論 文 の 内 容 の 要 旨

論文題目　Domain Adaptation Algorithms with Scarce Data
（不十分なデータに基づくドメイン適応手法）

氏　　名　石井　雅人

Machine learning has become one of the most important technologies to provide intelligent products and systems that can enrich people's lives. The latest machine learning algorithms have achieved human-level performance in many applications, for example, face recognition, speech recognition, and multi-lingual translation. One of the most important factors of such success is the availability of a large-scale training dataset. While it is absolutely essential, obtaining such a large-scale dataset requires an expensive cost, which will be a major barrier to apply machine learning technologies to a wide variety of applications.

In this dissertation, we tackle several types of problem settings of domain adaptation with scarce data. The domain adaptation is a technology to reuse existing datasets, and it can substantially reduce the cost of obtaining the training dataset for a new domain. It basically aims to match the distribution of source data (the data we want to reuse) to that of target data (the data on which the trained model should perform well). By properly conducting the domain adaptation, the model trained with the adapted source data can perform well in the target domain. For effective adaptation, we need a sufficient amount of source

and target data. However, due to the motivation of the domain adaptation, the amount of target data is often supposed to be small. Additionally, we cannot access source data in some practical cases, for example, due to data privacy issues. Therefore, how to effectively handle such data scarceness is quite important to make it possible to apply this technology to real-world problems. In this dissertation, we present three contributions to domain adaptation with scarce data.

In the first part of this dissertation, we consider the case in which we only have incomplete target data. In this setting, a certain subset of classes is missing in unlabeled target data, while all classes appear in labeled source data, and the goal is to discriminate all classes in the target domain. We call this problem setting partially zero-shot domain adaptation. To solve this problem, we utilize an adversarial training scheme and adopt instance weighting to estimate the loss related to unavailable target data in the missing classes. The instance weight is computed based on the prediction of deep neural networks, implying which instance would be similar to unseen data and having useful information for the loss estimation. This estimation makes it possible to explicitly consider all classes during the domain adaptation training even in the partially zero-shot setting, which leads to accurate adaptation between domains. Experimental results with several benchmark datasets validate the advantage of our method.

In the second part of this dissertation, we consider the most extreme case called zero-shot domain adaptation in which we do not have any target data for domain adaptation. If we do not have any information about the target data, the adaptation may be impossible. Therefore, we first clarify a possible scenario where we can assume availability of some knowledge instead of data, which enables us to effectively conduct the domain adaptation. We consider the situation where domain shift is caused by a prior change of a specific factor and assume that we know how the prior changes between the source and target domains. We call this factor an attribute, and reformulate the domain adaptation problem to utilize the attribute prior instead of target data. In our method, source data are reweighted with the sample-wise weight estimated by the attribute prior and the data itself so that they are useful in the target domain. We theoretically reveal that our method provides more accurate estimation of sample-wise transferability than a straightforward attribute-based reweighting approach. Experimental results with both toy datasets and benchmark datasets show that our method can perform well, though it does not use any target data.

In the third part of this dissertation, we consider another extreme case of domain adaptation called source-free domain adaptation in which we cannot access any source data during adaptation. In this setting, a model pretrained with source data is given instead of source data, and we aim to fine-tune this model with unlabeled target data. For distributional alignment between domains, we propose utilizing batch normalization statistics stored in the pretrained model to approximate the distribution of unobserved source data. This makes it possible to explicitly evaluate the distributional discrepancy between domains, and we conduct domain adaptation by minimizing this discrepancy. Experimental results with several benchmark datasets show that our method achieves competitive performance with state-of-the-art domain adaptation methods even though it does not require access to source data during adaptation.

In summary, this dissertation was devoted to increasing the applicability of domain adaptation. We have studied three kinds of problem settings for domain adaptation with scarce data. Our proposed algorithms are designed with a common perspective of matching data distributions between domains, and the experiments with several benchmark datasets have validated their advantages. Therefore, we conclude that we have succeeded to make it possible to apply domain adaptation to more diverse situations that is conceivable in real-world problems.