Pattern Recognition 71 (2017) 444–445

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Editorial of the Special Issue on Multi-instance Learning in Pattern Recognition and Vision

1. Introduction

Multi-instance learning (MIL) has served as an important tool for a wide range of applications in pattern recognition and computer vision, for instance, in drug activity prediction, text classification, image classification, object detection, and visual tracking, while understanding big visual data in mobile internet is a more recent trend. Though supervised learning methods, such as deep convolutional neural network, have made tremendous progress in many large-scale visual recognition tasks, the major bottleneck in their application is their heavy reliance on large volumes of human-annotated data that typically are non-trivial to obtain. Whereas the availability of finely labeled images and videos, for which pixel and bounding-box labels are provided, is very limited, coarse labels are more often readily available. There are, for instance, huge amounts of weakly labeled visual data that have image-level labels.

Multi-instance learning is a popular tool for exploring semantic information in weakly labeled data. In MIL, instead of being given the labels of each individual instance, the learner receives a set of labeled bags (e.g., images), each typically containing a large amount of instances (e.g., patches extracted from the images). MIL algorithms can not only learn bag classifiers, but are also able to infer the semantic label of instances. Some important progress has been made in this aspect. However, MIL algorithms still have to cope with challenges in both algorithm design and applications, such as handling large-scale data efficiently and effectively.

This special issue aims at presenting recent advancements in both algorithmic and application-driven research related to multi-instance learning. We have received 31 submissions and finally accepted 5 papers after a strict review process. Each submission has been reviewed by at least two reviewers and has went through at least two rounds of review. We hope the accepted papers to this special issue will provide a useful reference to researchers who are interested in the research of MIL algorithms or applications, especially in the field of pattern recognition and computer vision.

2. Overview of accepted articles

The accepted articles can be roughly divided into two categories: designing novel representations for MIL tasks (e.g., new multi-instance dictionaries) and applications of MIL methods (e.g., weakly supervised object detection).

The article “Diversified dictionaries for multi-instance learning” by Qiao et al. proposes a new approach to reduce the instance-level representations in MIL into bag-level representations such that single-instance classifiers can be readily applied to solve the MIL task. Class specific dictionaries are learned for each class, and a novel diversity regularizer (based on the determinantal point process) is enforced to promote diversity among the learned class specific dictionaries.

Another article related to dictionary learning is “Multi-instance dictionary learning via multivariate performance measure optimization” by Wang et al., which proposes the POD (Performance-Optimal multi-instance Dictionary) method to jointly learn dictionary and classifier for MIL tasks. POD is designed in response to the observation that different multivariate performance metrics require different optimal multi-instance dictionaries. The POD learning algorithm learns the metric-specific dictionary and classifier jointly, which achieves higher performance metric than methods that learn the dictionary and the classifier in separate learning steps.

Xu et al. aims at reducing the computational complexity of MIL methods in “SALE: Self-adaptive LSH encoding for multi-instance learning”. The proposed SALE representation is built on top of the locality sensing hashing (LSH) and self-adaptive reconstruction, which not only makes SALE scale to large datasets but also helps SALE focus on key information in the MIL bags. Experiments on MIL datasets with different sizes verify both the effectiveness and the efficiency of SALE.

The other two articles in this special issue present interesting applications of MIL in pattern recognition and computer vision.

“Weakly supervised vehicle detection in satellite images via multi-instance discriminative learning” by Cao et al. presents a novel application of multi-instance learning. Satellite images often contain huge number of cars, hence training data for supervised detectors (which require annotating every car in the training images) are very expensive to obtain. Cao et al. only ask a user to annotate whether a large region contains cars or not. This weakly supervised setup is exactly a multi-instance learning task. The proposed MIL-based framework has achieved state-of-the-art performance in car detection in satellite images.

The article “Deep patch learning for joint object classification and discovery” by Tang et al. is another application of MIL algorithms. By treating an image as a bag and image patches as instances, this article proposes an end-to-end weakly supervised deep CNN framework for simultaneous image classification and object discovery, and shows that these two tasks are complementary. The CNN loss contains two terms: for classification and discovery, respectively; and the discovery loss is formulated in the MIL

http://dx.doi.org/10.1016/j.patcog.2017.05.026

setup. With only bag/image labels available in the training phase, the proposed weakly supervised/MIL object discovery method has achieved state-of-the-art object discovery results.

Acknowledgments

We thank all the authors for their contributions to this special issue and the reviewers for their careful and insightful reviews. We also would like to thank Prof. Ching Y. Suen, then Editor-in-Chief of the *Pattern Recognition* Journal for proving this special issue, and the Elsevier staff for managing and supporting the organization of this special issue.

Jianxin Wu*
Nanjing University, China

Xiang Bai
Huazhong University of Science and Technology, China

Marco Loog
Pattern Recognition Laboratory, Delft University of Technology, The Netherlands
The Image Group, University of Copenhagen, Denmark

Fabio Roli
University of Cagliari, Italy

Zhi-Hua Zhou
Nanjing University, China

*Corresponding author.
E-mail address: wujx2001@nju.edu.cn
(J. Wu)

Dr. Jianxin Wu received his Ph.D. degree from the Georgia Institute of Technology in computer science in 2009. From 2009 to 2013, he was an assistant professor in the School of Computer Engineering, Nanyang Technological University, Singapore. Since fall 2013, he has been a professor in the Department of Computer Science and Technology and the National Key Laboratory for Novel Software Technology in Nanjing University, China. Dr. Wu's research interests include computer vision and machine learning.

Dr. Xiang Bai received the B.S., M.S., and Ph.D. degrees from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2003, 2005, and 2009, respectively, all in electronics and information engineering. He is currently a professor with the School of Electronic Information and Communications, HUST. His main research interests include object recognition, shape analysis, scene text recognition, and intelligent systems.

Dr. Marco Loog received an M.Sc. degree in mathematics from Utrecht University and a Ph.D. degree from the Image Sciences Institute. Subsequently, he moved to Copenhagen where he acted as an assistant and, eventually, associate professor, next to which he worked at Nordic Bioscience. Currently, Marco is affiliated to the Pattern Recognition Laboratory at Delft University of Technology. He is also an honorary professor in pattern recognition at the University of Copenhagen. Marco's principal research interest is with supervised pattern recognition in all sorts of shapes and sizes.

Dr. Fabio Roli received his M.S. degree, with honors, and Ph.D. degree in electronic engineering from the University of Genoa, Italy. He is now professor of computer engineering and director of the Research Lab on Pattern Recognition and Applications (http://pralab.diee.unica.it). He is a member of the governing boards of the International Association for Pattern Recognition and of the IEEE Systems, Man and Cybernetics Society. He is fellow of the IEEE and fellow of the International Association for Pattern Recognition.

Dr. Zhi-Hua Zhou received his B.Sc., M.Sc. and Ph.D. degrees in computer science from Nanjing University, China, in 1996, 1998 and 2000, respectively, all with the highest honor. He joined the Department of Computer Science & Technology of Nanjing University as an assistant professor in 2001, and at present he is a professor, standing deputy director of the National Key Lab for Novel Software Technology, and founding director of LAMDA. He is a fellow of the ACM, AAAS, AAAI, IEEE, IAPR, IET/IEE, and CCF.