J. Vis. Commun. Image R. 33 (2015) 134–148

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Robust domain adaptation image classification via sparse and low rank representation ☆

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

[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 9 July 2014
Accepted 9 September 2015
Available online 25 September 2015

Keywords:
Robust domain adaptation learning
Low rank representation
Maximum mean discrepancy
Sparse representation
Transfer learning
Semi-supervised learning
Image classification
Robustness

ABSTRACT

Domain adaptation image classification addresses the problem of adapting the image distribution of the source domain to the target domain for an effective learning task, where the classification objective is intended but the data distributions are different. However, corrupted data (e.g. noise and outliers, which exist universally in real-world domains) can cause significant deterioration of the practical performance of existing methods in cross-domain image classification. This motivates us to propose a robust domain adaptation image classification method with sparse and low rank representation. Specifically, we first obtain an optimal Domain Adaptation Sparse and Low Rank Representation (DASLRR) for all the data from both domains by incorporating a distribution adaptation regularization term, which is expected to minimize the distribution discrepancy between the source and target domain, into the existing low rank and sparse representation objective function. Formulating an optimization problem that combines the objective function of the sparse and low rank representation, constrained by distribution adaptation and local consistency, we propose an algorithm that alternates between obtaining an effective dictionary, while preserving the DASLRR to make the new representations robust to the distribution difference. Based on the obtained DASLRR, we then provide a flexible semi-supervised learning framework, which can propagate the labels of labeled data from both domains to unlabeled data from In-Sample as well as Out-of-Sample datasets by simultaneously learning a prediction label matrix and a classifier model. The proposed method can capture the global mixture of the clustering structure (by the sparseness and low rankness) and the locally consistent structure (by the local graph regularization) as well as the distribution difference (by the distribution adaptation) of the domains data. Hence, the proposed method is robust for accurately classifying cross-domain images that may be corrupted by noise or outliers. Extensive experiments demonstrate the effectiveness of our method on several types of images and video datasets.

1. Introduction

With the development of computer network and storage technologies, there has explosive growth of web images. In the field of multimedia and computer vision, many researchers have recently proposed a variety of machine learning and data mining algorithms for image classification also termed image annotation [1–3]. While these works have shown promising achievements in overcoming the well-known semantic gap by applying machine learning algorithms to image classification, explosive amounts of emerging image data have brought a dilemma of data deluge and label scarcity to the task of image classification with traditional methods [4]. In other words, on the one hand, expensive and time-consuming human labor is required for the collection of labels of new emerging images. On the other hand, there exists a large amount of outdated labeled images for previous tasks. While exploiting the vast amount of unlabeled data directly (e.g., via the semi-supervised learning (SSL) paradigm [5]) is valuable in its own right, it is beneficial to leverage labeled data of relevant categories across data sources. For example, it is increasingly popular to enrich our limited collection of training samples with those from the Internet. One problem with this strategy, however, comes from the possible misalignment of the target domain of interest and the source domain that provides the auxiliary data and labels. This misalignment corresponds to the shift in data distribution in a certain feature space. To be precise, the marginal distribution of the

☆ This paper has been recommended for acceptance by Prof. M.T. Sun.

* Corresponding author.
E-mail addresses: jianwen_tao@aliyun.com (J. Tao), wensht@nit.net.cn (S. Wen).

http://dx.doi.org/10.1016/j.jvcir.2015.09.005
1047-3203/

samples in the source domain and that in the target are different. This makes it harmful to incorporate data from the source domain into the target domain directly: in theory, the disparity violates the basic assumption underpinning supervised learning; in practice, the resulting performance degrades considerably on the target test samples [4].

The above theoretical and practical paradox has inspired recent research efforts into the domain adaptation learning (DAL) problem [6] in computer vision and machine learning [7]. Domain adaptation (or cross-domain) image classification emerges as one of the major techniques to accommodate to the above-mentioned dilemma, which aims to adapt the feature distribution in the source domain to the target domain. In general, a domain refers to data of a certain type, from a certain source, or generated in a certain period of time, etc. Unlike conventional image classification methods, a related source (or auxiliary) domain is provided in domain adaptation image classification to assist the classification process in the target domain and the distinction of data distribution between source and target domains should be minimized in the process of adaptation. Recently, many cross-domain learning techniques have been proposed to solve the problem of distribution mismatch in the field of image or video concept detection and classification [4,8–15].

Existing solutions to DAL vary in setting and methodology. Depending on how the source information is exploited, the division is between classifier-based and representation-based adaptation. The former advocates implicit adaptation to the target distribution by adjusting a classifier from the source domain (e.g., [14,15]), whereas the latter attempts to achieve alignment by adjusting the representation of the source data via learning a transformation [9,10,13,16]; e.g., Pan et al. [16] proposed to extract a "good" feature representation through which the probability distributions of labeled and unlabeled data are drawn close, which achieves much better classification performance by explicitly reducing distribution divergence. Orthogonal to this, the extant methods can also be classified into supervised (e.g., [8,12,14,15]) and unsupervised (e.g., [11,16]) adaptation, based on whether labels have been exploited during the adaptation.

While the effectiveness and efficiency of cross-domain image classification make it of particular use in practice, it also brings a new challenge, i.e., how to handle the errors (e.g., noise and corruption) which may exist in the training data. Since the image data may be randomly obtained from the Internet or other open source websites such as Flicker and YouTube, noise and outliers may by nature abound in the training data [17]. The common issues with the existing cross-domain image classification methods are twofold. First, during the adaptation, they typically deal with source samples separately, without accounting for the mixed (both local and global) structures of data. This may (either implicitly or explicitly) cause the adapted distribution to be arbitrarily scattered around [18] and any structural information beyond single data of the source data may become undermined. Second, they blindly translate all image data, including the noise and particularly possible outliers, from the source domain to the target [4]. The latter can lead to significantly distorted or corrupted models when the classification models are learned. Hence, how to guarantee the robustness of cross-domain image classification by handling data that may not strictly follow domain distribution structures, is an important challenge to robust domain adaptation image classification tasks.

Note that image representation is a crucial procedure for robust image processing and understanding [19]. With this viewpoint, in this paper, we study the robust cross-domain image classification problem using robust feature representation such as sparse representation (SR) [20,21] and low-rank representation (LRR) [22–24]. To this end, by exploiting the advances in LRR [23,24] and DAL [4], we propose a Robust Domain Adaptation Learning framework (RDAL) based on the obtained Domain Adaptation Sparse and Low Rank Representation (DASLRR) of domains data for cross-domain image classification. Specifically, we first aim to represent whole of the domains data as a linear combination of the learned bases, where the representation coefficients should be sparse and low rank, and locally consistent as well as robust to distribution difference. Distribution adaptation ensures that DASLRR is effective for data drawn from different distributions. Lowest rankness ensures that DASLRR can better capture the global cluster structures of the data and is more robust to noise and outliers, while sparsity and local consistency capture the mixed (global and local) discriminative information. We formulate a constrained optimization problem that incorporates the minimum distribution distance criterion, which makes the new representations of both domains close to each other, and ensures local consistency in the objective function of low rank and sparse representation [24]. Subsequently, we further propose a robust semi-supervised learning (SSL) framework based on the newly-obtained DASLRR, thus smoothly implementing label propagation from source domain to target domain [8]. Our method to learn an optimal DASLRR is markedly different from previous works in that we propose to jointly optimize the DASLRR with the constructed over-complete dictionary and constrain the optimal DASLRR coefficients to be adaptive to distribution divergence. Moreover, the proposed classification framework also intrinsically differentiates it from existing works [25,26] in such aspects as Laplacian regularization term construction and feature representation. Extensive experiments on public databases for various cross-domain image classification tasks demonstrate that the proposed method can significantly improve the performance of domain adaptation image classification. These results clearly verify that the proposed framework is more robust and discriminative than conventional methods.

The main contributions of this paper can be summarized as follows:

(1) We present a robust domain adaptation image representation method termed DASLRR using sparse and low-rank representation regularized by both local consistency and distribution adaptation. It extends conventional low-rank and sparse representation into cross-domain learning scenarios where the training data may be drawn from a different distribution. In addition, we alternate learning an optimal dictionary and obtaining the DASLRR of the entire dataset. The dictionary learned by our method has good reconstruction and discriminative capabilities. With this high-quality dictionary, we are able to learn an optimal DASLRR and further a robust classification model.
(2) Based on the obtained DASLRR, we further provide a semi-supervised learning framework, which can propagate the labels of labeled data to unlabeled data from In-Sample as well as Out-of-Sample datasets by jointly learning a prediction label matrix and a classifier model in the framework.
(3) To the best of our knowledge, this is the first-ever to consider simultaneously both global and local discriminative information as well as distribution adaptation in the LRR objective and form a unified optimization function. The proposed method can capture the global mixture of clustering structure (by the sparseness and low rankness) and the local intrinsic structure (by the local graph regularization) as well as distribution difference (by the distribution adaptation) of the data.
(4) We have conducted extensive experiments on public databases for various cross-domain image classification tasks. In many of these experiments, we see that the proposed method can significantly improve the performance of DAL