Guest Editor's Introduction to the Special Issue on Domain Adaptation for Vision Applications

Dong Xu · Rama Chellappa · Trevor Darrell · Hal Daumé III

Domain adaptation is an emerging research topic in computer vision. In some vision applications, the domain of interest (i.e., the target domain) contains very few or even no labelled samples, while an existing domain (i.e., the auxiliary/source domain) is often available with a large number of labelled examples. For example, millions of loosely labeled Flickr photos or YouTube videos can be readily obtained by using keywords (also called tags) based search. On the other hand, users may be interested in retrieving and organizing their own multimedia collections of images and videos at the semantic level, but may be reluctant to put forth the effort to annotate their photos and videos by themselves. This problem becomes more challenging because the feature distributions of training samples from the web domain and consumer domain may significantly differ in statistical properties. In order to effectively utilize the training samples from different domains, domain adaptation techniques aim to explicitly cope with variations in feature distributions.

We received many high-quality submissions, from which we selected 9 articles to include in this special issue. These articles cover a wide range of domain adaptation topics for various vision applications such as object recognition, face recognition, action and event recognition. A brief summary of these articles is provided below.

The article "Learning Kernels for Unsupervised Domain Adaptation with Applications to Visual Object Recognition" by Gong et al. first proposes an unsupervised learning approach for domain adaptation, in which the so-called geodesic flow kernel (GFK) is calculated by interpolating an infinite number of subspaces to bridge the source and target domains. They also propose a supervised learning approach by combining multiple base GFKs. Their approaches achieve state-of-the-art results on the benchmark datasets.

In the article "Asymmetric and Category Invariant Feature Transformations for Domain Adaptation", Hoffman et al. first propose a distance metric learning method for domain adaptation by learning asymmetric transformations. They also propose a new domain adaptation approach to learn the classifiers based on the max-margin framework. Extensive experiments demonstrate the effectiveness of their algorithms.

In "Weakly-Supervised Cross-Domain Dictionary Learning for Visual Recognition", Zhu and Shao propose a new dictionary learning method by using auxiliary domain knowledge to expand the intra-class diversities of original training data. In this work, dictionary learning is on samples from both auxiliary and target domains by using a transformation matrix. Comprehensive experiments for various vision applications including action recognition, event recognition

D. Xu (✉)
School of Computer Engineering, Nanyang Technological University, Blk N4, 2A-29, Nanyang Avenue, Singapore 639798, Singapore
e-mail: dongxu@ntu.edu.sg

R. Chellappa
Department of Electrical and Computer Engineering and Center for Automation Research, University of Maryland, College Park, MD 20742-3275 , USA
e-mail: rama@umiacs.umd.edu

T. Darrell
Electrical Engineering and Computer Sciences, College of Engineering, University of California, Berkeley, CA, USA
e-mail: trevor@eecs.berkeley.edu

H. Daumé III
Department of Computer Science, University of Maryland, College Park, MD 20742, USA
e-mail: hal@umiacs.umd.edu

and image classification demonstrate the effectiveness of this approach.

The article "Harnessing Lab Knowledge for Real-world Action Recognition" by Ma et al. proposes a multi-task learning framework for action recognition in the real-world videos by borrowing the knowledge from the lab videos. Their framework can still explore the shared knowledge between real-world and lab datasets even when the action categories are different. Promising results are reported for real-world action recognition.

The article by Shao et al., entitled "Generalized Transfer Subspace Learning through Low-Rank Constraint," presents a new subspace learning method by using the low-rank constraint to bridge the source and target domains in the low-dimensional space. With the low-rank constraint, the knowledge can be transferred from the source domain to the target domain only when the samples from both domains are aligned. This work is evaluated for different vision applications such as face recognition, kinship verification and object recognition and the newly proposed methods achieve better results when compared with several existing algorithms.

The article "Domain Adaptation for face recognition: Targetize Source domain images bridged by Common subspace" by Kan *et al.* proposes a new face recognition method by using a sparse set of target domain images to reconstruct each source domain image in the image space. They also propose to learn a common subspace in order to preserve the structures of two domains and simultaneously reduce the mismatch between the two domains. Extensive experiments demonstrate their approach outperforms several existing techniques for face recognition under different scenarios.

In the article "Model-Driven Domain Adaptation on Product Manifolds for Unconstrained Face Recognition", Ho and Gopalan propose to derive a latent domain by representing each subject as a point on a product of Grassmann manifolds. They also develop a new kernel discriminant analysis based approach and a probabilistic approach for image and video classification. The effectiveness of their approach is demonstrated by comprehensive evaluations.

Yamada et al. develop in "Domain Adaptation for Structured Regression" a new semi-supervised domain adaptation method based on Twin Gaussian Processes to reduce large structural biases. They also propose a new approach to measure domain similarity between the source and target domains. This work achieves the state-of-the-art performances for 3D head and 3D human pose estimation.

In "Exploring Transfer Learning Approaches for Head Pose Classification from Multi-view Surveillance Images", Rajagopal et al. present new transfer learning methods for multi-view head pose classification. To cope with challenging situations in this application (e.g., the range of head poses in the images from different domains may be different), the authors propose a new boosting-based transfer learning algorithm and new adaptive weights learning methods. Extensive experiments demonstrate the proposed methods outperform the state-of-the-art algorithms for multi-view head pose classification.