<u>Multi Source Domain Adaptation</u>

The goal of this project was to create a deep learning architecture that can learn from multiple different but similar sources of input to make classifications on an unlabeled domain in a semi supervised manner. Conventional unsupervised learning setting assumes that our source and target domains are generally from the same distribution. In a real world scenario, this is generally not true. For this project, we assume that we have the data from multiple domains at training time, although not all of them are labelled. This unlabelled data is what we called the target domain. The idea to minimize the domain gap between the source and transfer domains and match the distributions of the domains. We decided to train our model on the VISDA dataset, which was the secondary contribution of this paper. VISDA is a dataset with 0.6 million images which was specifically curated for the task of domain adaptation.

We begin by recreating the approach in the “Moment Matching for Multi-Source Domain Adaptation”. The model in this paper was trained on a digit dataset from different domains. In this case the domains were different datasets. We trained our model on the VISDA dataset as it was specifically created or the purpose of Domain Adaptation. This offer is a challenge in terms of the preprocessing required of the images and the number of features to be learnt. Since the images in the VISDA dataset are real world objects as compared to digits, it is much harder for the model to learn the feature representation. The increase in number of classes also required us to change certain hyperparameters of the model, although the underlying architecture remained the same. The Visda Dataset contains 0.6 million images from 6 distinct domains which are Real, sketch, quickdraw, painting, infographic and clipart. At any point we could have a maximum of 5 source domains and 1 target domain. The motivation between the learning methodology we use was to not only focus on aligning the source domains with the target domain, but also to align the source domains among themself. Another contribution we made in extending the work in this paper was to change the distance metric used to determine the similarity between domains. We used the original Euclidean distance as baselines. We further experimented using the Mahalanobis distance metric and dynamic partial distance function. Euclidean distance or any Lk norm suffers with large dimensions. In high dimensions, the ratio between the nearest and farthest points approaches 1, i.e. the points essentially become uniformly distant from each other. The premise is to find the features that are closer, but as the features become uniformly distant, the distance metric loses its values. One approach we use is the Dynamic Partial Distance metric wherein we select a value ‘p’ such that only the ‘p’ closest features are used to determine the distance. We chose ‘p’ features such that these features are not uniformly distant. The other metric we used is a widely accepted distance metric used for higher dimension spaces; the Mahalanobis distance.

This project led to an architecture that can work in a real world setting for image classification in a semi supervised manner. The lack of training labels made it harder to train but it also explains the need of more research methods in unsupervised learning as the cost of annotation is very high. Helped me develop the skills required to review literature and to be able to recreate an existing work, identify rooms for improvement and provide extensions to existing work.