzed in the setting of multi-modal emotion recognition using data from the audio/visual emotional challenge 2011. It is shown that the Kalman filter is well-suited for real-time non-stationary classifier fusion. Combining the available sequential uni- and multi-modal decisions does not only result in a consistent continuous stream of decisions, but also leads to significant improvements compared to the input decision performance.

### **Gender Classification Using Mixture of Experts from Low Resolution Facial Images**

*Yomna Safaa El-Din (Ain Shams University, Egypt)  
Mohamed N. Moustafa (American University in Cairo, Egypt)  
Hani Mahdi (Ain Shams University, Egypt)*

In this study, we propose a novel two-stages mixture of experts scheme estimating gender from facial images. The first stage combines a couple of complementary gender classifiers with a third arbiter in case of decision discrepancy. Experimentally, we have verified the common thinking that one appearance-based (Haar-features cascade) classifier with another shape-based (landmarks positions metrology with SVM) classifier form a complementary couple. Subsequently, the second stage in our scheme is a Bayesian framework that is activated only when the arbiter cannot take a confident decision. We demonstrate that the proposed scheme is capable of classifying gender reliably from faces as small as 16x16 thumbnails on benchmark databases, achieving 95% gender recognition on FERET database, and 91.5% on the Labeled Faces in the Wild dataset.

## **Binary Decision Trees for Melanoma Diagnosis**

*Yu Zhou (University of Leeds, UK)*

*Zhuoyi Song (University of Sheffield, UK)*

Although computer aided diagnosis of melanoma is an active research area for more than two decades, its clinical application is still just on horizon. To speed up its clinical application, two critical challenges need to be solved: the data gap and the decision-making gap. Ideally, these two issues shall be attacked simultaneously. However, in the literature, most current methods designing melanoma diagnosis classifiers adopt a biased approach by either focusing on the data gap or on the decision-making gap while neglecting the other. In this article, we present one prototype system covering both the data gap and the decision-gap. Performance of this new method is presented and comparisons with respect to alternative approaches, including the conventional one, are also included.



# Conference Organization

## **MCS 2013 Conference Committee**

Chairs:

Zhi-Hua Zhou (Nanjing University, China)

Josef Kittler (University of Surrey, U.K.)

Fabio Roli (University of Cagliari, Italy)

Local Arrangement Chair:

Ming Li (Nanjing University, China)

Publicity Chairs:

Yang Yu (Nanjing University, China)

Giorgio Fumera (University of Cagliari, Italy)

## **PSL 2013 Conference Committee**

Chairs

Zhi-Hua Zhou (Nanjing University, China)

Friedhelm Schwenker (University of Ulm, Germany)

Local Arrangement Chair

Ming Li (Nanjing University, China)

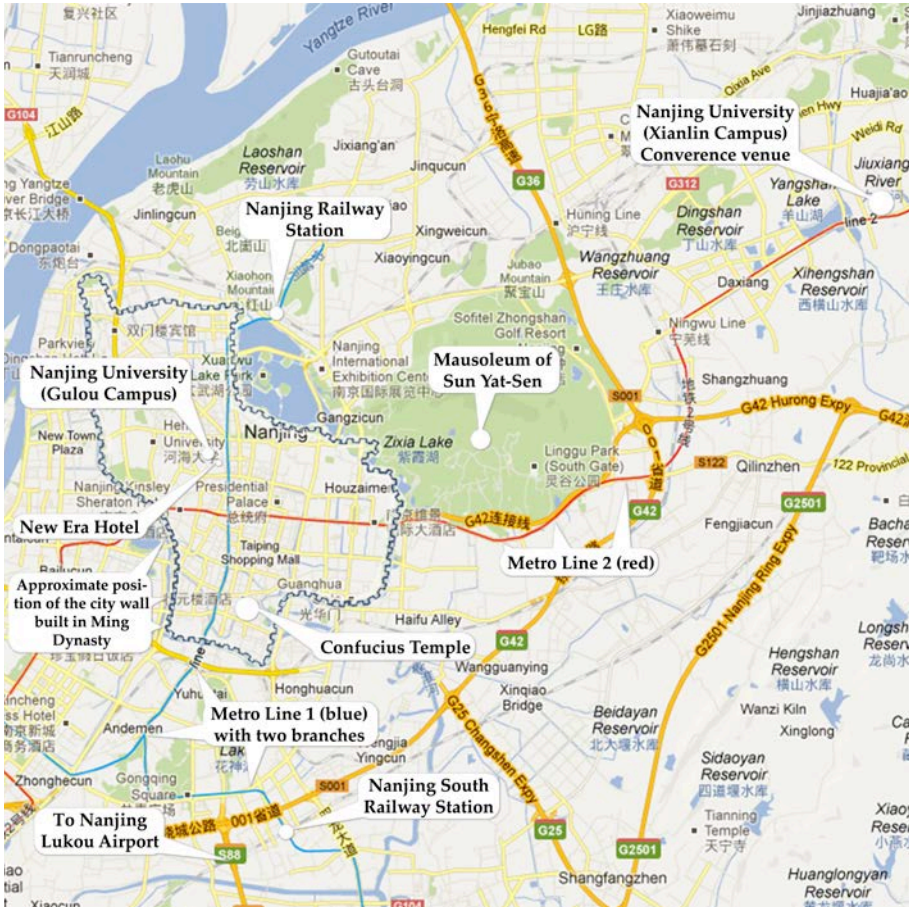
Publicity Chairs

Yang Yu (Nanjing University, China)

Michael Glodek (University of Ulm, Germany)

# Participant Guidelines

## Nanjing Mini-Map



\* The city center is around the crossing station of the two metro lines.

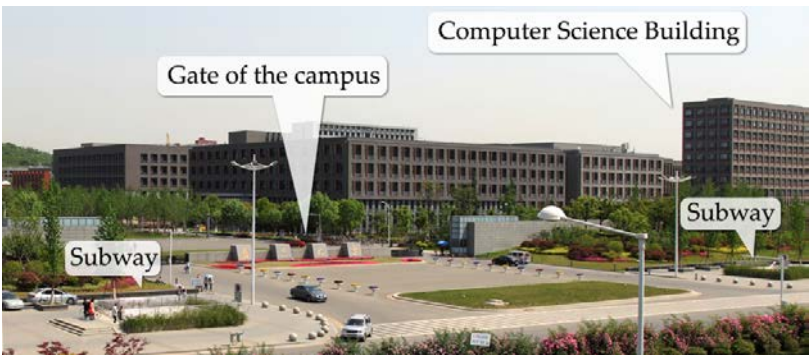
\* About 20km long fractions (but not all) of the city wall built in Ming Dynasty (about 600 years ago) is preserved.

## New Era Hotel Surrounding Map



- \* Walking from the metro entry 1 to the New Era Hotel takes about 3 minutes.
- \* Shuttle buses to the conference venue will be waiting in the front yard of New Era Hotel.
- \* You can easily find food services around.

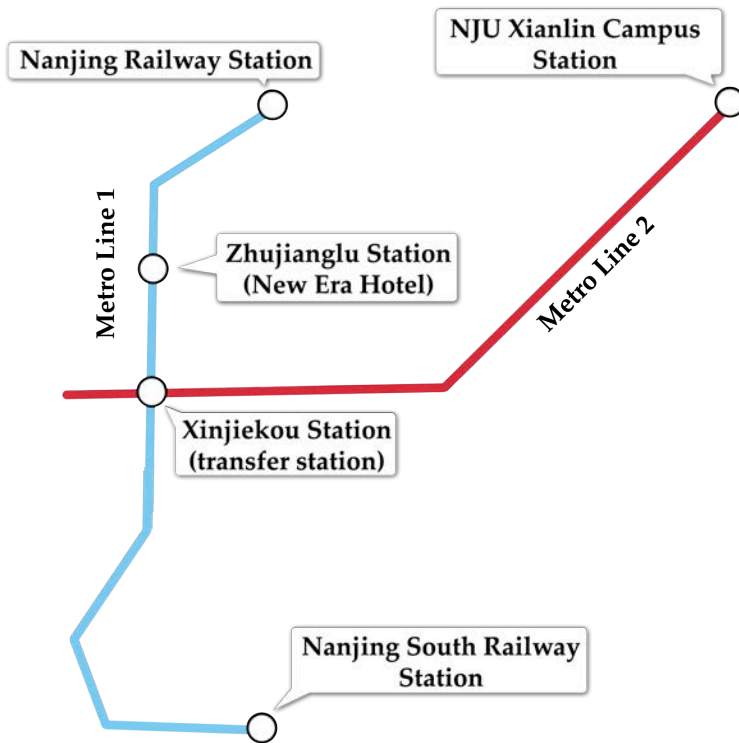
# Venue Surrounding Map



Gate of the campus

- \* When arrival at the metro station, please follow the people stream to go across the road through a subway.
- \* All sessions will be in Room 233 of the CS Building

# A Simple Metro Map



- \* The last metro train dispatches at 11:00pm.
- \* There are two railway stations for passengers in Nanjing: Nanjing Railway Station and Nanjing South Railway Station. They are connected through Metro Line 1.
- \* When taking the Metro Line 1 to Nanjing South Railway Station, please be aware that the line has two branches. Nanjing South Railway Station lies on the branch heading to “China Pharmaceutical University (CPU)”.

# Shuttle Bus Schedule

May 14

7:30 Dispatch from New Era Hotel

16:30 Dispatch from CS Building (Xianlin Campus)

May 15

7:30 Dispatch from New Era Hotel

12:30 Dispatch from CS Building (Xianlin Campus)

21:00 Dispatch from the banquet restaurant

May 16

7:30 Dispatch from New Era Hotel

20:00 Dispatch from the dinner restaurant (Xianlin Campus)

May 17

7:30 Dispatch from New Era Hotel

16:30 Dispatch from CS Building (Xianlin Campus)

\* When missing the bus at the hotel, you may seek help from the volunteers for transferring.



### Sun Yat-Sen Mausoleum

Dr. Sun Yat-sen (1866-1925) was a great forerunner of the Chinese democratic revolution and led by Dr. Sun the Chinese people brought down the corrupt rule of the Qing Dynasty (1644-1911) and ended 2000 years of the feudal monarchy system, which led the Chinese peoples into a new age.



Covering an area of 80,000 square meters (about 20 acres), Dr. Sun Yat-sen's Mausoleum is located in the Zhong Mountain Scenic Area in the east suburb of Nanjing City, Jiangsu Province. As the mausoleum of Dr. Sun Yat-sen, the father of the Republic of China, it is considered the Holy land of Chinese people both home and abroad. With deep historical significance, magnificent architecture and beautiful scenery, it is a must see when traveling in Nanjing.

### Ming Xiaoling Mausoleum

Ming Xiaoling Mausoleum was recognized by the World Heritage List of Cultural Sites as part of Ming and Qing Imperial Mausoleums in July, 2003.

Ming Xiaoling Mausoleum is one of the biggest imperial tombs in China. It lies in the eastern suburbs of Nanjing City at the southern foot of Zhongshan (Purple) Mountain. Emperor Chengzu, Zhu Yuanzhang, the first emperor of the Ming Dynasty (1368-1644) and Queen Ma were buried there.



Construction of the mausoleum began in 1381 and was completed in 1431. In 1384, Queen Ma died and was buried there. Emperor Chengzu had bestowed upon her the title 'Queen of Xiao Ci' which means 'Queen of Filial Piety and Kindness.' Hence, the name Mingxiaoling derives from her title.

## Confucius Temple

Nanjing Confucius Temple is for consecrating and worshipping Confucius, the greatest thinker and educator of ancient China. Nanjing Confucius temple was built in Song Dynasty (A.D. 1034) and expanded in East Jin Dynasty. These buildings were ruined and rebuilt for several times. The Confucius Temple once was rebuilt in Qing Dynasty (A.D. 1869), but was seriously fired and destroyed by the Japanese aggressor troops in 1937. In order to protect the old city, the Chinese Government appropriated funds to rebuild and fix the Confucius Temple. Nowadays it has re-expressed its splendid civilization to the world and has been regarded the special scenery of Nanjing. In 1991, the Confucius Temple was elected the "Forty Best Scenery Spots".



## Exhibition Hall of the History of Jiangnan Examination School

Founded in 1168 (the Song Dynasty), the school was used to administer the rigorous civil service exams used to choose officials during China's imperial age. At the height of its prosperity, the complex contained 20,644 examination cells, each just one square meter in area. Candidates spent a total of nine days in their cell--the wooden desk plank turned into a bed at night--without the freedom to leave. Success was rare as only 200 of the 20,000 candidates passed. The modern exhibit contains forty examination cells plus a small museum that claims to be the only specialized museum in China dedicated to the imperial examination system.





# Living in Nanjing

*Language:* Local language is Chinese (Mandarin). A proportion of youth can speak English more or less.



*Volunteer:* There are about 20 student volunteers serving for the conference. You may ask them for help. Volunteers are with orange cards which look like the right-side image.

*Emergency:* The emergency calls are 110 (ask for police), 119 (for fire), and 120 (for ambulance). Emergency calls can be made from any local phone for free. You can also call Dr. Ming Li for emergency: +86-133-9090-9609



logo of Nanjing metro

*Traffic:* Metro lines open from 7:00am~11:00pm. A metro ticket is 2.00~4.00RMB, depending on the distance. Taxi charges according to the taximeter, with starting cost 11.00RMB. Public bus ticket is 1.00RMB or 2.00RMB per person.



*Electric Power:* 220V/50Hz, the most common electrical outlets look like the right image.

*Currency:* RMB (Chinese CNY) only. You can exchange money in banks either at airport or around the hotel.

*Credit Cards/ATM:* Retail outlets only accept card from Union Pay of China. By foreign credit cards, you need withdraw cash in RMB from ATMs with logo of your card company, such as VISA, MasterCard and American Express.

*Temperature:* In May, the temperature is usually around 17°C~30°C, or 62°F~86°F.

*Holidays and weekends:* Banks and retail outlets are open on holidays and weekends.

# Food Adventure

The Chinese dining culture begins with its eight traditional cuisines known as the “Eight Culinary Traditions of China”, namely: Cantonese, Hunan, Shandong, Fujian, Jiangsu, Zheijiang, Anhui, Szechuan. In Nanjing, you can find dishes from all the cuisines.

Besides, Nanjing also serves particular local snacks, including



Salted duck



Rice noodle with duck blood



Steamed bun



Fried dumpling

and more.

You may seek recommendations on restaurants and dishes from conference volunteers. Information can be also discovered from web sites like TripAdvisor: [http://www.tripadvisor.com/Restaurants-g294220-Nanjing\\_Jiangsu.html](http://www.tripadvisor.com/Restaurants-g294220-Nanjing_Jiangsu.html)

## Useful Sentences in Chinese

Please go to New Era Hotel:

请到新纪元大酒店（中山路）

Please go to Nanjing Railway Station:

请到南京火车站

Please go to Nanjing South Railway Station:

请到南京南站

Please go to the airport:

请到机场

Please go to Nanjing University Xianlin Campus:

请到南京大学仙林校区（地铁二号线南大仙林校区站对面）

Please go to Shanghai Railway Station:

请到上海火车站

Please go to Shanghai Hongqiao Railway Station:

请到上海虹桥火车站

Please go to Shanghai Pudong Airport:

请到上海浦东机场

Could you please make a phone call for me:

可以帮我打个电话吗

Is there any washroom around:

请问附近有厕所吗

Could you tell me how to get to the nearest metro station:

请问最近的地铁站怎么走



## **National Key Laboratory for Novel Software Technology**

The Laboratory of Novel Software Technology at Nanjing University was established in 1978, and it was officially recognized as one of the National Key Laboratories in 1990, and opened to researchers all over the world since then. The Laboratory has passed four evaluations organized by the Ministry of Science and Technology of China. It was ranked in 2007 and 2012 as the 1st among all National Key Laboratories in Computer Science and honored as an "Excellent" National Key Laboratory.

The laboratory holds the thought of 'pursuing originality, overcoming technological problems, improving industry by achievements' as its guild. While paying attention to basic scientific problems in the field, the laboratory focuses on the research and application of novel software technology, which can be divided into three levels: the exploration of new software methods and technologies, the research and development of software technologies and the application of new technologies and software product development. Major research areas include software engineering and methodology, machine learning and data mining, pervasive and distributed computing, multimedia and computer graphics.

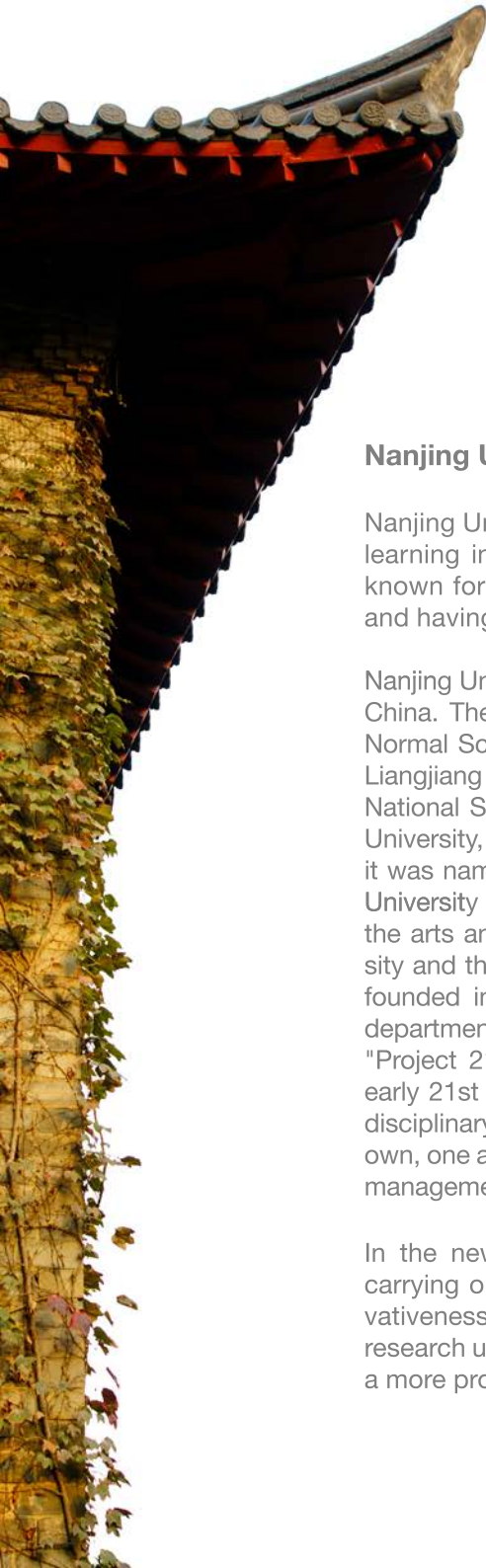
### **Department of Computer Science and Technology**

The Department of Computer Science and Technology of Nanjing University is a key constructing discipline of national "211 Project", "985 Project", and one of the 11 "Top Priority" of Jiangsu Province. Our top-level discipline, Computer Science and Technology, is a national key discipline.

Research of computer science in Nanjing University started in 1958. In 1978 the department was formally set up. The department had made remarkable achievements before and at the initial stage of her construction: in the 1960s, assembled the first computer among all the colleges affiliated to the State Ministry of Higher Education; developed China's first advanced language compiling program; in the 1970s, presided over the research and production of medium-sized computer DJS-210 and XT-1 operating system and other software systems; in the 1980s, developed the first distributed system ZCZ within China; cultivated the first PhD of computer software on China mainland.

By now, the Department has primarily constructed a domestically and internationally influential personnel training and scientific research center, covering a wide range of subjects, also catering to the need of both international IT development and domestic economic and social development





## Nanjing University

Nanjing University, one of the oldest institutions of higher learning in China, is located in Nanjing, a beautiful city known for being China's capital of six ancient dynasties and having a rich cultural heritage.

Nanjing University is one of the top universities of mainland China. The University dates back to 1902 when Sanjiang Normal School was founded. The School then evolved as Liangjiang Normal School, Nanjing Higher Normal School, National Southeast University, National Fourth Zhongshan University, and National Central University. In August 1949, it was named National Nanjing University. Today's Nanjing University is formed mainly on the basis of the merging of the arts and sciences schools of National Nanjing University and the University of Nanking (a well-known university founded in 1888) during the nationwide institutional and departmental restructuring in 1952. In implementing the "Project 211", Nanjing University is building itself by the early 21st century into an internationally influential, multi-disciplinary socialist university with characteristics of its own, one among the best in the world in teaching, research, management and resources.

In the new century, NJU will deepen its reforms while carrying on its traditions. Aiming at excellence and innovativeness, it is headed for a world-class comprehensive research university with a global vision, and thus it will play a more prominent role in China's rise in the 21st century.