Information Sciences 418–419 (2017) 242–257

Contents lists available at ScienceDirect

Information Sciences

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

Domain class consistency based transfer learning for image classification across domains

Lei Zhang[a,*], Jian Yang[b], David Zhang[c]

[a] College of Communication Engineering, Chongqing University, Chongqing 400044, China
[b] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
[c] Department of Computing, The Hong Kong Polytechnic University, Hong Kong

ARTICLE INFO

Article history:
Received 17 October 2016
Revised 30 July 2017
Accepted 4 August 2017
Available online 8 August 2017

Keywords:
Transfer learning
Representation learning
Subspace learning
Kernel learning
Domain adaptation

ABSTRACT

Distribution mismatch between the modeling data and the query data is a known domain adaptation issue in machine learning. To this end, in this paper, we propose a $l_{2,1}$-norm based discriminative robust kernel transfer learning (DKTL) method for high-level recognition tasks. The key idea is to realize robust domain transfer by simultaneously integrating *domain-class-consistency* (DCC) metric based discriminative subspace learning, kernel learning in reproduced kernel Hilbert space, and representation learning between source and target domain. The DCC metric includes two properties: *domain-consistency* used to measure the between-domain distribution discrepancy and *class-consistency* used to measure the within-domain class separability. The essential objective of the proposed transfer learning method is to maximize the DCC metric, which is equivalently to minimize the *domain-class-inconsistency* (DCIC), such that domain distribution mismatch and class inseparability are well formulated and unified simultaneously. The merits of the proposed method include (1) the robust sparse coding selects a few valuable source data with noises (outliers) removed during knowledge transfer, and (2) the proposed DCC metric can pursue more discriminative subspaces of different domains. As a result, the maximum class-separability is also well guaranteed. Extensive experiments on a number of visual datasets demonstrate the superiority of the proposed method over other state-of-the-art domain adaptation and transfer learning methods.

1. Introduction

One basic assumption of machine learning is that the training data and testing data should hold similar probability distribution, i.e. independent identical distribution (*i.i.d*) which shares the same feature subspace. However, in many real applications, machine learning faces with the dilemma of insufficient labeled data. For learning a robust classification model, researchers have to “borrow” more data from other domains for training. One problem of the borrowed data is that the distribution mismatch between *source* domain and *target* domain violates the basic assumption of machine learning. Specifically, domain mismatch often results from a variety of visual cues or abrupt feature changes, such as camera viewpoint, resolution (e.g. image sensor from webcam to DSLR), illumination conditions, color correction, poses (*e.g.* faces with different angles), and background, etc. Physically, such distribution mismatch or domain shift is common knowledge in vision

* Corresponding author.
E-mail address: leizhang@cqu.edu.cn (L. Zhang).

http://dx.doi.org/10.1016/j.ins.2017.08.034

Fig. 1. Examples of object images from 4 sources: Amazon (1st row), DSLR (2nd row), Webcam (3rd row) and Caltech (4th row).

problems. With this violation, significant performance degradation is suffered in classification [2]. For example, given a typical object recognition scenario in computer vision, users often recognize a given query object captured by a mobile phone via a well-trained model using the labeled training data from an existing object dataset, such as Caltech 256 [14] or web images. However, these training data may be sampled under different ambient visual cues from the query image. As a result, a failure will be encountered during users' testing process. Some example images of objects from different domains are shown in Fig. 1, which explicitly shows the domain shifts/bias.

In order to deal with such domain distribution mismatch issues, transfer learning and domain adaptation based methods have been emerged [4,13,16,20,32,33,40,41,42], which can be generally divided into two categories: classifier-based and feature-based. Specifically, the classifier based methods advocate learning a transfer classifier on the source data, by leveraging a few labeled data from the target domain simultaneously [1,4,5,6,40,42]. The "borrowed" target data implies the role of regularization, which can trade-off the decision boundary, such that the learned decision function (e.g. SVM) is posed the transfer capability and can be used for classification of domains with bias. The idea of classifier based techniques is straightforward and easy to understand, however, during the decision boundary determination, a number of labeled data are necessary, which may increase the cost of data labeling. Essentially, the classifier based methods attempt to learn a generalized decision function without mining the intrinsic visual drifting mechanism, thus they cannot solve the distribution mismatch fundamentally.

Further, the feature based representation and transformation methods [9,12,13,43,44] aim at aligning the domain shift by adapting features from the source domain to target domain without training classifiers. Although these methods have been proven to be effective for domain adaptation, two issues still exist. First, for representation based adaptation, the noise and outliers from source data may also be transferred to target data due to overfitting of naïve transformation, which leads to significantly distorted and corrupted data structure. Second, the learned subspace is suboptimal, due to the fact that the subspace and the representation (e.g. global low-rank, local sparse coding etc.) are learned independently, which limits the transfer ability. Third, nonlinear transfer often happens in real application, and cannot be effectively interpreted by using linear reconstruction. Therefore, subspace learning and kernel learning that help most to representation transfer and nonlinear transfer should be conducted and integrated simultaneously.

Additionally, Long, et al. [24,25] proposed class-wise adaptation regularization method (ARTL) which learns an adaptive classifier by jointly optimizing the structural risk and distribution matching between both marginal and conditional distribution for transfer learning. Considering the labeling cost of target domain, unsupervised domain adaptation methods have been proposed [11,26]. By leveraging the strong learning capability of deep learning, with the convolution neural network (CNN) and maximum mean discrepancy (MMD) criteria, deep transfer learning methods such as residual transfer network (RTN) [27], deep adaptation network (DAN) [28,29], and joint CNN model [37,38] have also been proposed. Deep transfer learning depends on pre-trained knowledge network on a larger dataset (e.g. ImageNet), so that the transfer performance is greatly improved. In this paper, the proposed method is essentially a shallow transfer learning model, therefore, for comparing with deep transfer models, the CNN based deep features (e.g. DeCAF) are exploited in this paper.

As described in Fig. 2, in this paper, we propose a novel model which targets at learning a discriminative subspace $\mathbf{P}$ by using a newly proposed *domain-class-consistency* metric, a reproduced kernel Hilbert space, and a $l_{2,1}$-norm constrained representation. This work is an extension of the IJCNN conference paper [45], by adding more detailed algorithmic deduction and discussion throughout the paper, conducting new experiments on benchmark datasets, introducing parameter sensitive analysis, empirical comparison of computational time, and comparing with more deep transfer learning methods. The proposed method has three merits:

(1) It can learn a discriminative subspace for each domain and guarantee the maximum separability of different classes (i.e. c_1, c_2, c_3.) within the same domain. In the model, we formulate to maximize the inter-class distance within the same domain, such that the inter-class difference within a domain can cover the between-domain discrepancy. In