Neural Networks 76 (2016) 135–151

Contents lists available at ScienceDirect

Neural Networks

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

Multi-source adaptation joint kernel sparse representation for visual classification

JianWen Tao[a,*], Wenjun Hu[b], Shiting Wen[a]

[a] School of Information Science and Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China

[b] School of Information and Engineering, Huzhou Teachers College, Huzhou 313000, China

ARTICLE INFO

Article history:
Received 12 December 2014
Received in revised form 22 November 2015
Accepted 18 January 2016
Available online 3 February 2016

Keywords:
Multi-source domain adaptation
Joint kernel sparse representation
Graph-based semi-supervised learning
Maximum mean discrepancy

ABSTRACT

Most of the existing domain adaptation learning (DAL) methods relies on a single source domain to learn a classifier with well-generalized performance for the target domain of interest, which may lead to the so-called negative transfer problem. To this end, many multi-source adaptation methods have been proposed. While the advantages of using multi-source domains of information for establishing an adaptation model have been widely recognized, how to boost the robustness of the computational model for multi-source adaptation learning has only recently received attention. To address this issue for achieving enhanced performance, we propose in this paper a novel algorithm called multi-source **A**daptation **R**egularization **J**oint **K**ernel **S**parse **R**epresentation (ARJKSR) for robust visual classification problems. Specifically, ARJKSR jointly represents target dataset by a sparse linear combination of training data of each source domain in some optimal Reproduced Kernel Hilbert Space (RKHS), recovered by simultaneously minimizing the inter-domain distribution discrepancy and maximizing the local consistency, whilst constraining the observations from both target and source domains to share their sparse representations. The optimization problem of ARJKSR can be solved using an efficient alternative direction method. Under the framework ARJKSR, we further learn a robust label prediction matrix for the unlabeled instances of target domain based on the classical graph-based semi-supervised learning (GSSL) diagram, into which multiple Laplacian graphs constructed with the ARJKSR are incorporated. The validity of our method is examined by several visual classification problems. Results demonstrate the superiority of our method in comparison to several state-of-the-arts.

1. Introduction

Traditional machine learning methods usually work well when sufficient training data are available. However, since manually labeling data is both expensive and time-consuming, it is desirable to have new techniques to learn a classifier with high accuracy but from only a limited number of labeled training data. While semi-supervised learning (SSL) (Zhu, 2005) exploits unlabeled data to remedy the lack of labeled training data to some extent, it requires that the unlabeled data be sampled under the same distribution as the labeled. To address this issue, many domain adaptation learning (DAL) methods have been proposed to learn robust classifiers with only a few or even no labeled data from the target domain of interest by leveraging a large amount of labeled training data from other domains referred to as source/auxiliary domains (Bruzzone & Marconcini, 2010; Pan & Yang, 2010; Tao, Fu-Lai Chung, & Wang, 2012a). The need for domain adaptation is prevalent in many real-world visual applications such as image classification (Duan, Tsang, Xu, & Maybank, 2009; Tao, Hu, & Wang, 2014), video concept detection (Duan, Tsang, & Xu, 2012; Duan, Tsang, Xu, & Maybank, 2009), face recognition, and image notation (Geng, Tao, & Xu, 2011), to name just a few. As is known, the brute-force domain transfer without the selection of source domains, some of which may be useless or even harmful for DAL, would degrade the classification performance of DAL (Duan, Xu, & Tsang, 2012; Seah, Tsang, & Ong, 2010). This is a well-known open problem termed as negative transfer (Rosenstein, Marx, & Kaelbling, 2005). For this end, recently, several multi-source domain adaptation methods (Duan, Tsang, Xu et al., 2009; Duan, Xu et al., 2012; Luo, Zhuang, Xiong, Xiong, & He, 2008; Mansour, Mohri, & Rostamizadeh, 2009; Rosenstein et al., 2005; Seah et al., 2010; Tao, Chung, & Wang, 2012b; Yang, Yan, & Hauptmann, 2007) have been proposed to learn classifiers with training data

* Corresponding author.
E-mail address: jianwen_tao@aliyun.com (J. Tao).

http://dx.doi.org/10.1016/j.neunet.2016.01.008
0893-6080/© 2016 Elsevier Ltd. All rights reserved.

from multiple source domains. In this paper, we will mainly focus on robust domain adaptation learning from multiple sources to transfer.

In recent years, the theories of sparse representation (SR) (Patel & Chellappa, 2011) have emerged as powerful tools for efficient processing of data in signal processing, statistics, and pattern recognition (Wright, Ma, Mairal et al., 2010; Wright, Yang, Sastry, & Ma, 2009). Wright et al. (2010) proposed the sparse representation-based classification (SRC) algorithm for face recognition, which showed that by exploiting the inherent sparsity of data, one could obtain improved recognition performance over traditional methods especially when data were contaminated by various artifacts such as illumination variations, disguise, occlusion, and random pixel corruption (Wright et al., 2010). Nagesh and Li (2009) presented an expression-invariant face recognition method using distributed compressive sensing and joint sparsity models. Very recently, Fan, Gu, Qiao, and Zhang (2011) proposed a robust sparse regularization semi-supervised classification algorithm based on sparse representation (SR), and Cheng, Liu, and Yang (2009) also proposed a novel label propagation algorithm based on sparsity induced similarity measure strategy. The idea of joint sparsity has been explored recently for image classification (Yuan & Yan, 2010; Zhang, Nasrabadi, Zhang, & Huang, 2011) and segmentation (Cheng, Liu, Wang, Huang, & Yan, 2011). For example, Yuan and Yan (2010) proposed a multi-task sparse linear regression model for image classification, in which they used group sparsity to combine different features of an object for classification. Zhang et al. (2011) proposed a joint dynamic sparse representation model for object recognition, the essential goal of which was to recognize the same object viewed from multiple observations. Shekhar et al. (2012); Shekhar, Patel, Nasrabadi, and Chellappa (2014) also presented a robust feature level fusion algorithm for multi-biometric recognition via joint sparse representation, which was more general in that it could deal with both multimodal as well as multivariate sparse representations. For a survey of applications of SR algorithms, see Elhamifar and Vidal (2011), Patel and Chellappa (2011), Wagner et al. (2012), Wright et al. (2010) and the references therein.

However, when labeled and unlabeled images are sampled from different distributions, SR-based methods may quantize them into different visual features of the dictionary and encode with different representations (Quanz, Huan, & Mishra, 2011). In this case, the dictionary learned from the labeled images cannot effectively encode the unlabeled images with high fidelity, and also the unlabeled images may reside far away from the labeled images under the new representation. This distribution difference will greatly challenge the robustness of existing SR algorithms for cross-distribution image classification problems. To this end, inspired by recent progress in SR and transfer learning, Long et al. (2013) proposed a Transfer Sparse Coding (TSC) algorithm to construct sparse representations for classifying cross-domain images by minimizing the distribution divergence between labeled and unlabeled images using a nonparametric distance measure. Lastly, Quanz et al. (2011) have also explored sparse coding to extract features for knowledge transfer. However, their method adopts a kernel density estimation technique to estimate the probability density functions of distributions and then minimizes the Jensen–Shannon divergence between them, which may be prone to over-fitting (Long et al., 2013). Besides, these SR-based knowledge transfer methods focus only on preprocessing of features for single-source domain transfer problems. More importantly, even when the data to be analyzed is a set of images which are from the same class and sharing common (correlated) features, these methods would still be performed for each input signal independently. This may cause a large computational burden on those algorithms, especially in the case of large-scale samples.

Inspired by recent progress in joint sparse representation and domain adaptation learning, we propose in this paper a novel multi-source domain adaptation method via multi-source Adaptation Regularization Joint Kernel Sparse Representation (ARJKSR). Specifically, ARJKSR jointly represents the dataset of target domain by a sparse linear combination of training data of each source domain in some optimal reproduced kernel Hilbert space (RKHS) recovered by minimizing the inter-domain distribution discrepancy criterion, whilst constraining the observations from both target and each source domains to share their sparse representations. Moreover, we additionally incorporate the graph Laplacian term of sparse coefficients (Zheng et al., 2011) into our objective function, which can discover more discriminating representations for classification tasks. Thus, we take into account correlations as well as coupling information among multiple source domains and thereby make the representation robust for further cross-domain visual classification problems. By using ARJKSR, we then learn a robust label prediction matrix for the unlabeled data of target domain based on the classical graph-based semi-supervised learning (GSSL) diagram (Liu, Wang, & Chang, 2012; Wang, Wang, Zhang, Shen, & Quan, 2009; Wang & Zhang, 2008; Zhou, Bousquet, Lal, Weston, & Schölkopf, 2004), into which multiple Laplacian graphs constructed with ARJKSR are incorporated. Therefore, we may suppose that the labels of data from both target and source domains could be smoothly propagated to the unlabeled data of target domain by forcing the different data to interact through their sparse coefficients (Tao et al., 2014). In the whole, we make the following distinctive contributions:

(1) We present a robust multi-source adaptation regularization joint kernel sparse representation (ARJKSR) framework, into which two regularization terms, i.e., sparse representation maximum mean discrepancy regularization and sparse representation Laplacian regularization, are incorporated. This framework inhibits and extends the potential advantages of joint sparse representation into the area of DAL, and can be easily generalized to handle multi-kernel multi-source adaptation learning problems.
(2) Under the framework ARJKSR, we further propose a multi-source adaptation label propagation model for visual classification, regularized by multi-graph Laplacian terms constructed using ARJKSR. This model is generally superior to traditional GSSL methods based on nearest neighbor criterion, especially for high-dimensional data. As a result, different samples will get different neighborhood sizes, which is more adaptive to complex data distributions.
(3) We introduce a multi-source adaptation joint sparsity concentration index (SCI) criterion to remove the noise points and recover the incomplete data objects from datasets. And we further present a multi-kernel regression Out-of-Sample extension to the ARJKSR-based classification model.
(4) We conduct a serial of extensive experiments on several real-world visual applications to validate the robustness and effectiveness of our method.

The remainder of this paper is organized as follows. In Section 2, we briefly review several related works. Section 3 details our method and its optimization algorithms. In Section 4, we give discussions about the robustness and the extension to handle out-of-sample dataset. Experimental results on several real-world visual datasets from different domains are reported in Section 5 involving parameter selection and results analysis. Finally, Section 6 concludes the paper.