Supplementary Material
CAFT: Class Aware Frequency Transform for Domain Gap Reduction

This document provides additional qualitative results of transformed source samples and t-SNE plots. We organize the main contents of this supplementary material in following sections.

1. Latent Space Visualization with t-SNE (Section 1)
2. Class aware frequency transformed source images (Section 2)
3. Qualitative Analysis of Source transformed samples (Section 3)

1. Latent Space Visualization with t-SNE

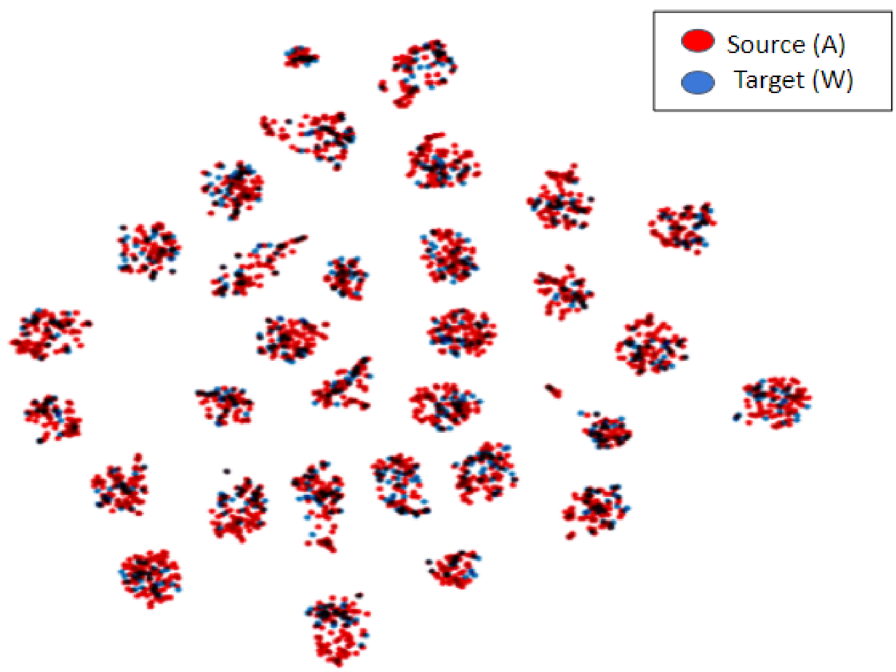

Figure 1. t-SNE Plot for **Amazon** to **Webcam** adaptation with CAFT

Using t-SNE, We analyse the discriminative feature alignment for source and target sample features taken from adapted model feature extractor. We experiment with Amazon to Webcam source-target pair of Office-31 dataset on CAFT with RevGrad adaptation method. We can observe the alignment of discriminative features in Fig 1, which looks separable in projected space. A better clustering of features in the latent space implies well separation of class-discriminative features. It should result in better model performance.

2. Class Aware Frequency Transformed Source Images

In this section, we provide qualitative results using class aware Fourier transformation. Fig. 2 shows the source images taken from Amazon domain of Office-31 dataset. Fig. 3 shows target images taken from Webcam domain and Fig. 4 shows transformed source samples using CAFT. We can observe that transformed Amazon source looks closer to the target compared to the original Amazon source.

Figure 2. Source (***Amazon***) images

Figure 3. Target (***Webcam***) images

Figure 4. Source (***Transformed Amazon***) images

3. Qualitative Analysis of Source Transformed Samples

In this section, we show the transformed source samples by varying the percentage of low frequency component of the source to be transferred. Fig. 5 demonstrate the qualitative results for source transformation samples. We observe that if we keep increasing the percentage of low frequency component to be swapped, transformed source sample starts loosing the class-discriminative features, which will reduce the classification performance.

Figure 5. **Frequency Transformation Effect for different window sizes**. The top row represents the source domain images taken from Amazon dataset. The second row represents the target domain image taken from DSLR dataset. Each row from row 3 onwards represent transformed source image with increasing window size from 1% up to 10% with step size of 1%. As evident, for very low frequency window size, transformation is very weak, while for very high value (near 10%) undesirable artifacts start to appear in the transformed image. Hence we need to find and intermediate window size for best results. We found 4% window size gives best results for Office-31