Dyadic Transfer Learning for Cross-Domain Image Classification

Hua Wang, Feiping Nie, Heng Huang, Chris Ding
Department of Computer Science and Engineering,
University of Texas at Arlington, Arlington, TX 76019, USA
huawangcs@gmail.com, feipingnie@gmail.com, heng@uta.edu, chqding@uta.edu

Abstract

Because manual image annotation is both expensive and labor intensive, in practice we often do not have sufficient labeled images to train an effective classifier for the new image classification tasks. Although multiple labeled image data sets are publicly available for a number of computer vision tasks, a simple mixture of them cannot achieve good performance due to the heterogeneous properties and structures between different data sets. In this paper, we propose a novel nonnegative matrix tri-factorization based transfer learning framework, called as Dyadic Knowledge Transfer (DKT) approach, to transfer cross-domain image knowledge for the new computer vision tasks, such as classifications. An efficient iterative algorithm to solve the proposed optimization problem is introduced. We perform the proposed approach on two benchmark image data sets to simulate the real world cross-domain image classification tasks. Promising experimental results demonstrate the effectiveness of the proposed approach.

1. Introduction

Automatically organizing and indexing multimedia content becomes increasingly important as the online images and videos continue to be a vital resource in everyday life. Consequently, devising effective visual category models has attracted considerable attention in computer vision area in recent years. When sufficient labeled training images are available, traditional classification methods usually work well. However, because image labeling requires expensive and time-consuming human labors, it is not likely that we always have enough training data to achieve satisfactory performance for new computer vision tasks. Therefore, it is desirable in image classification to leverage labeled images from previous learning tasks, as well as abundant unlabeled images from online resources.

Semi-supervised learning methods are able to find structures in available unlabeled data and use the structures to improve the performance of a supervised task, *e.g.*, [12,14]. However, these methods assume that both unlabeled data and labeled data are drawn from a same distribution, and do not generally exploit knowledge from different distributions, such as the data learned from previous supervised tasks. A simple mixture of data sets from different tasks cannot help the new computer vision task. As a result, transfer learning methods that discover useful knowledge from previous tasks to make a future related learning task possibly only using a small number of samples have been found useful in many real world applications, such as video concept detection [4, 13], sentiment classification [6], natural language learning processing [1], and many others [7, 8].

Considering the data of a computer vision task forming a domain, transfer learning learns the classifiers via a limited number of available labeled data from the *target domain* by taking advantage of a large amount of (labeled) data from other domains, referred to as *source domains* (also called as auxiliary domains in some research papers). The main problems to utilize transfer learning in image classifications are what knowledge should be transferred? and how to transfer them cross-domain?

To address these problems, in this paper, we propose a novel Nonnegative Matrix Tri-factorization (NMTF) based transfer learning framework to extract and transfer knowledge by two ways from the source domain to help the classification in the target domain, *i.e.*, both unsupervised and supervised ways. As illustrated in Figure 1, we use the unsupervised way to transfer the native structural information of the source data to the target data by sharing feature clustering structures. For example in image data, green color (*e.g.*, tree) and blue color (*e.g.*, sky) could appear together to characterize outdoor scenes, while the car class could be abstracted as a combination of several round shapes (features) together with two trapezoid shapes (features). Meanwhile, we employ the supervised way to transfer the association between feature clusters and semantic labels. Because the proposed approach uses two different paths to transfer knowledge, we call it as Dyadic Knowledge Transfer (DKT) approach. A new unified objective is proposed to combine both unsupervised and supervised cross-domain knowledge

978-1-4577-1102-2/11/$26.00

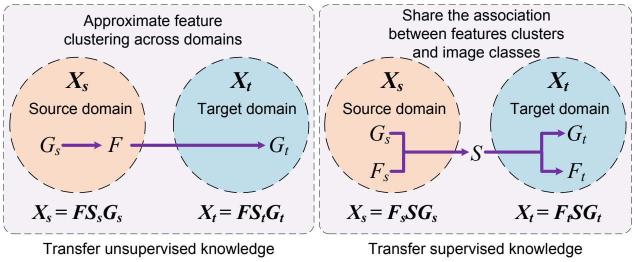


Figure 1. Dyadic Knowledge Transfer— bridging domains in two paths: (left) sharing common feature clustering structures encoded in F; and (right) sharing the association between feature clusters and semantic image classes by S.

transferring. An efficient iterative algorithm to solve the proposed optimization objective is derived. Promising results in extensive experiments demonstrate the effectiveness of our method.

2. NMTF based dyadic knowledge transfer

In this section, we first briefly review NMTF, based on which we formalize the problem of cross-domain image classification. Then we will introduce the objective of the proposed DKT approach, followed by the derivation of a new efficient algorithm to solve the objective.

Review of NMTF. NMTF [2, 3] aims at approximating the nonnegative data matrix $X \in \mathbb{R}_+^{m \times n}$ of an image data set by three nonnegative factor matrices:

$$\min_{F \geq 0, S \geq 0, G \geq 0} \|X - FSG^T\|^2, \tag{1}$$

where $\|\cdot\|$ denotes the Frobenius norm of a matrix, $F \in \mathbb{R}_+^{m \times k_1}$ and $G \in \mathbb{R}_+^{n \times k_2}$ are two side factor matrices and $S \in \mathbb{R}_+^{k_1 \times k_2}$ absorbs the different scales of the X, F and G. The true power of NMTF in statistical learning lies in its equivalence to simultaneous k-means clustering of both features and data points, called as *co-clustering* [2, 3]. Specifically, each row of F is the soft clustering indication for a feature (*e.g.*, a codeword), and each row of G is the soft clustering indication for a data point (*e.g.*, image). S hence describes the associations between the feature clusters and data clusters. In other words, F contains the native structural information of X which is unsupervised; and S contains supervised information of X due to the human annotations on the source data points. In the proposed model, we transfer knowledge across domains using both F and S.

Due to its mathematical elegance and promising practical results in statistical learning, NMTF has been further developed to deal with transfer learning problems [6, 15]. However, these methods are only able to transfer one type of source knowledge, either unsupervised [6] or supervised [15]. In order to explore the full potential of the source data, we design our DKT approach to transfer the both types of source knowledge to the target data.

Problem Formalization. For a cross-domain image classification task, we have two image data sets, one in source domain $X_s = \left[\mathbf{x}_s^1, \ldots, \mathbf{x}_s^{n_s}\right] \in \mathbb{R}^{m \times n_s}$ and the other in target domain $X_t = \left[\mathbf{x}_t^1, \ldots, \mathbf{x}_t^{n_t}\right] \in \mathbb{R}^{m \times n_t}$. We assume that the both data sets use a same codebook with m codewords: if the codebooks differ, we may simply pad zero in the feature vectors and re-express them under the same unified codebook such that the indices of the feature vectors from both data sets correspond to the same codeword.

Typically a number of images in the source domain are manually labeled to capture the domain specific annotations. The partial annotation information can be described by an indication matrix $Y_s \in \mathbb{R}^{n_s \times k_2}$ such that $Y_{s(ik)} = 1$ if $\mathbf{x}_s^i$ belongs to the k-th class, and $Y_{s(ik)} = 0$ otherwise. Sometimes, though not always, we also have a limited amount of image annotations in the target domain. We similarly describe them using $Y_t \in \mathbb{R}^{n_t \times k_2}$ such that $Y_{t(ik)} = 1$ if $\mathbf{x}_t^i$ belongs to the k-th class, and $Y_{t(ik)} = 0$ otherwise. Our goal is to predict labels for the unannotated images in the target domain.

We assume that the two data sets share a same set of classes. If not, we may pad the zero columns to Y_s or Y_t, or both, such that the column indices of the both matrices correspond to the same classes. In addition, we encode the difference between the two sets of classes by two matrices, one for the source domain and the other for the target domain: $Q_s \in \mathbb{R}^{k_2 \times k_2}$ is a diagonal matrix with $Q_{s(ii)} = 1$ if the i-th class comes from the source data set, and $Q_{s(ii)} = 0$ otherwise; and $Q_t \in \mathbb{R}^{k_2 \times k_2}$ is a diagonal matrix with $Q_{s(ii)} = 1$ if the i-th class comes from the target data set, and $Q_{t(ii)} = 0$ otherwise. Note that, we need the both matrices because one class could appear in the both data sets.

Throughout this paper, we denote $\mathbb{R}$ as real numbers and $\mathbb{R}^+$ as positive real numbers. The entry at the i-th row and j-th column of a matrix M is denoted as $M_{(ij)}$. Some frequently used notations are summarized in Table 1.

2.1. Objective of the DKT approach

Given the image data X_s in the source domain and the corresponding annotations Y_s, we formulate the following optimization problem [5, 6, 15]:

$$\begin{aligned} \min_{F_s \geq 0, S_s \geq 0, G_s \geq 0} & \|X_s - F_s S_s G_s^T\|^2 \\ & + \alpha \, \mathbf{tr} \left[Q_s \left(G_s - Y_s\right)^T C_s \left(G_s - Y_s\right) \right], \end{aligned} \tag{2}$$

where $\mathbf{tr}(\cdot)$ denotes the trace of a matrix. Here, $\alpha > 0$ is a parameter that determines to which extent we enforce the prior labeling knowledge in the source domain $G_s \approx Y_s$, $C_s \in \mathbb{R}^{n_s \times n_s}$ is a diagonal matrix whose entry $C_{s(ii)} = 1$

Table 1. Some frequently used notations.

X_s	data matrix of the image data set in the source domain
X_t	data matrix of the image data set in the target domain
n_s	number of images of the source data set
n_t	number of images of the target data set
F	codeword cluster indicator matrix
S	the matrix associating codeword clusters and image classes
G_s	image class indicator matrix of the source data set
G_t	image class indicator matrix of the target data set
Y_s	label matrix of the source data set
Y_t	label matrix of the target data set
Q_s	class sharing indication matrix of the source data set
Q_t	class sharing indication matrix of the target data set
C_s	image annotation indication matrix of the source data set
C_t	image annotation indication matrix of the target data set

if image $\mathbf{x}_s^i$ is annotated by the i-th row of Y_s, and $C_{s(ii)} = 0$ otherwise. Note that, if $C = I$, all the images in the source domain are fully annotated and specified by Y_s. Q_s is used to enforce the prior knowledge only for the classes that belong to the source data set.

Solving Eq. (2), we obtain F_s^* and S_s^*, which contain the unsupervised and supervised information of source data, respectively. We hence transfer the knowledge from the source domain to the target domain by both of them.

Transfer *unsupervised* knowledge via sharing feature clusters by F_s^*. First, we attempt to transfer the unsupervised knowledge in the source domain, which is the native structural information of the source data set. We achieve this by solving the following optimization problem in the target domain:

$$\begin{aligned} \min_{S_t \geq 0, G_t \geq 0} \; & \|X_t - F_s^* S_t G_t^T\|^2 \\ & + \alpha \, \mathbf{tr} \left[Q_t \left(G_t - Y_t \right)^T C_t \left(G_t - Y_t \right) \right] \quad . \end{aligned} \tag{3}$$

The second term in Eq. (3) acts same as that in Eq. (2), which enforces label information in the target domain if it is available. Here, $C_t \in \mathbb{R}^{n_t \times n_t}$ is a diagonal matrix whose entry $C_{t(ii)} = 1$ if image $\mathbf{x}_t^i$ is annotated by the i-th row of Y_t, and $C_{t(ii)} = 0$ otherwise. When image labels in the target data set are not available, $C_t = 0^{n_t \times n_t}$ is a zero matrix. The key part the first term, where we force the feature clustering indication of the target data set to be same as that of the source data set. As a result, the unsupervised structural information of the source data X_s is transferred to the label assignments G_t of the target data X_t via the feature cluster indications F_s^*, which is schematically shown in Figure 2(a) (the image features can be color moment, SIFT, *etc*).

Transfer *supervised* knowledge via sharing the association between the feature clusters and image classes by S^*. Compared to feature clusters, the associations between feature clusters and image classes are more reliable to convey semantic relationships across different domains [15]. Formally, we solve the following optimization problem on the target data set:

$$\min_{F_t \geq 0, G_t \geq 0} \; \|X_t - F_t S_s^* G_t^T\|^2 \quad . \tag{4}$$

As a result, S_s^*, learned from source data set, is used as supervision to classify target data. Namely, S_s^* bridges the source domain and the target domain so that the prior labeling (supervised) knowledge can be transferred from the former to the latter. Figure 2(b) illustrates an example to label a new image using the transferred associations of classes and features.

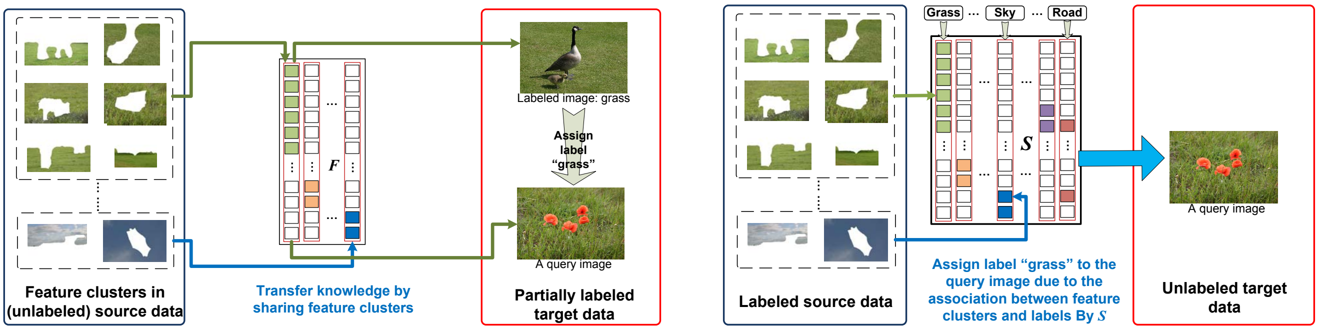
Our optimization objective. Finally, instead of solving the three optimization problems in Eqs. (2–4) separately, we propose to solve the following joint optimization problem:

$$\begin{aligned} \min_{\substack{F \geq 0, S \geq 0, \\ G_s \geq 0, G_t \geq 0}} \; J = & \|X_s - F S G_s^T\|^2 \\ & + \|X_t - F S G_t^T\|^2 \\ & + \alpha \, \mathbf{tr} \Big[Q_s \left(G_s - Y_s \right)^T C_s \left(G_s - Y_s \right) \\ & + Q_t \left(G_t - Y_t \right)^T C_t \left(G_t - Y_t \right) \Big] \, . \end{aligned} \tag{5}$$

In this formulation, both F and S are shared in the two matrix factorizations for both the source data and the target one. The former connects the two domains to transfer unsupervised knowledge, while the latter is used as a bridge for supervised knowledge transformation from the source domain to the target domain.

Note that, when the source data set and the target data set do not share any features, the first term in Eq. (3) for matrix factorization will be decoupled into two subproblems, one for each of the two data sets. Consequently, the results of the subproblem for the source domain will have no effect on the target data set, and no unsupervised knowledge transfer will be performed. Similarly, if the source data set and target data set do not share common classes, there will be no supervised knowledge transformation in the optimization problem Eq. (4), because it is decoupled into two independent subproblems, one for source image data and the other for target image data. However, these two cases rarely happen at the same time, such that we can always transfer knowledge by Eq. (5) through sharing either feature clusters or the the association between feature clusters and image classes, or both. Because Eq. (5) transfers knowledge across domains by two different paths, we call it as the proposed *Dyadic Knowledge Transfer (DKT)* approach.

Solving Eq. (5), we may classify the unlabeled data points $\mathbf{x}_t^i$ in the target data set. If the target data set is a single-label data set where each data point belongs to



(a) Transfer unsupervised information.

(b) Transfer supervised information.

Figure 2. Transfer knowledge by two different paths for both unsupervised information as in Figure 2(a) and supervised information as in Figure 2(b).

only one class, we assign label using the following rule [3] $l\left(\mathbf{x}_t^i\right) = \arg\max_k \; G_{t(ik)}$. If the target data set is a multi-label data set, we classify $\mathbf{x}_t^i$ using the correlative threshold method introduced in [11].

2.2. Optimization algorithm

In the rest of this section, we will derive the solution to Eq. (5) via an alternating scheme to optimize the objective. Specifically, we will optimize one variable while fixing the rest variables. The procedure repeats until convergence.

First, we expand the objective in Eq. (5) as follows:

$$\begin{aligned} &J\left(F, S, G_s, G_t\right) = \\ &\quad \mathbf{tr}\left[-2X_s^T FSG_s^T + G_sS^TF^TFSG_s^T \right.\\ &\quad - 2X_t^T FSG_t^T + G_tS^TF^TFSG_t^T \\ &\quad + \alpha\left(Q_sG_s^TC_sG_s - 2Q_sG_s^TC_sY_s \right.\\ &\quad \left.\left. + Q_tG_t^TC_tG_t - 2Q_tG_t^TC_tY_t\right)\right], \end{aligned} \tag{6}$$

where constant terms are discarded.

Computation of F. For the constraint $F \geq 0$, following standard theory of constrained optimization, we introduce the Lagrangian multiplier $U \in \mathbb{R}^{m\times k_1}$, thus the Lagrangian function is

$$L\left(F\right) = J - \mathbf{tr}\left(UF^T\right) \quad . \tag{7}$$

Setting $\partial L\left(F_s\right)/\partial F_s = 0$, we obtain

$$\begin{aligned} U = &-2X_sG_sS^T + 2FSG_s^TG_sS^T \\ &- 2X_tG_tS^T + 2FSG_t^TG_tS^T \quad . \end{aligned} \tag{8}$$

Using Karush-Kuhn-Tucker (KKT) condition we have $U_{(ij)}F_{(ij)} = 0$, which is

$$\begin{aligned} &\left(-2X_sG_sS^T + 2FSG_s^TG_sS^T\right. \\ &\quad \left.- 2X_tG_tS^T + 2FSG_t^TG_tS^T\right)_{(ij)} F_{(ij)} = 0, \end{aligned} \tag{9}$$

which leads to the following updating formula:

$$F_{(ij)} \leftarrow F_{(ij)} \sqrt{\frac{\left(X_sG_sS^T + X_tG_tS^T\right)_{(ij)}}{\left(FSG_s^TG_sS^T + FSG_t^TG_tS^T\right)_{(ij)}}} \quad . \tag{10}$$

Computation of S, G_s and G_t. Following the same derivations as in Eqs. (7–10), we obtain the updating rules for the rest variables of J as following:

$$S_{(ij)} \leftarrow S_{(ij)} \sqrt{\frac{\left(F^TX_sG_s + F^TX_tG_t\right)_{(ij)}}{\left(F^TFSG_s^TG_s + F^TGSG_t^TG_t\right)_{(ij)}}} \tag{11}$$

$$G_{s(ij)} \leftarrow G_{s(ij)} \sqrt{\frac{\left(X_s^TFS + \alpha C_sY_sQ_s\right)_{(ij)}}{\left(G_sS^TF^TFS + \alpha C_sG_sQ_s\right)_{(ij)}}} \tag{12}$$

$$G_{t(ij)} \leftarrow G_{t(ij)} \sqrt{\frac{\left(X_t^TFS + \alpha C_tY_tQ_t\right)_{(ij)}}{\left(G_tS^TF^TFS + \alpha C_tG_tQ_t\right)_{(ij)}}} \tag{13}$$

The above iterative updating procedures to optimize Eq. (5) are summarized in Algorithm 1, whose convergence can be proved by following the similar way to [2, 3, 5, 15].

3. Experimental Results

In this section, we experimentally evaluate the proposed DKT approach in cross-domain image classification tasks.

3.1. Data preparation

We experiment with the following two image data sets, which are broadly used in computer vision studies.

TRECVID 2005 data set[1] contains 61901 video sub-shots labeled with 39 LSCOM-Lite concepts.

MSRC data set[2] is provided by the computer vision group at Microsoft Research Cambridge, which contains 591 images annotated by 23 classes.

[1] http://www-nlpir.nist.gov/projects/trecvid/

[2] http://research.microsoft.com/en-us/projects/objectclassrecognition

Algorithm 1: Procedures to optimize Eq. (5).

Data: 1. Data matrix X_s of the source image data set,
2. Data matrix X_t of the target image data set,
3. Ground truth annotation Y_s in source domain,
4. Optional ground truth annotation Y_t in target domain,
5. Trade-off parameters α.

Result: Labels assigned to the unannotated images $\mathbf{x}_t^i$ in target data set.

1. Initialize F, S, G_s, and G_t following [15];

repeat

2. Update F using Eq. (10),
3. Update S using Eq. (11),
4. Update G_s using Eq. (12),
5. Update G_t using Eq. (13),

until *converges*

6. Predict labels for $\mathbf{x}_t^i$ using $l\left(\mathbf{x}_t^i\right) = \arg\max_k\ G_{t(ik)}$ for single-label data or the correlative threshold method introduced in [11] for multi-label data.

These two data sets are considered to come from different domains, because the former is from video broadcasts while the latter is from digital photos in daily life. They share the following 10 semantic classes: "building", "sky", "mountain", "airplane (aeroplane)", "waterfront_waterscape (water)", "face", "car", "road", "person (body)", "boat_ship".

We extract dense SIFT (DSIFT) [9] features for the images in these two data sets. Following [10], we resize the images to 256×256 and extract features with grid size of 5 pixels. As a result, 2601 DSIFT features of are exacted for every image. Our method can also be applied to other image features. We use DSIFT features for demonstration.

Using these two image data sets, we construct the following test data set to simulate real world cross-domain image classification tasks, in which the source and target data sets share classes and use a same codebook. We use all the classes of the two data sets, and end up with $23+39-10 = 42$ classes. For TRECVID 2005 data set, following [11], we randomly select 100 images for each of its classes. For MSRC data set, we use all its images. Following [10], a unified 240-dimensional codebook is created for all the images selected from both of the two data sets, where k-means clustering is used to find the codewords from the corresponding DSIFT descriptors. Then feature vector of each image is constructed using the histogram of its DSIFT descriptors based upon the computed codebook [10]. We should also notice that if we directly perform the typical classification methods on the mixed data sets, it is equivalent to do classifications on the simple mixture data without transferring knowledge.

3.2. Effectiveness of knowledge transfer in cross-domain image classification

Because the main purpose of the proposed DKT approach is to transfer knowledge from a source domain to another target domain that does not have sufficient labeled images, we evaluate the knowledge transfer capability of the proposed approach. We compare the proposed DKT approach against its degenerate version as following:

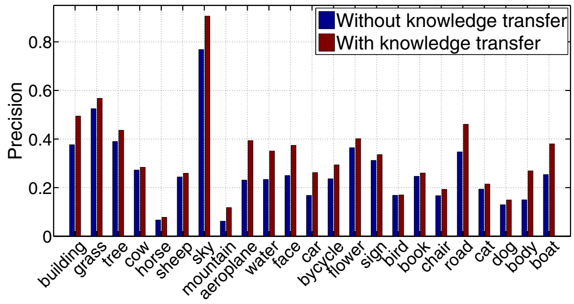
$$\min_{F_t \geq 0, S_t \geq 0, G_t \geq 0} J = \|X_t - F_t S_t G_t^T\|^2 + \mathbf{tr}\left[\alpha Q_t \left(G_t - Y_t\right)^T C_t \left(G_t - Y_t\right)\right], \tag{14}$$

where the knowledge transfer terms in Eq. (5) are removed. Therefore, Eq. (14) is a semi-supervised learning method performed on the target data, whose similar form was ever proposed in [5].

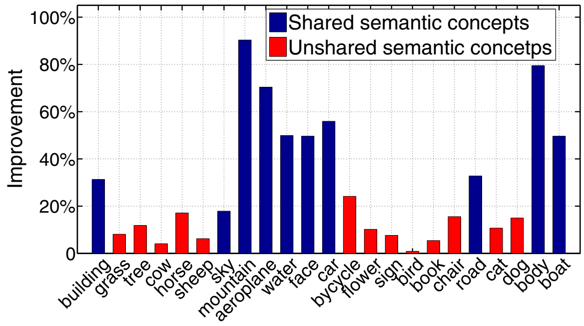
We perform the evaluation on the constructed test data set, where we use TRECVID 2005 data set as source data and MSRC data set as target data. Because in real applications not all the images in source domain are annotated, for the source data set we randomly pick up 70% images from each class as labeled data, and use the rest as unlabeled data. On the other hand, because in real applications the images in target domain are mostly not annotated, we randomly picking up 20% images from each class of the target data and set them as labeled image. Our task is to predict labels for the rest 80% unannotated images in the target image data set. We repeat the experiments for 50 times. The average classification performance measured by precision for each class is reported in Figure 3(a).

From the results we can see that our approach using knowledge transfer outperforms its degenerate version without using knowledge transfer in all the classes. By a more careful analysis on the results, we can see that the improvements due to knowledge transfer (computed by "(Precision with knowledge transfer - Precision without knowledge transfer) / Precision without knowledge transfer") for the shared classes between the two data sets (shown as blue color in Figure 3(b)) are much greater than those of unshared classes (shown as red color in Figure 3(b)) in general. These observations concretely demonstrate the usefulness of knowledge transfer in the task of cross-domain image classification.

In order to evaluate the detailed impact of training information in source data on the classification performance on target data, we vary the amount of labeled images in the source data set and examine the corresponding classification performance on the target data set of our approach. The average precisions over all the classes for different amount of labeled source images are reported in Figure 4, which show that the more labeled data we have in the source data set, the better classification performance we can achieve on the



(a) Performance comparison of the proposed approach (Eq. (5)) and its degenerate version (Eq. (14)).



(b) Performance improvement due to knowledge transfer.

Figure 3. The effectiveness of the proposed DKT approach.

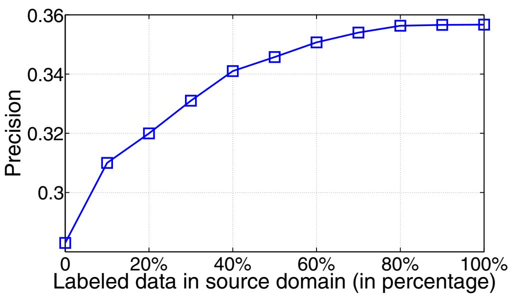


Figure 4. Average precisions over all the classes on the target data set when the amount of labeled images in the source data set varies.

target data set. This is consistent with the theoretical analysis, and again confirms the effectiveness of our approach to transfer knowledge in cross-domain image classification.

4. Conclusions

In this paper, we proposed a novel NMTF based transfer learning approach for cross-domain image classification. Our new approach is flexible to make use of either unlabeled source images by sharing feature cluster structures or labeled source images via the associations between feature clusters and classes, or the both. In addition, labeled images in the target data set, though often unavailable, can also be exploited. We introduced an efficient algorithm to solve the proposed objective. Extensive empirical studies have demonstrated promising results that validate our new approach.

Acknowledgement.

This research was supported by NSF-IIS 1117965, NSF-CCF-0830780, NSF-DMS-0915228, NSF-CCF-0917274.

References

[1] H. Daumé. Frustratingly easy domain adaptation. In *ACL*, 2007.

[2] C. Ding, X. He, and H. Simon. On the equivalence of nonnegative matrix factorization and spectral clustering. In *SDM*, 2005.

[3] C. Ding, T. Li, W. Peng, and H. Park. Orthogonal nonnegative matrix tri-factorizations for clustering. In *SIGKDD*, 2006.

[4] L. Duan, I. Tsang, D. Xu, and S. Maybank. Domain transfer SVM for video concept detection. In *CVPR*, 2009.

[5] Q. Gu and J. Zhou. Transductive Classification via Dual Regularization. In *ECML/PKDD*, 2009.

[6] T. Li, V. Sindhwani, C. Ding, and Y. Zhang. Knowledge transformation for cross-domain sentiment classification. In *SIGIR*, 2009.

[7] A. Quattoni, M. Collins, and T. Darrell. Transfer learning for image classification with sparse prototype representations. In *CVPR*, 2008.

[8] K. Saenko, B. Kulis, M. Fritz, and T. Darrell. Adapting visual category models to new domains. In *ECCV*, 2010.

[9] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms, 2008.

[10] C. Wang, D. Blei, and F. Li. Simultaneous image classification and annotation. In *CVPR*, 2009.

[11] H. Wang, C. Ding, and H. Huang. Multi-label linear discriminant analysis. In *ECCV*, pages 126–139, 2010.

[12] H. Wang, H. Huang, and C. Ding. Image annotation using multi-label correlated green's function. In *ICCV*, 2009.

[13] J. Yang, R. Yan, and A. Hauptmann. Cross-domain video concept detection using adaptive svms. In *ACM MM*, pages 188–197. ACM, 2007.

[14] X. Zhu. Semi-supervised learning literature survey. *Technical report, University of Wisconsin-Madison*, 2006.

[15] F. Zhuang, P. Luo, H. Xiong, Q. He, Y. Xiong, and Z. Shi. Exploiting associations between word clusters and document classes for cross-domain text categorization. In *SDM*, 2010.