Domain Adaptation Using Small Group Learning

Aadil Ahamed
UC San Diego
aahamed@ucsd.edu

John O'Boyle
UC San Diego
jboboyle@ucsd.edu

Alex Tung
UC San Diego
aktung@ucsd.edu

Abstract

Domain adaptation is an unsupervised deep learning technique aimed to extract a meaningful representation of a target domain $\mathcal{D}^T$ using a model previously trained on a source domain $\mathcal{D}^S$. The method relies on the assumption that features common to both $\mathcal{D}^S$ and $\mathcal{D}^T$ have invariant properties. The existence of such features allows for disentanglement of the underlying factors of variation across the source and target domains. Different approaches to domain adaption have been implemented with varying degrees of success. In the current work, we will explore the small group learning approach in domain adaptation and evaluate its performance using metrics derived from existing domain adaptation methods.

1 What research problem are you aiming to solve?

Transfer learning broadly describes the application of a trained neural network to another dataset. This technique can be classified by whether the adaptation is performed over the learning domain (i.e. feature space and distribution) or the learning task (i.e. label space, objective function)[1].

We are interested in exploring a sub-problem within transfer learning known as domain adaptation. Domain adaptation transfers a model trained to a set task $\mathcal{T}$ on a source domain $\mathcal{D}^S$ to a target domain $\mathcal{D}^T$ with the same feature space but a different marginal distribution of features [2].

2 Why is this problem important?

Domain adaptation is important as source and target distributions can differ in practice even when performing a conventional learning task with a constant dataset. Though network performance can be predicted by portioning training data into validation and test groups, the network may still suffer if deployed in an environment that diverges from the entire set. Thus, optimization of a network using domain adaptation methods can improve its robustness and external validity. With regard to supervised learning research in new fields, it can be expensive to label large amounts of data from the target domain if there are no extant datasets to train from – a problem that domain adaptation can also address.

Small group learning (SGL hereafter) is a newly proposed search method in which multiple networks, termed learners, are first independently trained on a labeled dataset, then cross-trained with datasets with pseudo-labels from other learners, and finally optimized on a validation set [3]. Following the terminology given above, this procedure is considered a semi-supervised and transductive form of domain adaptation because the intermediate step involves labels generated locally by each learner.

Given the nature of domain adaptation in exploring datasets that a given network has not been exposed to, any expansion in the understanding of SGL and its implementation would develop the relatively young field. In addition, the original authors implemented this topology in the context of a DARTS architecture search [4], but it would be interesting to explore different weight initializations to examine whether individual learners arriving at local loss minima can further converge towards a

Preprint. Under review.

global minimum when cross-trained in a small group. This is of particular interest when optimizing non-convex problems.

3 Related Work

Domain adaptation was originally applied to the problem of sentiment classification across different text reviews [5], but the principle has since expanded to multiple approaches. In their survey, Wilson et al. broadly classify these approaches into the following categories: domain invariant feature learning (DIFL hereafter), domain mapping, normalization statistics and ensemble methods [1]. We will focus primarily on DIFL as it relates closely to the discussions in SGL structure. DIFL attempts to extract features from the input that generalize well to both source and target domains [1]. The following sections summarize the three main feature extraction methods used in DIFL.

The divergence method aims to minimize the distance between source and target distributions based on some metric. Common metrics used in this approach are maximum mean discrepancy, correlation alignment, contrastive domain discrepancy and the Wasserstein metric [1] [6].

The reconstruction method attempts to build a feature representation that can both classify labels against data from the source domain and reconstruct the original input from the source and target domains. There have been broad efforts made to adapt autoencoders to different domains altogether without necessarily keeping the feature spaces constant [7].

The adversarial learning method consists of a feature extractor and domain classifier acting as adversaries, as the name suggests. The domain classifier tries to classify whether an input sample originated from the source or target domain and the feature extractor tries to extract domain invariant features that make it difficult for the domain classifier to classify correctly [8].

4 Proposal

We propose to try applying the SGL learning framework to the problem of domain adaptation. Specifically, this learning framework can be used in combination with some existing models that have worked well for domain adaptation. For example, we could create a two-learner system, where the first learner uses a model based on divergence and the second learner uses a model based on reconstruction. Alternatively, both learners can use divergence but implement different divergence metrics.

5 Novelty

Our proposal will be the first attempt to tackle the problem of domain adaptation using SGL. The original authors have shown that SGL can perform better than state of the art models in tasks such as image classification in conjunction with architecture search and it is plausible that it can provide strong results for different domains as well. Depending on the progress of research, it may be interesting to evaluate the performance of SGL in a strictly unsupervised transductive sense by using the network to extract features of other image-like inputs (e.g. spectrograms, GIS data).

6 Tentative Timeline

Month	Jan				Feb				March		
Academic Week	1	2	3	4	5	6	7	8	9	10	11
Align project objectives and scope											
Research existing DA and SGL methods											
Select domain/dataset											
Implement DA algorithm in Python											
Integrate SGL methods into DA model											
Compare results with other DA Methods											
Test on various domains											
Report Findings											

References

[1] Wilson, G., D. J. Cook. A survey of unsupervised deep domain adaptation, 2020.

[2] Wang, M., W. Deng. Deep visual domain adaptation: A survey, 2018.

[3] Du, X., P. Xie. Small-group learning, with application to neural architecture search, 2020.

[4] Liu, H., K. Simonyan, Y. Yang. Darts: Differentiable architecture search, 2019.

[5] Glorot, X., A. Bordes, Y. Bengio. Domain adaptation for large-scale sentiment classification: A deep learning approach. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, page 513–520. Omnipress, Madison, WI, USA, 2011.

[6] Olkin, I., F. Pukelsheim. The distance between two random vectors with given dispersion matrices. *Linear Algebra and its Applications*, 48:257 – 263, 1982.

[7] Ghifary, M., W. B. Kleijn, M. Zhang, et al. Domain generalization for object recognition with multi-task autoencoders. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 2551–2559. 2015.

[8] Goodfellow, I. J., J. Pouget-Abadie, M. Mirza, et al. Generative adversarial networks, 2014.