

# Few-Shot Learning Method Based on Data Enhancement and Transfer Learning

**Bo Liu**

Information Engineering College, Capital Normal University, Beijing, China

1221003115@cnu.edu.cn

**Abstract.** With the significant progress of deep learning technology in many fields, the dependence of model training on a large amount of labeled data is increasingly prominent. However, in many practical application scenarios, especially in tasks with high labeling costs, data scarcity often occurs. This realistic challenge has promoted the rise of Few-Shot Learning (FSL) technology, which seeks to achieve effective learning of models with extremely limited samples. This article provides a comprehensive overview of the theoretical background, key technologies of FSL, to explore its potential and effectiveness in solving the problem of small-sample learning. In the method overview section, this article pays special attention to FSL strategies based on data augmentation and transfer learning. By reviewing and analyzing these methods, this article aims to provide some theoretical support and technical references for further exploration in this field and hopes to contribute to solving the problem of data scarcity and promote the sustainable development of this field.

**Keywords:** Data enhancement, Transfer learning, Supervised contrast learning, self-supervised learning.

## 1. Introduction

Considering the quick development of deep learning and artificial intelligence, the improvement of model performance depends on a large number of high-quality labeled data. In practical applications, the cost of data acquisition and labeling can be extremely high, especially in certain fields (such as medical image analysis, rare disease diagnosis, industrial testing, etc.), where the number of labeled samples is very limited. At this time, the concept of Few-Shot Learning (FSL) came into being. FSL aims to achieve effective learning and generalization of the model through a very small number of training samples, even one or several samples, to maintain good performance in the environment of sample scarcity [1, 2].

The emergence of FSL brings new challenges and opportunities for the development of deep learning. Different from traditional deep learning models, FSL models can extract useful features and achieve task objectives in the case of scarce samples through transfer learning, meta-learning, and data enhancement. Its core goal is to improve the generalization ability of the model so that it can still have good classification, prediction, or recognition ability in the case of very small amounts of data.

FSL technology has a wide range of application prospects, especially in areas where data collection is difficult and labeling costs are high. In medical image analysis, FSL can be applied to the diagnosis of rare diseases, as these cases often lack sufficient training samples. In addition, in emerging fields such

as virtual assistants and automatic customer service systems, FSL can greatly reduce the amount of data required for training, thus accelerating the development and application of the system.

This paper reviews FSL from the perspective of data enhancement and transfer learning methods. First of all, data enhancement helps the model learn effectively in the case of scarce samples by expanding the diversity of the original data set. This paper will discuss in detail several common data enhancement methods, including image enhancement, adversarial generative network (GAN), and other cutting-edge technologies, as well as their applications and effects in FSL.

At the same time, transfer learning, as an important part of FSL, can improve the model's learning ability in new tasks through knowledge transfer. This paper discusses the FSL method based on transfer learning and introduces how to improve the efficiency of FSL by transferring the parameter or feature representation of the pre-trained model.

Through the multi-angle analysis of FSL techniques based on data enhancement and transfer learning methods, this paper aims to provide a more comprehensive perspective and technical reference for researchers in this field and help the future model training and application scenario expansion under the environment of data scarcity.

## **2. Method of Few-Shot Learning**

The workflow of FSL is usually divided into two stages, namely the basis learning stage and the model training and optimization (Meta-Learning Phase). In the basic learning phase, the model learns an initial set of common features on a large-scale training set. The training data for this stage is not necessarily the same as the data for the target task, but there is a certain correlation. By learning these general basics, the model can be better adapted to the target task in the subsequent meta-learning phase. In the model training and optimization phase, once the representations of features are obtained, the model can be trained using these representations. Data enhancement and transfer learning are two important strategies for FSL. Data enhancement enhances the generalization ability of the model by artificially expanding the data set (such as rotation, flipping, scaling, and other image processing techniques) to increase the training sample of the model. Transfer learning is the process of transferring knowledge from a similar domain or task so that the model can learn on a new task with less data.

### *2.1. Data enhancement*

Lee et al. [3] proposed a method for data enhancement using supervised comparison. The learning task is divided into two stages, the pre-training stage and the FSL stage. In the pre-training phase, the authors used supervised contrastive learning. In the first stage, the model is pre-trained using supervised contrastive learning (SCL). The image to be input is preprocessed first, then the image is converted into normalized embedding by an encoder network, and the embedding is converted into low-dimensional embedding by attaching a projection network. The supervised contrast loss (TIM) is calculated by pulling the positive samples with the same label closer and pushing away the negative samples. The second step in the training process is to use a new classifier to refine the encoder network and reject the projection network. After pre-training with SCL, the model is fine-tuned through a small sample task, where only a small number of labeled samples are used. The SCL phase learns to represent the feature space used to initialize the model so that the model can be better generalized in the face of a small number of labeled samples. This method of pre-training by supervised contrast learning in the first stage and then fine-tuning effectively improves the efficiency of FSL while maintaining some accuracy on large data sets. In the small-sample task learning stage, the model uses the pre-training features obtained through supervised contrast learning (SSL) [4] in the first stage to quickly adapt to the new small-sample task. At this stage, the model is exposed to only a small number of new class samples (usually only a few images per class), which is consistent with the core goal of FSL. Through the first stage of training, the model has learned a good feature representation that can distinguish between categories. From there, the model is fine-tuned based on a small number of samples. Since the pre-training stage has made the model learn features with strong generalization ability, the model can quickly adapt to new small-sample

tasks with the help of these small samples, reduce the risk of overfitting, and achieve good results in classification tasks.

When FSL is used for classification problems, a large number of labeled data sets are required, but these data sets are not necessarily available promptly. Donahue et al. [5] proposed an unsupervised [6] meta-learning algorithm. They used clustering methods to group unlabeled and unclassified data and unsupervised clustering methods to quickly generalize a small number of labeled samples to new tasks. Specifically, the authors treat the results of the clustering as pseudo-tags, and the resulting pseudo-tag sample set will be used to construct different tasks by various data enhancement methods. For example, samples can be transformed by random clipping, rotation, scaling, etc., and different support sets and query sets can be constructed to simulate the task structure in FSL. At the same time, the author proposes the CUMCA method, which is to enhance the data by clustering the samples. Specifically, the cluster embedding method is used to guarantee that each sample of one-time data comes from a different class, and to ensure that every sample with the same one-hot label originates from the same class between the inner and outer loops, the data improvement function is utilized. The advantage of CUMCA is that it generates pseudo-classes through clustering, making full use of unsupervised data so that the model can be trained without real labels. By using the cross-task adaptation mechanism of meta-learning, the model can learn and generalize more quickly on new tasks. CUMCA provides an effective unsupervised meta-learning approach that utilizes clustering and cross-task adaptation to improve generalization performance in FSL.

Yao et al. [7], in their study of low-lens natural language understanding (NLU), found that most previous enhancement methods only brought marginal gains, and in many cases, the use of data enhancement resulted in erratic performance and even failure mode. So his team came up with FlipDA, an efficient and effective data enhancement approach specifically for FSL scenarios. The core idea of FlipDA is to generate new samples using simple image flipping operations (horizontal flipping, vertical flipping, and two-way flipping) and further enhance the model's perception of flip types through self-supervised learning tasks. FlipDA also adds a flip classification task that asks the model to recognize the flip type of the image during training. FlipDA was evaluated on 8 tasks using 2 pre-trained models. Experimental results show that the proposed method significantly improves the accuracy of learning tasks and proves its effectiveness in FSL. Compared with other complex data enhancement methods, FlipDA can achieve performance improvement through simple operations and has the advantages of low computational overhead and easy implementation. The application of data enhancement in FSL is not limited to traditional geometric transformation or color adjustment, and enhancement methods such as FlipDA show that by combining self-supervised tasks, a model can further improve its ability to represent small sample data. Future data enhancement studies can continue to explore the direction of automatically searching optimization enhancement strategies by combining generative adversarial networks (GANs) [8] and other generative models, to provide more flexible and effective solutions for FSL.

## 2.2. Transfer learning

Transfer learning [9, 10] is also a commonly used method when faced with a small sample set. Applying the information gained from one work (the target task) to another similar but distinct task is the goal. Unlike traditional machine learning methods, transfer learning does not require training the model on the target task from scratch, but instead uses the knowledge already learned on the source task, thereby improving the learning efficiency and performance of the model, especially when the training data for the target task is limited.

In 2020, Yu et al. [11] proposed TransMatch, a scheme that combines transfer learning and semi-supervised learning, aiming to solve the problem of FSL. TransMatch is trained on large data sets to learn common feature representations. These pre-trained feature representations are then transferred to the FSL task in the target task. This migration not only speeds up the learning process of the model but also reduces the dependence on the target task-labeled data. To further enhance the model's functionality, TransMatch introduces semi-supervised learning, which uses unlabeled data in the target task for

training. The introduction of unlabeled data solves the problem of labeled data scarcity in tasks with small samples. The first step in TransMatch is to pre-train the feature extractor on the underlying class data. This component uses a lot of training data to learn common feature representations. The pre-trained feature extractor provides high-quality feature representations for new categories in subsequent small sample tasks, thus reducing the need for the model to label data for new categories. After the pre-training is complete, TransMatch will use the pre-trained feature extractor to initialize the classifier weights for the new class. Transfer learning reduces the risk of overfitting on small sample classes and ensures that models can effectively learn from limited new class data. The final component of TransMatch is to further update the model using semi-supervised learning methods to take advantage of unlabeled data. The model's capacity to adjust to new categories is improved through semi-supervised learning; in particular, the model's classification performance and generalization ability are improved by fully utilizing unlabeled input.

Medina et al. [9] proposed a transfer learning framework that combines self-supervised learning and prototype networks, aiming to improve the accuracy and generalization ability of small-sample classification. The core of this method consists of two parts, Self-Supervised Contrastive Learning, and Prototypical Networks, which are used together to improve feature learning and classification performance in small-sample classification tasks. To obtain a rich visual representation, the authors first apply unsupervised meta-learning and self-supervised contrastive learning to a large amount of unlabeled data. In this process, model learning brings together positive sample pairs (different enhanced views of the same image) while pulling away negative sample pairs (enhanced views of different images). In this way, through the self-supervised learning phase, the model can learn more generalized representations. After the pre-training of self-supervised learning, the model is transferred to the FSL task. The transfer learning is combined with the prototype network, and the features obtained by the pre-training are used for classification. At the same time, they extended the prototype nearest neighbor classifier ProtoNet. ProtoNet works by taking the average feature representation of each class of samples as a prototype for that class and classifying new samples according to the distance between them. This method enables FSL to have a strong classification ability even when each type of sample is very limited. The experimental findings demonstrate that in numerous small sample classification problems, the suggested approach outperforms the current mainstream methods. Especially in the case of very few samples (only 1 to 5 samples per class), the model can better perform classification tasks by transferring self-supervised pre-trained features and showing strong cross-task generalization ability.

### **3. Existing limitations and prospects**

Despite its numerous obstacles, FSL is a popular area of study in the field of artificial intelligence due to its enormous application potential and research opportunities. The main goal of FSL is to learn and generalize from a very small amount of data, but the generalization ability of FSL models is often limited in practical applications. Even after training on a large number of related tasks, the model may still have overfitting problems when faced with new tasks, especially when there are large differences between tasks. Current FSL techniques rely more on transfer learning on similar tasks, so the performance of models is significantly reduced when dealing with situations that vary widely across domains or tasks. **Data enhancement** This enhancement method based on manual rules can only increase the diversity of samples to a certain extent and is mostly limited to low-level features (such as geometric transformations in images). For higher-level semantic features, such as complex scenes or sentence structure in natural language processing, traditional enhancement methods have limited effectiveness. In addition, generative models such as GANs can generate new samples, but the generated samples are often easily constrained by training data, resulting in low quality or lack of diversity of the generated samples. Especially in a small sample environment, the training of GAN may be unstable, which will affect the overall performance of FSL. While the goal of FSL is to achieve efficient learning with a small amount of data, the models that achieve this are often very complex, involving a lot of task training and model fine-tuning.

Although FSL faces many challenges, its huge application potential and research prospects make it still a hot research direction in the field of artificial intelligence. Future research and technological developments are expected to overcome existing limitations and push FSL to achieve breakthroughs in more fields. In the future, FSL needs stronger cross-domain generalization ability to adapt to scenarios with large differences between tasks. This requires the development of more general learning methods that allow models to transfer knowledge more efficiently across domains. One potential direction is Adaptive Learning, which enables models to dynamically adjust their learning strategies according to the characteristics of new tasks. In addition, exploring Cross-Domain Meta-Learning will also be an important way to improve the generalization ability of FSL. At the same time, future data enhancement methods need to go beyond the traditional geometric transformation and the extension of low-level features to deeper semantic features at higher levels. By enhancing the semantic information, the model can obtain richer feature representation on fewer samples. For example, in the field of image, the enhancement method can simulate the pose and scene structure of the object, while in the field of natural language processing, data enhancement can be realized by syntactic structure, semantic substitution, and other technologies. Improving the computational efficiency of the FSL model will also be an important research direction in the future.

#### 4. Conclusions

This paper reviews FSL from the perspective of data enhancement and transfer learning, focusing on the limitations and prospects of these two methods and the existing FSL methods. In terms of limitations, although transfer learning can make use of large-scale pre-trained models, the transfer effect is often not ideal in the case of large differences between tasks, and the model is difficult to generalize. Most of the existing data enhancement methods rely on low-level geometric transformation, which cannot effectively extend high-level semantic features, and limit the performance of models in complex scenes. In addition, generative models such as GANs are also prone to instability in the case of small samples, which affects the quality of enhanced samples. The research direction of FSL will aim to overcome these limitations. Smarter high-level semantic data enhancement methods will be key, with models learning richer features from less data through enhanced semantic understanding. At the same time, transfer learning methods will also continue to evolve, especially in cross-domain and cross-task environments, enabling more efficient knowledge transfer. In addition, improving the stability of generative models such as GANs and meta-learning strategies that combine self-supervised learning and contrast learning will also improve the effectiveness of FSL. Through these innovations, FSL is expected to achieve breakthroughs in more complex application scenarios.

#### References

- [1] Wang, Y., & Yao, Q. (2019). Few-shot learning: A survey. CoRR, abs/1904.05046.
- [2] Wang, Y., Yao, Q., & Kwok, J. T. (2020). Generalizing from a few examples: A survey on few-shot learning. ACM Computing Surveys (CSUR), 53(3), 1-34.
- [3] Lee, T., & Yoo, S. (2021). Augmenting few-shot learning with supervised contrastive learning. IEEE Access, 9, 61466-61474.
- [4] Khosla, P., Teterwak, P., Wang, C., et al. (2020). Supervised contrastive learning. Advances in Neural Information Processing Systems, 33, 18661-18673.
- [5] Donahue, J., Krähenbühl, P., & Darrell, T. (2017). Adversarial feature learning. In 5th International Conference on Learning Representations (ICLR), Toulon, France, April 24-26, 2017.
- [6] Ren, Z., Yan, J., Yang, X., Yuille, A. L., & Zha, H. Unsupervised learning of optical flow with patch consistency and occlusion estimation.
- [7] Yao, Y., Yu, Z., Zhang, L., & Sun, S. (2021). FlipDA: Effective and robust data augmentation for few-shot learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 12147-12156).

- [8] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.
- [9] Medina, C., Devos, A., & Grossglauser, M. (2020). Self-supervised prototypical transfer learning for few-shot classification. *arXiv preprint arXiv:2006.11325*.
- [10] Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques* (pp. 242-264). IGI Global.
- [11] Yu, Z., Chen, L., Cheng, Z., et al. (2020). Transmatch: A transfer-learning scheme for semi-supervised few-shot learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12856-12864).