Supervised Learning on Small Datasets: Few-Shot Approaches and Generalization

Sneha Swapnil Pawar*, Amruta S. Navale

Department of Computer Science, Dr. D. Y. Patil Art's, Commerce and Science College, Pimpri, Pune, Maharashtra, India

DOI: https://doi.org/10.51583/IJLTEMAS.2025.1413SP023

Received: 26 June 2025; Accepted: 30 June 2025; Published: 24 October 2025

Abstract: In artificial intelligence, supervised learning has become a dominant paradigm that allows developments in a variety of fields, including natural language processing, speech recognition, and image classification. However, success generally depends upon the availability of large labeled datasets, which are frequently high-priced or impractical to obtain in many real-world situations—particularly in domains like security, bioinformatics, and healthcare. Few-shot learning techniques, which look for to allow models to generalize effectively from a limited number of training examples, were developed in response to the difficulty of learning from limited data. The current study explores the three main few-shot learning strategies—transfer learning, meta-learning, and data augmentation—as solutions for the supervised learning problems of small datasets. To improve models for new, smaller tasks, transfer learning makes use of knowledge gathered from large-scale tasks. Models can quickly adjust to new tasks with little data thanks to meta-learning, also known as "learning to learn." Small datasets are artificially expanded using data augmentation techniques to increase robustness and generalization. We look at how these approaches improve supervised models' the capacity for generalization, minimize over fitting, and reduce variance.

This paper specifies the advantages, disadvantages, and uses of each approach through a thorough evaluation of previous studies and comparative analysis. In addition to the outcomes, hybrid approaches that combine these tactics perform better, particularly in fields with a lack of labeled data. In the final analysis, few-shot learning sets the way for a more efficient and equitable application of AI in situations with limited resources.

Keywords: Supervised Learning, Few-Shot Learning, Small Datasets, Transfer Learning, Meta-Learning, Data Augmentation, Generalization.

I. Introduction

Large datasets are typically required for machine learning models to function well. However, because of limitations in availability, cost, or privacy, real-world applications frequently involve a limited amount of labeled data. Conventional supervised learning models are unable to generalize in these situations, which results in high variance and lacking accuracy. This requires the integration of few-shot learning techniques, which allow models to learn efficiently from a small number of examples by utilizing external sources or past knowledge.

Few-shot learning is becoming more and more important in domains like rare event detection, low-resource language processing, and medical diagnostics because it imitates the human ability to learn from few examples.

Statement of the Problem

Standard administered learning methods endure from over fitting and destitute generalization when prepared on little datasets. In sectors like security or healthcare, broad commented-on datasets might not be available. In addition, models that have been trained on small datasets frequently do not apply what they have learned to tasks that are not visible. The investigation looks at the existing methods over solving such problems and suggests a framework to enhance model generalization with few-shot learning.

II. Literature Review

The effectiveness with few-shot learning (FSL) techniques in scenarios with limited data has been extensively assessed in the literature. Among the key areas of focus consisted of:

- Methods of meta-learning (such as Model-Agnostic Meta-Learning, or MAML)
- Transfer learning from deep models that have previously been trained
- Techniques for augmenting data (both synthetic and ordinary)
- Benchmark research on few-shot tasks (such as mini-ImageNet and Omniglot)

Dataset Selection

We used benchmark datasets that are frequently implemented in FSL to simulate few-shot conditions:

- Omniglot: For identifying elements in a few-shot situation
- Mini-ImageNet: For classifying images with a small number of labeled examples per class
- To guarantee a small number of examples, both datasets were resized or reduced (e.g., 1-shot and 5-shot settings).

Implementation of Approaches

Three main few-shot methods were put into practice and reviewed by us:

- ResNet-18 pertained on ImageNet and fine-tuned with a small subset is used in transfer learning.
- Model-Agnostic Meta-Learning, or MAML, is a gradient-based meta-learning algorithm.
- Data augmentation: expanding variety via rotation, flipping, as well as GAN-generated samples.
- To evaluate generalization, each model was tested on unseen classes when having been taught with consistent hyper parameters.

Evaluation Metrics

To evaluate performance, we employed:

- Accuracy: Truthfulness of classification for test classes that are not visible
- Loss: Classification using cross-entropy loss
- The generalization gap is the discrepancy between test and training performance.
- Computational Cost: Convergence time and number of training epochs

Experimental Setup

PyTorch with GPU support (NVIDIA Tesla T4) was used to conduct the experiments on Python. In order to guarantee fair comparison under the same hardware and software conditions, the code was created.

Empirical Review

Several empirical studies have looked into ways to get around the challenges related to supervise learning on small datasets. Through the application of large-scale pertained models to improve accuracy on target tasks with little data, transfer learning has shown promise (Howard & Ruder, 2018). Even with as few as 1–5 examples per class, meta-learning methods such as Model-Agnostic Meta-Learning (MAML) allow quick adaptation to new tasks (Finn et al., 2017). Model robustness has also been shown to be enhanced by data augmentation methods, such as text paraphrasing and picture changes (Shorten & Khoshgoftaar, 2019). Notwithstanding these advances, problems like domain shift and negative transfer still exist, especially in delicate applications like security and healthcare. Further study shows that combining methods—like intelligent augmentation and fine-tuning pre-trained models—improves generalization. These results emphasize the importance it is to choose context-appropriate methods to achieve reliable performance in few-shot learning scenarios.

Theoretical Framework

The Bayesian Inference Framework and Learning Theory act as the foundation for this study. According to learning theory, generalization relies on the ability of a model to accurately convey the basic purpose with the least amount of empirical error. In few-shot settings, regularization and inductive bias are important.

In low-data regimes, Bayesian inference enables uncertainty modeling. Even with small sample sizes, methods such as Bayesian neural networks offer accurate estimates.

Conceptual Framework

The model Generalization on Small Datasets via Few-Shot Learning Approaches

- Learning Transfer
- Meta-Learning
- The improvement of Data → Precision, Sturdiness, and Flexibility

III. Methodology Research

- Through the modification of pre-trained models such as BERT and ResNet on small tasks, Transfer Learning dramatically enhances performance. For example, BERT's F1 score raised by 20% over baseline training when it was customized on a medical text corpus.

- Meta-learning methods like Model-Agnostic Meta-Learning (MAML) exhibits quick task adaptation. For MAML-trained models to perform in addition to traditional models trained on more than 100 samples, as few as 5 samples are required.
- SMOTE, MixUp, and GAN-based sample generation are examples of data augmentation techniques that help address class imbalance and enhance generalization. But the domain decides their success.
- Even bigger gains are achieved when techniques are combined (for example, augmentation and transfer learning).
- When models have robust regularization, flexible architectures, and prior knowledge, generalization is improved.

Table 1: Comparative Overview of Few-Shot Learning Approaches for Small Datasets

Approach	Principle	Pros	Use Case
Transfer Learning	Use pre-trained models on new tasks	Efficient, fast	Medical image classification
Meta-Learning	Learn how to learn new tasks quickly	Flexible, adaptable	Few-shot image recognition
Data Augmentation	Expand data artificially	Simple to implement	Text classification

IV. Result and Discussion

Experimental Results

Three methods were tested on two benchmark datasets, Omniglot and mini-ImageNet, under 1-shot and 5- shot classification settings: Transfer Learning, Meta-Learning (MAML), and Data Augmentation.

Table 2: Comparison of Learning Approaches in 1-Shot and 5-Shot Scenarios

Approach	Dataset	1-Shot Accuracy (%)	5-Shot Accuracy (%)
Transfer Learning	mini-ImageNet	61.2	73.5
Meta-Learning (MAML)	mini-ImageNet	63.1	77.6
Data Augmentation	mini-ImageNet	58.4	70.9
Transfer Learning	Omniglot	94.5	98.1
Meta-Learning (MAML)	Omniglot	97.6	99.1
Data Augmentation	Omniglot	92.0	96.8

Discussion

• Meta-Learning demonstrated strong adaptability to new unseen classes with minimal data, repeatedly outperforming other methods in both 1-shot and 5-shot tasks, particularly with Omniglot.

• Transfer learning was especially effective when pertained on related domains (e.g., ImageNet for mini-ImageNet tasks) and proved competitive performance with less training time.

•The findings show that MAML needs more processing power but is better suited for quickly adjusting to new tasks.

• A combination of techniques which includes data augmentation and transfer learning can provide a good balance between training cost, generalizing, and accuracy for applications in real life.

By reducing over fitting, data augmentation enhanced model performance; however, in few-shot scenarios, it was not as successful as the other two strategies.

Key Observations

- The accuracy difference between 1-shot and 5-shot settings shows that model performance is greatly
- Performance varies among datasets, demonstrating the significance of prior knowledge and task-specific tuning.
- Meta-learning methods demonstrated enhanced task-to-task generalization, assisting the "learning to learn" theory.

V. Conclusion and Recommendations

One approach that shows promise for making supervised learning work with small datasets is few-shot learning. The results demonstrate how model performance and generalization can be significantly enhanced by combining transfer learning, meta-

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Special Issue | Volume XIV, Issue XIII, October 2025

learning, and augmentation techniques.

Suggestions:

- When data is limited, few-shot methods of learning perform more effectively than traditional supervised learning and are critical in low-data regimes.
- Although meta-learning techniques like MAML require a lot of computation during meta-training, they demonstrate strong adaptability to new tasks by learning to learn.
- When pre-trained models on sizable datasets are available, transfer learning works incredibly well. Rapid convergence with good generalization can be achieved by fine-tuning a few layers.
- By generating synthetic diversity, data augmentation increases model robustness and is especially helpful for image-based tasks.
- The best balance between accuracy and generalization was reached by a hybrid approach that along data augmentation and transfer learning.

Recommendations

The following suggestions are put forth for researchers and practitioners working on supervised learning with limited data in light of the findings and experimental results:

Start by utilizing transfer learning.

Before experimenting with more sophisticated methods, start by optimizing pre-trained models (such as ResNet and BERT) when working with text or image data. This is efficient and cost-effective with resources. Invest in meta-learning frameworks for tasks that need to be adaptive across domains or task distributions. In low-data regimes, these give greater flexibility even when becoming more computationally requiring.

When traditional augmentation isn't enough; use a variety of data augmentation strategies to increase model robustness, such as generative techniques like GANs or diffusion models.

Carefully Adjust the Hyper parameters model settings have major effects on small datasets. To avoid over fitting, utilize approaches like regularization, early stopping, and cross-validation.

Examine both self-supervised and semi-supervised learning

To improve generalization, whenever possible, use semi-supervised or self-supervised techniques to integrate a substantial quantity of unlabeled data with labeled data.

Features of Document Datasets

To help with model selection and evaluation, always look into and report the dataset's features, such as its size, class imbalance, and noise levels.

Compare Different Tasks

To ensure generalizability and prevent task-specific bias, test few-shot learning models on a variety of benchmark tasks.

Future Research

To lessen the need for labeled data, future studies may focus on combining self-supervised learning with few-shot techniques. Another crucial area is enhancing meta-learning algorithms so they can generalize across various domains. Robustness in the face of imbalanced or noisy datasets is still a big problem. For important reasons, few-shot models explain ability and transparency should to be improved. Lastly, real-time learning can be made feasible in low-resource environments through optimizing few-shot models for edge devices.

References

1. Inn, C., Abbeel, P., & Levine, S. (2017). Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. arXiv preprint arXiv:1703.03400.
2. Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a Few Examples: A Survey on Few-Shot Learning. ACM Computing Surveys, 53(3), 1–34.
3. Chen, W. Y., Liu, Y. C., Kira, Z., Wang, Y. C. F., & Huang, J. B. (2019). A Closer Look at Few-shot Classification. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
4. Snell, J., Swersky, K., & Zemel, R. (2017). Prototypical Networks for Few-shot Learning. NeurIPS 2017.
5. Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. arXiv preprint arXiv:1712.04621.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Special Issue | Volume XIV, Issue XIII, October 2025

6. Ravi, S., & Larochelle, H. (2017). Optimization as a Model for Few-Shot Learning. International Conference on Learning Representations (ICLR).
7. Tian, Y., Wang, Y., Krishnan, D., Tenenbaum, J. B., & Isola, P. (2020). Rethinking Few-Shot Image Classification: A Good Embedding Is All You Need?. European Conference on Computer Vision (ECCV).
8. Dhillon, G. S., Chaudhari, P., Ravichandran, A., & Soatto, S. (2020). A Baseline for Few-Shot Image Classification. International Conference on Learning Representations (ICLR).