Supplementary Material of Ad^{2}mix: Adversarial and Adaptive Mixup for Unsupervised Domain Adaptation

Lei Zhu [1] Yanyu Xu [2] Yong Liu [1] Rick Siow Mong Goh [1] Xinxing Xu [3]
[1]Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, #16-16 Connexis, Singapore 138632, Republic of Singapore
[2]The Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR), Shandong University, Shandong, China
[3]Microsoft Research Asia, Singapore

1. More Details on Cataract Dataset

Cataract consists of 1,428 eye images of 2 classes collected with different devices: Slit-Lamp (S) and Camera (C). The Slit-Lamp dataset is a private dataset of eye images captured with specialized slit-lamp device in a local hospital for cataract detection. The Camera dataset[1] is a public Kaggle competition dataset of eye images captured with camera device. Table 1 shows the detailed data statistics for the two datasets. We construct the Cataract dataset to study domain adaptive cataract detection from well-captured eye images with specialized slit-lamp device in a clinical environment to eye images captured with camera device in a less restricted environment. In Fig. 1, we display representative images in each dataset. Note, the displayed slit-lamp images are public images downloaded online but they are similar to the images in our private dataset. We do not display images from our Slit-Lamp dataset due to privacy issues. As can be observed, due to the different devices and procedures in capturing eye images, the two datasets display quite different appearance, resulting in large domain gap across domains.

Table 1. Data Statistics of Slit-Lamp and Camera Datasets

Dataset	Normal	Cataract	Total
Slit-Lamp	408	408	816
Camera	306	306	612
Total	714	714	1428

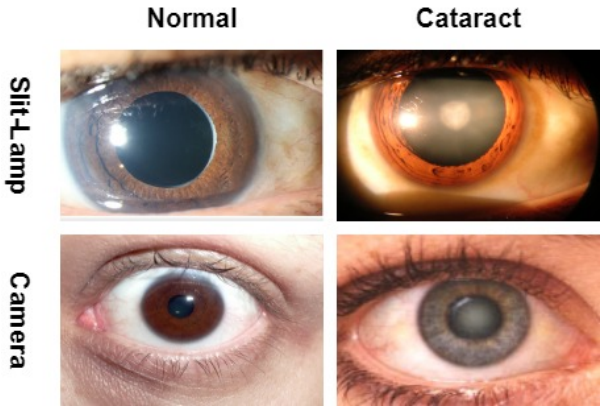

Figure 1. Comparison of images from Slit-Lamp and Camera datasets.

2. Visualization Analysis

Fig. 2 presents the T-SNE visualization results of the Swin-Base baseline method, the PMTrans method, and our proposed Ad^{2}mix method on pnt→info and info→pnt transfer tasks on DomainNet. As can be observed, at the center of the embedding spaces of both Swin-Base and PMTrans methods, there are large amount of target data features which cluster together and do not align well with existing source data class clusters. Our proposed Ad^{2}mix method better aligns the target data points towards the source data class clusters and better matches the data distributions across domains.

3. Sensitivity Analysis

Fig. 3 presents the post experiment sensitivity analysis on hyper-parameter γ_{mix}, γ_{adv}, and γ_{unsup} on transfer task Art→Clipart on Office-Home. As can be observed, our method is generally robust to hyper-parameter changes in a wide range.

[1]https://www.kaggle.com/datasets/nandanp6/cataract-image-dataset

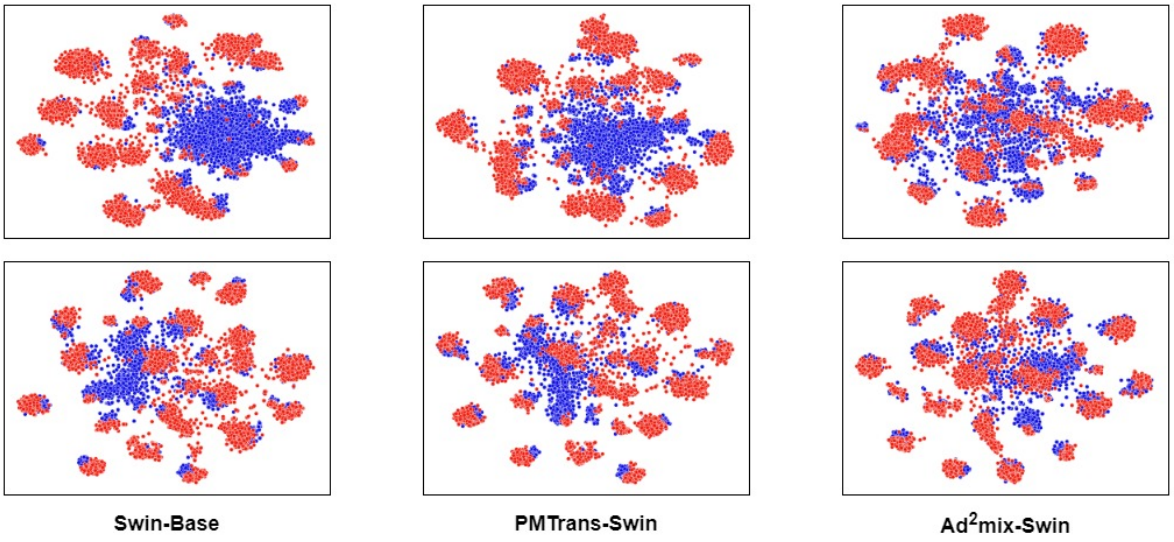

Figure 2. T-SNE visualization of pnt→info (top row) and info→pnt (bottom row) transfer tasks with twenty-five classes on DomainNet. Source and target instances are shown in red and blue colors respectively.

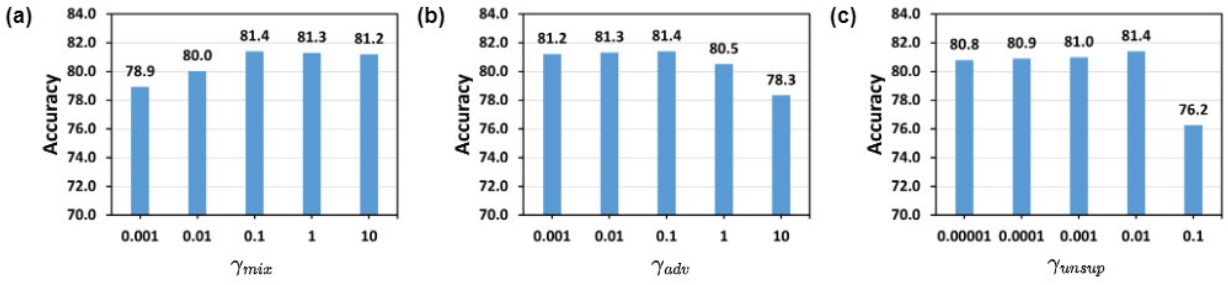

Figure 3. Sensitivity analysis on (a) γ_{mix}, (b) γ_{adv}, and (c) γ_{unsup} for Art→Clipart transfer task on Office-Home.