ITM Web of Conferences **80**, 01042 (2025) https://doi.org/10.1051/itmconf/20258001042
ACAAI 2025

Comparative Analysis of Core Methods in the Few-Shot Learning Area

Zuoyou Dang[*]

Mathematics Department and Statistics and Data Science Department, University of California, Los Angeles, Los Angeles, 90095, California, United States

Abstract. Few-shot learning (FSL) aims to address the critical challenge of models struggling to learn from very limited examples. It becomes extremely important when data collection is difficult or costly. This paper provides a comprehensive comparative analysis of core FSL methods, focusing on meta-learning approaches and data augmentation strategies. The paper examines metric-based meta-learning algorithms-including Matching Networks, Prototypical Networks, and Relation Networks-alongside augmentation-based methods such as traditional data augmentation and generative models. The comparative study reveals that metric-based approaches leverage learned embedding spaces and distance metrics to achieve strong classification performance with minimal data. Meanwhile, augmentation-based techniques significantly improve model generalization by synthetically expanding the support data, with generative augmentation yielding notable accuracy gains. Our findings highlight the strengths and limitations of each method under various conditions. In summary, this work underscores the value of combining meta-learning with data augmentation to tackle data scarcity, and it outlines key challenges, such as domain shift, scalability, or theoretical gaps, to guide future few-shot learning research and applications.

1 Introduction

Living in an era where artificial intelligence has become an integral part of people’s daily lives. Users increasingly rely on AI systems to encounter various issues they have encountered. These systems are trained based on large-scale data and machine learning algorithms, allowing them to learn and make predictions. While these systems have an excellent capability to handle common problems with vast amounts of training data, they still struggle with rare or uncommon scenarios, providing vague or irrelevant responses. This limitation highlights the importance of Few-Shot Learning (FSL). It enables AI models to learn effectively and precisely from minimal examples, then enhances their adaptability and accuracy when data is insufficient.

In these years, many scientists have dedicated themselves to this potential area. One outstanding branch is utilizing meta-learning. Meta-learning is an approach that allows models to adapt to new tasks with scarce data. Timothy et al. conclude that it is learning to

Corresponding author: ooctoo@g.ucla.edu

learn. Based on this, metric-based, memory-based, and learning-based methods are created [1]. Metric-based methods by Wang et al. achieved classification by learning a feature space and measuring distances between inputs and queries [2]. Their models showed strong performance in the sensor technology area. Another significant part is learning the optimizer. Furthermore, Deng et al. proposed a method based on a dynamic-memory-based prototypical network to address the few-shot difficulties in event detection [3]. Such memory mechanisms allow the system to improve both adaptation speed and accuracy in low-data settings by recalling relevant information related to given tasks. Despite these directions having a progressive achievement, issues such as computational complexity and overfitting in cross-domain scenarios still exist.

This study is affected by prior work and provides a comparative analysis of the strengths and limitations of few-shot learning methods developed in recent years. It aims to offer clear and practical guidance for effectively utilizing various methods of FSL to handle all kinds of situations with severe data constraints.

2 Theoretical background

2.1 Few-shot learning definition and motivation

Starting from the elementary concepts, Few-Shot learning (FSL) refers to learning how to classify from only a small number of examples (typically 1-5 per class). This pattern is defined compared to traditional deep learning, which usually requires a number of samples for each class for good performance. The large calculative costs and vast amount of data requirements become core motivations for FSL systems. One insightful thought that brings learning theory to a new level is that conventional generalization bounds presume large sample sizes, which “do not fully explain why few-shot learning is at all possible”. So recent theoretical work by Tyukin et al. addresses this gap by proving that if a model’s decision space is sufficiently high-dimensional and if certain data non-concentration conditions hold, even one or few-shot learning can complete the learning process with meaningful generalization [4]. This dimensional idea provides a mathematical foundation explaining how FSL is possible [4]. Based on this foundation, several branch categories were created and developed over the years. Song et al. emphasize that FSL has emerged as an effective approach when data collection is difficult, and they position meta-learning as a key concept enabling rapid learning-to-learn from limited data [5].

2.2 Meta-learning

As mentioned above, a dominant approach to FSL is meta-learning, often described as learning to learn [1]. In other words, rather than training a model for a single task, meta-learning trains a learner across many tasks and areas so it can quickly adapt to new tasks with only learning a small amount of data. This kind of model typically involves two parts of the training process: an outer loop that learns across tasks and an inner loop that learns within each task. Through training on simulated few-shot tasks, the model obtains learning approaches that enhance fast adaptation. Meta-learning methods can be broadly categorized into three families in terms of their design philosophy. To start with, the most popular one is metric-based methods, which learn an embedding space and a distance metric to compare samples. The meta-learner produces a feature representation and uses a fixed distance function to classify samples by comparing them to support examples. Moreover, Model-based (Memory-based) methods, which equip neural networks with external memory or architecture modifications that allow quick updates. Finally, Optimization-based methods, which learn how to optimize or initialize models for quick adaptation. It includes algorithms like MAML, which optimize for initial parameters that are adaptable [6]. All these branch types share the same goal of extracting transferable knowledge from previous experience, which allows models to learn a new task with only a few examples and minimal training.

3 Comparative analysis

3.1 Metric-based meta-learning

3.1.1 Matching networks

Matching Networks (MatchNets) introduced an episodic training strategy and an attention-based metric for one-shot learning. Key components include an embedding function and an attention mechanism that computes a weighted vote based on the support set. Instead of updating weights at test time, the model finds the closest example in the support. This procedure combines the benefits of non-parametric models with deep learned representations. Vinyals et al. also emphasized training the network in conditions that are similar to the test to apply it in the one-shot testing scenario [7]. Matching Networks achieved significant improvement on one-shot image classification benchmarks. One-shot accuracy on miniImageNet went from ~88% to ~93% in their experiments [7]. They were among the first meta-learning models to demonstrate the power of metric-learning for FSL on tasks like miniImageNet. With a fixed metric, its accuracy has reached an outstanding level. A limitation comes out at the same time, that it does not explicitly learn a relation behind the learning process. This gap was addressed by later methods.

3.1.2 Prototypical networks

Prototypical Networks (ProtoNets) simplify the metric-learning approach by representing each class in the support set with the mean embedding of that class's support examples. A query is classified to the class with the closest prototype in the embedding space calculated by Euclidean distance, which Snell et al. found is more effective than cosine in this context [8]. The only learning component is the embedding function, and prediction involves a simple similarity computation. Prototypical Networks achieved excellent results. On the 5-way miniImageNet benchmark, ProtoNets reached about 49.4% mean accuracy in the 1-shot setting and 68.2% in 5-shot, better than Matching Networks and even some more complex meta-learning methods of the time [8]. The success of ProtoNets brings out a point that learning a good feature extractor is perhaps the most important part of few-shot classification. However, the one-shot accuracy of ProtoNets and MatchNets was similar. ~49% vs ~43% on miniImageNet, under comparable conditions, because in the 1-shot case, the prototype is just that one example, making ProtoNets and MatchNets equivalent in this scenario [8].

3.1.3 Relation Networks

Relation Networks (RNs) learn the metric by introducing a small relation module that computes a similarity score between query and support features. The Relation Network framework has two main components: an embedding module that produces feature vectors for images, and a relation module that inputs a pair of feature vectors and outputs a scalar relation score. What makes RNs special is that Relation Networks generalize the metric. Rather than output distances in feature space, they let a neural network learn the comparison function that works best for the task. Sung et al. found that Relation Net achieved around 50.4% accuracy on 5-way 1-shot miniImageNet and about 65% on 5-way 5-shot, which was competitive with or better than ProtoNets of the time [9]. Relation Networks demonstrated that metric learning can be improved by self-learning, inspiring later variants that incorporate, such as graph neural networks or cross-attention for comparing sets of features [9].

3.2 Augmentation-based method

Other than metric-based methods, another way is augmentation. In few-shot learning, it is to augment the scarce training data, either by applying transformations or by synthesizing new examples. The research discusses two categories: data augmentation and generative models that produce new data instances. These approaches are complementary to meta-learning and can be used in combination with metric-based or optimizer-based methods.

3.2.1 Data augmentation

In FSL, data augmentation creates additional training examples from the few data. Traditional augmentation techniques, such as random flips, crops, and rotations for images, are considered a baseline strategy. There are a few strategies that researchers have proposed. One example is the Adversarial Feature Hallucination Network (AFHN) by Li et al. [10]. AFHN uses a conditional Wasserstein GAN to generate additional features for the classes, augmenting the support set in feature space [10]. It reported that compared with the other four data augmentation-based methods, AFHN beats Δ-encoder by more than 8% for the 5-shot setting and Dual TriNet by more than 3% for the 1-shot setting [10]. For MetaGAN, which is also based on GAN, AFHN achieves about 10% improvements for both the 1-shot and 5-shot settings [10]. Instead of generating features, another approach to data augmentation in FSL is to generate additional images using techniques like image mixing. Mangla et al. and Zhang et al. use meta-learners to produce slight deformations of support images to create new samples [11, 12]. The goal across these methods is similar, which is to increase the effective sample size. These results highlight the contribution of using a conditional Wasserstein GAN to generate new features in feature space, effectively expanding the support set, enhancing model generalization by increasing the effective sample size under limited data conditions [10].

3.2.2 Generative models

Recent work by Xu and Le exhibits generative models. They propose generating representative samples using a conditional variational autoencoder (CVAE) [13]. In their method, semantic information is used as a condition to generate feature vectors. What stands out is that Xu & Le make sure that the VAE focuses on generating representative samples by filtering out weird outcomes during training [13]. Xu and Le found that accuracy has critically improved by adding VAE-generated samples [13]. They found 6.8% ~ 10% and 1-shot accuracy improvements and 0.5, ~ 2.7% performance on 5-shot for benchmarks like miniImageNet and tieredImageNet [13]. Compared to adversarial feature augmentation, such as AFHN, the CVAE approach is more explicitly focused on coverage of the sample space. The challenge for generative models in FSL is avoiding overfitting with only a few provided samples [13]. Xu & Le proposed a textual description of classes and so on. Overall, generative models like CVAE demonstrate strong potential in few-shot learning by improving class representation and enhancing model generalization despite limited data availability [13].

4 Current challenges and possible improvements

Despite this progress, few-shot learning techniques still face several challenges, including domain shift, scalability, and model complexity. The following contexts will focus on discussing potential improvements or research directions.

4.1 Domain shift and generalization

One main challenge is the domain shift between the training data and the target application domain, specifically in the meta-learning area. Recently, many FSL methods make an unrealistic assumption that there is a similar data distribution between base classes and the target classes. In fact, during the process of cross-domain few-shot learning (CDFSL), two domains are significantly different. This will lead to a drop in performance of standard FSL methods [14]. Guo et al. discovered that simple fine-tuning can perform better than sophisticated meta-learning in some cross-domain scenarios [14]. This suggests that meta-learners often overfit to the base domain, and it is hard to transfer an existing model to a new category. Tseng et al. and Wang & Deng suggested domain-agnostic feature learning, which means using feature normalization or adversarial training to learn embeddings invariant to domain-specific styles [15,16]. Another possible direction is meta-learning with diverse domains or simulated shifts so that the model is explicitly trained to handle domain changes. Generally, combining meta-learning with transfer learning or unsupervised domain adaptation techniques may work well in this area [15,16].

4.2 Scalability and incremental learning

Another challenge is scaling few-shot learning to handle a growing number of classes or sequential learning of new tasks. The standard FSL setting considers a static, small set of novel classes, but what happens in the real world often requires continually adding new classes. This kind of problem was defined by Tao et al. as Few-Shot Class-Incremental Learning (FSCIL) [17]. In FSCIL, a model must incrementally learn new classes from limited data, without forgetting the previously learned classes. Tao et al. addressed this problem with a method called TOPIC. It uses a topology-preserving network to gradually expand the representation space for new categories and minimize changes to old categories [17]. Their approach has an outstanding performance on incremental benchmarks compared to plain fine-tuning or naive extensions of meta-learning [17]. This research highlighted the importance of preservation mechanisms when adding the number of classes. Scalability also includes computational and model complexity when the number of tasks grows. Many meta-learning algorithms are computationally intensive due to episodic training, which limits their usability as the task complexity increases. Recent research is exploring more efficient meta-learning to maintain scalability, such as optimizing the selection of episodes or using lightweight adaptation modules. However, how to systematically improve and trust these capabilities is an open question. Also, in the vision area, techniques like meta-distillation and better training routines are being used to scale few-shot learning to larger datasets. In summary, improving few-shot learners' scalability not only requires class-incremental learning methods but also considers computational scalability. Integrating ideas of continual learning to address forgetting and efficient meta-learning seems to be a promising path [17].

4.3 Model complexity and simplicity of representations

Few-shot research has produced very complex models with meta-learners, memory modules. However, Tian et al. proposed that much of this complexity might be unnecessary by showing a simple baseline [18]. This baseline trains a deep network on the base classes can perform better than many sophisticated meta-learning algorithms [18]. In their work, a high-quality supervised embedding with a bit of fine-tuning can achieve strong results on standard benchmarks [18]. This result shows the critical importance of feature space. Many meta-learning methods were trying to learn good features, but most of the time, a simple transfer learning approach might satisfy requirements in many cases. Similarly, Chen et al. also made a similar point with a “Baseline++” method [19]. These ideas show that meta-learning methods must demonstrate clear advantages, like faster adaptation or the ability to handle tasks that fine-tuning is not able to deal with. Ongoing research is taking advantage of both methods. They use a strong pre-trained encoder and then apply meta-learning on top for precise adaptation [18, 19].

4.4 Theoretical limitations and future directions

Few-shot learning also presents theoretical challenges. While Tyukin et al. set up a theoretical framework for few-shot success, there remains an unknown in understanding why current methods work, and how to guarantee their performance [4]. Classical learning theory struggles with the few-shot scenarios due to traditional VC-bound or PAC analyses, which require large samples to generalize. New theories, based on high-dimensional geometric properties, have excellent performance but are still abstract [4]. One specific limitation is that these theories assume the data distributions are independent, but this condition does not always hold in practice. Moreover, what makes a good task distribution for meta-training, or how tasks relate to each other in a way that a meta-learner can utilize. These are still open questions. As a possible improvement, researchers are exploring PAC-Bayesian approaches to meta-learning, which try to derive generalization bounds for the transfer from training tasks to new tasks [20]. Another possible theoretical way out is understanding the process of how the higher-dimensional representation space affects the few-shot learner so dramatically. Lastly, an important future direction noted in surveys is that many current benchmarks assume all novel classes are equally balanced with, say, 5 examples each, and tasks are homogeneously sampled [2, 5]. In real applications, some classes might have 10 examples, others only 1, and tasks might have varying ways or even require unsupervised adaptation. Developing theory and algorithms for such unbalanced or heterogeneous few-shot problems is a challenge on the horizon [2, 5].

5 Conclusion

Few-shot learning (FSL) has developed into one of the most powerful directions for addressing the problem of data scarcity. This comparative study discussed two major families of methods, metric-based meta-learning and augmentation-based learning. Metric-based methods such as Matching Networks, Prototypical Networks, and Relation Networks rely on learning a shared embedding space, which allows models to classify samples by comparing their distances to labeled examples. These approaches have achieved success on benchmarks like miniImageNet, where accuracies often exceed 50% in one-shot tests. However, their performance has a limitation that training and testing domains share similar distributions.

In contrast, augmentation-based approaches focus on enriching the training set. Traditional transformations like cropping, flipping, or rotating can only provide limited improvements. More advanced strategies, such as Adversarial Feature Hallucination Networks (AFHN) and Conditional Variational Autoencoders (CVAE), generate synthetic samples to expand the data sets and enhance model performance. These methods have shown 5–10% accuracy improvement in one-shot or few-shot benchmarks, proving the potential of generative augmentation. Importantly, augmentation strategies are complements to meta-learning, integrating both often outperform either approach alone.

Despite this progress, FSL still faces several challenges. One main challenge is cross-domain generalization. Models often overfit to the base domain and fail to transfer well when it comes to new areas. Solutions such as domain-invariant embedding learning or cross-domain meta-training may address this issue. Another obstacle is scalability and incremental learning. Real-world cases must continually incorporate new classes without forgetting previous ones. The TOPIC model with topology-preserving and memory-conserving designs has provided promising steps forward.

Finally, the theoretical understanding of FSL lags behind empirical success. Although geometric and PAC-Bayesian analyses offer partial reasonable explanations, the field still lacks a complete framework describing how small-sample learning generalizes effectively.

References

1. T. M. Hospedales, A. Antoniou, P. Micaelli, A. J. Storkey, Meta-learning in neural networks: A survey. IEEE Trans. Pattern Anal. Mach. Intell. **44**, 5149–5169 (2022)
2. Y. Wang, Q. Yao, J. T. Kwok, L. M. Ni, Generalizing from a few examples: A survey on few-shot learning. ACM Comput. Surv. **53**, 63 (2020)
3. S. Deng, N. Zhang, J. Kang, Y. Zhang, W. Zhang, H. Chen, Meta-learning with dynamic-memory-based prototypical network for few-shot event detection, in Proceedings of the 13th ACM International Conference on Web Search and Data Mining (WSDM), 151–159 (2020)
4. I. Y. Tyukin, A. N. Gorban, M. H. Alkhudaydi, Q. Zhou, Demystification of few-shot and one-shot learning, in International Joint Conference on Neural Networks (IJCNN), 1–7 (2021)
5. Y. Song, T. Wang, S. K. Mondal, J. P. Sahoo, A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. arXiv preprint arXiv:2205.06743 (2023)
6. C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, in Proceedings of the 34th International Conference on Machine Learning (ICML), 1126–1135 (2017)
7. O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, D. Wierstra, Matching networks for one shot learning, in Advances in Neural Information Processing Systems (NeurIPS) **29**, 3630–3638 (2016)
8. J. Snell, K. Swersky, R. S. Zemel, Prototypical networks for few-shot learning, in Advances in Neural Information Processing Systems (NeurIPS) **30**, 4078–4088 (2017)
9. F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. Torr, T. M. Hospedales, Learning to compare: Relation network for few-shot learning, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1199–1208 (2018)
10. K. Li, Y. Zhang, K. Li, Y. Fu, Adversarial feature hallucination networks for few-shot learning, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 13467–13476 (2020)
11. P. Mangla, M. Singh, A. Sinha, N. Kumari, V. N. Balasubramanian, Charting the right manifold: Manifold mixup for few-shot learning, in Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), 2207–2216 (2020)
12. C. Zhang, Few-shot image augmentation via meta-learning, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)

13. J. Xu, H. Le, Generating representative samples for few-shot classification, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 9003–9013 (2022)
14. Y. Guo, N. C. Codella, L. Karlinsky, J. V. Codella, J. R. Smith, K. Saenko, T. Rosing, R. Feris, A broader study of cross-domain few-shot learning, in European Conference on Computer Vision (ECCV), 124–141 (2020)
15. H. Y. Tseng, H. Y. Lee, J. B. Huang, M. H. Yang, Cross-domain few-shot classification via learned feature-wise transformation, in International Conference on Learning Representations (ICLR) (2020)
16. H. Wang, Z. H. Deng, Cross-domain few-shot classification via adversarial task augmentation, in Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI), 1075–1081 (2021)
17. X. Tao, X. Hong, X. Chang, S. Dong, X. Wei, Y. Gong, Few-shot class-incremental learning, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 12183–12192 (2020)
18. Y. Tian, Y. Wang, D. Krishnan, J. B. Tenenbaum, P. Isola, Rethinking few-shot image classification: A good embedding is all you need? in European Conference on Computer Vision (ECCV), 266–282 (2020)
19. W. Y. Chen, Y. C. Liu, Z. Kira, Y. C. Wang, J. B. Huang, A closer look at few-shot classification, in International Conference on Learning Representations (ICLR) (2019)
20. T. M. Hospedales, A. Antoniou, P. Micaelli, A. J. Storkey, PAC-Bayesian approaches to meta-learning. arXiv preprint arXiv:2103.00243 (2022)