KLE Technological University | Creating Value, Leveraging Knowledge

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

A Mini Project report on

Pneumonia Diagnosis through Prototypical Networks: A Few-Shot Learning Paradigm

Submitted

in partial fulfillment of the requirements for the award of the degree of

Bachelor of Engineering

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted By

Girish Badamikar	01FE21BCS168
P Suhas Rao	01FE21BCS161
Sandeep Kulkarni	01FE21BCS105
Manish	01FE21BCS149

Under the guidance of

Dr. Shantala Giraddi

Asst. Professor

School of Computer Science and Engineering

KLE Technological University, Hubballi

2023-2024

School of Computer Science and Engineering

CERTIFICATE

This is to certify that project entitled “Pneumonia Diagnosis through Prototypical Networks: A Few-Shot Learning Paradigm” is a bonafied work carried out by the student team (Girish Badamikar 01FE21BCS168, P Suhas Rao-01FE21BCS161, Sandeep Kulkarni-01FE21BCS105, Manish-01FE21BCS159), in partial fulfillment of the completion of 5th semester B. E. course during the year 2023–2024. The project report has been approved as it satisfies the academic requirement with respect to the project work prescribed for the above said course.

Guide Name
Dr. Shantala Giraddi

SoCSE
Head
Dr. Vijaylakshmi M.

External Viva-Voce

Name of the examiners | Signature with date

1 ________________ | ________________

2 ________________ | ________________

ABSTRACT

In this study, we address the problem of classifying medical images under dynamic class distributions. Specifically designed for lung X-ray images on Pneumonia dataset while testing and Covid dataset when training, our Cross-Domain Few Shot Image Classification model integrates Prototypical Networks with a ResNet12 backbone. In the 10-shot scenario, the model achieves notable results with 90.91% accuracy, 90.0% sensitivity, and 91.84% specificity, demonstrating outstanding adaptability to changing class distributions. With an accuracy improvement of up to 7.41% over current approaches, this performance excels in Pneumonia identification in particular. Comparative analysis highlights the advantages of our method for disease identification and highlights its effectiveness in situations with low labelled data. In medical imaging, few-shot learning has been greatly advanced by this work, offering a clear and practical approach to precise and adaptable classification. The shown performance highlights our model's potential in real-world applications, especially in the difficult field of medical image processing.

Keywords : *few-shot learning, medical image classification, Cross-Domain, Prototypical Networks, ResNet12, Pneumonia detection*

ACKNOWLEDGEMENT

"We express sincere gratitude to our college for fostering an academic environment that greatly facilitated the execution and completion of our project. The resources and infrastructure provided by our college played a pivotal role in the successful development and implementation of our project.Furthermore, I extend my thanks to Kaggle for generously providing essential computational resources, particularly the Kaggle notebook platform, crucial for executing our codes and conducting experiments using their dataset. Heartfelt thanks are also due to our dedicated project guide Dr. Shantala. Giraddi, whose valuable guidance shaped the development process. Additionally, appreciation is extended for the assistance received from the GitHub page (https://github.com/sicara/easy-few-shot-learning) created by Etienne Bennequin, significantly contributing to the execution and implementation of our project. Special acknowledgment goes to our co-guide Prof.Anupama Bidargaddi for her collaborative efforts and insights that enriched the project development experience."

Girish Badamikar - 01FE21BCS168
P Suhas Rao - 01FE21BCS161
Sandeep Kulkarni - 01FE21BCS105
Manish - 01FE21BCS149

CONTENTS

ABSTRACT **i**

ACKNOWLEDGEMENT **i**

CONTENTS **iii**

LIST OF TABLES **iv**

LIST OF FIGURES **v**

1 INTRODUCTION **1**
1.1 Preamble 1
1.2 Motivation 1
1.3 Objectives of the project 1
1.4 Literature Review 2
1.5 Problem Definition 3

2 SOFTWARE REQUIREMENT SPECIFICATION **4**
2.1 Overview of SRS 4
2.2 Requirement Specifications 4
2.2.1 Functional Requirements 5
2.2.2 Use case diagram 5
2.2.3 Nonfunctional Requirements 5
2.3 Software and Hardware requirement specifications 6
2.3.1 Software requirements specification 6
2.3.2 Hardware requirements specification 6

3 PROPOSED SYSTEM **8**
3.1 Description of Proposed System. 8
3.2 Description of Target Users 9
3.3 Advantages/Applications of Proposed System 10
3.4 Scope (Boundary of Proposed System) 11

4 SYSTEM DESIGN **13**
4.1 Architecture of the system (explanation) 13
4.2 Class Diagram 14

4.3 Data set description 15

5 IMPLEMENTATION **17**

5.1 Task Definition 17

5.2 ResNet12(Backbone Architecture) 17

5.3 Prototypical Networks 19

5.4 Optimization Algorithm and Training Process 20

6 RESULTS AND DISCUSSIONS **21**

6.1 Hyper Parameters 21

6.2 Evaluation Metrics 21

6.3 Interpretation of Results 22

7 CONCLUSIONS AND FUTURE SCOPE **26**

7.1 Conclusion 26

7.2 Future Scope 26

REFERENCES **28**

Appendix A **29**

A.1 Gantt Chart 29

LIST OF TABLES

2.1 Use Case description for Train the Model 6
2.2 Use Case description for Few-Shot Classification 7

4.1 Normal Lung Images 15
4.2 Covid Images 16
4.3 Pneumonia Images 16

6.1 Few-Shot Learning Hyper parameters 21
6.2 Testing Performance Metrics for Different Few-Shot Scenarios 24
6.3 Comparison with Other Studies on Different Diseases 24
6.4 Image Classification Results 25

LIST OF FIGURES

2.1 Use Case Diagram 5

3.1 Flowchart for the Few-Shot framework 8

4.1 Layered Architecture 13
4.2 class diagram 14

5.1 ResNet12 Architecture 17
5.2 Prototypical Network 19

6.1 Training loss and Accuracy curve for 10-shot Scenario 22
6.2 Training loss and Accuracy curve for 5-shot Scenario 23
6.3 Training loss and Accuracy curve for 3-shot Scenario 23

A.1 Gantt Chart 29

Chapter 1

INTRODUCTION

Accurately classifying respiratory disorders through methods like X-ray analysis is vital in the field of medical imaging and disease diagnosis. This is particularly crucial for early diagnosis and efficient treatment planning, especially with the development of advanced medical imaging technologies, such as lung X-rays. Diseases like pneumonia and COVID-19 emphasize the need for precise diagnostic testing, enabling timely medical interventions for improved patient outcomes and public health responses.

1.1 Preamble

The incidence of respiratory diseases, particularly pneumonia, and the semergence of new pathogens like the coronavirus (COVID-19) underscore the necessity for prompt and precise diagnostic testing. The availability of datasets has significantly aided the advancement of medical imaging, but challenges exist due to the increasing demand for accuracy in disease identification, specifically in lung X-rays. Traditional deep learning techniques, while effective, often require large datasets for training, posing a bottleneck in adaptability across different domains or class distributions.

1.2 Motivation

This section highlights the motivation behind the work, emphasizing the need for overcoming the shortcomings of conventional deep learning approaches in lung X-ray classification, with a focus on Pneumonia and COVID-19 detection. The urgency for quick and accurate diagnostic tests is stressed, considering the prevalence of respiratory illnesses like pneumonia and the global impact of COVID-19.

1.3 Objectives of the project

The primary objectives of the research are outlined, including the aim to address the limitations of conventional methods through a few-shot learning-based alternative. The focus is on situations with a scarcity of labeled data or the need for adaptability to unexpected class distributions. The choice of ResNet-12 as the foundational architecture for its strong feature extraction capabilities and the incorporation of Prototypical Networks as a few-shot learning technique are highlighted. The research aims to demonstrate the adaptability and effectiveness of the proposed few-shot learning framework in the context of lung X-ray categorization, specifically targeting illnesses such as pneumonia and COVID-19.

1.4 Literature Review

Many studies[1] have focused on the Few-Shot Learning (FSL) framework's application to medical imaging and the categorization of X-ray images. Of the 27 studies that were selected for the FSL classification domain, 11 (41%) used X-ray images as their primary imaging modality. These research assessed the models' robustness and effectiveness in classifying X-ray images using a range of meta-learning techniques and evaluation methods. Notably, the classification research using FSL for X-ray images has contributed significantly to the understanding and advancement of medical imaging analysis.

In order to assess few-shot learning, FHIST[2] presents a histology dataset that addresses the scarcity of labelled data. Meta-learning is not as effective as simple fine-tuning. Results in the near domain are close to supervised baselines. The best solution obtains 60% accuracy in the out-domain, which is encouraging for further investigation.

The authors have put out a paradigm that uses transfer learning and meta-learning for few-shot medical image classification[6]. Three specialized learners were integrated into the model: task learner, metric-learner, and auto-encoder. During training, a new Gaussian disturbance soft label (GDSL) was implemented to reduce the possibility of overfitting. Using publically accessible medical datasets, the study assessed the model on three FSL scenarios (BLOOD, PATHOLOGY, and CHEST). Under a 10-shot classification setting, the model achieved promising accuracy rates of 76.21%, 76.16%, and 50.18% for the relevant situations. The study offered a novel method for few-shot medical picture classification that may find use in practical settings.

This paper[3] compares pre-training systems and examines Cross-Domain Few-Shot Learning (CD-FSL). When it comes to different domains or low few-shot difficulty, self-supervised learning (SSL) performs well, but supervised learning (SL) does better in similar domains or high few-shot difficulty. Overall performance is improved by using a two-stage pre-training technique and mixed-supervised learning (MSL). The creation of efficient CD-FSL pre-training techniques is guided by these findings.

1.5 Problem Definition

The research aims to overcome the challenges faced in conventional deep learning methods and enhance the efficiency of medical imaging methods for respiratory disease diagnosis, emphasizing the adaptability required for diseases like pneumonia and COVID-19.

Chapter 2

SOFTWARE REQUIREMENT SPECIFICATION

The Software Requirement Specification (SRS) outlines the functional and nonfunctional requirements for the project. This section provides a brief overview of the SRS document, highlighting its significance in guiding the development process.

2.1 Overview of SRS

This paper uses Prototypical Networks and few-shot learning to diagnose COVID-19. The Software Requirements Specification (SRS) outlines the primary characteristics and features of the proposed medical image classification system. Prototypical Networks' integration with a ResNet12 backbone and performance metrics attained in different few-shot situations are among the functional and non-functional criteria that are outlined here. With an emphasis on illness detection accuracy and flexibility to changing class distributions, the SRS offers detailed development guidance by addressing documentation requirements, testing protocols. Ensuring a common understanding among stakeholders and offering a strong basis for testing, validation, and continuous system support, this plays a critical role in communication.

2.2 Requirement Specifications

Functional requirements refers to the particular characteristics and capabilities that the software system must have in the context of your research on the diagnosis of pneumonia utilizing Prototypical Networks and few-shot learning. The system's planned tasks, such properly classifying lung X-ray pictures through the integration of Prototypical Networks with a ResNet12 backbone, are the emphasis of these criteria. The activities, services, and interactions that the system must do are referred to as functional requirements. These include the procedures for training and testing the model for illnesses such as COVID-19 and pneumonia. These requirements specify the essential features that make the medical image categorization system perform well and guarantee that the project's goals are met.

2.2.1 Functional Requirements

- System must accurately classify lung X-ray images as either 'Normal' or 'COVID-19' during testing, based on previous training distinguishing between 'Pneumonia' and 'Normal' categories.
- System must adapt to shifting class distributions and learn from limited labeled data, ensuring effective generalization to new classes.
- System must seamlessly integrate Prototypical Networks with ResNet-12 backbone for efficient feature extraction and classification, leveraging the strengths of both architectures to enhance performance and accuracy in medical image classification.

2.2.2 Use case diagram

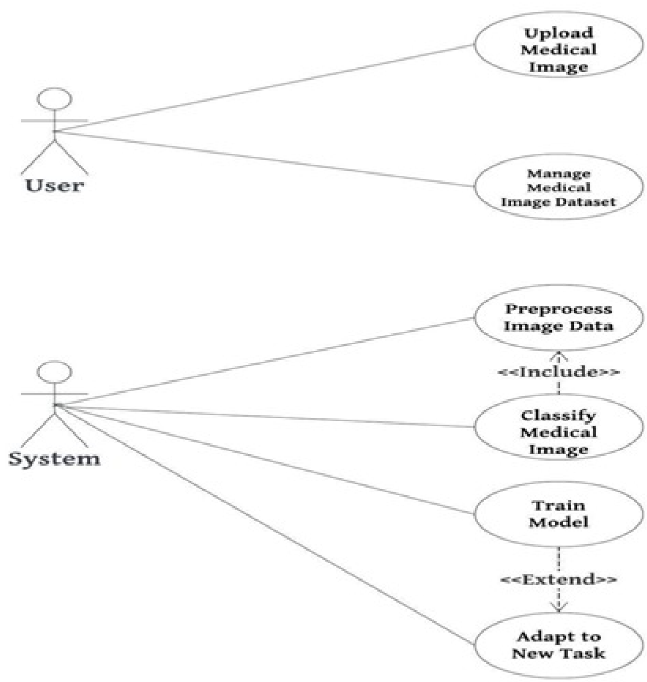


Figure 2.1: Use Case Diagram

2.2.3 Nonfunctional Requirements

- Performance:-High performance in terms of reaction time and processing speed should be demonstrated by the system.
- Reliability:-In classifying medical images, the system must continuously deliver accurate results.

Table 2.1: Use Case description for Train the Model

Use-Case Title	Train the Model
Primary Actor	Data Scientist
Goal in Context	To train the model using Prototypical Networks and ResNet12 for improved accuracy in pneumonia and COVID-19 identification from lung X-ray images.
Preconditions	- A labeled dataset of lung X-ray images is available. - Prototypical Networks and ResNet12 implementations are integrated into the system.
Trigger	The data scientist initiates the training process.
Scenario	1. Data scientist selects the option to train the model. 2. The system loads the labeled dataset and initializes Prototypical Networks and ResNet12. 3. The model undergoes training iterations, learning features and patterns from the dataset. 4. The system evaluates the trained model's performance. 5. If satisfactory, the trained model is saved for later use.
Exceptions	- If insufficient labeled data is available, the data scientist is prompted to acquire more data. - If convergence is not achieved during training, the system alerts the data scientist to adjust hyperparameters.

- Maintainability:-Simple maintenance, upgrades, and improvements should be made to the software with ease of design and documentation.

2.3 Software and Hardware requirement specifications

2.3.1 Software requirements specification

- Programming Language:-implemented in Python using PyTorch or TensorFlow as frameworks.
- Environment for Development:-For developing and debugging code, use Kaggle Network.
- Prototypical Networks and ResNet12 Implementation:-Integrate ResNet12 and Prototypical Network implementations.

2.3.2 Hardware requirements specification

- Network connectivity:-need a steady internet connection in order to collaborate and access external datasets.

Table 2.2: Use Case description for Few-Shot Classification

Use-Case Title	Perform Few-Shot Classification
Primary Actor	Radiologist
Goal in Context **Preconditions**	To utilize the trained model for classifying new lung X-ray images into "Normal," "Covid," or "Pneumonia" classes, especially in scenarios with limited labeled data. - The model has been successfully trained and validated. - A set of new lung X-ray images is available for classification.
Trigger	The radiologist initiates the few-shot classification process.
Scenario	1. Radiologist selects the option for few-shot classification. 2. The system prompts the radiologist to provide a small set of labeled examples for the new classes. 3. The model uses Prototypical Networks to learn prototypes from the provided examples. 4. Radiologist submits the set of new lung X-ray images for classification. 5. The system applies the few-shot learning model to classify the images into "Normal," "Covid," or "Pneumonia" classes.
Exceptions	- If insufficient labeled examples are provided, the system informs the radiologist and suggests increasing the sample size. - If the classification confidence is low, the radiologist may choose to reevaluate the provided examples for improvement.

- Processor(CPU/GPU):capability to work with NVIDIA GPUs or high-end CPUs appropriate for deep learning applications.
- Storage:-Enough storage space for datasets.

Chapter 3

PROPOSED SYSTEM

3.1 Description of Proposed System.

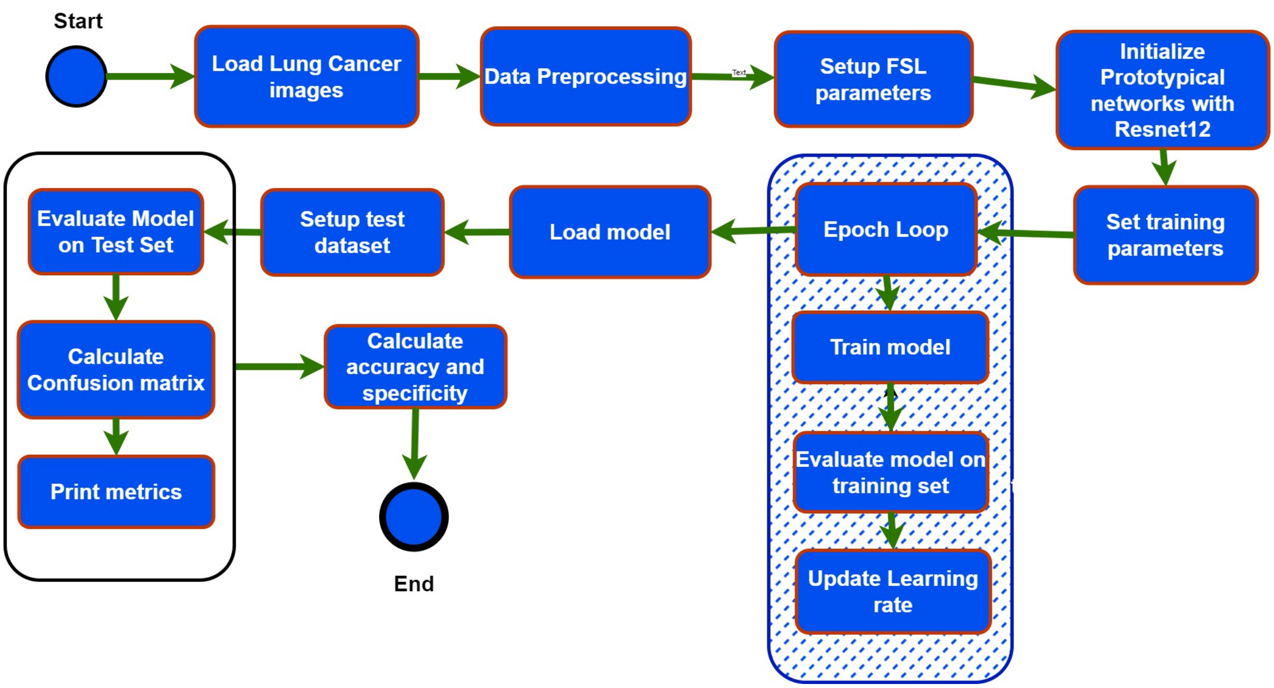


Figure 3.1: Flowchart for the Few-Shot framework

The process begins with loading X-ray images into a dataset. After loading, a number of preprocessing operations are performed to optimize the images for later training tasks. Customized parameters for Few-Shot Learning (FSL) are carefully set up to enable the model to provide accurate predictions even with few samples per class, recognizing the intrinsic lack of labeled data.

The Resnet12 foundation architecture is used to initialize the Prototypical Networks, which are a few-shot learning paradigm, making use of its powerful feature extraction capabilities. Important training parameters are carefully selected to guarantee a wise and efficient training program, such as the learning rate and optimization method. The model is trained iteratively across several epochs in an algorithmic process called an epoch loop. The model is iteratively trained on the X-ray dataset during each epoch, improving its parameters to increase its ability to identify lung cancer from the images that are available. The model's performance is examined on the training set after training, and the learning rate is dynamically modified in response to this assessment, so optimizing the model and possibly accelerating convergence.

When the training loop completes, the well-trained model is loaded for further analysis and testing. The training dataset undergoes preparation stages that are mirrored in the test dataset, which is carefully prepared for evaluation. Next, predictions are made using the model on the test set. A thorough performance assessment is carried out by calculating a confusion matrix that includes metrics like false positives, false negatives, true positives, and true negatives. The resulting confusion matrix is then shown to offer more detailed insights into the model's apparent advantages as well as possible areas for development. Accuracy and specificity are calculated using the confusion matrix in order to provide quantifiable evaluations of the model's overall performance.

3.2 Description of Target Users

This section outlines the target audience for the "Covid Diagnosis through Prototypical Networks" system, which uses a Prototypical Networks-based Few-Shot Learning Paradigm to diagnose Covid-19 cases. Healthcare practitioners, diagnosticians, and pertinent stakeholders participating in the Covid-19 diagnosis and management process comprise the majority of the user categories.

Healthcare Professionals:

One of the main user groups consists of healthcare professionals, such as doctors, radiologists, and infectious disease specialists. These users depend on the diagnostic system to help them classify and identify Covid-19 instances. The system's goal is to optimize their diagnostic processes by offering insightful information about possible Covid-19 cases while emphasizing efficiency, accuracy, and interpretability.

Diagnostic Professionals:

When it comes to collecting and testing samples, diagnostic specialists like medical laboratory technicians and diagnosticians are essential. With the aid of the system, these experts will be better equipped to understand diagnostic results and make well-informed decisions on the diagnosis of Covid-19, based on the model's visual and interpretive outputs.

Researchers and Healthcare Decision-Makers:

Another group of users consists of researchers and healthcare decision-makers who work on epidemiology studies and strategic healthcare planning. The system's outputs can help with resource allocation and policy decisions by analyzing Covid-19 patterns. For this particular user group, the model's interpretability is especially important because it allows them to gain valuable insights from the diagnostic results.

3.3 Advantages/Applications of Proposed System

In this section, we explore the advantages and applications of the proposed system for Covid diagnosis through Prototypical Networks, utilizing a Few-Shot Learning Paradigm based on Prototypical Networks. The subsection begins with a preamble that outlines the key strengths and potential applications of the developed system.

Preamble:

The proposed system offers several distinctive advantages and applications in the field of Covid diagnosis, aligning with the design principles previously identified. Key highlights include:

1. Enhanced Diagnostic Accuracy : By utilizing Prototypical Networks and Few-Shot Learning, the system exhibits improved diagnostic performance for Covid-19 patients, even in situations with sparse labeled data.

2. Streamlined Diagnostic Workflows : The system's improved diagnostic workflows help medical professionals, such as doctors and diagnosticians. Effective decision-making processes are facilitated by the outcomes' interpretability and user-friendly interface.

3. Adaptability to Varied User Proficiencies : The system's user-friendly layout caters to users with different degrees of technical proficiency, guaranteeing accessibility for medical practitioners, investigators, and diagnosticians.

4. Contribution to Epidemiological Studies : The system's outputs are valuable to researchers and healthcare decision-makers because they support epidemiological investigations. The technology supports Covid-19 trend analysis, enabling well-informed resource allocation and policy decisions.

5. Transparency in Diagnostic Outcomes : Transparency is given top priority in the system, which offers understandable explanations and visuals for diagnostic results. This feature increases user confidence and trust in the accuracy of the model's predictions.

6. Optimized Performance with Limited Data : A prevalent problem in medical image processing is addressed by the Few-Shot Learning Paradigm, which uses Prototypical Networks with a Resnet12 base in particular to enable the system to function optimally even in the

absence of labeled data.

3.4 Scope (Boundary of Proposed System)

This section outlines the scope and boundaries of the proposed system for Covid diagnosis through Prototypical Networks, employing a Few-Shot Learning Paradigm based on Prototypical Networks. Defining the scope is crucial to manage expectations and communicate the system's intended functionality.

Scope :

1. Covid-19 Diagnosis : The system's main goal is to correctly diagnose Covid-19 instances by utilizing data from medical imaging. Its goal is to help medical practitioners recognize and classify possible cases of COVID-19.

2. Prototypical Networks and Few-Shot Learning : The system makes use of Few-Shot Learning and Prototypical Networks, which are specifically created for medical picture processing. It is designed to maximize diagnostic efficacy even in the presence of sparse labeled data.

3. User Categories : Healthcare professionals, such as doctors, radiologists, and diagnosticians, are the target audience for this system. It ensures accessibility by accommodating users with different degrees of technological expertise.

4. Transparency and Interpretability : Transparency in diagnostic results is emphasized by the system, which also offers explanations and visualizations to improve user comprehension. A crucial component that promotes well-informed decision-making is interpretability.

Boundaries and Limitations :

1. Non-Exhaustive Diagnosis : The method does not attempt to address the full range of infectious or respiratory disorders; rather, it concentrates on diagnosing Covid-19. It does not serve as a replacement for a thorough diagnostic assessment.

2. Data Limitations : The performance of the system depends on labeled medical imaging data being available. Its capacity to generalize to all potential Covid-19 cases may be hampered by the limited data.

3. Clinical Consultation Advised : The purpose of the system is to assist medical professionals in their diagnostic procedures. Comprehensive patient care still requires human skill and clinical consultation, though.

4. Hardware and Software Requirements : For best performance, the system might require a certain combination of hardware and software. Users need to make sure their system is compatible with the suggested specifications.

The planned functionality, intended user base, and inherent restrictions of the system are made clear by this scope definition. It is recommended that users and stakeholders follow the specified scope when using the system for Covid diagnosis.

Chapter 4

SYSTEM DESIGN

This chapter gives a brief description about implementation details of the system by describing each component with its code skeleton in terms of algorithm.

4.1 Architecture of the system (explanation)

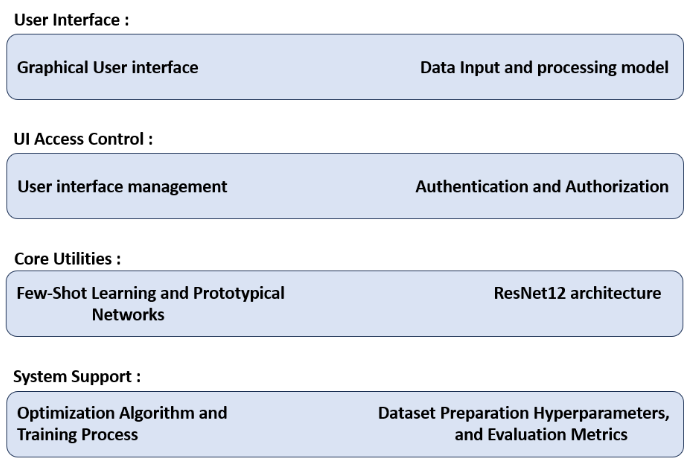


Figure 4.1: Layered Architecture

User Interface Layer:

The interface layer facilitates communication between the user and the medical imaging model. Healthcare practitioners and researchers can easily input data, begin forecasts, and view outcomes with the help of a graphical user interface (GUI). This layer makes sure that users and the system communicate effectively by giving priority to the user experience.

UI Access Control :

Optimizing the model's interaction and enabling a seamless user experience are key components of managing the user interface. Authorization and authentication strengthen security by putting in place access rules that limit model usage to individuals who are authorized. This layer guarantees sensitive medical data privacy in addition to user ease.

Core Utilities :

The model's essential functions are stored in the Core Business Logic Layer, which is the central component of the system. Prototypical networks and few-shot learning create the fundamental logic of picture categorization, improving overall performance and adaptability. Essential features are extracted by the ResNet12 architecture, which is crucial to the model's precision in pneumonia diagnosis.

System Support :

The foundational infrastructure required for the model to function is provided by the System Support Layer. The system is used to execute training procedures and optimization algorithms, which enhance the model's capacity for learning. The compilation of datasets, tweaking of hyperparameters, and performance evaluation metrics are database-related tasks that are essential to guaranteeing the model's solid performance and ongoing enhancement.

4.2 Class Diagram

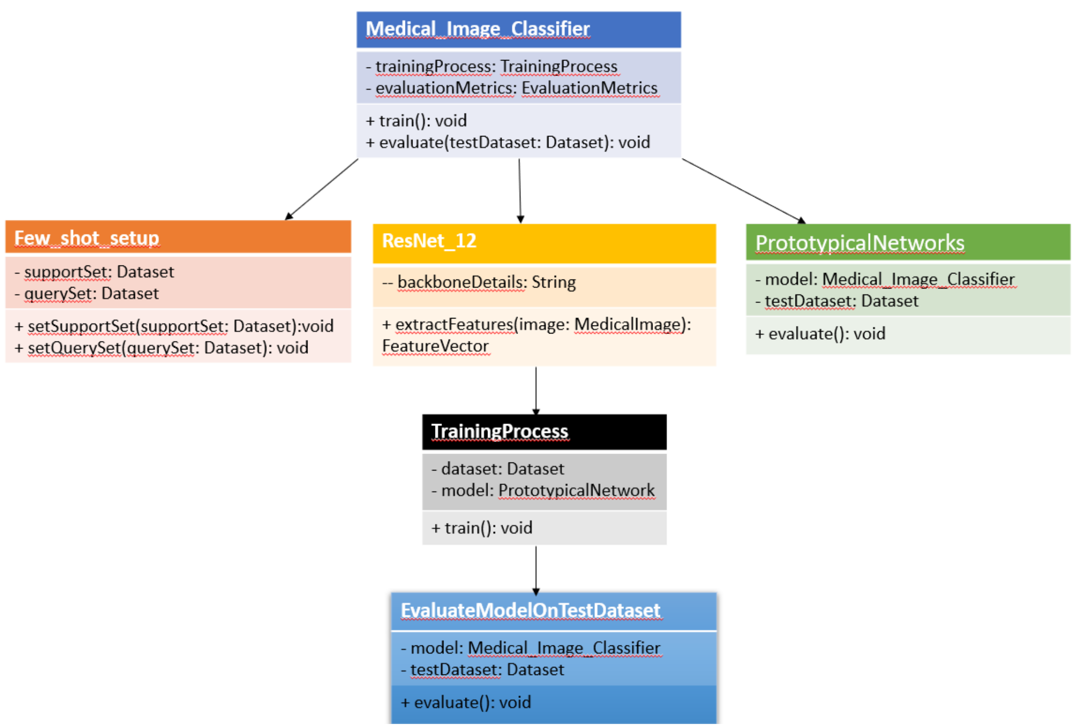


Figure 4.2: class diagram

4.3 Data set description

The traning dataset consisted of nearly 600 images. There are two kinds of lung X-ray images included in this dataset ”Normal” and ”Pneumonia” The testing dataset consisted of nearly 300 images,it has two classes ”Normal” and ”Covid”. All the images were set to a resolution of 224 x 224 pixels while training and testing.

Table 4.1: Normal Lung Images

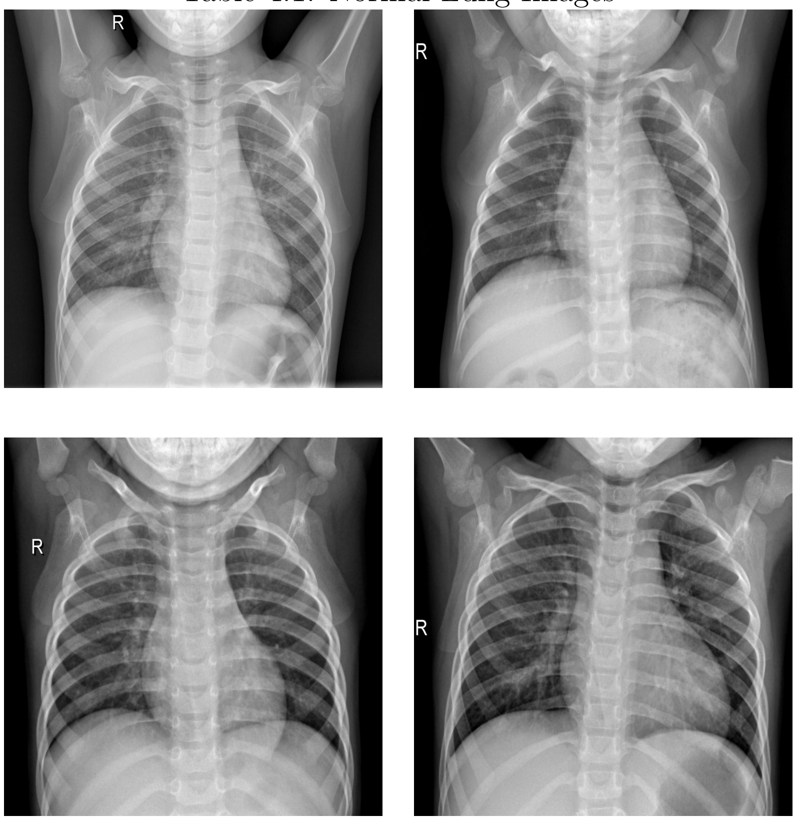


Table 4.1 shows some of the images from the dataset which belong to the class "Normal". Table 4.2 and Table 4.3 represent the images belonging to the classes "Covid" and "Pneumonia".

Table 4.2: Covid Images

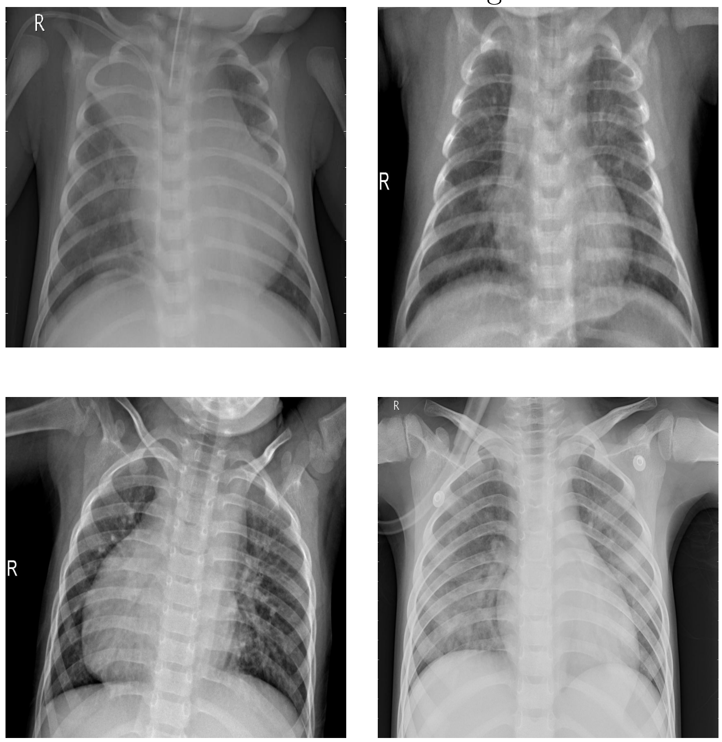
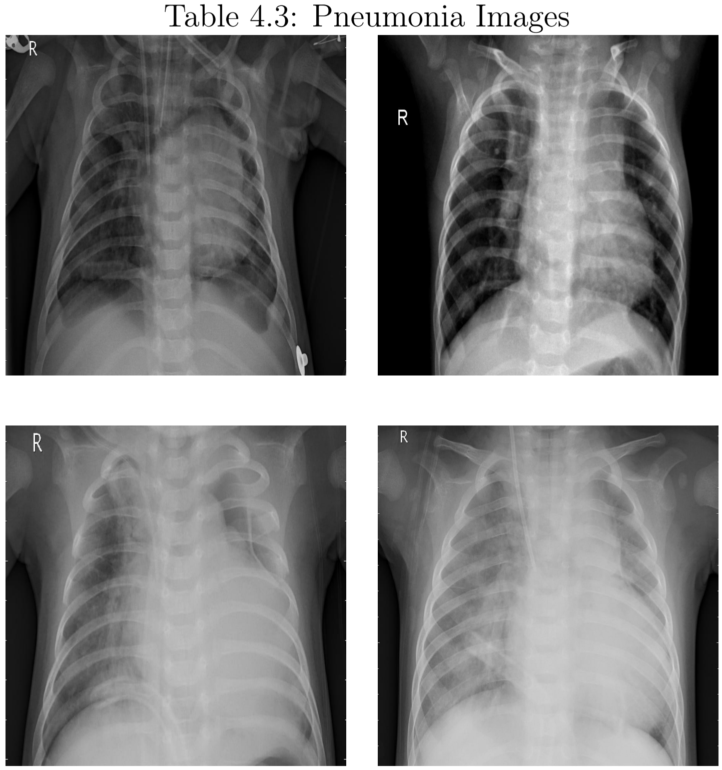


Table 4.3: Pneumonia Images



Chapter 5

IMPLEMENTATION

Fundamentally, FSL enables the model to learn from a small number of instances each class, improving performance and adaptability in scenarios where conventional deep learning methods are inadequate. Few-shot learning (FSL) using prototypical networks involves training a model to recognize and generalize to new classes or tasks with very limited labeled examples per class. The FSL consists of following steps

5.1 Task Definition

We train the model with lung X-ray images with two classes, "Normal" and "Covid" and then test the model on a novel dataset ,"Normal" and "Pneumonia". The training dataset consists of lung x ray images with classes Covid and Normal. The testing dataset consists of lung x ray images with classes pneumonia and Normal

5.2 ResNet12(Backbone Architecture)

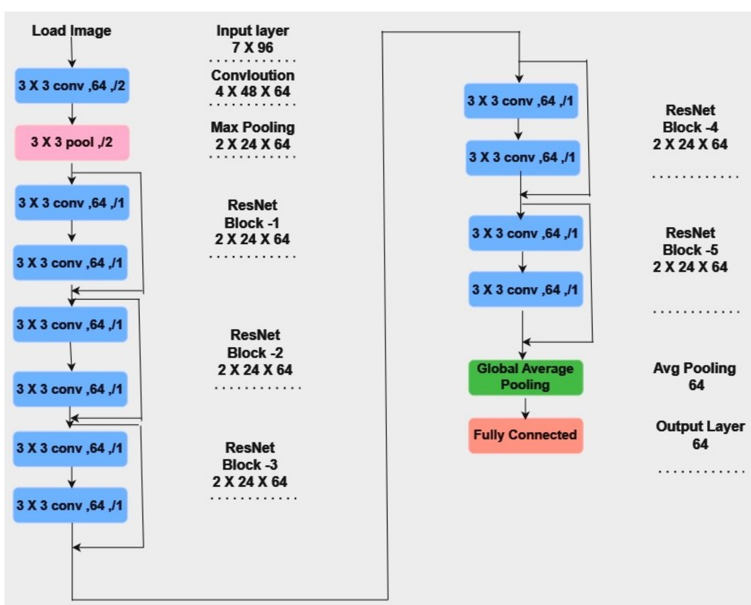


Figure 5.1: ResNet12 Architecture

In our work, the fundamental architecture for extracting important features from lung X-ray images is a pretrained ResNet12. Pretraining suggests that before the ResNet12 model is optimised for your particular application, it has been trained on a sizable dataset for general image understanding.

ResNet12's architecture (refer Fig. 1) consists of several residual blocks, each of which is intended to extract complex information from the X-ray images. These blocks make use of rectified linear unit (ReLU) activation functions[14], which come after convolutional layers. By utilising residual connections, problems like the vanishing gradient problem are mitigated and complex mappings can be learned effectively.

$$F(x) = W_2\sigma(W_1 x) \tag{5.1}$$

Here, x represents the input data or feature vector. W_1 represents weight matrix associated with the first layer. σ is the activation function, typically the rectified linear unit (ReLU) applied element-wise. W_2 is the weight matrix associated with the second layer.

Global Average Pooling (GAP) is a crucial ResNet12 component. The GAP layer, which comes after the convolutional layers, determines the average value of each feature map in order to condense the spatial information. Through this technique, the retrieved characteristics are condensed into a representation that highlights the most important data.

The following represents the mathematical representation of the Global Average Pooling (GAP) operation for a feature map F_i with height H and width W. The formula for $\mathrm{F}_i(j,k) is$

$$\mathrm{GAP}(F_i) = \frac{1}{H \times W} \sum_{j=1}^{H} \sum_{k=1}^{W} F_i(j,k) \tag{5.2}$$

The value of the feature map at position (j, k) is indicated by $F_i(j, k)$. The end product is a compressed representation that highlights the most important data for further processing by capturing the average contribution of each characteristic across the spatial dimensions. The condensed representation functions as a reliable and insightful input for the neural network's subsequent layers.

The pretrained model's[15] use highlights its versatility, which makes it especially useful for tasks such as lung X-ray picture categorization. This architecture generates embeddings that provide a rich and informative foundation for subsequent few-shot learning, improving the model's flexibility to accommodate a wide range of class distributions in medical imaging applications.

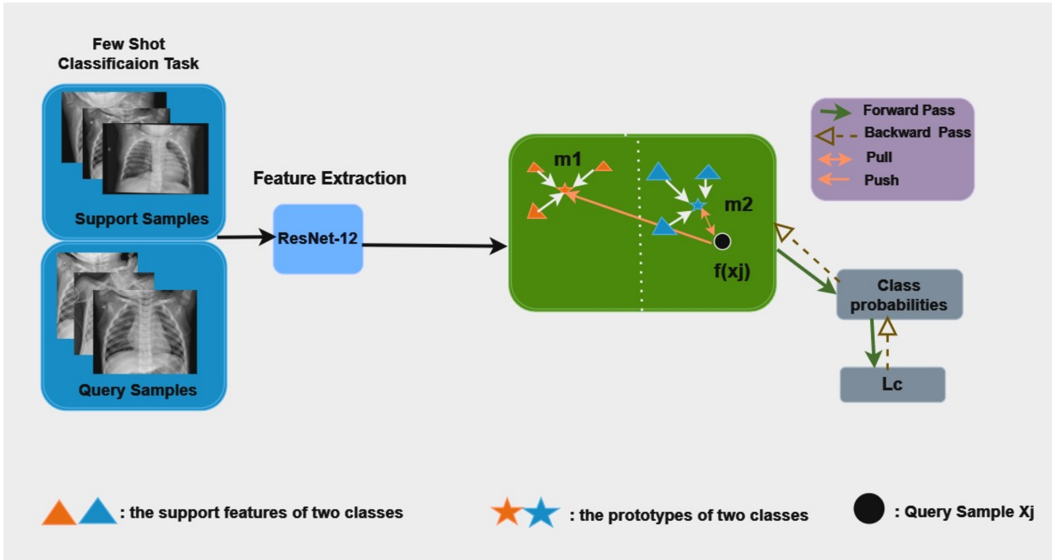


Figure 5.2: Prototypical Network

5.3 Prototypical Networks

Feature extraction and prototype computation are the two stages of a process that define prototypical networks[16]. As the foundation for feature extraction in our implementation, we use a ResNet12. A rich representation for later few-shot learning is produced by this pretrained neural network, which excels at extracting hierarchical characteristics from input images.

Prototypical Networks[17][20] serve as the foundation for our few-shot learning strategy, as seen in Fig. 2. The network extracts features from lung X-ray pictures by using a pretrained ResNet12 as its backbone architecture. The fundamental principle of Prototypical Networks is to classify query images by computing prototypes for each class in the support set and using the distance metric in the feature space. This procedure is important to our methodology and improves the model's performance on sparsely labelled data.

Prototypes (P_c) for each class (c) in a prototypical network are calculated as the average of the support features (S_c) for that class. The prototype can be computed as follows given n_c support examples for class c:

$$P_c = \frac{1}{n_c} \sum_{i=1}^{n_c} F(x_i) \tag{5.3}$$

where $F(x_i)$ is the matching feature vector and x_i is a support example for class c.

The Euclidean distance between the query feature (Q) and class prototypes (P_c) is used by the model to identify the most likely class when classifying query photos. One can obtain the classification scores as follows:

$$\text{Euclidean\_Distance}(Q, P_c) = \sqrt{\sum_{i=1}^{d}(Q_i - P_{c_i})^2} \tag{5.4}$$

Our framework, Prototypical Networks, makes use of the cross-entropy loss function. A common option for classification tasks is the cross-entropy loss, which quantifies the difference between the genuine probability distribution (one-hot encoded ground truth labels) and the anticipated probability distribution (the model's output). The cross-entropy loss[19] for a single training example can be represented mathematically as follows:

$$\text{CrossEntropyLoss}(\mathbf{p}, \mathbf{q}) = -\sum_{i} p_i \log(q_i) \tag{5.5}$$

where q is the expected probability distribution and p is the genuine probability distribution (ground truth). This loss function effectively directs the training process towards accurate classification by helping the model minimise the discrepancy between the true and predicted class probabilities.

5.4 Optimization Algorithm and Training Process

We used Stochastic Gradient Descent (SGD)[18] as the optimisation procedure to improve our model's generalisation capacity by optimising its parameters. Based on the calculated gradients of the loss function with respect to the parameters, SGD modifies the model's parameters.

A learning rate scheduler with a step size of 5 and a gamma of 0.1 was used to modify the learning rate during training. The learning rate was set to 0.001.

The model was trained on the prepared dataset for thirty epochs. The training set was used to train the model during each epoch, and the parameters were updated to minimise the cross-entropy loss.

Additionally, a patient early termination mechanism was used to avoid overfitting. If the training accuracy did not increase after a predetermined number of successive epochs (referred to as the patience parameter) which was set to 5, the training process would terminate prematurely. This made it possible to guarantee the model's good generalisation to new data.

Chapter 6

RESULTS AND DISCUSSIONS

6.1 Hyper Parameters

Table 6.1: Few-Shot Learning Hyper parameters

No of ways	2
No of shots in Support set	3,5,10
No of shots in Query set	30
Learning rate	0.01

The important parameters in a few-shot learning scenario are listed in the Table 6.1. The model is intended to function with two classes in this specific configuration (No of ways = 2). The number of shots for the support set changes and is 3, 5, and 10. A small subset of examples utilised for learning in each class is called the support set. In addition, during the assessment phase, the model attempts to categorise or generalise its knowledge to 30 examples per class in the query set. The learning rate was set to 0.01.

6.2 Evaluation Metrics

Accuracy, sensitivity, and specificity are the three main criteria we used to assess the performance of our Cross-Domain Few Shot Image Classification model. In order to evaluate how well machine learning models perform in classification, these measures are frequently used.

The overall correctness of the model's predictions is represented by accuracy. It is computed as the ratio of the total number of predictions to the sum of true positive (TP) and true negative (TN) predictions:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{6.1}$$

where the numbers TP, TN, FP, and FN represent the true positive, true negative, false positive and false positive, respectively.

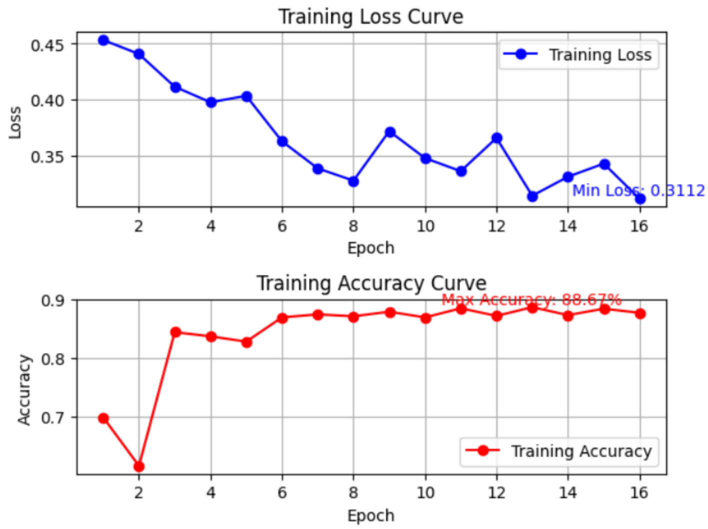


Figure 6.1: Training loss and Accuracy curve for 10-shot Scenario

Sensitivity measures the model's accuracy in identifying positive examples. The ratio of true positives (TP) to the total of false negatives and true positives (FN) is used to compute it.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100 \tag{6.2}$$

Specificity gauges how well the model can recognise negative examples. The ratio of true negatives (TN) to the total of false positives and true negatives (FP) is used to compute it.

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100 \tag{6.3}$$

6.3 Interpretation of Results

We report the results of using our few-shot learning model to classify images. A curated dataset was used to train and assess the model, with an emphasis on ensuring a balanced distribution of both normal and lung opacity instances. As a few-shot learning strategy, we combined a pretrained ResNet12 backbone architecture with prototypical networks.

Fig. 6.1 showcases the learning curve for 10-shot scenario, illustrating the evolution of both training loss and accuracy across multiple epochs. The maximum training accuracy observed was 88.67% and the minimum loss was found to be 0.3112. These measures highlight the model's strong performance in a range of few-shot scenarios, demonstrating its ability to accurately classify the images in a variety of support scenarios. Fig. 6.2 and Fig. 6.3 showcases the learning curve for 5-shot and 3-shot scenarios respectively. The maximum training accuracy obtained was 86.96% and 85.19% for 5-shot and 3-shot scenarios respectively. The minimum loss obtained were 0.3536 and 0.3587.

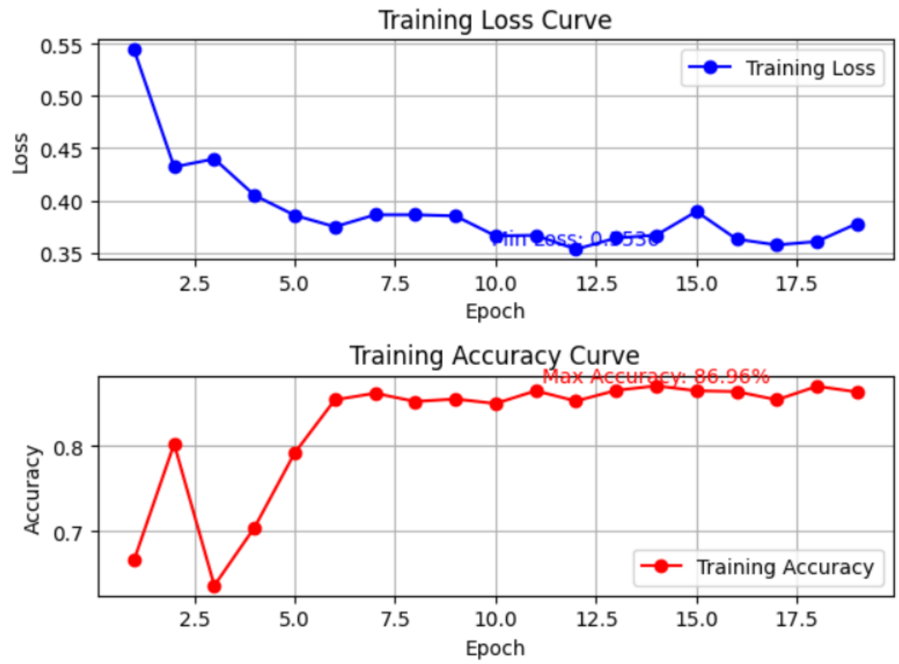


Figure 6.2: Training loss and Accuracy curve for 5-shot Scenario

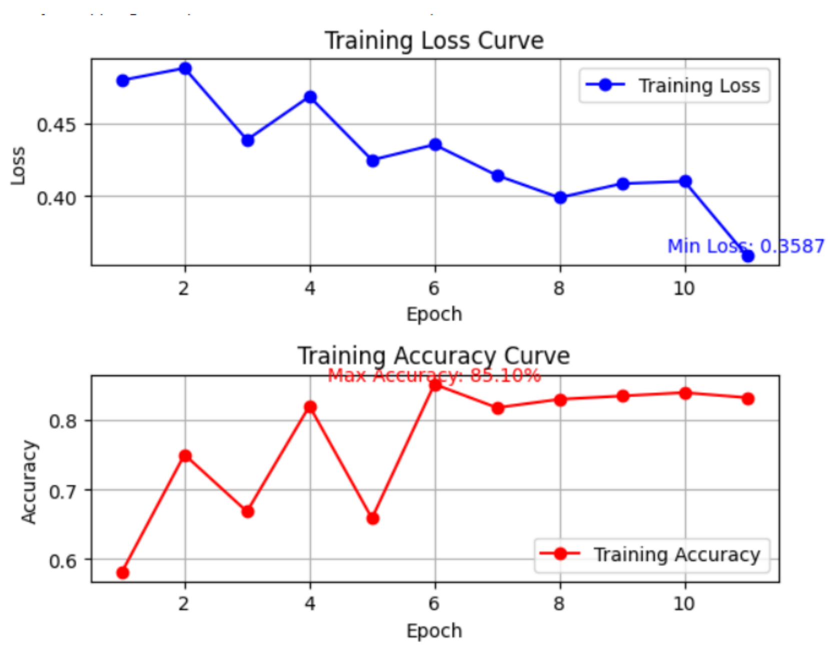


Figure 6.3: Training loss and Accuracy curve for 3-shot Scenario

Within the framework of the proposed Pneumonia detection model, the testing performance metrics for various few-shot scenarios are displayed in the Table 6.2. The model's accuracy was 75.24% in the 3-shot scenario, in which three samples are provided for each class. The model's sensitivity, which measures how well it can identify normal cases, was 78.67%, and its specificity, which gauges how well it can identify Covid cases, was 68.90%. With five support samples per class in the 5-shot scenario, the accuracy increased to 83.84%, while the sensitivity and specificity increased to 87.27% and 79.55%, respectively. The model showed even more improvement in the 10-shot scenario, which had ten support samples for every class, with 90.91% accuracy, 90.0% sensitivity, and 91.84% specificity.

Table 6.2: Testing Performance Metrics for Different Few-Shot Scenarios

Scenario	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
3-shot	76.76	78.67	68.90
5-shot	83.84	87.27	79.55
10-shot	**90.91**	**90.0**	**91.84**

Table 6.3: Comparison with Other Studies on Different Diseases

Study	**Disease**	**Methodology**	**Performance Metrics**
L-MED[7]	COVID-19	ResNet-152x4	Accuracy 83.00%
Momentum FSL[8]	COVID-19	ResNet-50(Baseline)	Accuracy 88.5%
MAML[9]	Lung	Self-attention augmented MAML	Accuracy 73.4%
VGG-MAML[10]	Lung, COVID-19	8 block VGG+, MAML++	Accuracy 85.7%
Proposed Model	**Pneumonia**	**Few-shot learning(Prototypical networks), ResNet12**	**Accuracy 90.91%, Sensitivity 90.0%, Specificity 91.84%**

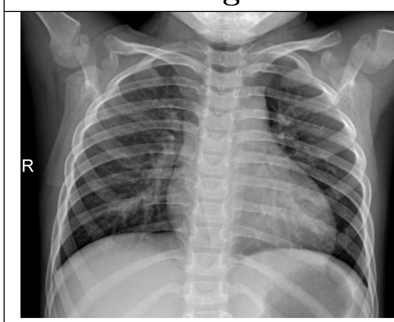
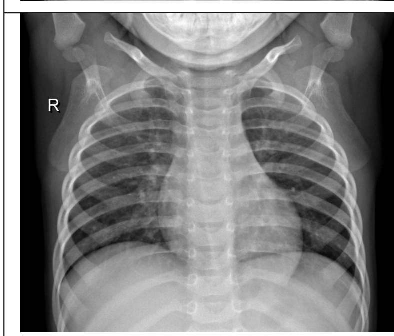
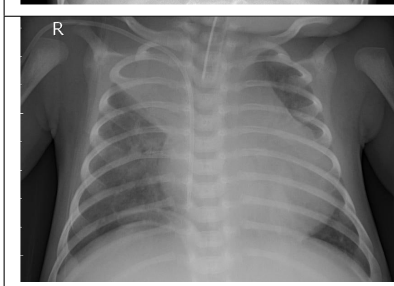
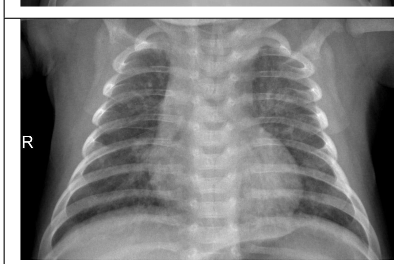
A thorough comparison of our suggested approach with existing research on disease detection, specifically with regard to COVID-19, is provided in Table 6.3. A unique study is represented by each entry in the table, which includes information about the disease being studied, the methodology used, and the performance indicators that go along with it. Remarkably, L-MED [7] uses ResNet-152x4 and achieves 83.00% accuracy in COVID-19 detection. Using ResNet-50 as the baseline, Momentum FSL [8] achieves an astounding 88.5% accuracy rate in COVID-19 detection. A self-attention augmented MAML approach is used by MAML [9] to investigate lung disease detection, with an accuracy of 73.4%. VGG-MAML [10] uses an 8-block VGG+ architecture with MAML++ to diagnose COVID-19 and lung illnesses, with an accuracy of 85.7%. With an accuracy of 90.91%, sensitivity of 90.0%, and specificity of

91.84%, our model highlighted in the table introduces a unique few-shot learning strategy with ResNet12 and Prototypical network for COVID-19 detection and exhibits exceptional performance. This demonstrates how well our model performs in comparison to previous research in the field.

When comparing our model's performance to previous research, it becomes clear that the ResNet12 backbone architecture in conjunction with the Few-Shot Learning technique produces competitive outcomes for COVID-19 identification. The remarkable results attained in terms of accuracy, sensitivity, and specificity confirm the effectiveness of our suggested approach.

Table 6.4 demonstrates the classification results for the lung X-ray images.

Table 6.4: Image Classification Results

Image	**Predicted Class**
	Normal
	Normal
	Covid
	Covid

Chapter 7

CONCLUSIONS AND FUTURE SCOPE

7.1 Conclusion

In summary, our suggested Cross-Domain Few Shot X-Ray Image Classification model has strong performance in adjusting to shifts in class distributions by utilising Prototypical Networks with a ResNet12 backbone. The model outperformed current techniques in disease identification, especially in detecting COVID-19 cases, with remarkable accuracy, sensitivity, and specificity. With minimal labelled data, our study demonstrates the potential of few-shot learning for medical image classification, providing a useful and efficient way to adjust to dynamic class fluctuations.

7.2 Future Scope

In future endeavors, it is crucial to enhance the performance and versatility of the model by expanding the dataset to include diverse lung X-ray images. This diversification will contribute to a more comprehensive understanding of respiratory conditions, thereby improving the model's adaptability. Additionally, iterative refinement of the model through experimentation with various techniques and architectures is essential to optimize its predictive capabilities. Collaborating closely with radiologists will provide valuable insights and foster human-machine collaboration, ensuring the model's alignment with clinical expertise. Furthermore, evaluating the model's robustness to external factors and varying conditions is imperative for establishing its reliability in real-world scenarios. To extend the utility of the model, efforts should be directed towards detecting a broader range of respiratory conditions, thereby transforming it into a more inclusive and effective diagnostic tool.

REFERENCES

[1] Eva Pachetti, Sara Colantonio. "A Systematic Review of Few-Shot Learning in Medical Imaging," arXiv preprint arXiv:2309.11433 (2023).

[2] Fereshteh Shakeri, Malik Boudiaf, Sina Mohammadi, Ivaxi Sheth, Mohammad Havaei, Ismail Ben Ayed, Samira Ebrahimi Kahou, "FHIST: A Benchmark for Few-shot Classification of Histological Images," arXiv:2206.00092 (2022)

[3] Jaehoon Oh, Sungnyun Kim, Namgyu Ho, Jin-Hwa Kim, Hwanjun Song, Se-Young Yun "Understanding Cross-Domain Few-Shot Learning Based on Domain Similarity and Few-Shot Difficulty," arXiv:2202.01339 (2022).

[4] Covid XRay Dataset

https://www.kaggle.com/datasets/ahemateja19bec1025/covid-xray-dataset/data

[5] Manu Siddhartha "chest xray pneumo res"

https://www.kaggle.com/datasets/sid321axn/chest-xray-pneumo-res/data

[6] Hongyang Jiang , Mengdi Gao, Heng Li , Richu Jin, Hanpei Miao and Jiang Liu "Multi-Learner Based Deep Meta-Learning for Few-Shot Medical Image Classification"(2023)

[7] Mehdi Cherti, Jenia Jitsev "Effect of Pre-Training Scale on Intra- and Inter-Domain Full and Few-Shot Transfer Learning for Natural and Medical X-Ray Chest Images" (2022) International Joint Conference on Neural Networks (IJCNN), 2022, pp. 1-9

[8] Chen, X.; Yao, L.; Zhou, T.; Dong, J.; Zhang, Y. Momentum contrastive learning for few-shot covid-19 diagnosis from chest ct images. Pattern Recognit. 2021, 113, 107826.

[9] Achraf Ouahab, Olfa Ben-Ahmed, and Christine Fernandez-Maloigne. A Self-attentive Meta-learning Approach for Image-Based Few-Shot Disease Detection. In Resource-Efficient Medical Image Analysis: First MICCAI Workshop, REMIA 2022, Singapore, September 22, 2022, Proceedings, pages 115–125. Springer, 2022.

[10] Tarun Naren, Yuanda Zhu, and May Dongmei Wang. COVID-19 diagnosis using model agnostic meta-learning on limited chest X-ray images. In Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, pages 1–9, 2021.

[11] Wang, Yaqing, et al. "Generalizing from a few examples: A survey on few-shot learning." ACM computing surveys (csur) 53.3 (2020): 1-34.

[12] Snell, Jake, Kevin Swersky, and Richard Zemel. "Prototypical networks for few-shot learning." Advances in neural information processing systems 30 (2017).

[13] Choi, Hyungeun, Ryu, Seunghyoung, Kim, Hongseok (2018). Short-Term Load Forecasting based on ResNet and LSTM. 1-6. 10.1109/SmartGridComm.2018.8587554.

[14] Agarap, A.F. "Deep learning using rectified linear units (relu)". (2018) arXiv preprint arXiv:1803.08375.

[15] Han-Jia Ye, Hexiang Hu, De-Chuan Zhan, Fei Sha. "Few-Shot Learning via Embedding Adaptation with Set-to-Set Functions". arXiv:1812.03664 (2020)

[16] Qiang Lyu, Weiqiang Wang "Compositional Prototypical Networks for Few-Shot Classification." arXiv:2306.06584 (2023)

[17] Hao Quan, Xinjia Li, Dayu Hu, Tianhang Nan, Xiaoyu Cui. "Dual-channel Prototype Network for few-shot Classification of Pathological Images." arXiv:2311.07871 (2023)

[18] Liu, Yanli, Yuan Gao, and Wotao Yin. "An improved analysis of stochastic gradient descent with momentum." Advances in Neural Information Processing Systems 33 (2020): 18261-18271.

[19] Zhang, Zhilu, and Mert Sabuncu. "Generalized cross entropy loss for training deep neural networks with noisy labels." Advances in neural information processing systems 31 (2018).

[20] Liu, Jinlu, Liang Song, and Yongqiang Qin. "Prototype rectification for few-shot learning." Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16. Springer International Publishing, 2020.

[21] Ibrahim, Abdullahi Umar, et al. "Pneumonia classification using deep learning from chest X-ray images during COVID-19." Cognitive Computation (2021): 1-13.

Appendix A

A.1 Gantt Chart

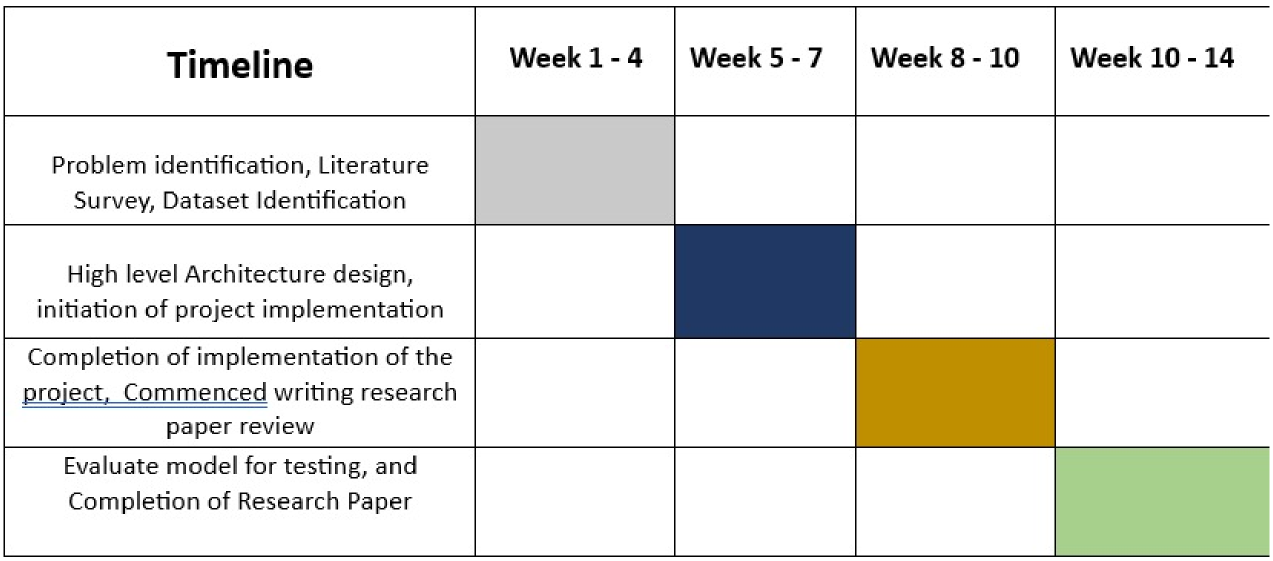


Figure A.1: Gantt Chart