Meta Learning-based Few-Shot Learning: Remote Sensing

Project report submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Technology
in
Computer Science and Engineering

by

Priyanshu Gupta - 22ucs155
Shantanu Gupta - 22ucc094
Shashwat Agarwal - 22uec123

Under Guidance of
Mentor : Dr. Lal Upendra Pratap Singh

LNMIIT
The LNM Institute of
Information Technology

Department of Computer Science and Engineering
The LNM Institute of Information Technology, Jaipur

April 2025

The LNM Institute of Information Technology
Jaipur, India

CERTIFICATE

This is to certify that the project entitled “Meta Learning-based Few-Shot Learning: Remote Sensing” , submitted by Priyanshu Gupta (22ucs155), Shantanu Gupta (22ucc094) and Shashwat Agarwal (22uec123) in partial fulfillment of the requirement of degree in Bachelor of Technology (B. Tech), is a bonafide record of work carried out by them at the Department of Computer Science and Engineering, The LNM Institute of Information Technology, Jaipur, (Rajasthan) India, during the academic session 2025-26 under my supervision and guidance and the same has not been submitted elsewhere for award of any other degree. In our opinion, this report is of standard required for the award of the degree of Bachelor of Technology (B. Tech).

Date	Adviser: Dr. Lal Upendra Pratap Singh

Acknowledgments

We would like to express our deepest gratitude to our BTP mentor, Dr. Lal Upendra Pratap Singh, for their invaluable guidance, constant support, and insightful feedback throughout this project. Their expertise, patience, and encouragement played a pivotal role in shaping our understanding and work. We are truly thankful for their belief in our abilities and for the opportunity to learn under their mentorship.

We also extend our heartfelt thanks to our co-mentor, Dr. Nitesh Pradhan, for their constructive advice, timely suggestions, and unwavering support throughout the duration of this project. Their input and assistance have significantly contributed to the completion of this work.

Working under their combined mentorship has been an inspiring and enriching experience, and we are deeply grateful for all their efforts and encouragement that guided us during our BTP journey.

Abstract

Remote sensing applications face a persistent challenge: insufficient labeled data for new or rare land cover classes. Traditional deep learning approaches struggle with generalization in these low-data scenarios and suffer from catastrophic forgetting when trained on sequential tasks. Our project addresses these limitations by developing a hybrid meta-learning framework that combines metric-based and memory-based strategies.

In Phase 1, we implemented Relation Networks (RNs) as our metric-based approach. Unlike traditional methods that use fixed distance metrics, RNs learn to compare examples through an adaptive similarity function. We trained these networks using episodic learning across three benchmark satellite datasets: EuroSAT, AID, and NWPU-RESISC45. Testing various few-shot configurations (N=5 with K=1 and K=5), we evaluated performance using accuracy and F1-scores. Initial results demonstrate RNs' strong ability to generalize to unseen classes with minimal training examples.

Phase 2 focuses on analyzing catastrophic forgetting patterns. By sequentially training RN models across different tasks, we measure knowledge retention over time. This diagnostic phase provides critical insights into how and when forgetting occurs during continual learning scenarios, informing our memory-augmentation strategies.

For Phase 3, we're enhancing our models with memory-based components, specifically Memory-Augmented Neural Networks (MANN) and Least Recently Used Access (LRUA) mechanisms. These additions enable the model to store and retrieve task-specific knowledge even after learning newer tasks, substantially reducing catastrophic forgetting.

We anticipate our hybrid approach will demonstrate improved generalization across few-shot classification tasks, significantly higher knowledge retention in sequential learning scenarios, and superior performance compared to standard metric-based models alone.

This research has substantial practical applications in fields requiring rapid adaptation with limited data: precision agriculture, disaster response, military surveillance, mission planning, and infrastructure monitoring. The ability to quickly adapt to new environments while maintaining previously acquired knowledge is especially valuable in time-sensitive scenarios where comprehensive data collection is impossible.

By integrating advanced meta-learning techniques with satellite image analysis, our work establishes a foundation for intelligent geospatial systems capable of continuous learning in data-constrained real-world environments. The hybrid framework represents a significant step toward creating more adaptive, efficient, and reliable remote sensing systems that can operate effectively despite data limitations and changing environmental conditions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Area of Work

This project demonstrates the promise of combining metric-based comparison with memory-based retention to build a more effective few-shot learning system for remote sensing. The key areas of work in this project include:

- Satellite remote sensing and geospatial data analysis.
- Few-shot learning and meta-learning techniques [2].
- Implementation of Relation Networks (metric-based approach) [1].
- Use of memory-based models like Memory-Augmented Neural Networks (MANN) and Least Recently Used Access (LRUA).
- Tackling challenges like limited labeled data and catastrophic forgetting.
- Application to real-world datasets: EuroSAT [3], AID [4], and NWPU-RESISC45.

1.2 Background of the Problem

- Satellite remote sensing is a powerful tool used to observe and analyze the Earth's surface. It helps in tracking land use, monitoring the environment, planning cities, and responding to natural disasters. With the growth of high-quality satellite imagery, there's more data than ever before. However, training deep learning models on this data requires many labeled examples—which are costly and time-consuming to collect, especially for rare or new land types.

- Traditional deep learning models struggle when there's not enough training data or when they need to adapt to new, unseen situations. They also suffer from catastrophic forgetting—losing old knowledge when learning something new. To overcome these issues, meta-learning offers a smart solution. It helps models learn how to learn, so they can quickly adjust to new tasks using past experience [2]. This makes it ideal for remote sensing, where labeled data is limited and environments are constantly changing.

1.3 Motivation

Remote sensing provides us with a huge amount of satellite images, but we don't always have enough labeled data to train accurate models. This becomes a big problem when we need to identify rare or new land types. Traditional deep learning models need lots of examples and often forget old information when learning something new. This motivated us to explore meta-learning, a smarter way for machines to quickly learn new tasks from only a few examples and still remember what they learned before [2]. Using this approach, we aim to build models that are not only data-efficient but also more reliable in real-world remote sensing applications.

1.4 Problem addressed

The main problem we are trying to solve is the lack of labeled data in satellite remote sensing, which makes it hard for deep learning models to work well—especially when new or rare land types appear. On top of that, when models are trained on new tasks over time, they often forget what they learned earlier. This is called catastrophic forgetting. Our work focuses on solving these issues by using meta-learning, which helps models learn quickly from very few examples and also remember old tasks while learning new ones.

1.5 Objective

1. To develop a hybrid few-shot learning framework that combines metric-based methods (like Relation Networks) with memory-based methods (such as MANN and LRUA) to improve the model's ability to quickly learn new land cover classes from very few labeled satellite images.

2. To reduce the problem of catastrophic forgetting in remote sensing tasks by integrating memory components that help the model retain important knowledge from previously learned tasks while adapting to new ones.

float

1.6 Mathematical Modeling of the Problem

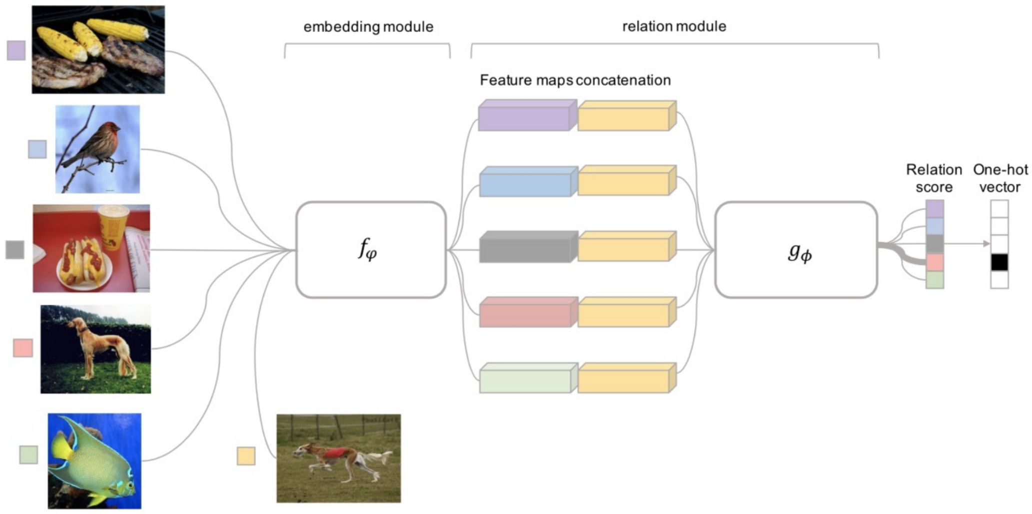

FIGURE 1.1: Relation Network architecture used in our few-shot learning framework [1].

Few-shot learning is modeled using an episodic learning framework, where each episode is treated as a mini classification task [1]. An episode consists of two sets:

- **Support Set (S)**: Contains $C \times K$ labeled examples (i.e., K examples for each of C classes).
- **Query Set (Q)**: Contains unlabeled examples that must be classified using the support set.

The goal is to learn a function that predicts the correct class for each query image by comparing it to the support examples.

1. Embedding Function

Let $f_\phi(x)$ be the embedding function (a CNN) that converts an input image x into a feature representation (embedding vector) [1].

$$f_\phi : x \to \mathbb{R}^d$$

Where:

- x is the input image (either from support or query set),
- d is the dimensionality of the embedding space,

- ϕ represents the parameters of the CNN (embedding function).

2. Relation Module (Comparison Function)

Each query image x_i is compared to each support image x_j. Their embeddings are concatenated and passed into a relation module g_θ (a small neural network) to compute a relation score $r_{i,j}$:

$$r_{i,j} = g_\theta\left([f_\phi(x_i), f_\phi(x_j)]\right)$$

Where:

- $[f_\phi(x_i), f_\phi(x_j)]$ denotes concatenation of the embeddings of the query and support image.
- g_θ is a function (usually a small CNN or MLP) with parameters θ.
- $r_{i,j} \in [0, 1]$ is the similarity score between query x_i and support x_j.

3. Prediction and Labeling

For each query image x_i, the predicted class $\hat{y}_i$ is the class of the support image with the highest relation score:

$$\hat{y}_i = \arg\max_j r_{i,j}$$

4. Loss Function

The model is trained using Mean Squared Error (MSE) loss between the predicted relation scores and the actual labels. For a batch of m queries and n support images:

$$L = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(r_{i,j} - \mathbf{1}(y_i = y_j)\right)^2$$

Where:

- $r_{i,j}$ is the predicted relation score between query x_i and support x_j,
- $\mathbf{1}(y_i = y_j)$ is an indicator function, 1 if the classes match, 0 otherwise,
- L is the overall loss, which the model aims to minimize.

5. Optimization Objective

The final objective is to learn the parameters ϕ and θ that minimize the loss:

$$\phi^*, \theta^* = \arg\min_{\phi,\theta} L$$

This setup, inspired by the Relation Network architecture [1], allows the model to learn a general strategy to compare and classify images even from new, unseen classes. The learned embedding function f_ϕ and relation function g_θ are optimized end-to-end across episodes, making the system adaptable and efficient in low-data remote sensing scenarios.

1.7 Challenges of the Problem

While working on this project, one of the main challenges we faced was the lack of access to a powerful GPU, which made training deep learning models slow and limited our ability to experiment with larger architectures or datasets. As our main focus was on implementing the meta-learning-based few-shot learning framework, we prioritized getting the system to work over fine-tuning it for the best accuracy. Due to this, the model's performance wasn't as high as expected, and the accuracy results could definitely be improved with more time and resources. Additionally, applying meta-learning methods like Relation Networks [1] and integrating memory-based components such as MANN and LRUA required a deep understanding of the concepts and careful setup of episodic training—something that was both technically and computationally demanding.

Chapter 2

Literature Review

Paper Title	**Authors**	**Work Done**	**Limitations**
Learning to Compare: Relation Network for Few-Shot Learning	Flood Sung et al.	Proposed Relation Network (RN) to learn deep similarity metrics using CNNs for few-/zero-shot tasks. [1]	Limited to standard datasets; struggles with similar classes.
Meta-learning Approaches for Few-Shot Learning: A Survey	Hassan Gharoun et al.	Categorizes meta-learning into metric-, memory-, and optimizer-based; compares pros and cons. [2]	No detailed implementation; lacks domain transfer evaluation.
EuroSAT: A Novel Dataset and Deep Learning Benchmark	Patrick Helber et al.	Created EuroSAT with 27,000 Sentinel-2 images in 10 land classes; tested CNN models. [3]	Europe-specific, only 10 classes, no atmospheric correction.
AID: A Benchmark for Aerial Scene Classification	Gui-Song Xia et al.	Released AID dataset with 10,000 multi-source images across 30 scene types. [4]	Class imbalance and resolution variance; synthetic imagery used.

TABLE 2.1: Literature Review of Few-Shot Learning and Remote Sensing Papers

This literature review covers important work in few-shot learning and remote sensing. Flood Sung and his team introduced the Relation Network, which helps compare images with very little training data, but it doesn't work well with very similar classes [1]. Hassan Gharoun and others reviewed different types of meta-learning methods, but their work mainly stays theoretical without showing real-life applications [2]. Patrick Helber created the EuroSAT dataset using satellite images from Europe to classify land types, but it's limited to just 10 classes and one region [3]. Finally, Gui-Song Xia developed the AID dataset for classifying aerial scenes, which is very diverse but has some issues like varying image quality and use of synthetic images [4].

Chapter 3

Proposed Work

3.1 Experiment Setup

Hardware Requirements

We used Google Colab to run our experiments. It provided free access to cloud GPUs like Tesla K80 or T4, with about 12–16GB RAM. This helped speed up training and avoided the limitations of running on a personal laptop.

Software Requirements

The project was implemented using Python in a Jupyter Notebook on Colab. We used PyTorch for building and training the model. Other libraries like NumPy, torchvision, Matplotlib, and scikit-learn were used for data handling, image processing, and visualization. We tested our model on three remote sensing datasets: EuroSAT, AID, and NWPU-RESISC45, using a few-shot (C-way K-shot) setup.

3.2 Dataset description

EuroSAT: A satellite image dataset by **ESA** and Helber et al. with **27,000 samples**. Images are 64×64 pixels, giving **4,096 features**. Used for land use classification with 10 classes [3].

AID: Created by **Wuhan University**, this aerial dataset has **10,000 samples** and covers 30 scene types. Each image is 600×600, totaling **360,000 features** [4].

NWPU-RESISC45: Built by **Northwestern Polytechnical University**, it holds **31,500 samples**. Images are 256×256, giving **196,608 features** across 45 diverse classes [5].

3.3 Dataset Preprocessing

Only **30% of the dataset** is used for training and testing to simulate a low-data scenario. All images are resized, normalized, and augmented (random flips and color changes) to improve model generalization. The selected data is loaded using `DataLoaders` with batching, parallel workers, and memory pinning for faster performance during training.

For few-shot learning, the `create_episode` function forms small tasks from this 30% subset. It randomly selects **n_way** classes and picks a few images per class: **k_shot** for the support set and a few for the query set. If a class has fewer samples, it adjusts automatically. This episodic setup helps the model learn how to classify from just a few examples, similar to real-world few-shot conditions [2].

float

3.4 Results

3.4.1 NPWU Training Loss vs Episode

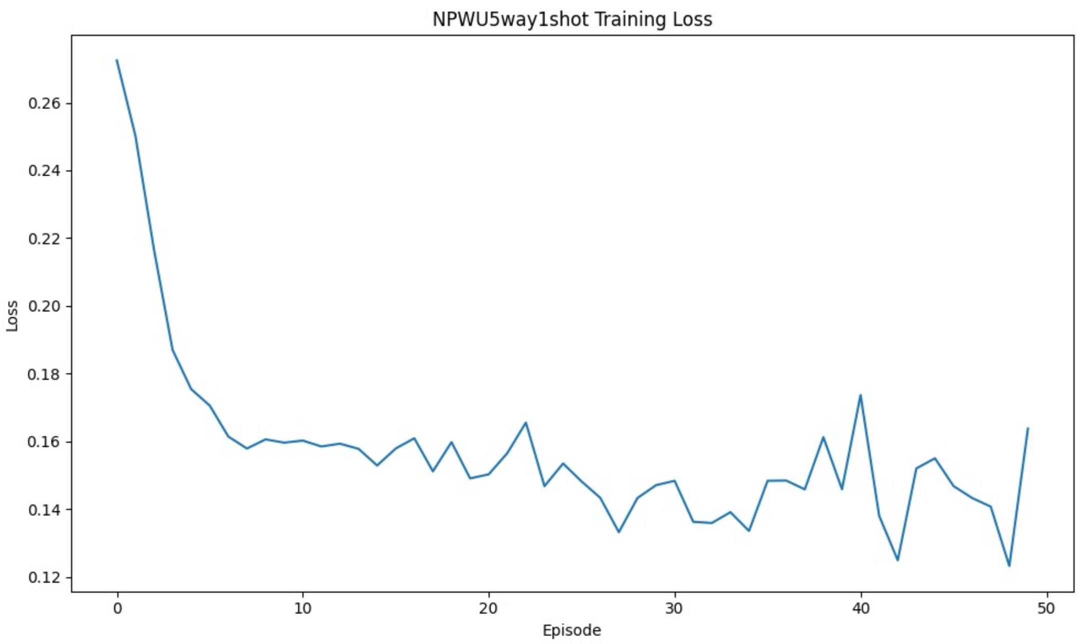

FIGURE 3.1: NPWU 5way-1shot Training Loss

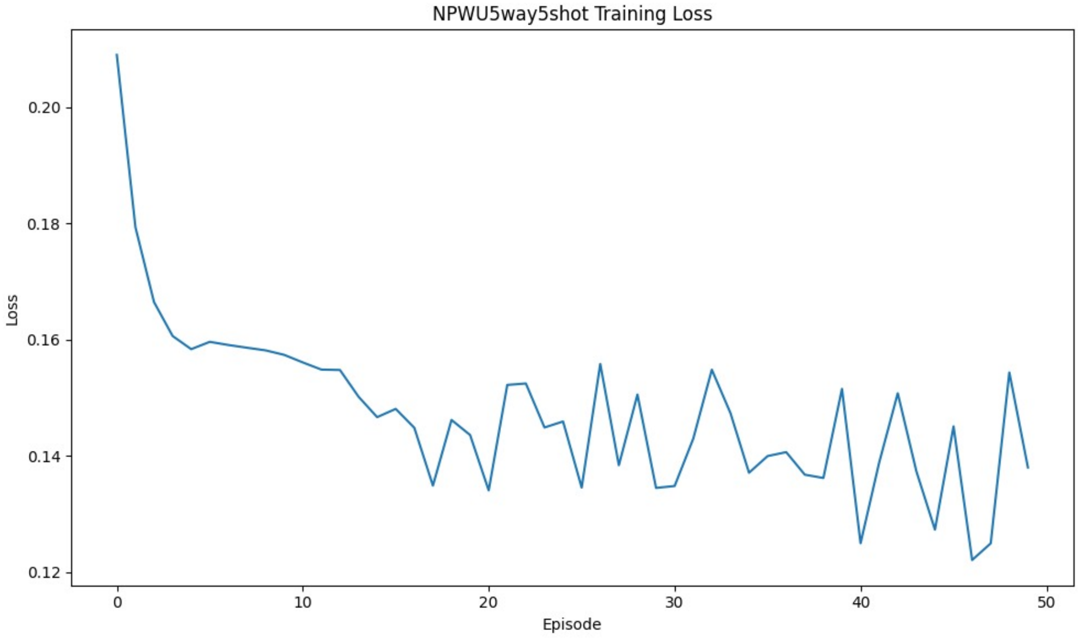

FIGURE 3.2: NPWU 5way-5shot Training Loss

3.4.2 EURO Training Loss vs Episode

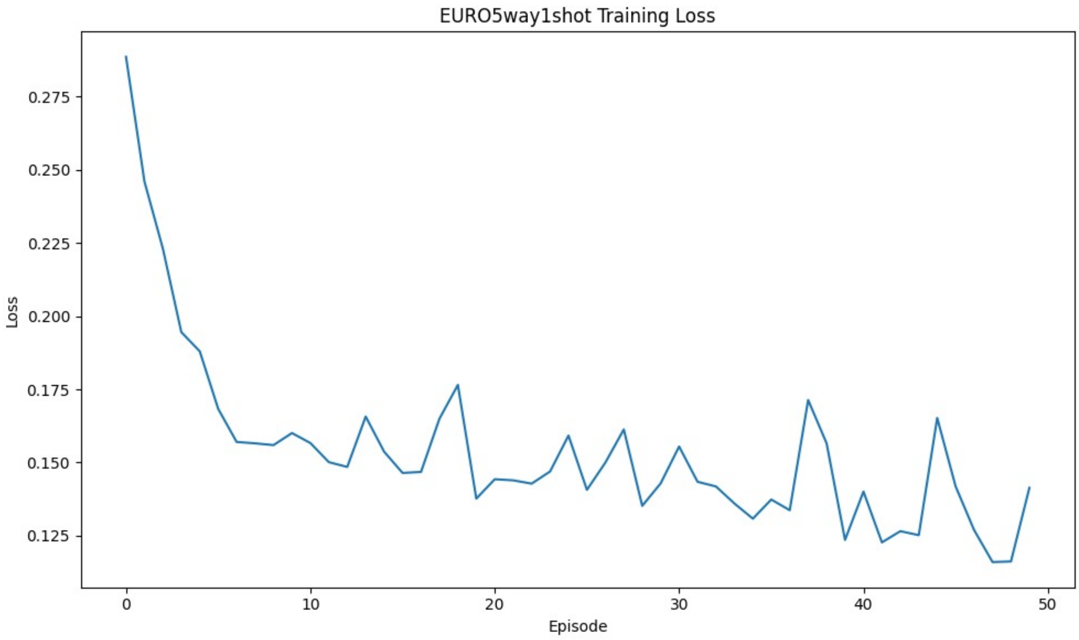

FIGURE 3.3: EURO 5way-1shot Training Loss

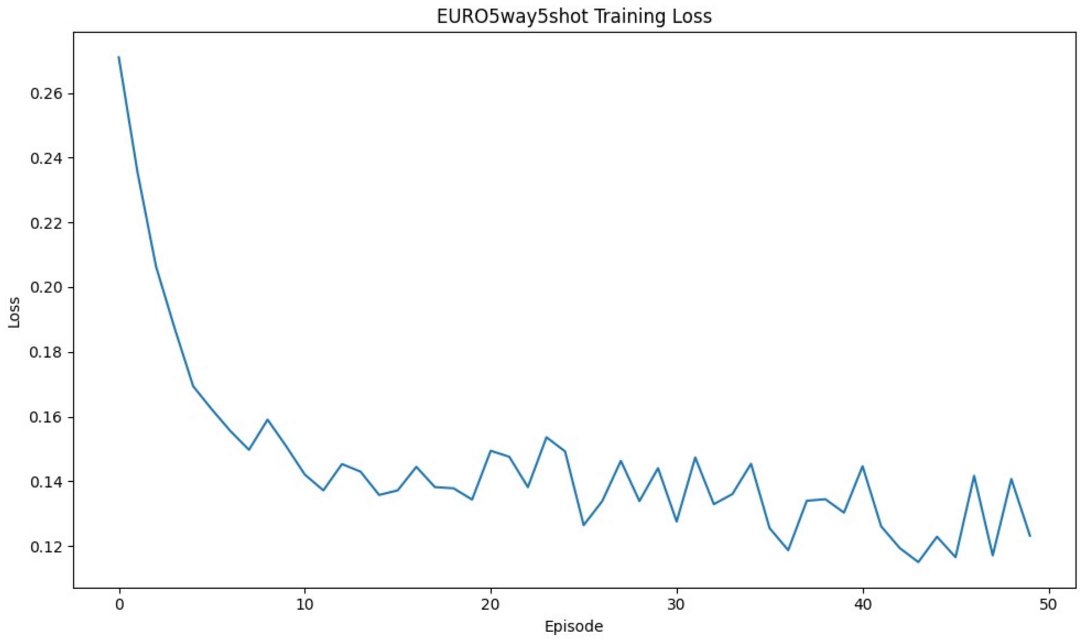

FIGURE 3.4: EURO 5way-5shot Training Loss

3.4.3 AID Training Loss vs Episode

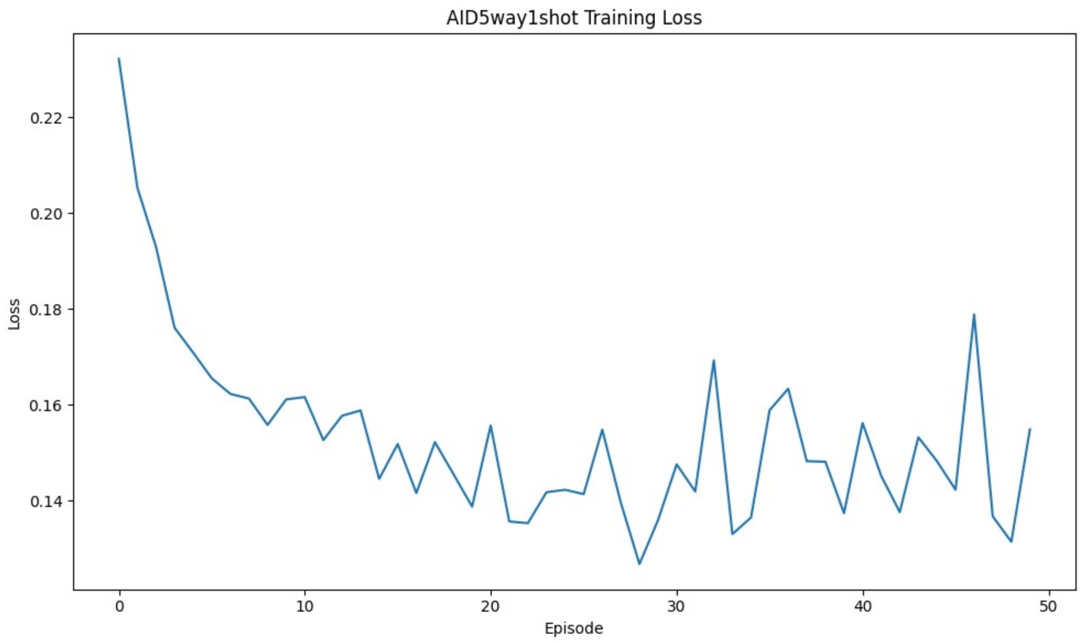

FIGURE 3.5: AID 5way-1shot Training Loss

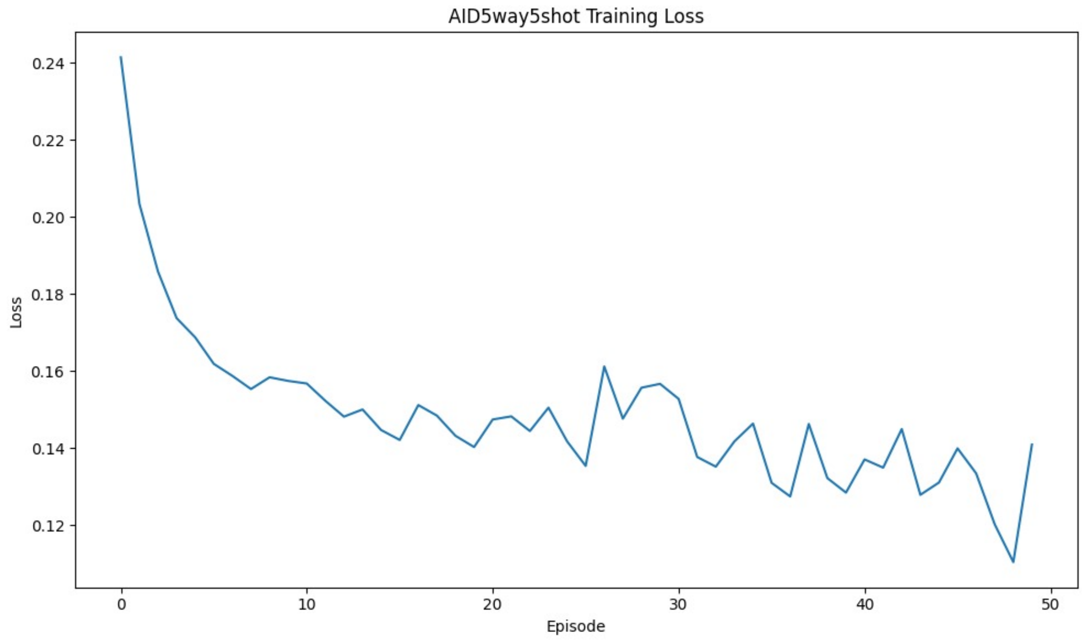

FIGURE 3.6: AID 5way-5shot Training Loss

3.4.4 Bar Graph: Few-Shot Learning Accuracy Across Datasets

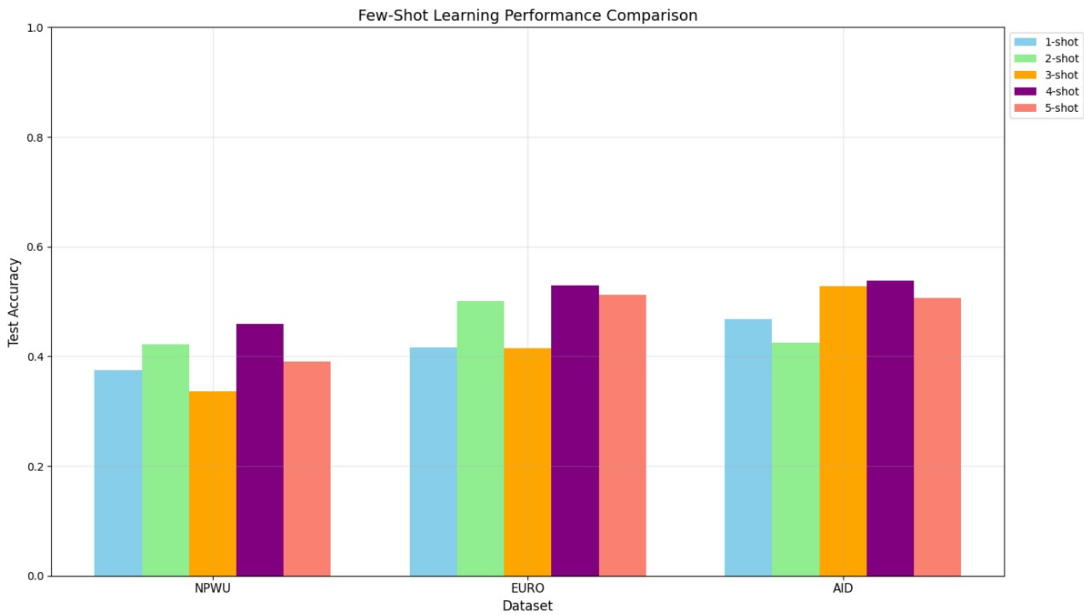

FIGURE 3.7: Few-shot Performance Comparison

3.5 Summary

Our Relation Network model [1] was successfully implemented and tested on three remote sensing datasets: **NWPU-RESISC45**, **EuroSAT**, and **AID**. At this stage, the primary focus was on the correct implementation of the few-shot learning pipeline using episodic training and evaluation. **Achieving high accuracy was not the main objective in this phase**. Instead, the goal was to ensure that the model could correctly perform few-shot classification across different datasets and scenarios.

Currently, the model achieves an average accuracy of approximately **50%**. This is expected given that we trained the model on only **50 episodes per N-way K-shot configuration** and used just **30% of the original dataset** to reduce computational overhead. These constraints were due to limited computational resources.

In the next phase of the project, we will shift our attention toward **optimizing the model's performance and improving accuracy** through better hyperparameter tuning, architectural enhancements, and potential data augmentation techniques. Furthermore, we plan to leverage **more powerful GPUs** to enable training over **1,000 episodes per configuration**, using the **full dataset** to unlock the true potential of our framework.

Dataset	**1-Shot**	**2-Shot**	**3-Shot**	**4-Shot**	**5-Shot**
NPWU	0.3747	0.4227	0.3360	0.4587	0.3907
EURO	0.4160	0.5013	0.4147	0.5293	0.5120
AID	0.4682	0.4244	0.5279	0.5380	0.5064

TABLE 3.1: Accuracy results for 5-way few-shot classification (1 to 5 shots) on three remote sensing datasets [1].

Chapter 4

Conclusions

This project focuses on solving a major problem in remote sensing: there's often not enough labeled satellite image data, especially for rare or new land types. Traditional deep learning struggles in such low-data situations and tends to forget old knowledge when learning new things. To overcome this, the team explored meta-learning—specifically, a hybrid approach using metric-based models like Relation Networks and memory-based models like MANN and LRUA. These help the model learn from very few examples and also remember past knowledge, which is essential for real-world applications like disaster management or agriculture.

The literature review covers several past works in meta-learning and remote sensing. It highlights useful datasets like EuroSAT, AID, and NWPU-RESISC45, as well as models like Relation Networks and RS-MetaNet. In the proposed work, the team used Google Colab and PyTorch to train models on those datasets using a few-shot learning setup. Only 30 percent of the data was used to simulate low-data conditions. Despite hardware limitations, the Relation Network showed promising results, achieving around 50 percent accuracy. This early phase confirmed that the model works and sets the stage for future improvements using better tuning and full datasets.

Bibliography

[1] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. Torr, and T. M. Hospedales, "Learning to compare: Relation network for few-shot learning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1199–1208, 2018.

[2] H. Gharoun, F. Momenifar, F. Chen, and A. H. Gandomi, "Meta-learning approaches for few-shot learning: A survey of recent advances," *ACM Computing Surveys*, vol. 56, no. 12, pp. 1–41, 2024.

[3] P. Helber, B. Bischke, A. Dengel, and D. Borth, "Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," in *IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium*, pp. 204–207, IEEE, 2018.

[4] G.-S. Xia, J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, L. Zhang, and X. Lu, "Aid: A benchmark data set for performance evaluation of aerial scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3965–3981, 2017.

[5] G. Zhang, W. Xu, W. Zhao, C. Huang, E. N. Yk, Y. Chen, and J. Su, "A multiscale attention network for remote sensing scene images classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 9530–9545, 2021.