Few-shot Learning: Methods and Applications

Jiaxiang Li[1*], *Mingyang* Li[2]

[1]Middle School Affiliated To Renmin University of China Tongzhou Campus, Beijing, China
[2]Shool of Computer and Artificial Intelligence, Beijing Technology and Business University, Beijing, China

Abstract. The Few-shot learning (FSL) approach distills meaningful features from a constrained sample set, allowing models to swiftly adjust to novel tasks and decreasing the dependency on extensive datasets. This approach leverages methods involving meta-learning, transfer learning, and data augmentation to boost the model's ability to recognize new categories. In many areas of artificial intelligence, obtaining large annotated datasets is often high in financial demand and extensively time-consuming, particularly in specialized fields with limited data availability. Therefore, the study of FSL is particularly critical. This paper first reviews the relevant methods of FSL, primarily categorizing them into model fine-tuning based FSL, data augmentation-based FSL, and transfer learning-based FSL. Model fine-tuning based FSL involves making slight adjustments to pre-trained models, allowing them to adapt to new tasks. Data augmentation-based FSL enhances the model's generalization ability by generating or expanding existing data. Transfer learning-based FSL transfers knowledge acquired by models from large datasets to smaller ones, enhancing the learning outcomes. Subsequently, this paper reviews the application areas of FSL and explores its impact in these fields. This paper aims to present the current state and prospects of this technique, providing valuable insights for researchers in related fields.

1 Introduction

With the advent of the information and intelligence era, large models in the field of deep learning have gradually become a research focus. The rich data sources and massive experimental data have provided these models with unprecedented training and optimization opportunities, leading to a gradual increase in model accuracy. Common tasks such as object detection and image classification can now be performed excellently [1]. However, in fields such as natural disasters and space remote sensing, where large-scale data acquisition is limited, existing deep learning models fail to function effectively. Consequently, Few-shot learning (FSL) models have been proposed. FSL addresses the issue of data scarcity to a specific extent by utilizing a small number of samples for sufficient learning.

In early studies, researchers commonly employed methods such as Support Vector Machines, Random Forests, and Naive Bayes for the classification of small datasets. These methods were based on traditional machine learning theories but required extensive manual

* Corresponding author: 2307010207@st.btbu.edu.cn

feature processing and annotated data. With the evolution of deep learning technologies, these issues have been gradually resolved.

Convolutional Neural Networks (CNNs) have been extensively applied across domains like healthcare, agriculture, and manufacturing [2, 3, 4]. Although CNNs perform well on large-scale datasets, their effectiveness on small-sample datasets remains inadequate. Traditional image classification methods include steps such as data preprocessing, feature extraction, model training, and parameter tuning. CNNs extract features through convolutional layers, use nonlinear activation functions to obtain complex feature representations, and employ pooling layers to reduce dimensionality and enhance feature abstraction. The stacking of convolutional and pooling layers enhances the network's expressive ability, ultimately extracting features further through fully connected layers to provide decision support for classification tasks. While this approach works well when data is sufficient, it fails to accurately learn effective features when data is insufficient, resulting in poor classification performance.

This paper delves into FSL models, aiming to address the challenges deep learning faces when data acquisition is limited. It summarizes previous research achievements in this field, not only detailing the intrinsic mechanisms of various FSL models but also providing valuable references for future research directions and applications. Given the potential of FSL to alleviate the problem of data scarcity, this research is of significant importance to theoretical development and offers practical solutions to challenges in real-world application scenarios, paving the way for the broader application of deep learning in various fields.

2 Methods of Few-shot learning

2.1 The Few-shot learning method based on model fine-tuning

The FSL paradigm, predicated on model fine-tuning, generally commences with the training of foundational models on extensive datasets, followed by the refinement of these models on a reduced, task-specific dataset. The fine-tuning methods can be broadly categorized into comprehensive model fine-tuning and parameter-efficient fine-tuning (PEFT). Comprehensive fine-tuning of large models involves adjusting all layers and parameters to adapt to specific tasks. This process generally employs a smaller learning rate and task-specific data, fully leveraging the general features of the pre-trained model, though it may require more computational resources. Parameter-efficient fine-tuning is to enhance the pre-trained model's capabilities on novel tasks through a reduction in the number of parameters adjusted and the computational overhead, thereby reducing the training burden of large pre-trained models. Even in resource-constrained environments, PEFT techniques can quickly adapt the pre-trained model to new tasks, enabling effective transfer learning. Thus, PEFT not only improves model performance but also significantly shortens training time and reduces computational costs, allowing more researchers to engage in deep learning research. PEFT includes various methods such as Low-Rank Adaptation (LoRA), Factorized Attention Regularization (FAR), Adapter, Prefix Tuning, Prompt Tuning, Implicit Prompt Tuning (IPT), and Differentiable Pruning (DiffPruning). The comparison of PEFT is shown in Table 1.

Table 1. Comparison of PEFT methods.

Methods	Description	Advantages	Disadvantages	Typical Use Cases
LoRA	Introduces low-rank matrices into matrix multiplication modules to adjust key dimensions.	Reduces computational complexity; efficient parameter adjustment.	Limited to specific model architectures; complex implementation.	Large language models, NLP tasks.
FAR	Uses factorized attention mechanisms to regularize the attention weights.	Improves model generalization; helps in preventing overfitting.	May require fine-tuning of regularization parameters.	NLP, vision tasks.
Adapter	Adds an adapter module to the Transformer architecture, tuning only this new component.	Minimal additional parameters; efficient; retains original model parameters.	Limited adaptability compared to full fine-tuning.	Transfer learning, domain adaptation.
Prefix Tuning	Prepends adaptable virtual tokens to the input, tuning these tokens during training.	Reduces training parameters; efficient.	Limited to Transformer-based models.	Text generation, classification.
Prompt Tuning	Embeds directive tokens in the input layer to guide model outputs towards specific goals.	Effective for large models; requires no changes to additional layers.	May not capture complex task specifics.	Text and image classification tasks.
IPT	Converts prompts into learnable embedding layers, processed with MLP and LSTM.	Greater flexibility; strong representation capabilities.	More complex setup; additional tuning required.	Advanced NLP, multi-layer tasks.
DiffPruning	Uses differentiable techniques to prune less important parts of the model.	Reduces computational and storage costs.	Can be complex to implement; may affect model performance.	Model optimization, efficiency improvement.

Comprehensive model fine-tuning is a deep learning method that enhances the model's performance on certain tasks by continuing training on a pre-trained model. This process involves selecting an appropriate pre-trained model, preparing task-specific data, adjusting the model architecture, training with a small learning rate, and evaluating performance. Its advantages include improved task performance and reduced training resources, but care must be taken to avoid overfitting. It is widely applied in fields such as NLP and computer vision, where frameworks like TensorFlow and PyTorch can simplify the fine-tuning process. Ethical and societal impacts must also be considered.

Here are six mainstream improvement methods for parameter-efficient fine-tuning of models. Prefix Tuning represents an effective strategy for model refinement, which involves prepending adaptable virtual tokens to the input and tuning these tokens throughout the training process. This method only updates the prefix parameters, keeping the rest of the Transformer model unchanged, consequently, this diminishes the volume of parameters needed for training and enhances the efficiency of the training process. Prompt Tuning embeds directive tokens in the input layer, steering the model's output toward task-specific goals without necessitating further modifications to any additional multi-layer perceptron (MLP) layers, and its efficacy rivals full fine-tuning in larger model scales. P-Tuning further develops the concept of prompts by converting them into learnable embedding layers and processing them with MLP and long short-term memory (LSTM), providing greater flexibility and stronger representation capabilities. P-Tuning v2 inserts prompt tokens into multiple layers of the model, increasing the number of learnable parameters and exerting a more direct influence on model predictions, thereby achieving better performance across different tasks and model scales. Adapter Tuning introduces an adapter module within the Transformer architecture, optimizing solely this new component and retaining the foundational model parameters unchanged, thus ensuring efficiency and adding a minimal set of parameters. LoRA simulates full-scale fine-tuning by introducing low-rank matrices into matrix multiplication modules, updating key low-rank dimensions in language models to achieve efficient parameter adjustment and reduce computational complexity. These methods are fine-tuning strategies designed for large pre-trained models, aiming to reduce computational costs while enhancing model performance on specific tasks, allowing researchers and developers to choose the most suitable fine-tuning strategy based on task requirements and resource constraints.

Although model fine-tuning methods address the issue of insufficient target data, a limited set of data cannot reflect the whole true distribution of large datasets, making the model prone to overfitting. To address the aforementioned overfitting problem, FSL methods based on data augmentation and transfer learning have been proposed.

2.2 Data augmentation-driven Few-shot learning

The purpose of data augmentation is to improve the model's generalization ability by increasing the quantity or quality of data, thereby reducing the model's tendency to overfit. Common methods include oversampling techniques and generative methods.

Oversampling techniques involve directly increasing the number of samples. Popular approaches include random oversampling and synthetic minority oversampling. Chao X et al. [5] proposed a hierarchical synthetic minority over-sampling technique (H-SMOTE) method, an improved version of the synthetic minority over-sampling technique (SMOTE). SMOTE typically balances datasets by creating additional synthetic samples for the minority class. H-SMOTE further optimizes this process to provide a more stable and reasonable data augmentation approach. SMOTE, proposed by Jia B et al. [6], is an oversampling technique designed to handle imbalanced datasets by generating synthetic samples between minority class samples. This technique avoids overfitting through feature space interpolation while maintaining the original data distribution. Although SMOTE can effectively improve classifier performance, attention must be given to parameter selection to reduce computational costs and avoid noise amplification. SMOTE is extensively applied in specific fields such as healthcare as well as finance. Zeng X et al. [7] introduced the Active PETs method, an FSL technique that improves classification accuracy by constructing multiple PET model committees. This method leverages model diversity to capture uncertainty and adopts a query strategy to select informative samples for labeling. It uses a weighting mechanism to allocate voting rights based on model complexity and balances label distribution through data oversampling, thereby enhancing model generalization ability and learning efficiency.

Generative methods involve learning from samples and generating new ones to expand the dataset. Common approaches include Bayesian methods, variational autoencoders, diffusion models, and Generative Adversarial Networks (GANs). Among these, GANs are widely used due to their adversarial learning framework, which produces more realistic data. Hong Y et al. [8] proposed a Few-shot generative adversarial network (F2GAN) technology for generating images using FSL. The generator in F2GAN A U-Net architecture incorporates skip connections, merging high-level features from multiple conditional images generated with random interpolation coefficients. A non-local attention fusion module fills in missing details, and the discriminator combines adversarial as well as classification losses to ensure the generated images' realism and category consistency. Additionally, mode-seeking loss combined with interpolation regression loss is introduced to enhance the diversity of the generated images. Experiments have precisely illustrated the effectiveness and superiority of this method in addressing Few-Shot image generation challenges. Fontanini T et al. [9] introduced the Meta-Learning Generative Adversarial Network (MetalGAN) method, designed to solve multi-domain unlabeled image synthesis problems. MetalGAN combines conditional GAN (cGAN) with meta-learning algorithms, enabling the generator (G) to effectively learn image reconstruction through nearly complete training and Few-Shot inference, and switch between different domains. The uniqueness of MetalGAN lies in its use of a single GAN to train across multiple tasks, forcing the underlying weight structures of the generator and discriminator networks to learn a universal and effective representation for various tasks. Unlike other methods such as star generative adversarial network (StarGAN), MetalGAN does not require target labels to adjust the generator and discriminator network outputs but instead provides conditions implicitly through meta-learning algorithms, allowing for more flexible and generalizable image synthesis capabilities. Gordon J et al. [10] introduced a meta-learning approach grounded in probabilistic inference using Bayesian decision theory. Their approach extends Bayesian decision theory, particularly in cases where the full predictive distribution is returned, by introducing a variational inference-based objective function for probabilistic inference in meta-learning. Specifically, during training, they employ multiple internal train/test splits to train the model, learning model parameters by minimizing the average expected loss of the tasks. This method is mainly applied to address FSL challenges, improving the model’s capacity for generalization on new tasks by effectively utilizing prior knowledge and statistical information during the learning process. Fig 1 and Fig 2 show the schematic diagram of GAN and cGAN.

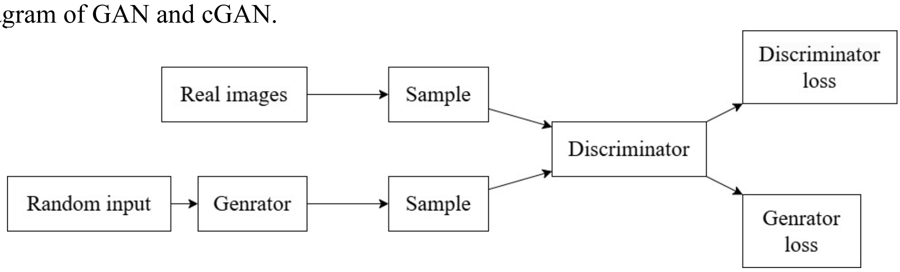

Fig. 1. Schematic diagram of GAN

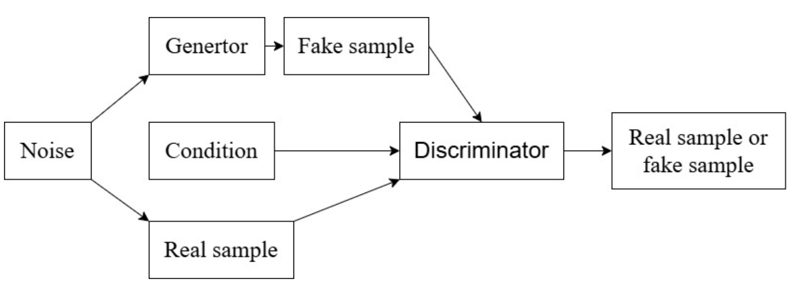

Fig. 2. Schematic diagram of cGAN

2.3 Few-shot learning employing transfer learning techniques

Transfer learning enhances the model's performance on small sample data by leveraging large labeled datasets for pre-training and then transferring the acquired knowledge to the target task. Typically, the dataset is partitioned into three distinct sets, a training set (source data), a support set (target domain training samples), and a query set (target domain test samples). Common methods include those based on metric learning, meta-learning, and graph neural networks. Metric learning-based methods: These methods involve learning a distance or similarity metric, making similar samples closer in the metric space. They are commonly used in clustering, image retrieval, and recommendation systems. Representative methods include the siamese network and the triplet network. Meta-learning-based methods: These methods focus on learning how to fastly adapt to new tasks, thereby enhancing the model's generalization ability. Popular techniques include model-agnostic meta-learning (MAML) and prototypical networks. Graph neural network-based methods: These methods are designed to handle graph-structured data, capturing complex relationships between nodes. They are widely used in social network analysis, bioinformatics, and recommendation systems. These approaches effectively learn and transfer knowledge from small sample data in various ways, thereby improving the model's performance on new tasks.

3 Classic scenario applications

3.1 Object recognition

FSL techniques have been extensively applied in the areas of object recognition, particularly in critical areas such as wildlife conservation and market management. In wildlife conservation, FSL enables precise species identification under limited sample conditions, effectively addressing the challenge of insufficient labeled data for rare or endangered species. In market management, FSL also demonstrates its advantages by efficiently identifying specific brands or product models under sample-restricted conditions, thereby enhancing the accuracy and practicality of recognition tasks.

3.2 Multi-label text classification

In the study of large-scale multi-label classification problems, traditional assessment methods often overlook rare and unseen labels. To address this deficiency, Anthony et al. [11] proposed an innovative neural network architecture. The system ensures precise categorization of Few-shot and zero-shot labels through a meticulous correlation of electronic health record discharge summaries with the pertinent feature vectors. Additionally, the study introduces a fine-grained assessment mechanism for labels, particularly focusing on frequent, Few-shot, and zero-shot labels, to deeply analyze the model's performance in handling uncommon labels.

3.3 Drug discovery

In the realm of medical research, the utilization of FSL techniques is rapidly becoming a focal point of interest for investigators. Han et al. refined the iterative refinement of long short-term memory (IR-LSTM) [12] innovatively and combined it with graph convolutional neural networks (GCNN) [13], significantly enhancing the capacity to acquire significant measures of separation among diminutive molecules. This innovative method has turned the tide in drug discovery, streamlining the data requirements for potent predictive analysis.

 https://doi.org/10.1051/itmconf/20257002012

Utilizing limited biological data, this method can accurately predict new compounds, optimize the selection of candidate molecules, and significantly improve the efficiency of drug development. This FSL approach provides an innovative solution for drug discovery, effectively overcoming the traditional deep learning methods' dependency on large datasets and opening up new avenues for medical research.

4 Conclusion

This study delves into the distinct methodologies and practical applications of FSL, underscoring its vital role in mitigating data limitations and improving the models' capacity for broader generalization, as gleaned from an in-depth review of the scholarly discourse. Initially, the paper reviews FSL methods from three perspectives: model fine-tuning, data augmentation, and transfer learning, demonstrating how FSL can train efficient models with limited data to achieve specific task objectives. Subsequently, The manuscript offers an in-depth critique of the efficacy of FSL in real-world applications, covering areas such as object recognition and drug discovery. By analyzing these applications, the paper showcases the effectiveness and potential of FSL in various tasks and discusses its future development directions and impact on the field of artificial intelligence. The paper anticipates that these Few-shot studies will promote breakthroughs in artificial intelligence technology across a broader range of application fields, providing new ideas and methods for solving real-world problems. As technology persistently evolves, it is reasonable to believe that FSL will play an even more critical role in the future trajectory of artificial intelligence.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

1. K. Simonyan. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, (2014).
2. J. Chen, Y. Lu, Q. Yu, et al. Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306, (2021).
3. A. Milioto, P. Lottes, C. Stachniss, Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs, in Proceedings of the 2018 IEEE international conference on robotics and automation (ICRA), (2018).
4. Y. Babakhin, A. Sanakoyeu, H. Kitamura, Semi-supervised segmentation of salt bodies in seismic images using an ensemble of convolutional neural networks, in Proceedings of the Pattern Recognition: 41st DAGM German Conference, DAGM GCPR 2019, Dortmund, Germany, September 10-13 (2019).
5. X. Chao, L. Zhang, Few-shot imbalanced classification based on data augmentation. Multimed. Syst, **29**, 2843 (2023).
6. B. Jia, Y. Tian, D. Zhao, et al. Bidirectional RNN-Based Few-Shot Training for Detecting Multi-stage Attack, in Proceedings of the Information Security and Cryptology: 16th International Conference, Inscrypt 2020, (2021).

7. X. Zeng, A. Zubiaga. Active PETs: active data annotation prioritisation for few-shot claim verification with pattern exploiting training. arXiv preprint arXiv:2208.08749, (2022).
8. Y. Hong, L. Niu, J. Zhang, et al. F2GAN: Fusing-and-Filling GAN for Few-Shot Image Generation, in Proceedings of the 28th ACM international conference on multimedia, (2020).
9. T. Fontanini, E. Iotti, L. Donati, et al. MetalGAN: Multi-domain label-less image synthesis using cGANs and meta-learning. Neural Netw, **131**, 185-200 (2020).
10. J. Gordon, J. Bronskill, M. Bauer, et al. Meta-learning probabilistic inference for prediction. arXiv preprint arXiv:1805. 09921 (2018).
11. A. Rios, R. Kavuluru, Few-shot and zero-shot multi-label learning for structured label spaces, in Proceedings of the Conference on Empirical Methods in Natural Language Processing, Conference on Empirical Methods in Natural Language Processing, (2018).
12. H. Altae-Tran, B. Ramsundar, A.-S. Pappu, et al. Low data drug discovery with one-shot learning. ACS Cent. Sci, **3**, 283 (2017).
13. N Wagh, Y Varatharajah, Eeg-gcnn: Augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network, in Proceedings of the Machine Learning for Health, (2020).