論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Adapting Pre-trained Models for Cross-Domain Few-shot Learning
著者(和文)	ZHENGHao
Author(English)	Hao Zheng
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第12937号, 授与年月日:2024年9月20日, 学位の種別:課程博士, 審査員:金﨑 朝子,篠田 浩一,村田 剛志,佐藤 育郎,井上 中順
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第12937号, Conferred date:2024/9/20, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

（博士課程）

Doctoral Program

論 文 要 旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	Computer Science 系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of (Philosophy)
学生氏名： Student's Name	ZHENG Hao	審査員主査： Chief Examiner	Asako KANEZAKI

要旨（英文 800 語程度）

Thesis Summary（approx.800 English Words ）

The demand for data in deep learning has sparked research interest in the field of few-shot learning, with cross-domain few-shot classification problems attracting widespread attention from researchers. This thesis begins by reviewing the development of deep learning and the history, practical applications, and academic necessity of few-shot learning. Stemming from the limitations of traditional small sample size classification problems, this chapter introduces the concept of cross-domain few-shot classification.

To better address the issue of insufficient inductive bias learning for target domain data in cross-domain few-shot classification, we adopt the strategy of incorporating pre-trained models to aid learning in downstream tasks. We studied and addressed three settings of cross-domain few-shot classification tasks. For the scenario where there are multiple source domains during training and the target domain is completely unseen, and where the domain gap between the source and target domains is relatively small, we proposed a meta-learning-based pre-training method and a DAC module. The principle is based on the hypothesis that a model's poor generalization performance on unseen domains is mainly due to inaccurate estimation of domain-specific features. Therefore, to improve generalization performance, it is only necessary to correct domain-specific features, rather than adjusting the overall features as done in previous studies. We also discovered that although the transferability of mid-level features is the best, placing the DAC module at different layers of ResNet to correct domain-specific features at various levels all result in beneficial effects.

This phenomenon has inspired us to recognize the necessity of multi-layer knowledge transfer in cross-domain few-shot classification tasks. We drew on this insight in addressing the challenge where a small amount of unlabeled target domain data is visible during training, and there is a significant domain gap between the source and target domains. This distillation mechanism not only allows for a comprehensive transfer of knowledge from the source domain, but it also incrementally enhances the feature extraction capabilities of the model in downstream tasks. This enables it to extract more discriminative and easily classifiable features on the target domain. In this chapter, we found that when the domain gap is large, the path of distillation is crucial. It is not advisable to directly distill the high-level features of the pre-trained model into the high-level of the target domain model, as this may lead to negative transfer. In the testing phase, we proposed an adaptive feature denoising method, which selectively sparsifies the original features during the testing phase. This method does not add any trainable parameters, reducing the risk of overfitting in the network and improving its performance on the test set.

Given that sparsified features during testing can improve the performance in cross-domain few-shot classification problems, it led us to ponder: can sparsified features during training also enhance the performance in cross-domain few-shot classification? We introduced Sparse and Efficient Feature Augmentation Regularization (SEFAR) to assist in fine-tuning pre-trained models on a small number of labeled samples from the target domain. Without altering the classification loss generated by the use of sparse features, SEFAR uses the classification results produced by sparse features to simultaneously self-distill the classification results generated by the complete features. We theoretically proved that Sparse and Efficient Feature Augmentation Regularization (SEFAR) can reduce the Rademacher complexity during the model learning process.

Empirical evidence demonstrated that the model could converge to a region of flat minima. Based on this chapter, we can draw the following key conclusions:

1） The model's poor performance on unseen domains is due to its inadequate estimation of domain-specific features. Correcting these biased estimations can lead to improved generalization capabilities.
2） When dealing with a large domain gap and using knowledge distillation to transfer knowledge from pre-trained models, choosing the appropriate distillation path is crucial. An improper selection of the distillation path can lead to negative transfer. This conclusion is highly thought-provoking, as it touches upon the two most fundamental questions in transfer learning: “what to transfer” and “where to transfer.”
3） Sparse features can be an effective method for addressing cross-domain few-shot classification problems, as they

reduce the complexity of the model and lower the risk of overfitting.

In the course of our experiments, we experimented with integrating the three methods: CLD, SEFAR, and DAC. For instance, we attempted to merge CLD and SEFAR. However, we observed potential interference arising from the complexity of numerous hyper-parameters. Moving forward, to effectively amalgamate these three methodologies, we are considering the application of reinforcement learning. This approach would be instrumental in fine-tuning the hyper-parameters and optimally positioning the DAC module.