SOURCE CLASS SELECTION WITH LABEL PROPAGATION FOR PARTIAL DOMAIN ADAPTATION

Qian Wang

Department of Computer Science
Durham University, United Kingdom
qian.wang173@hotmail.com

Toby P. Breckon

Department of Computer Science
Department of Engineering
Durham University, United Kingdom

ABSTRACT

In traditional unsupervised domain adaptation problems, the target domain is assumed to share the same set of classes as the source domain. In practice, there exist situations where target-domain data are from only a subset of source-domain classes and it is not known which classes the target-domain data belong to since they are unlabeled. This problem has been formulated as Partial Domain Adaptation (PDA) in the literature and is a challenging task due to the negative transfer issue (i.e. source-domain data belonging to the irrelevant classes harm the domain adaptation). We address the PDA problem by detecting the outlier classes in the source domain progressively. As a result, the PDA is boiled down to an easier unsupervised domain adaptation problem which can be solved without the issue of negative transfer. Specifically, we employ the locality preserving projection to learn a latent common subspace in which a label propagation algorithm is used to label the target-domain data. The outlier classes can be detected if no target-domain data are labeled as these classes. We remove the detected outlier classes from the source domain and repeat the process for multiple iterations until convergence. Experimental results on commonly used datasets Office31 and Office-Home demonstrate our proposed method achieves state-of-the-art performance with an average accuracy of 98.1% and 75.4% respectively.

Index Terms— Partial Domain Adaptation, Label Propagation, Domain Adaptation, Subspace learning, Locality Preserving Projection

1. INTRODUCTION

Traditional supervised learning requires a large amount of labeled data for training and the training data are assumed to be drawn from the same distribution as those for testing. In many real-world applications, we may not have access to sufficient training data for the task of interest since annotating data is time-consuming and cost-intensive. One promising solution to this issue of training data scarcity is domain adaptation. It aims to take advantage of abundant labeled data from a different but related domain (i.e. source domain) and transfer knowledge learned from the source-domain data to the domain where the task of interest resides in (i.e. target domain). Due to the data distribution shift between the source and target domains, domain adaptation methods are needed for the knowledge transfer.

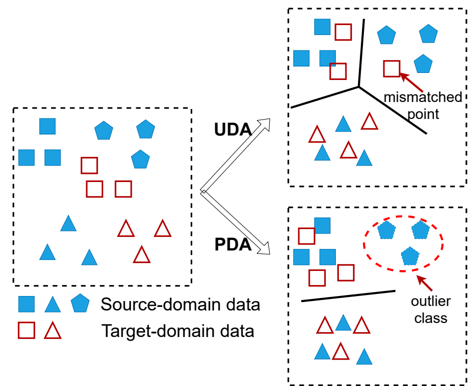

Fig. 1. The difference between PDA and UDA in terms of addressing the negative transfer issue. PDA finds out and removes the outlier classes in the source domain.

Among a variety of domain adaptation problems, Unsupervised Domain Adaptation (UDA) has been well studied in recent years [1, 2, 3]. UDA assumes the availability of labeled source-domain data and unlabeled target-domain data for training and the training data across two domains share the same label space (i.e. data are from the same set of classes). This is a strong assumption and restricts the usage of UDA from a practical perspective. A more realistic problem formulation is it is unknown which classes the unlabeled target-domain data belong to and we have labeled source-domain data from a large number of classes. The target-domain label space is a subset of the source-domain label space. This problem is well known as Partial Domain Adaptation (PDA) and has been studied in recent literature [4, 5, 6, 7, 8, 9].

The challenge of PDA is the *negative transfer* issue [4] which is caused by the label space mismatch in the source and target domains. As shown in Figure 1, traditional UDA approaches suffer from potentially disastrous misalignment

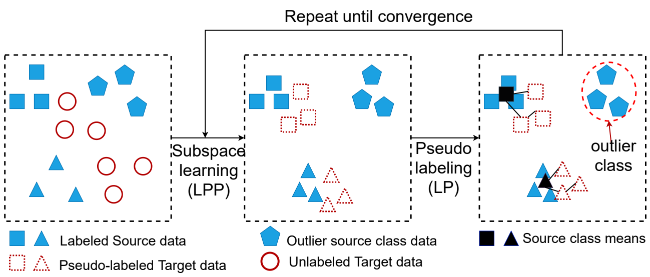

Fig. 2. The framework of our proposed approach SCS-LP (LPP: Locality Preserving Projection; LP: Label Propagation).

when they try to align the marginal data distributions of two domains. In PDA approaches, the negative transfer issue is addressed by down-weighting the contributions of all source data belonging to the outlier source label space in [4, 5]. The source class weights are computed based on the predicted probability of target-domain data belonging to each source class. In general, the target-domain data are unlikely to be predicted as outlier source classes hence these outlier classes will have lower weights and contribute less to the domain adaptation. Zhang et al. [6] extends the aforementioned idea by proposing the Importance Weighted Adversarial Nets (IWAN) which computes instance-level weights of contributing to the domain adaptation for the source-domain data. Example Transfer Network (ETN) [10] jointly learns domain-invariant representations across domains and a progressive weighting scheme to quantify the transferability of source examples.

The aforementioned methods share some common properties. On one hand, all of them aim to train a deep neural network with specially designed loss functions. There is no existing PDA approach based on deep feature transformation which has been proved superior to deep network learning in UDA problems [2]. On the other hand, they weight source-domain data on the instance-level and most outlier source-domain samples still contribute to the domain adaptation though with lower weights. In this work, we present a novel PDA approach based on deep feature transformation and source class selection (SCS). Specifically, our method is based on the state-of-the-art UDA approaches CAPLS [1] and SPL [2] and uses the Locality Preserving Projection (LPP) technique to learn a common subspace. In the subspace, we use label propagation (LP) to predict the labels of target-domain data. Based on the predicted labels of target-domain data, we can detect some outlier source classes. The source-domain data from the detected outlier source classes are subsequently removed from training data and we re-learn the common subspace with LPP using the updated training data (some outlier source class removed and pseudo-labeled target data included). The process is repeated until convergence. Our experiments demonstrate the process converges fast within 3 and 10 iterations for Office31 and Office-Home datasets respectively.

The contributions of this work can be summarized as follows: (1) we present a novel deep feature transformation based PDA approach using LPP and label propagation; (2) we demonstrate that the source class selection strategy is effective to address the negative transfer issue in PDA; (3) experimental results on two datasets show the superiority of our proposed approach (CSC-LP).

2. METHOD

Given a labelled dataset $\mathcal{D}^s = \{(\boldsymbol{x}_i^s, y_i^s)\}, i = 1, 2, ..., n_s$ from the source domain $\mathcal{S}$, $\boldsymbol{x}_i^s \in \mathbb{R}^d$ represents the feature vector of i-th labelled sample in the source domain, d is the feature dimension and $y_i^s \in \mathcal{Y}^s$ denotes the corresponding label. PDA aims to classify an unlabelled data set $\mathcal{D}^t = \{\boldsymbol{x}_i^t\}, i = 1, 2, ..., n_t$ from the target domain $\mathcal{T}$, where $\boldsymbol{x}_i^t \in \mathbb{R}^d$ represents the feature vector in the target domain. The target label space $\mathcal{Y}^t$ is a subset of the source label space $\mathcal{Y}^s$ and it is unknown which classes the target-domain data are from. It is assumed that both the labelled source domain data $\mathcal{D}^s$ and the unlabelled target domain data $\mathcal{D}^t$ are available for model learning.

The proposed method consists of two modules (common subspace learning and pseudo-labeling) within an iterative learning framework as shown in Figure 2. The first module aims to learn a projection matrix mapping data from both domains into a common subspace in which the second module predicts pseudo labels for all target-domain data. We employ the supervised LPP and label propagation algorithms for these two modules as enabling techniques respectively. Supervised LPP can learn a discriminative subspace by preserving the data structure in the original domains hence generalize better to out-of-distribution data than other dimensionality reduction techniques such as LDA [11, 2]. Label propagation can explore the target-domain data structure in the learned subspace in a self-supervised manner. The source classes with no target-domain data assigned to them are treated as outlier classes and removed from the training data set in the next iteration. Once some outlier classes are removed from the training data set, a better subspace can be learned by supervised LPP and label propagation results in more accurate pseudo labels for target-domain data. The iterative learning stops when no more outlier source classes are detected.

2.1. Common Subspace Learning

The goal of subspace learning is to find a lower-dimensional subspace in which the projected data from both domains are well aligned. To promote the class-wise alignment of two domains, we use the supervised locality preserving projection [12] as an enabling technique to learn a domain invariant yet discriminative subspace $\mathcal{Z}$ from $\tilde{\mathcal{X}}$.

The objective of SLPP is to learn a projection matrix $\mathbf{P}$ by minimizing the following cost function:

$$\min_{\mathbf{P}} \sum_{i,j} ||\mathbf{P}^T \boldsymbol{x_i} - \mathbf{P}^T \boldsymbol{x_j}||_2^2 \mathbf{M}_{ij}, \tag{1}$$

where $\mathbf{P} \in \mathbb{R}^{d \times d_1}$ and $d_1 \leq d$ is the dimensionality of the learned subspace; $\boldsymbol{x}_i$ is the i-th column of the labeled data matrix $\mathbf{X}^l \in \mathbb{R}^{d \times (n'_s + n_t)}$ and $\mathbf{X}^l$ is a collection of selected n'_s labeled source data and n_t pseudo-labeled target data. The similarity matrix $\mathbf{M} \in \mathbb{R}^{(n'_s+n_t) \times (n'_s+n_t)}$ is defined as follows:

$$\mathbf{M}_{ij} = \begin{cases} 1, & y_i = y_j, \\ 0, & otherwise. \end{cases} \tag{2}$$

The idea is that samples from the same class should be projected close to each other in the subspace regardless of which domain they are originally from. Following the treatment in [12, 2], the problem defined in Eq.(1) is equivalent to the following generalized eigenvalue problem:

$$\mathbf{X}^l \mathbf{D} \mathbf{X}^{lT} \boldsymbol{p} = \lambda (\mathbf{X}^l \mathbf{L} \mathbf{X}^{lT} + \mathbf{I}) \boldsymbol{p}, \tag{3}$$

where $\mathbf{L} = \mathbf{D} - \mathbf{M}$ is the laplacian matrix, $\mathbf{D}$ is a diagonal matrix with $\mathbf{D}_{ii} = \sum_j \mathbf{M}_{ij}$. Solving the generalized eigenvalue problem gives the optimal solution $\mathbf{P} = [\boldsymbol{p}_1, ..., \boldsymbol{p}_{d_1}]$ where $\boldsymbol{p}_1, ..., \boldsymbol{p}_{d_1}$ are the eigenvectors corresponding to the largest d_1 eigenvalues.

Learning the projection matrix $\mathbf{P}$ for domain alignment requires labeled samples from both source and target domains. To get pseudo-labels of target samples for projection learning, we describe pseudo-labeling methods via label propagation in the following sub-section.

2.2. Pseudo-Labeling and Outlier Class Detection

In the learned subspace, we follow [11, 2] and process the projections $\boldsymbol{z} = \boldsymbol{P}^T \boldsymbol{x}$ of data from both domains by mean subtraction and l_2 normalization. Subsequently, we compute the class means for all reserved source classes by:

$$\boldsymbol{m}_c = \frac{1}{n_c} \sum_{y_i^s = c} \boldsymbol{z}_i^s \tag{4}$$

where $c \in \mathcal{Y}^s \setminus \mathcal{Y}^{outlier}$, n_c is the number of source-domain samples from class c. The label propagation algorithm [13] is applied to the combination of all target-domain data projections $\boldsymbol{Z}^t \in \mathbb{R}^{d_1 \times n_t}$ and all such reserved source class means (i.e. the set $\{\boldsymbol{m}_c\}$).

The source class means are labeled and the target-domain data are unlabeled. A similarity graph over data $\boldsymbol{Z}^t \cup \{\boldsymbol{m}_c\}$ is constructed using K nearest neighbours in the subspace. Suppose there are $n = n_t + n_c$ nodes in the similarity graph, where $n_c = |\mathcal{Y}^s \setminus \mathcal{Y}^{outlier}|$ is the number of reserved source classes, and the similarity matrix is denoted as $\boldsymbol{S} \in \mathbb{R}^{n \times n}$, the labeling information is propagated from labeled nodes to unlabeled nodes by the following three steps.

Algorithm 1 Patial Domain Adaptation Using SCS-LP

Input: Labeled source data set $\mathcal{D}^s = \{(\boldsymbol{x}_i^s, y_i^s)\}, i = 1, 2, ..., n_s$ and unlabeled target data set $\mathcal{D}^t = \{\boldsymbol{x}_i^t\}, i = 1, 2, ..., n_t$, dimensionality of LPP subspace d_1, number of nearest neighbors K for label propagation.

Output: The predicted labels $\{\hat{y}^t\}$ for target samples.

Initialize $iter = 0$;
Learn the projection $\mathbf{P}_0$ using only source data $\mathcal{D}^s$;
Assign pseudo labels for all target data using label propagation;
while not converge **do**
  $iter \leftarrow iter + 1$;
  Detect the outlier source classes and update the training data set $\mathcal{S}_{iter} \cup \hat{\mathcal{D}}^t$ by removing the source data belonging to the outlier classes;
  Learn P_{iter} using updated training data;
  Update pseudo labels for all target data using label propagation.
end while

Firstly, we initialize a matrix $\boldsymbol{F}_0 \in \mathbb{R}^{n \times n_c}$. The first n_c rows correspond to the labeled source class means. Each row contains a 1 in the column corresponding to the true class label, and a 0 in every other column. The last n_t rows correspond to the unlabeled target-domain data, and contain a 0 in all columns. Secondly, at iteration t (starting with $t = 1$), we update the matrix $\boldsymbol{F}_t = \boldsymbol{P} \boldsymbol{F}_{t-1}$, where $\boldsymbol{P}_{ij} = \frac{\boldsymbol{S}_{ij}}{\sum_{k=1}^n \boldsymbol{S}_{ij}}$ is the probability of propagating label information from node i to node j. Note that the first n_c rows of $\boldsymbol{F}_t$ need to keep the same as those of $\boldsymbol{F}_0$ during the update. Finally, we repeat the second step until the $\boldsymbol{F}$ values converge.

After label propagation, we can obtain the pseudo labels of target-domain data from the matrix $\boldsymbol{F}$. We count the number of pseudo labels for each class and detect the outlier source classes if there is no data assigned to these classes. The source data belonging to the detected outlier source classes are removed in the next iteration of subspace learning and pseudo-labeling. The complete algorithm of SCS-LP is summarized in Algorithm 1.

3. EXPERIMENTS AND RESULTS

We conduct experiments on two commonly used datasets to evaluate the effectiveness of the proposed approach for PDA and present the experimental results in this section.

3.1. Dataset

Office-Home [16] consists of four different domains: Artistic images (A), Clipart (C), Product images (P) and Real-World images (R). There are 65 object classes in each domain with a total number of 15,588 images. We follow [4] and use images

Table 1. Classification Accuracy (%) on Office-Home dataset using either ResNet50 features or ResNet50 based deep models.

Method	A→C	A→P	A→R	C→A	C→P	C→R	P→A	P→C	P→R	R→A	R→C	R→P	Average
PADA [4]	52.0	67.0	78.7	52.2	53.8	59.0	52.6	43.2	78.8	73.7	56.6	77.1	62.1
IWAN [6]	53.9	54.5	78.1	61.3	48.0	63.3	54.2	52.0	81.3	76.5	56.8	82.9	63.6
SAN [5]	44.4	68.7	74.6	67.5	65.0	77.8	59.8	44.7	80.1	72.2	50.2	78.7	65.3
DRCN (soft) [9]	54.0	76.4	83.0	62.1	64.5	71.0	70.8	49.8	80.5	77.5	59.1	79.9	69.0
ETN [10]	59.2	77.0	79.5	62.9	65.7	75.0	68.3	55.4	84.4	75.7	57.7	84.5	70.5
SAFN [14]	58.9	76.3	81.4	70.4	73.0	77.8	72.4	55.3	80.4	75.8	60.4	79.9	71.8
RTNet$_{adv}$[7]	63.2	80.1	80.7	66.7	69.3	77.2	71.6	53.9	84.6	77.4	57.9	85.5	72.3
MCC [15]	63.1	80.8	86.0	**70.8**	72.1	80.1	**75.0**	**60.8**	**85.9**	78.6	**65.2**	82.8	75.1
w/o LPP	49.9	72.2	78.3	64.6	69.3	72.5	60.2	49.7	75.8	74.3	54.7	77.9	66.6
w/o SCS	57.9	80.7	88.3	69.0	79.6	**84.4**	69.6	51.2	82.8	**79.6**	57.1	**87.4**	74.0
SCS-1NN	51.9	72.7	80.9	63.1	68.3	77.6	66.2	50.1	78.2	73.6	55.9	80.7	68.3
SCS-LP (Ours)	**65.0**	**81.2**	**90.0**	70.0	**81.7**	81.7	70.2	55.4	82.8	79.2	60.3	**87.4**	**75.4**

Table 2. Classification Accuracy (%) on Office31 dataset using either ResNet50 features or ResNet50 based deep models.

Method	A→W	D→W	W→D	A→D	D→A	W→A	Avg
PADA [4]	86.5	99.3	100.0	82.2	92.7	**95.4**	92.7
IWAN [6]	89.2	99.3	99.4	90.5	95.6	94.3	94.7
SAN [5]	93.9	99.3	99.4	94.3	94.2	88.7	95.0
DRCN (hard) [9]	90.8	100.0	100.0	94.3	95.2	94.8	95.9
ETN [10]	94.5	100.0	100.0	95.0	**96.2**	94.6	96.7
RTNet$_{adv}$[7]	96.2	100.0	100.0	97.6	92.3	**95.4**	96.9
w/o LPP	98.3	97.3	98.7	98.1	93.9	85.3	95.3
w/o SCS	96.9	97.3	98.7	97.5	93.5	89.4	95.5
SCS-1NN	82.7	98.0	98.7	87.9	91.6	90.0	91.5
SCS-LP (Ours)	**99.0**	**100.0**	**100.0**	**100.0**	94.3	**95.4**	**98.1**

from the first 25 classes in alphabetical order as the target domain and images from all 65 classes as the source domain. Image features are extracted by the ResNet50 [17] model pre-trained on ImageNet [18] without fine-tuning. **Office31** [19] consists of three domains: Amazon (A), Webcam (W) and DSLR (D). There are 31 common classes for all three domains containing 4,110 images in total. Following the protocol in [4], all 31 classes are used for the source domain and the fixed 10 classes are used for the target domain. Image features are extracted by the ResNet50 [17] model pre-trained on ImageNet [18] without fine-tuning.

3.2. Implementation Details

The proposed approach is implemented in Matlab R2020b [1]. We set the dimensionality of learned subspace d_1 to 128 for both datasets. For Office-Home dataset, we also apply PCA to reduce the dimensionality of ResNet50 features from 2048 to 512 in pre-processing to speed up the computation. The values of K for similarity graph construction in label propagation are set to 15 and 10 for Office-Home and Office31 datasets respectively.

3.3. Experimental Results

We compare our proposed method with contemporary state-of-the-art methods for PDA. Specifically, we compare with methods designed for PDA problems such as PADA [4], IWAN [6], SAN [5], ETN [10], SAFN [14], RTNet$_{adv}$ [7], DRCN [9] and MCC [15].

The comparison results are shown in Tables 1 and 2. On the Office-Home dataset, our proposed method performs the best with the average accuracy of 75.4%. On the Office31 dataset, our method outperforms all the other methods on 5 out of 6 tasks and achieves the highest average accuracy of 98.1%. In particular, our method achieves 100% accuracy on three tasks.

We include an ablation study to investigate the contributions of different components to superior performance in Tables 1 and 2. Specifically, we investigate three variants of the proposed method: *without LPP* (i.e. the label propagation is done in the original feature space rather than the learned subspace), *without SCS* (i.e. the outlier source classes are not removed from the training data set) and *SCS-1NN* (i.e. the label propagation algorithm is replaced by 1 nearest neighbour for pseudo-labeling). The results shown in Tables 1 and 2 demonstrate: (1) both the subspace learning and label propagation are essentials for good performance of the proposed method; (2) source class selection can further improve the performance.

Our proposed is computational efficient and the iterative learning converges within less than 3 and 10 iterations for Office31 and Office-Home datasets. One limitation of the outlier source class detection scheme is that some source classes may be removed by mistake and all the target-domain data belonging to these classes will be misclassified. Such mistakes have been observed in our experiments on the Office-Home dataset but source class selection is generally beneficial to the overal performance as demonstrated in the ablation study.

4. CONCLUSION

We propose a novel deep feature transformation based approach to PDA and achieve state-of-the-art performance on establish benchmark datasets within this domain. Both the supervised LPP based subspace learning and label propagation based pseudo-labeling are essentials for the superior performance as illustrated by our supporting ablation study.

[1] https://github.com/hellowangqian/scs-lp-pda

5. REFERENCES

[1] Qian Wang, Penghui Bu, and Toby P. Breckon, "Unifying unsupervised domain adaptation and zero-shot visual recognition.," in *International Joint Conference on Neural Networks*, 2019.

[2] Qian Wang and Toby P Breckon, "Unsupervised domain adaptation via structured prediction based selective pseudo-labeling," in *AAAI Conference on Artificial Intelligence*, 2020.

[3] Qian Wang, Fanlin Meng, and Toby P Breckon, "Data augmentation with norm-vae for unsupervised domain adaptation," *arXiv preprint arXiv:2012.00848*, 2020.

[4] Zhangjie Cao, Lijia Ma, Mingsheng Long, and Jianmin Wang, "Partial adversarial domain adaptation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 135–150.

[5] Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Michael I Jordan, "Partial transfer learning with selective adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 2724–2732.

[6] Jing Zhang, Zewei Ding, Wanqing Li, and Philip Ogunbona, "Importance weighted adversarial nets for partial domain adaptation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8156–8164.

[7] Zhihong Chen, Chao Chen, Zhaowei Cheng, Boyuan Jiang, Ke Fang, and Xinyu Jin, "Selective transfer with reinforced transfer network for partial domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 12706–12714.

[8] Fariba Zohrizadeh, Mohsen Kheirandishfard, and Farhad Kamangar, "Class subset selection for partial domain adaptation.," in *CVPR Workshops*, 2019.

[9] Shuang Li, Chi Harold Liu, Qiuxia Lin, Qi Wen, Limin Su, Gao Huang, and Zhengming Ding, "Deep residual correction network for partial domain adaptation," *IEEE transactions on pattern analysis and machine intelligence*, 2020.

[10] Zhangjie Cao, Kaichao You, Mingsheng Long, Jianmin Wang, and Qiang Yang, "Learning to transfer examples for partial domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2985–2994.

[11] Qian Wang and Ke Chen, "Zero-shot visual recognition via bidirectional latent embedding," *International Journal of Computer Vision*, vol. 124, no. 3, pp. 356–383, 2017.

[12] Xiaofei He and Partha Niyogi, "Locality preserving projections," in *Advances in neural information processing systems*, 2004, pp. 153–160.

[13] Xiaojin Zhu and Zoubin Ghahramani, "Learning from labeled and unlabeled data with label propagation," *CMU CALD tech report CMU-CALD-02-107*.

[14] Ruijia Xu, Guanbin Li, Jihan Yang, and Liang Lin, "Larger norm more transferable: An adaptive feature norm approach for unsupervised domain adaptation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1426–1435.

[15] Ying Jin, Ximei Wang, Mingsheng Long, and Jianmin Wang, "Minimum class confusion for versatile domain adaptation," in *European Conference on Computer Vision*. Springer, 2020, pp. 464–480.

[16] Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, and Sethuraman Panchanathan, "Deep hashing network for unsupervised domain adaptation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5018–5027.

[17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[18] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.

[19] Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell, "Adapting visual category models to new domains," in *European Conference on Computer Vision*. Springer, 2010, pp. 213–226.