A Unified Framework for Distance-Aware Domain Adaptation

Fei Wang, Youdong Ding, Huan Liang, Yuzhen Gao and WenqiChe
Department of Film and Television Engineering
Shanghai Engineering Research Center of Motion Picture Special Effects
Shanghai University, Shanghai 200072, P.R. China
Email: {s852989587; ydding; Lianghuan; gyz18723239; chewenqi}@ shu.edu.cn

Introduction

In this paper, we propose a novel domain adaptation method, which is called distance-aware domain adaptation (DADA), to address the deficiencies in existing works. The research objective of this study is to improve the performance of domain adaptation using the following two approaches. First, DADA aims to fuse domain-level and class-level discriminative distance. In domain-level discriminative distance, we constrain the same class samples from different domains to move closer while samples from different classes to move further away during training. In class-level distance, we cluster the samples in both domains. Therefore, the discriminative power of source domain can pass on to the target domain. Second, DADA incorporates the statistics distance and geometry structure into a unified framework.

Proposed Method

Cross-Domain Distance

$$\min \left\| \frac{1}{n_s}\sum_{i=1}^{n_s} A^T x_i - \frac{1}{n_t}\sum_{j=1}^{n_t} A^T x_j \right\|^2$$

$$\min \sum_{c=1}^{C} \left\| \frac{1}{n_s^{(c)}} \sum_{x_i\in X_s^{(c)}} A^T x_i - \frac{1}{n_t^{(c)}}\sum_{x_j \in X_t^{(c)}} A^T x_j\right\|^2$$

JDA [Long et al, ICCV'13]

$$\max \sum_{c=1}^{C} \left\| \frac{1}{n_s^{(c)}} \sum_{x_i\in D_s^{(c)}} A^T x_i - \frac{1}{\sum_{r\in\{\{1...C\}-\{c\}\}} n_t^{(r)}} \sum_{x_j\in D_t^{(r)}} A^T x_j \right\|^2$$

$$\max \sum_{c=1}^{C} \left\| \frac{1}{n_t^{(c)}} \sum_{x_i\in D_t^{(c)}} A^T x_i - \frac{1}{\sum_{r\in\{\{1...C\}-\{c\}\}} n_s^{(r)}} \sum_{x_j\in D_s^{(r)}} A^T x_j \right\|^2$$

repulsive force distance

Intra-Class Distance

$$\min \sum_{c=1}^{C}\sum_{i=1}^{n_s^c} \left\| A^T y_s^{ci} \boxed{\left(x_s^i - \mu_{s,c}\right)} \right\|^2$$

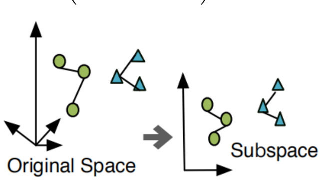
$$\min \sum_{c=1}^{C}\sum_{i=1}^{n_t^c} \left\| A^T y_t^{ci} \boxed{\left(x_t^i - \mu_{t,c}\right)} \right\|^2$$

increase the compactness of both domains

Statistics & Geometry Distance

$$\min \mathrm{tr}\left(A^T X_s H_s X_s^T A - A^T X_t H_t X_t^T A\right)$$

$$\min \mathrm{tr}\left(A^T X L X^T A\right)$$



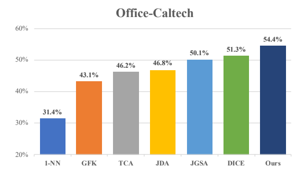
Laplacian Regularization

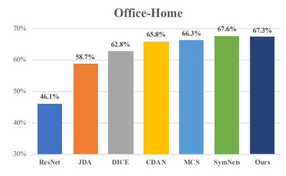
Overall Objective Function

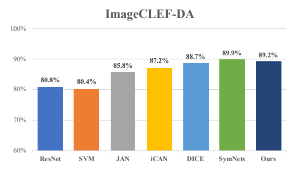
$$\min \mathrm{tr}\left(A^T X\left(M_a + L + \alpha Q\right) X^T A\right) + \lambda \| A \|_F^2$$

$$\text{s.t. } A^T\left(XHX^T + \beta S_b\right) A = I$$

Experiments





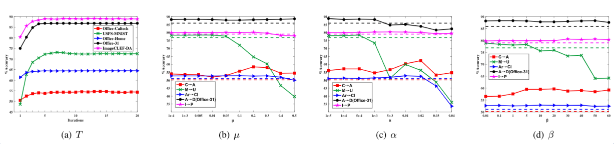


We conduct massive experiments on several famous datasets with shallow and deep features, including Office-Caltech, ImageCLEF-DA and Office-Home.

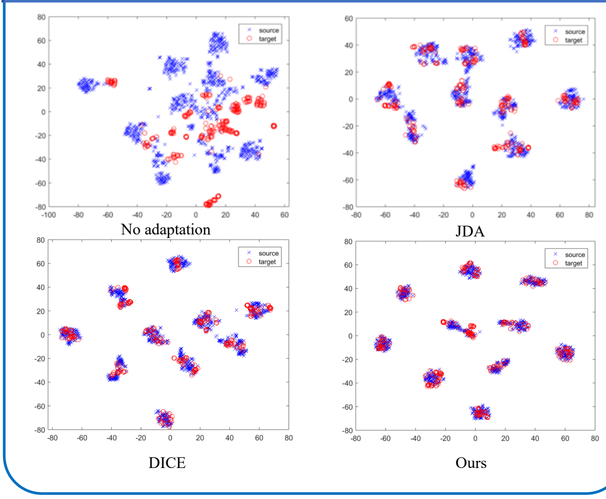
As figures show, Our method outperform all traditional methods in three datasets and can perform against deep methods. Although deep methods have better results,

Our method only contains several parameters that can easily be set by human experience or cross-validation. It indicates the excellent classifier performance and efficiency of our method.

Parameter Sensitivity



t-SNE Visualization



Acknowledgment

National Natural Science Foundation of China (No. 61303093, 61402278)