ARTICLE IN PRESS
JID: PR [m5G;April 24, 2017;11:38]
Pattern Recognition 000 (2017) 1–15
Contents lists available at ScienceDirect
Pattern Recognition
journal homepage: www.elsevier.com/locate/patcog

Semi-supervised transfer subspace for domain adaptation

Luís A.M. Pereira*, Ricardo da Silva Torres
Institute of Computing, University of Campinas – UNICAMP, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 16 November 2016
Revised 19 February 2017
Accepted 12 April 2017
Available online xxx

Keywords:
Cross-domain knowledge transfer
Cross-dataset classification
Dataset bias
Metric learning
Semi-supervised learning

ABSTRACT

Domain shift is defined as the mismatch between the marginal probability distributions of a source (training set) and a target domain (test set). A successful research line has been focusing on deriving new source and target feature representations to reduce the domain shift problem. This task can be modeled as a semi-supervised domain adaptation. However, without exploiting at the same time the knowledge available on the labeled source, labeled target, and unlabeled target data, semi-supervised methods are prone to fail. Here, we present a simple and effective Semi-Supervised Transfer Subspace (SSTS) method for domain adaptation. SSTS establishes pairwise constraints between the source and labeled target data, besides it exploits the global structure of the unlabeled data to build a domain invariant subspace. After reducing the domain shift by projecting both source and target domain onto this subspace, any classifier can be trained on the source and tested on target. Results on 49 cross-domain problems confirm that SSTS is a powerful mechanism to reduce domain shift. Furthermore, SSTS yields better classification accuracy than state-of-the-art domain adaptation methods.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Domain shift is defined as the mismatch between the marginal probability distributions of a source (training set) and a target domain (test set). This is a prevalent problem in machine learning mainly in real-world applications. In computer vision, for instance, domain shift occurs essentially because visual data are often captured by different devices and under varied imaging conditions such as scene, pose, and illumination [30]. Across visual datasets the domain shift – also known as dataset bias – can be severe [35] (Fig. 1). Speech and language processing are also subjected to domain shift [3,9]. Under such scenarios, however, conventional classifiers often fail to achieve desirable performances at test time [30] because they assume a stationary environment, *i.e.*, source and target domain are supposedly drawn from the same probability distribution. The limitation of this assumption has motivated the development of domain adaptation methods to reduce the domain shift and increase the classifier performance (refer to [16,29] for a complete literature review on domain adaptation).

Such methods are proposed either in semi-supervised or in unsupervised settings. The semi-supervised methods use a fully labeled source data and a partially labeled target data to guide the domain adaptation (*e.g.*, [30,36,38,40]). In contrast, unsupervised methods use a fully labeled source along with a fully unlabeled target data (*e.g.*, [5,12,14,22,23]). Usually, both fully labeled source and fully unlabeled target data are in plenty. Labeled target data, in turn, are more scarce. Generally, semi-supervised methods perform better than unsupervised methods even in the presence of few amount of labeled target data [14,15] – besides, better results can be achieved if the unlabeled target data is properly exploited [36,38,40].

Several semi-supervised domain adaptation methods were proposed as max-margin classifier extensions (*e.g.*, [4,11,18]), whose goal is to learn the model parameters on the source domain and then transfer them to the target domain. The assumption is that the target model is a perturbed version of the source model. Without enough representative labeled target instances, these methods are prone to perform poorly due to the importance given to class labels. Hence, several studies advocate for feature-based methods, whose goal is to reduce the domain shift by approximating the source and target feature distributions [29]; therefore, any classifier can be trained on the source domain and tested on the target domain.

A prominent feature-based method (which inspired several works including ours) was proposed by Saenko et al. [30]. The method uses information theoretic metric learning to learn a regularized transformation, *i.e.*, an explicit domain-invariant metric. This metric is then applied to map source features to target features. A key concept of this algorithm is to establish inter-domain pairwise constraints between the source and labeled target in-

* Corresponding author.
E-mail addresses: luis.pereira@ic.unicamp.br (L.A.M. Pereira), rtorres@ic.unicamp.br (R. da S. Torres).

http://dx.doi.org/10.1016/j.patcog.2017.04.011
0031-3203/© 2017 Elsevier Ltd. All rights reserved.

Please cite this article as: L.A.M. Pereira, R. da S. Torres, Semi-supervised transfer subspace for domain adaptation, Pattern Recognition (2017), http://dx.doi.org/10.1016/j.patcog.2017.04.011

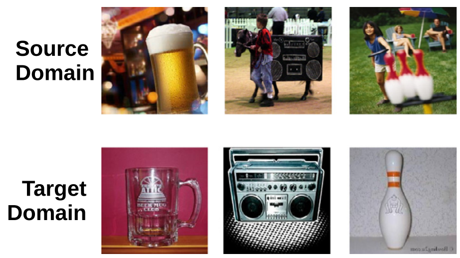

Fig. 1. Source (training set) and the target (test set) domain may contain images belonging to same class though captured under different scenes, pose, and/or illumination. This imaging conditions cause mismatch (dataset bias) between the source and target distributions.

stances to preserve the target discriminative structure. Unlike other feature-based algorithms, this method does not take any advantage of unlabeled target data (generally in plenty). This may justify its poor performance when compared to subspace-based algorithms (*e.g.*, [14,15]), which somehow exploit the information available on the unlabeled target data.

Semi-supervised subspace-based methods are feature-based algorithms designed under the assumption of a common domain-invariant subspace (*i.e.*, an implicit domain-invariant metric) in which the domain shift between source and target distributions is reduced. The state-of-the-art subspace-based methods exploit the information available on both source and target domain. The Laplacian Embedding framework is the common choice to exploit the underlying local structure of the unlabeled target data (*e.g.*, [6,26,36,39,40]). Despite its theoretical appeal, such approach is graph-dependent, *i.e.*, the better the graph is induced, the better it captures the data local structure; nevertheless, the opposite is also true [1,37]. Determining the best graph type (*e.g.*, k-nn or ϵ-ball) is a complex task [37,41] and, depending on the choice, the complexity of the domain adaptation can increase and its performance can also be negatively affected.

Here, we present a simple and effective semi-supervised method for domain adaptation. Our method – referred to as Semi-Supervised Transfer Subspace (SSTS) – exploits properties of the source and labeled/unlabeled target data to yield a domain-invariant subspace. Essentially, SSTS uses two processes to accomplish this task. First, SSTS takes advantage of the data global structure, (*i.e.*, the data variance), including the target unlabeled data. This allows enhancing the domain adaptation [14,15] without the need of handling the inconveniences of the local structure preserving approach. Second, SSTS establishes interdomain pairwise constraints between the source and labeled target instances to preserve discriminative properties (*i.e.*, the classes separability) in the domain-invariant subspaces. Mathematically, the combination of these two processes leads to a (non)linear domain-invariant metric (as we show in Section 5). After projecting both source and target domain onto the domain-invariant subspaces to reduce the domain shift, any classifier can be trained on the source and tested on target. A schematic illustration of SSTS is displayed in Fig. 2.

We carried out extensive experiments on 49 real-world visual cross-domain problems. Besides the standard manually-labeled cross-domain problems, we also evaluated SSTS in the weakly-labeled scenario, which has Internet photos retrieved by keyword-based image search engines. Results on these cross-domain problems confirm that the SSTS is a powerful mechanism to reduce domain shift. Furthermore, SSTS yields better classification accuracy than state-of-the-art domain adaptation methods.

2. Notations

We consider that our data come from two domains:

(i) a fully labeled source domain $X^{\mathcal{S}} = [\boldsymbol{x}_1^{\mathcal{S}}, \boldsymbol{x}_2^{\mathcal{S}}, \ldots, \boldsymbol{x}_{n_{\mathcal{S}}}^{\mathcal{S}}] \in \mathbb{R}^{\mathcal{D}\times n_{\mathcal{S}}}$; and
(ii) a partially labeled target domain with its labeled part denoted as $X^{\mathcal{T}} = [\boldsymbol{x}_1^{\mathcal{T}}, \boldsymbol{x}_2^{\mathcal{T}}, \ldots, \boldsymbol{x}_{n_{\mathcal{T}}}^{\mathcal{T}}] \in \mathbb{R}^{\mathcal{D}\times n_{\mathcal{T}}}$, and the unlabeled part denoted as $X_{\mathcal{U}}^{\mathcal{T}} = [\boldsymbol{x}_1^{\mathcal{U}}, \boldsymbol{x}_2^{\mathcal{U}}, \ldots, \boldsymbol{x}_{n_{\mathcal{U}}}^{\mathcal{U}}] \in \mathbb{R}^{\mathcal{D}\times n_{\mathcal{U}}}$, with $n_{\mathcal{T}} \ll n_{\mathcal{S}}$ and $n_{\mathcal{T}} \ll n_{\mathcal{U}}$;

such that $P(X^{\mathcal{S}}) \neq P(X^{\mathcal{T}} \cup X_{\mathcal{U}}^{\mathcal{T}})$, while $P(Y^{\mathcal{S}}|X^{\mathcal{S}}) \approx P(Y^{\mathcal{T}}|X^{\mathcal{T}} \cup X_{\mathcal{U}}^{\mathcal{T}})$ with $Y^{\mathcal{S}}$ and $Y^{\mathcal{T}}$ standing for the source and the target set of labels, respectively. Additionally, let $\tilde{\boldsymbol{X}} = [X^{\mathcal{S}}|X^{\mathcal{T}}|X_{\mathcal{U}}^{\mathcal{T}}] = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{\mathcal{D}\times N}$ be the concatenated matrix of instances from both domains with $N = n_{\mathcal{S}} + n_{\mathcal{T}} + n_{\mathcal{U}}$. Note that an instance $\boldsymbol{x} \in \tilde{\boldsymbol{X}}$ is a column vector denoted without any superscript.

Let $\mathcal{M}, \mathcal{C} \subset X^{\mathcal{S}} \times X^{\mathcal{T}}$ be two sets of pairwise constraints such that

- $\mathcal{M} = \{(\boldsymbol{x}_i^{\mathcal{S}}, \boldsymbol{x}_j^{\mathcal{T}}) \mid \boldsymbol{x}_i^{\mathcal{S}}$ and $\boldsymbol{x}_j^{\mathcal{T}}$ are similar$\}$ is the set of inter-domain must-link constraints; and
- $\mathcal{C} = \{(\boldsymbol{x}_i^{\mathcal{S}}, \boldsymbol{x}_j^{\mathcal{T}}) \mid \boldsymbol{x}_i^{\mathcal{S}}$ and $\boldsymbol{x}_j^{\mathcal{T}}$ are dissimilar$\}$ is the set of inter-domain cannot-link constraints.

The words “similar” and “dissimilar” mean “same class,” and “different class,” respectively.

3. Description of the proposed method

Here, we describe in detail the SSTS linear version. In Section 4, we show how this linear version can be extended to handle non-linear feature deformations.

3.1. Encoding discriminative knowledge

To exploit the discriminative knowledge available in the source domain (in large amount) and in the target domain (in a small amount), inter-domain pairwise constraints were encoded into two different functionals. The first functional concerns to *maximize* the squared induced distance between instances belonging to the *different* class and from *different* domains. Mathematically, this functional is defined as

$$\mathcal{F}_{\mathcal{C}}(Q) = \frac{1}{2} \sum_{\forall (x_i^{\mathcal{S}}, x_j^{\mathcal{T}}) \in \mathcal{C}} W_{ij}^{\mathcal{C}} ||Q^{\top}\boldsymbol{x}_i^{\mathcal{S}} - Q^{\top}\boldsymbol{x}_j^{\mathcal{T}}||_2^2, \tag{1}$$

in which

$$W_{ij}^{\mathcal{C}} = \begin{cases} \frac{1}{|\mathcal{C}|} & \text{if } (\boldsymbol{x}_i^{\mathcal{S}}, \boldsymbol{x}_j^{\mathcal{T}}) \in \mathcal{C}, \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

is the matrix of weights for cannot-link constraints. $Q \in \mathbb{R}^{\mathcal{D}\times d}$ is the transformation matrix.

Simplifying the functional in Eq. (1) implies that

$$\mathcal{F}_{\mathcal{C}}(Q) = \mathrm{Tr}(Q^{\top}\tilde{\boldsymbol{X}}\mathcal{L}_{\mathcal{C}}\tilde{\boldsymbol{X}}^{\top}Q), \tag{3}$$

in which $\mathrm{Tr}(\cdot)$ stands for the trace operator.[1] The matrix $\mathcal{L}_{\mathcal{C}} \equiv D^{\mathcal{C}} - W^{\mathcal{C}}$ is the Laplacian matrix [7], such that $D^{\mathcal{C}}$ is a diagonal matrix whose elements in the principal diagonal are defined as $D_{ii}^{\mathcal{C}} = \sum_j W_{ij}^{\mathcal{C}}$ (i.e., the column sum of the matrix $W^{\mathcal{C}}$).

The second functional attempts to *minimize* the squared induced distance between instances belonging the *same class* and from *different* domains. Analytically, this functional is defined as

$$\mathcal{F}_{\mathcal{M}}(Q) = \frac{1}{2} \sum_{\forall (x_i^{\mathcal{S}}, x_j^{\mathcal{T}}) \in \mathcal{M}} W_{ij}^{\mathcal{M}} ||Q^{\top}\boldsymbol{x}_i^{\mathcal{S}} - Q^{\top}\boldsymbol{x}_j^{\mathcal{T}}||_2^2, \tag{4}$$

[1] Note that the superscript $^{\top}$ denotes the transpose operator, while the superscript $^{\mathcal{T}}$ indicates the target domain.

Please cite this article as: L.A.M. Pereira, R. da S. Torres, Semi-supervised transfer subspace for domain adaptation, Pattern Recognition (2017), http://dx.doi.org/10.1016/j.patcog.2017.04.011

Download English Version:

https://daneshyari.com/en/article/6939685

Download Persian Version:

https://daneshyari.com/article/6939685

Daneshyari.com