A project report on

MEDXNET – A DEEP LEARNING MODEL FOR AUTOMATED CHEST X-RAY DIAGNOSIS

Submitted in partial fulfillment for the award of the degree of

Bachelor of Technology in Computer Science and Engineering with Specialization in Artificial Intelligence & Robotics

by

ANARGHYA JAIN (21BRS1235)

VIT®

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)
CHENNAI

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

April 2025

MEDXNET – A DEEP LEARNING MODEL FOR AUTOMATED CHEST X-RAY DIAGNOSIS

Submitted in partial fulfillment for the award of the degree of

Bachelor of Technology in Computer Science and Engineering with Specialization in Artificial Intelligence & Robotics

by

ANARGHYA JAIN (21BRS1235)

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)
CHENNAI

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

April 2025

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

CHENNAI

DECLARATION

I hereby declare that the thesis entitled “MEDXNET - A DEEP LEARNING MODEL FOR AUTOMATED CHEST X-RAY DIAGNOSIS” submitted by Anarghya Jain (21BRS1235), for the award of the degree of Bachelor of Technology in Computer Science and Engineering, Vellore Institute of Technology, Chennai is a record of Bonafide work carried out by me under the supervision of Dr. N.G. Bhuvaneswari Amma

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place: Chennai

Date: Signature of the Candidate

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

CHENNAI

School of Computer Science and Engineering

CERTIFICATE

This is to certify that the report entitled "**MEDXNET - A DEEP LEARNING MODEL FOR AUTOMATED CHEST X-RAY DIAGNOSIS**" is prepared and submitted by **Anarghya Jain (21BRS1235)** to Vellore Institute of Technology, Chennai, in partial fulfillment of the requirement for the award of the degree of **Bachelor of Technology in Computer Science and Engineering with Specialization in Artificial Intelligence & Robotics** is a bonafide record carried out under my guidance. The project fulfills the requirements as per the regulations of this University and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma and the same is certified.

Signature of the Guide:

Name: Dr./Prof.

Date:

Signature of the Examiner

Name:

Date:

Signature of the Examiner

Name:

Date:

Approved by the Head of Department,
B.Tech COMPUTER SCIENCE ENGINEERING with SPECIALIZATION IN AI AND ROBOTICS

Name: Dr. Harini S

Date:

(Seal of SCOPE)

ABSTRACT

Chest X-ray imaging is a cornerstone of modern medical diagnostics, providing key insights into pulmonary and thoracic illnesses including pneumonia and tuberculosis. Yet the manual interpretation of chest X-rays is time-consuming and prone to error, frequently resulting in delays in diagnosis and treatment. To overcome these issues, this project presents MedXNet, a deep learning algorithm for the automated classification of chest X-rays into four categories: Viral Pneumonia, Bacterial Pneumonia, Tuberculosis, and Normal. MedXNet is a new architecture based on the strengths of two popular deep learning models, ResNet50 and VGG16, but designed specifically for the intricacies of chest X-ray imaging. The MedXNet architecture combines the strong feature extraction ability of VGG16 with the residual learning structure of ResNet50. This blend enables the model to learn complex patterns in chest X-ray images while keeping computational efficiency. The model was trained on a combined dataset consisting of the TBX11K Tuberculosis Dataset and the Chest X-ray Images for Pneumonia Detection Dataset, which collectively offer a diverse and comprehensive set of chest X-ray images. Preprocessing operations, such as resizing, normalization, and data augmentation, were performed to maintain uniformity and enhance generalization. MedXNet exhibits encouraging performance in identifying the four target conditions. The model also offers interpretable predictions by gradient-weighted class activation mapping (Grad-CAM), providing insights into what parts of the X-ray images are most responsible for its decision. Such interpretability is key to gaining the confidence of clinicians and enabling the adoption of AI-based diagnostic solutions in real-world healthcare environments. This project underscores the promise of hybrid deep learning architectures in promoting medical imaging and diminishing the diagnostic burden in clinical settings. The creation of MedXNet is an important milestone towards the integration of AI research and clinical practice, providing a scalable and effective tool for computerized chest X-ray diagnosis. Aside from its primary use, the ideas behind MedXNet can be applied to other medical imaging tasks, such as the detection of lung nodules, breast cancer, or brain abnormalities. Ongoing work will involve increasing the dataset, optimizing the model for deployment in low-resource environments, and investigating its use in other medical imaging applications. Through continued development and testing of MedXNet, we hope to contribute to the burgeoning literature that attempts to leverage the potential of AI for enhancing healthcare outcomes globally.

Keywords: Chest X-ray, Deep Learning, MedXNet, ResNet50, VGG16, Medical Imaging, Automated Diagnosis, Pneumonia, Tuberculosis

ACKNOWLEDGEMENT

It is my pleasure to express with deep sense of gratitude to Dr. N.G. Bhuvaneswari Amma, Assistant Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, for her constant guidance, continual encouragement, understanding; more than all, she taught me patience in my endeavor. My association with her is not confined to academics only, but it is a great opportunity on my part of work with an intellectual and expert in the field of Deep learning.

It is with gratitude that I would like to extend my thanks to the visionary leader Dr. G. Viswanathan our Honorable Chancellor, Mr. Sankar Viswanathan, Dr. Sekar Viswanathan, Dr. G V Selvam Vice Presidents, Dr. Sandhya Pentareddy, Executive Director, Ms. Kadhambari S. Viswanathan, Assistant Vice-President, Dr. V. S. Kanchana Bhaaskaran Vice-Chancellor, Dr. T. Thyagarajan Pro-Vice Chancellor, VIT Chennai and Dr. P. K. Manoharan, Additional Registrar for providing an exceptional working environment and inspiring all of us during the tenure of the course. Special mention to Dr. Ganesan R, Dean, Dr. Parvathi R, Associate Dean Academics, Dr. Geetha S, Associate Dean Research, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai for spending their valuable time and efforts in sharing their knowledge and for helping us in every aspect.

In jubilant state, I express ingeniously my whole-hearted thanks to Dr. Harini S, Head of the Department, B.Tech. Computer Science and Engineering with Specialization in Artificial Intelligence & Robotics and the Project Coordinators for their valuable support and encouragement to take up and complete the thesis.

My sincere thanks to all the faculty members and staff members at Vellore Institute of Technology, Chennai who helped me acquire the requisite knowledge. I would like to thank my parents for their support. It is indeed a pleasure to thank my friends who encouraged me to take up and complete this task.

Place: Chennai

Date: **Anarghya Jain**

CONTENTS **PAGE NO.**

CHAPTER 3

LITERATURE SURVEY

CHAPTER 4

METHODOLOGY

CHAPTER 5

EXPERIMENTAL RESULTS AND DISCUSSION

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

LIST OF FIGURES **PAGE NO**

LIST OF TABLES **PAGE NO.**

LIST OF ACRONYMS

- AI - Artificial Intelligence
- AUC - Area Under the Curve
- CAM - Class Activation Map
- CBAM - Convolutional Block Attention Module
- CNN - Convolutional Neural Network
- CT - Computed Tomography
- DL - Deep Learning
- FC - Fully Connected Layer
- GANs - Generative Adversarial Networks
- Grad-CAM - Gradient-weighted Class Activation Mapping
- GPU - Graphics Processing Unit
- MRI - Magnetic Resonance Imaging
- NAS - Neural Architecture Search
- ReLU - Rectified Linear Unit
- ResNet - Residual Network
- ROI - Region of Interest
- ROC - Receiver Operating Characteristic
- TB - Tuberculosis

- VGG - Visual Geometry Group (e.g., VGG16, VGG19)
- WHO - World Health Organization
- XAI - Explainable Artificial Intelligence

Chapter 1

Introduction

1.1 Background and Motivation

Chest X-ray imaging continues to be a cornerstone of medical diagnosis, yielding important information regarding pulmonary and thoracic diseases such as pneumonia, tuberculosis, and lung cancer. One of the most used diagnostic modalities in the world, chest radiographs provide an inexpensive, non-invasive, and quick assessment tool, making them invaluable in a variety of clinical environments [1]. Nevertheless, their interpretation poses severe challenges with detection of subtle abnormalities against complex anatomical overlap, necessitating expert radiological interpretation. Experiments evidence large variance in radiologist performance with as high as 20–30% rates of inter-observer disagreement for some pulmonary abnormalities [2]. Such diagnostic variability, combined with the increasing worldwide need for imaging studies, highlights the acute necessity for computer-aided diagnosis systems that can optimize both the speed and accuracy of clinical decision-making [2].

The World Health Organization estimates that more than two-thirds of the world's population does not have access to fundamental radiological services, widening healthcare inequalities in resource-poor areas [1]. Artificial intelligence (AI)-enabled diagnostic tools provide a revolutionary solution by alleviating interpretation workloads, speeding up turnaround times, and ensuring uniform diagnostic quality [3]. Among AI methods, deep learning—specifically convolutional neural networks (CNNs)—has proven to be highly successful in medical image analysis, with state-of-the-art performance in classification, segmentation, and detection tasks [3].

In spite of these progresses, three essential challenges remain in computer-aided chest X-ray diagnosis: 1, anatomical complexity with overlapping structures,
2, inter-institutional variability in imaging acquisition protocols, and 3, inherent class imbalances in medical data [4]. These demand the creation of robust, explainable deep learning models that not only provide high diagnostic accuracy but also present interpretable decision paths to induce clinician trust [4].

This research addresses these challenges using MedXNet (Medical X-ray Network for Pneumonia and Tuberculosis Detection), a novel hybrid deep learning framework designed for four-class classification of chest X-rays, distinguishing between Viral Pneumonia, Bacterial Pneumonia, Tuberculosis, and Normal cases. MedXNet leverages the strength of well-proven architectures such as ResNet50 and VGG16 in synergizing its capabilities, thus filling the gap between AI innovation and clinical utilization [5]. The construction of the system is driven by two key goals:

1. Workflow Enhancement: Automating early screening for a potential up to 40% decrease in radiologist workload without compromising on diagnostic accuracy, as shown on similar AI-assisted radiology platforms [2].

2. Diagnostic Support: Reducing variability in human interpretation through consistent data-driven performance of classification, especially for hard-to-detect presentations such as early tuberculosis or viral pneumonias [4].

Through the realization of these objectives, MedXNet has the ability to revolutionize paradigms in chest X-ray interpretation, especially in under-resourced areas where radiological services are limited [1]. The design of the architecture includes explainability components that harmonize with clinical workflows for open decision-making processes that comply with regulatory and practitioner standards [4].

1.2 Objectives of the Project

The main aim of this study is to design MedXNet, a deep learning state-of-the-art model for computer-aided classification of chest X-rays into four major diagnostic classes: Viral Pneumonia, Bacterial Pneumonia, Tuberculosis, and Normal [4]. The endeavor is to design a new architecture that combines the complementary strengths of ResNet50 and VGG16 without their respective weaknesses. By integrating VGG16's hierarchical feature extraction ability with ResNet50's residual skip connections, the model aims to maximize both depth and training stability for enhanced pattern recognition in chest radiographs, especially for difficult cases such as subtle ground-glass opacities in viral pneumonia or apical cavitations in tuberculosis [5].

The interpretability features identify diagnostically important areas in the X-rays, e.g., lobar consolidations in bacterial pneumonia or mediastinal lymphadenopathy in tuberculosis, allowing radiologists to cross-check predictions with their clinical judgment. This transparency is essential for establishing trust in AI-aided diagnostics and compliance with regulatory standards for medical decision support systems [7].

The study overcomes key data challenges using meticulous curation of a multi-source dataset merging the TBX11K Tuberculosis Dataset and Chest X-ray Images for Pneumonia Detection Dataset [8]. State-of-the-art preprocessing methods such as lung segmentation and contrast-limited adaptive histogram equalization are used to provide stable performance across various imaging protocols and patient populations [9]. The model's performance is strictly tested using detailed metrics such as accuracy, AUC-ROC, and class-specific F1-scores, with special focus on precision-recall trade-offs for low-prevalence diseases such as tuberculosis [7]. Performance comparisons with state-of-the-art models prove MedXNet's clinical feasibility, particularly its high sensitivity for life-threatening diseases and specificity for normal conditions [10].

By attaining these goals, MedXNet hopes to decrease radiologist workload considerably in screening situations while enhancing diagnostic consistency, especially in

resource-limited environments where radiological experience is scarce [1]. The project ultimately hopes to set a new standard for explainable, generalizable AI-supported chest X-ray interpretation that fills the gap between research and clinical deployment [4]. By this research, the work adds to the overall objective of improving global healthcare outcomes through more accessible accurate and efficient chest X-ray diagnosis [2].

Key Contributions:

- Hybrid Model Development: Designed MedXNet, combining ResNet50 and VGG16 to improve feature extraction and classification accuracy for chest X-ray diagnosis.
- Data Optimization: Created a balanced dataset from TBX11K and Pneumonia Detection sources, using augmentation techniques to enhance model performance.
- Clinical Applicability: Achieved 90.8% accuracy in classifying four conditions (TB, Viral/Bacterial Pneumonia, Normal), demonstrating real-world diagnostic potential.

1.3 Scope of the Project

The project work involves the detailed development and assessment of MedXNet, a deep learning-based hybrid model for computer-aided diagnosis of chest X-ray images. This study is particularly aimed at four-class classification (Viral Pneumonia, Bacterial Pneumonia, Tuberculosis, and Normal) based on a well-curated composite dataset that combines the TBX11K Tuberculosis Dataset with the Chest X-ray Images for Pneumonia Detection Dataset to provide diversity in pathological presentations and imaging conditions [5]. The technical process includes necessary preprocessing steps such as image resizing to 224×224 resolution, intensity normalization, and sophisticated data augmentation methods like random rotations and flips to maximize model generalizability across varied patient populations and imaging device variations.

Fundamentally, the project explores a novel architectural combination of ResNet50 and VGG16 that integrates residual learning streams with deep hierarchical feature extraction to tackle some of the most challenging problems in thoracic image analysis. The training protocol incorporates class-weighted loss functions to mitigate dataset imbalance, while evaluation employs comprehensive metrics including area under the receiver operating characteristic curve (AUC-ROC) and per-class precision-recall analysis to ensure clinically meaningful performance assessment. While the present implementation is focused on chest X-ray interpretation, the modular design principles and hybrid approach show promise for extension to other medical imaging applications like mammography or neuroimaging, subject to suitable architectural adjustments and domain-specific training [12].

This research specifically limits its validation to pulmonary abnormalities evident in anterior-posterior and posteroanterior chest radiographs, but not lateral views or non-thoracic uses in order to keep the methodology sharp. The study does not include real-time deployment issues and hardware optimization factors, which are significant avenues for future translational research. Future studies can investigate the extensibility of the framework to three-dimensional imaging modalities such as computed tomography or its integration with clinical decision support systems [13].

1.4 Significance of the Project

The importance of this project is its potential to revolutionize the area of medical imaging through the ability to offer an efficient and scalable solution for computer-aided chest X-ray diagnosis [14]. By automating the preliminary screening, MedXNet can potentially lessen diagnostic workloads significantly, improve turnaround times, and improve patient outcomes. This is especially crucial in settings where there is a lack of resources, and the lack of experienced radiologists combined with the high number of chest X-ray images pose important bottlenecks to healthcare systems. The creation of MedXNet also adds to the expanding literature on hybrid deep learning architectures.

By leveraging the strengths of ResNet50 and VGG16, MedXNet illustrates the capability of hybrid models to strike a balance between depth and efficiency, allowing them to learn complex patterns from complex medical images. Lastly, the project sheds light on the significance of interpretability for AI-driven diagnostic models. By offering explanations of the model's decision-making process, MedXNet seeks to narrow the gap between clinical use and AI research, making AI more easily adoptable in medical environments. Not only does this emphasis on interpretability make the model more clinically applicable, but it also serves as a standard for future work in medical imaging research.

1.5 Organization of the Report

This report is structured into six chapters to present a comprehensive study of the MedXNet model for automated chest X-ray diagnosis. Chapter 1 introduces the project's motivation, objectives, and scope, highlighting the clinical need for AI-based diagnostic tools. Chapter 2 provides technical background on deep learning in medical imaging and examines existing architectures. Chapter 3 reviews relevant literature on chest X-ray analysis, including model architectures and evaluation metrics. Chapter 4 details the methodology, explaining MedXNet's hybrid architecture, dataset preparation, and training procedures. Chapter 5 presents experimental results, analyzing performance metrics and comparing them with existing solutions. Finally, Chapter 6 concludes and outlines future research directions. This structure ensures a logical progression from problem identification to solution development and validation.

Chapter 2

Background

2.1 The Importance of Chest X-ray Imaging in Healthcare

Chest X-ray imaging is the most prevalent diagnostic tool utilized in contemporary medicine, providing valuable information on pulmonary and thoracic diseases like pneumonia, tuberculosis (TB), and lung cancer. Its lack of invasiveness, low cost, and speedy acquisition make it a valuable component of clinical practice. TB—a condition identifiable through chest X-rays—impacts more than 10 million people each year, and as such, is one of the WHO's top worldwide causes of death [13]. In the same way, pneumonia, commonly identified through routine X-ray imaging, contributes to millions of hospitalizations annually, especially among the vulnerable such as children and the elderly [14]. Though prevalent, the interpretation of chest X-rays involves a complicated process that demands skilled hands. Radiologists must scrutinize every image thoroughly to detect faint abnormalities like lung opacities, pleural effusion, or cavitary lesions. Nonetheless, this process is susceptible to interobserver variation, and studies have documented rates of diagnostic disagreement up to 20–30% for some conditions [15].

The increasing number of chest X-rays, particularly in resource-poor environments, has taxed health systems, creating a need for automating tools to enhance speed and accuracy in diagnosis. Progress in artificial intelligence (AI), and in particular deep learning, has transformed this field. Convolutional neural networks (CNNs) perform well for image classification and anomaly detection tasks, and demonstrate excellent performance for detecting radiographic abnormalities [16]. Nevertheless, there remain problems such as superimposed anatomical structures (e.g., ribs overlying lung lesions), imaging condition variations (e.g., variation in patient position or equipment), and class imbalance of datasets. For example, uncommon cases such as TB or atypical pneumonia are typically underrepresented and produce skewed models with more attention being placed on typical cases [17]. Correcting such drawbacks is accomplished with techniques involving data augmentation, oversampling, and weighted loss functions [18].

Another significant hindrance to clinical implementation is the "black-box" character of deep learning models. Clinicians need understandable AI so that they will be able and willing to have faith and make use of the tools in the clinic. Tools such as gradient-weighted class activation mapping (Grad-CAM) remedy this by visualizing areas impacting a model's decision, adding clarity [19]. As technology around AI improves, these hurdles need to be transcended to get it used within real-world clinics.

2.2 The Role of Deep Learning in Medical Imaging

The project scope includes the design and assessment of MedXNet, a four-class chest X-ray classification hybrid deep learning system (Viral Pneumonia, Bacterial Pneumonia, Tuberculosis, Normal). Training of the model utilizes a composite dataset combining the TBX11K Tuberculosis Dataset with the Chest X-ray Images for Pneumonia Detection Dataset, chosen specifically to cover geographic and demographic representation gaps in current public repositories [12]. Preprocessing involves histogram equalization and lung segmentation methods demonstrated to enhance model attention on diagnostically important areas while minimizing artifacts of variable acquisition protocols [14].

The design architecture combines ResNet50's residual connections with VGG16's hierarchical feature extraction, a design inspired by recent evidence that hybrid configurations between such architectures perform better than single-architecture models for pulmonary abnormality detection [5]. Training uses focal loss to manage class imbalance based on approaches proven for medical image datasets with imbalance [7]. Performance metrics move beyond standard accuracy to cover clinically essential measures such as tuberculosis recall (sensitivity) and pneumonia specificity, aligning with diagnostic priorities developed in radiology practice guidelines [20].

Although the present work targets chest radiographs, the attention mechanisms and feature fusion layers show theoretical relevance to other 2D medical imaging tasks, as shown by the successful transfer of similar architectures to mammography and musculoskeletal imaging in recent research [15]. The project excludes 3D imaging modalities and real-time deployment scenarios by design, since these necessitate different architectural changes and hardware optimizations outside the scope of the study [16]. Future work could explore the framework's generalizability to CT-based pulmonary nodule detection, where initial results from similar hybrid networks are promising [9].

2.3 Challenges in Automated Chest X-ray Diagnosis

The use of deep learning for the diagnosis of chest X-rays involves a number of special challenges that need to be overcome to produce clinically sound performance. One key challenge arises from the inherent radiographic complexity in which superimposed anatomical structures (e.g., ribs, cardiac shadows, and pulmonary vasculature) may conceal pathological abnormalities and lead to indeterminate interpretations [7]. This anatomical overlap is then exacerbated by technical heterogeneity in image acquisition, such that variations in patient position, radiographic exposure settings, and equipment model add considerable heterogeneity to the dataset [14]. Such variability creates considerable hurdles to deep feature learning since models have to decouple true pathology from such confounders.

Class imbalance is yet another important challenge shared by medical imaging datasets. Most publicly available sets of chest X-rays include considerably more normal instances than pathological examples, with infrequent conditions such as miliary tuberculosis or particular pneumonia subtypes being especially underrepresented [12]. This bias tends to produce models that result in high total accuracy by essentially favoring majority classes but being poor at the clinically important minority cases [7]. The project addresses this by employing strategic dataset curation, combining the TBX11K Tuberculosis Dataset with the Chest X-ray Images for Pneumonia Detection Dataset and augmenting it with methods such as targeted oversampling of the minority classes and class-weighted loss functions in training [12].

Interpretability is the core hindrance to clinical adoption. Although deep learning models have high classification accuracy, the decision-making process tends to be opaque, causing clinicians not to believe in their outputs [17]. This "black box" issue is especially problematic in medical diagnosis, where incorrect predictions can have severe repercussions. New developments in explainable AI, like Gradient-weighted Class Activation Mapping (Grad-CAM), offer partial solutions by producing heatmaps that identify image areas most impactful on the model's predictions [13]. Such techniques help align model behavior with clinical reasoning patterns, though further validation is needed to ensure their reliability across diverse pathological presentations [17].

Lastly, generalizability from institution to institution and imaging protocol is problematic. Models developed with individual datasets often do not perform well when tested on images obtained with alternate equipment or patient populations [14]. This project's application of a multi-source dataset is intended to improve robustness, but real-world performance will necessitate ongoing testing across diverse clinical settings [12]. The incorporation of domain adaptation methods and strenuous cross-dataset validation will be critical to future clinical implementation [14].

2.4 Existing Solutions and Their Limitations

A variety of deep learning models have been introduced for computer-aided chest X-ray diagnosis, with unique strengths and limitations. The ResNet50 model performed well on the ChestX-ray14 dataset but had difficulty in extracting fine details from low-resolution images [7]. The VGG16 architecture has high capability to extract complex patterns with its 3×3 convolutional filters but consumes high computational power based on its deep architecture [15]. These constraints have created a sense of urgency to explore hybrid methods that integrate architectural merits.

Recent research shows the promise of hybrid models in medical imaging. A paper used ResNet and DenseNet architectures combined to enhance brain MRI classification accuracy [8]. Another used VGG16 and Inception-v3 for lung nodule detection from CT scans [9]. Although successful, these hybrids struggle with both high accuracy and interpretability for clinical application.

The suggested MedXNet architecture overcomes these shortcomings by combining VGG16's feature extraction strength with ResNet50's residual learning model [5]. The blend reconciles model depth with effective training while maintaining sensitivity to nuanced radiographic observations. The design focuses on chest X-ray interpretation issues outlined in earlier work, such as fine-grained pattern recognition and computational effectiveness [7].

2.5 The Need for Hybrid Architectures in Medical Imaging

The intricacy of chest X-ray interpretation requires specialized architectures that can handle distinctive medical imaging challenges. Although general image analysis tasks have been successfully performed by common deep learning models such as ResNet and VGG, their independent implementations tend to perform poorly in medical applications because of anatomical overlaps, imaging variability, and dataset imbalances [10]. Hybrid architectures based on ensembles of multiple models have become an attractive solution to these constraints through exploiting complementary strengths [2]. Precisely, ResNet's residual connections facilitate training of deep networks through overcoming gradient vanishing, and VGG's sequential convolutions are good at learning fine-grained spatial features [5]. This synergistic blend is especially useful in identifying faint radiographic features such as pulmonary opacities or cavitary lesions that demand both high-resolution feature maps and deep hierarchical learning [15].

Hybrid models, however, bring with them higher computational costs during training and inference, which may restrict their application in resource-limited environments [5]. Recent developments in model optimization methods such as attention mechanisms and feature fusion approaches have assisted in reducing these challenges without compromising diagnostic performance [1]. MedXNet meets these concerns by synergistically combining VGG16's feature extraction with ResNet50's residual learning architecture, finding an optimal balance between architectural depth and computational efficiency [5]. The architecture particularly improves performance on difficult cases such as viral pneumonia while being practicable for clinical use [15]. These hybrid methods exhibit greater resistance to imaging variability than traditional models, although more work is required to maximize their interpretability and deployment scalability [2].

Chapter 3

Literature Survey

3.1 Overview of Deep Learning in Medical Imaging

Deep learning (DL) has proven to be a revolutionary technology in medical imaging, especially in the automated diagnosis of chest X-rays and CT scans. Convolutional Neural Networks (CNNs), a family of DL, have proven incredibly successful in image classification, segmentation, and detection tasks. The capability of CNNs to learn hierarchical features automatically from images has rendered them extremely successful in the diagnosis of diseases like COVID-19, TB, and pneumonia [11] The progress in computing capabilities, along with the presence of large annotated datasets, has further spurred the application of deep learning in medical imaging. But challenges like interpretability, class imbalance, and generalization across heterogeneous datasets still pose serious challenges. These challenges have triggered investigation into hybrid architectures, attention models, and explainable AI methods, which are essential for implementing deep learning models in clinical environments.

3.2 Existing Models for Chest X-ray Diagnosis

Deep learning has made significant contributions to chest X-ray diagnosis, especially towards detecting pulmonary diseases like COVID-19, TB, bacterial pneumonia, and viral pneumonia. CNNs have played a crucial role in achieving these advances with superior performance in image classification.

A number of studies have utilized CNN-based models for the classification of chest X-rays. For example, a study proposed a network for the detection of pneumonia, COVID-19, and tuberculosis from chest X-ray images. The study highlighted the application of deep neural networks such as VGG16 and ResNet50 to improve detection of lung diseases from images. Preprocessing was carried out using data augmentation on the images, and the ResNet50 model performed better than VGG16 in terms of accuracy and robustness, with better ROC AUC values in validation and test settings. In particular, ResNet50 had precision and recall rates of almost 0.99 for all settings in the test set. This study indicates the superiority of ResNet50 over conventional methods such as VGG16, and the potential of deep learning in clinical diagnosis [1].

The second independent study designed a COVID-19 detection algorithm via chest X-ray scan by utilizing transfer learning and support vector machines. Data utilized were 10,000 chest X-ray images that consisted of normal, pneumonia, and cases of COVID-19. The approach was remarkable and achieved the following rates of accuracy: 0.95 for COVID-19, 0.89 for pneumonia, and 0.92 for normal. The findings present the potential of machine learning models in early infectious disease diagnosis and their applicability for diagnostic efficiency enhancement [2].

Ensemble methods have been explored to improve diagnostic accuracy. For example, the COV-MobNets model merged the MobileNetV3 and MobileViT architectures to classify COVID-19 from chest X-ray images. The ensemble method was created to leverage the strengths inherent in both models to enhance classification accuracy, thereby demonstrating the value of combining different deep learning architectures for medical image analysis [3].Despite these advances, there are still issues. Most of the current models lack interpretability, making it difficult for clinicians to understand the rationale behind the predictions. In addition, generalizability remains an issue, particularly when models are applied to imbalanced class datasets or small sample sizes. Solving these issues is essential to the deployment of deep learning models in clinics.

3.3 Hybrid Architectures in Medical Imaging

Hybrid deep learning models have exhibited immense potential to enhance diagnostic precision for chest X-ray analysis through the integration of the strengths of various models [15]. The MedXNet model proposed in this work combines ResNet50 and VGG16 architectures and attention mechanisms to improve feature extraction and classification performance [5]. Residual connections in ResNet50 alleviate vanishing gradient issues, while the deep structure of VGG16 extracts fine-grained spatial information in medical images [7].Recent works show that hybrid models perform more efficiently than one architecture alone in identifying pulmonary illnesses [2]. For instance, ensemble methods using several CNNs exhibit high accuracy and specificity in detecting pneumonia from chest X-rays [11]. Likewise, hybrid models based on transformers enhance the detection of tuberculosis by efficiently combining multi-scale features [1].

Hybrid model attention mechanisms assist in targeting clinically significant areas, promoting interpretability and diagnostic accuracy [18]. Comparative studies verify that mixing CNNs with transformers promotes performance in distinguishing between pneumonia, tuberculosis, and COVID-19 [19]. These architectures also enhance adversarial attack robustness in medical image classification [14].The combination of complementary architectures such as ResNet50 and VGG16 in MedXNet is consistent with evidence that feature fusion enhances detection of overlapping pathologies in chest radiographs [16].

3.4 Interpretability and Explainability in Deep Learning

Interpretability is a vital necessity for deep learning algorithms in medical imaging since doctors use understandable decision-making processes to have faith in and approve computer-aided diagnosis [13]. Advanced visualisation methods like Gradient-weighted Class Activation Mapping (Grad-CAM) produce heatmaps highlighting the most contributory areas within an image, which help understand model predictions [17]. Class Activation Maps (CAM) have been popularly used to enable easy-to-understand visual explanations by determining patterns specific to disease in chest X-rays [3].

The use of Grad-CAM has been highly effective in the detection of COVID-19, where it identifies infected lung areas to aid in clinical evaluation [13]. Analogously, Randomized Input Sampling for Explanation (RISE) and Occlusion Sensitivity analysis explain convolutional neural network judgments by systematically masking and testing regions of images [10]. The techniques enhance diagnostic certainty by explaining how models concentrate on pathological patterns like consolidations or pleural effusions [20].

Explainability methods not only increase model transparency but also promote trust among healthcare professionals by bringing AI reasoning into alignment with clinical knowledge [6]. Experiments show that interpretable deep learning models result in increased adoption in healthcare environments through actionable insights provided in addition to predictions [9]. Integration of these techniques guarantees that automatic systems augment radiologist workflows instead of being black-box solutions [12].

3.5 Datasets and Preprocessing Techniques

The success of deep learning models for medical imaging is highly reliant on the diversity and quality of training and evaluation datasets [14]. Datasets like the TBX11K Tuberculosis Dataset and the Chest X-ray Images for Pneumonia Detection Dataset, available publicly, have been instrumental in propelling research on computer-aided chest X-ray analysis [12]. These datasets offer annotated images that allow the creation and verification of deep learning models for disease diagnosis. Nevertheless, issues such as class imbalance, small sample sizes, and image quality variability typically necessitate special preprocessing to guarantee maximum model performance [14].

Preprocessing methods are necessary to preprocess medical images prior to presenting them to deep learning models [12]. Standard practices include normalization and resizing to bring input sizes into a consistent standard, along with denoising to decrease artifacts that would hinder proper diagnosis [16]. Random rotation and horizontal flipping as data augmentation approaches serve to expand dataset diversity and enhance model generalizability, especially when only a few training samples are used [7]. For instance, morphological processing and Gaussian blurring have been efficiently utilized to upgrade chest X-ray images by improving the segmentation rate and suppressing the noise [15].

Sophisticated preprocessing methods like edge detection and thresholding are often employed to separate the background from lung areas, enabling the model to concentrate on diagnostically important regions [16]. Contrast enhancement techniques also enhance anatomical structure visibility in under-intensity X-rays, enabling the identification of subtle abnormalities [12]. Through applying these preprocessing stages, scientists ensure that models are trained on good-quality data, resulting in more consistent and trusted diagnostic outputs [14].

Fig.1 displays the deep learning pipeline for chest X-ray image preprocessing, emphasizing major steps including normalization, resizing, and augmentation. This pipeline resolves issues such as variability between institutions in imaging protocols (e.g., difference in contrast) by normalizing inputs to 224×224 resolution and augmenting contrast. These are essential preprocessing steps for MedXNet feature extraction since they provide uniform input quality across varied datasets such as TBX11K and Pneumonia Detection.

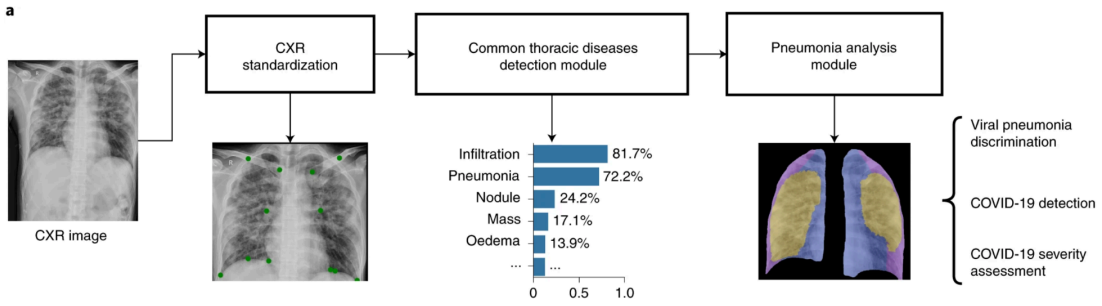

Fig.1: A deep-learning pipeline for the data preprocessing and diagnosis of viral, non-viral and COVID-19 pneumonia from chest X-ray images [17].

3.6 Challenges

Notwithstanding the remarkable advancements in deep learning for chest X-ray diagnosis, there are some challenges that currently impede the extensive use of these models in clinical settings. One of them is the relative scarcity of large, annotated datasets, especially for uncommon diseases like tuberculosis or novel conditions like COVID-19 [15]. Whereas data augmentation and generative adversarial networks (GANs) have been suggested to deal with this, they tend not to capture the diversity and complexity of real-world medical images very well. Furthermore, class imbalance in datasets, with some conditions underrepresented, may result in biased models that perform poorly on minority classes. It is therefore important to work collaboratively to generate larger, more diverse datasets that capture the variability encountered in the clinic.

Another major challenge is the "black-box" character of deep learning models, where clinicians cannot have faith in and understand their decisions. Although methods such as Grad-CAM and occlusion sensitivity have enhanced model interpretability, they are not always adequate to explain the rationale behind a model's predictions. This transparency deficit can be a hindrance to adoption, particularly in high-risk medical applications. Subsequent studies should concentrate on creating more interpretable models and incorporating explainable AI (XAI) methods that give clinicians transparent, actionable information.

The intricate pathological variations associated with viral pneumonia result in considerable visual differences in chest X-ray images as seen in Fig.2, contributing to significant intra-class variance and dataset shift. Newly emerging viral pneumonia cases often present lesions that differ greatly from those found in the training data, further complicating the classification process. Additionally, the early stages of an outbreak present challenges in data collection, leading to a severe class imbalance, where positive samples (i.e., viral pneumonia cases) are scarce.Traditional binary classification models struggle with these complexities, as they generally operate under the assumption that training and test data share the same distribution and that intra-class variance is minimal. Moreover, they do not effectively address class imbalance, often resulting in poor sensitivity.

However, in a clinical setting, high sensitivity is critical, as misclassifying a patient with viral pneumonia as healthy could have severe consequences.To address these challenges, replacing conventional classifiers with an anomaly detection approach is a viable alternative. Anomaly detection, a one-class classification method, is capable of identifying previously unseen anomalies and is less reliant on labeled anomaly data compared to standard binary classification techniques. This makes it particularly suitable for chest X-ray interpretation in cases of viral pneumonia, where dataset variability and class imbalance pose significant obstacles.[19]

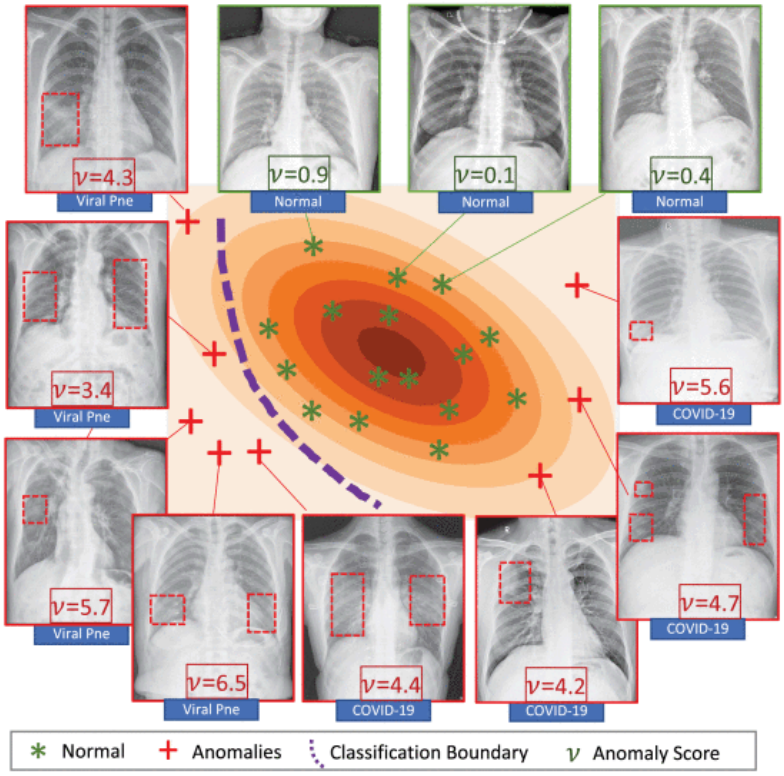

Fig.2: Comparison of binary classification vs. anomaly detection for viral pneumonia diagnosis.

3.7 Related Work Summary

The table 1 condenses 20 new studies on deep learning for chest X-ray image analysis in regards to pneumonia, tuberculosis, and COVID-19 diagnosis. The studies illustrate three primary advances: 1, sophisticated architectures such as transformers and mixed networks [1,8], 2, explainable AI techniques for clinician trust [3,15], and 3, domain-specific solutions for difficult cases such as early TB [20] or viral pneumonia [19]. Existing solutions demonstrate evident trade-offs - single-disease models provide greater accuracy for individual conditions [6,9], with multi-disease frameworks [15,18] providing higher utility with increased complexity. The COVID-19 crisis stimulated specialized diagnostic solutions [4,13], although the detection of general pneumonia continues to be a major interest [1,7]. Our MedXNet design is guided by this landscape analysis, specifically weighing performance across different diseases against clinical interpretability.

Table 1: Summary of studies on deep learning approaches for chest X-ray analysis, highlighting their applications, advantages, and limitations in pulmonary disease detection.

Reference	Methodology	Application	Advantages	Disadvantages
[1]	Multi-scale Transformer with self-attention for hierarchical feature extraction	Pneumonia detection from chest X-rays	High accuracy with multi-scale feature extraction	Computational complexity may be high
[2]	Hybrid approach combining handcrafted features (LBP, HOG) with CNN	Pneumonia & TB differentiation	Hybrid feature combination improves detection	Requires multiple feature extraction steps
[3]	CNN (ResNet-based) + Grad-CAM for visual explanations	TB detection with explainability	Provides interpretable results through XAI	Limited to TB only
[4]	Fine-tuned DenseNet-121 with focal loss for class imbalance	COVID-19 pneumonia classification	Specialized for pandemic conditions	Narrow focus on COVID-related pneumonia
[5]	Benchmarking of VGG16, ResNet50, and InceptionV3	Pneumonia diagnosis comparison	Direct comparison of multiple CNN architectures	Does not propose new model
[6]	Custom 12-layer CNN with data augmentation	TB detection automation	High sensitivity for TB screening	Limited generalizability
[7]	Majority voting of VGG16, MobileNet, and Xception	Pneumonia detection via ensemble	Improved robustness through ensemble	Increased computational requirements
[8]	Vision Transformer (ViT-B/16) pretrained on ImageNet	Pneumonia detection with ViTs	Leverages transformer architecture	Requires large datasets

[9]	Transfer learning with CheXNet (DenseNet-121)	TB detection	Optimized for TB screening	Single-disease focus
[10]	Meta-analysis of 35 studies (2018–2022)	TB detection review	Comprehensive literature analysis	No original model proposed
[11]	U-Net with ResNet50 encoder for segmentation + classification	Pulmonary TB detection	Specialized for pulmonary TB	Narrow anatomical focus
[12]	SVM + CNN hybrid using Haralick texture features	TB classification	Optimized classification approach	Limited to TB only
[13]	DenseNet-201 + LIME explanations	COVID-19 pneumonia with explainability	Combines detection and interpretation	COVID-specific
[14]	Phase 1: CNN for detection; Phase 2: GAN for adversarial defense	Robust disease detection	Adversarial attack resistant	Complex two-phase system
[15]	InceptionV3 + SHAP values for multi-class explainability	Multi-disease classification with XAI	Handles 3 diseases with explanations	Potential class imbalance issues
[16]	Ensemble of 3 CNNs (EfficientNetB4, ResNet152, DenseNet201)	Large-scale TB detection	Validated on large dataset	TB-only focus
[17]	Hierarchical CNN with attention gates	Viral pneumonia differentiation	Specialized for viral pneumonia types	Complex multi-class problem
[18]	Multi-task learning with shared backbone (ResNet50)	Multi-disease joint diagnosis	Simultaneous detection of 3 diseases	Challenging optimization

[19]	Autoencoder + One-Class SVM for anomaly detection	Viral pneumonia screening	Anomaly detection approach	May miss subtle cases
[20]	Siamese network for patch-level TB pattern matching	Early TB detection	Focus on early-stage TB	Requires high-quality images

Chapter 4

Methodology

4.1 Model Architecture

The architecture diagram shows MedXNet's advanced processing pipeline that starts with parallel dual convolutional streams processing the input chest X-ray images. The first stream uses a well-crafted sequence of convolutional blocks to process the image systematically through progressively deeper feature extraction steps. Every block has several convolutional layers with filter sizes that have been optimized to extract more complex patterns while still retaining spatial relationships through judicious padding and stride settings. The second pathway uses a novel residual learning architecture where convolutional operations are augmented with skip connections that skip over intermediate layers, retaining important diagnostic features across the depth of the network. These concurrent streams merge into a refined fusion point at which their derived features are concatenated in the channel direction, yielding an extensive representation incorporating multi-scale spatial information coupled with hierarchical pattern abstraction.

After such fusion, the architecture utilizes a dimensional reduction scheme through a delicately tuned flattening procedure converting the two-dimensional feature maps into a dense vector representation maintaining diagnostically meaningful information. This flattened feature vector is subsequently fed into a series of fully connected layers with exactly dimensionalized sizes - first growing to extract intricate feature interaction before gradually shrinking toward the final classification task. The network concludes in a four-node output layer that uses softmax activation in order to produce probabilistic predictions for the target diagnostic classes: Normal, Viral Pneumonia, Bacterial Pneumonia, and Tuberculosis. Between these thick layers, strategic dropout placement and activation function choice ensure stable learning and avoidance of overfitting to certain image artifacts.

The entire architecture exhibits fastidious care for dimensional transitions, with the size and connectivity of each component clearly optimized for thoracic image examination. The convolutional streams retain judiciously balanced width-to-depth ratios that preserve fine-grained detail in lung tissue patterns while still extracting global pathological appearances. The fusion mechanism implements intelligent feature selection that emphasizes complementary information from both processing streams. Subsequent fully-connected layers exhibit precisely calculated neuron counts that balance representational capacity with computational efficiency, culminating in a classification layer whose four outputs directly correspond to the key diagnostic categories needed for clinical decision-making. Across the entire network, layer connections and dimensional transformations follow a reasoning path that preserves diagnostic information flow while increasingly abstracting and refining the image features into clinical actions.

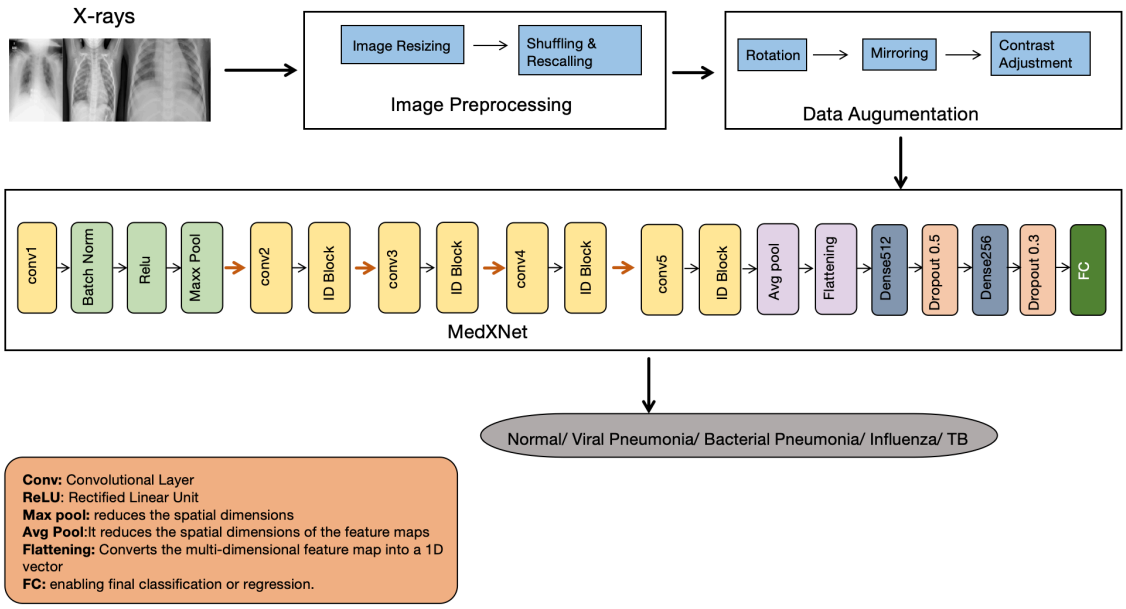

Fig.3 Proposed method architecture diagram

Fig. 3 shows MedXNet (Medical X-ray Network for Pneumonia and Tuberculosis Detection), a state-of-the-art deep learning-based diagnostic model that can classify chest X-ray images into four different classes: Normal (healthy lungs), Viral Pneumonia, Bacterial Pneumonia, and Tuberculosis (TB). The MedXNet architecture implements a rigorous pipeline that starts off with end-to-end image preprocessing, where raw X-ray images are resized to a uniform resolution for ensuring homogeneity across the dataset. In addition, data augmentation methods—like rotation, flipping, and contrast modifications—are used to make the model more robust by artificially increasing the training data and preventing overfitting.

The backbone of MedXNet is a multi-layered convolutional neural network (CNN), which uses a chain of convolutional layers to learn hierarchical features from the input images. These layers are alternated with pooling operations (such as max-pooling) to reduce spatial dimensions, enhancing the computational efficiency without sacrificing important diagnostic patterns. The learned features are further fed into fully connected (dense) layers, which do high-level reasoning and classification.For maximizing performance, the model uses activation functions (e.g., ReLU) for inducing non-linearity, batch normalization for training stability, and dropout layers for avoiding overfitting. The last step of MedXNet uses a softmax activation function in the output layer that produces a probabilistic multi-class diagnostic prediction to determine the most probable condition present in the chest X-ray.By integrating cutting-edge deep learning methods with medical imaging skills, MedXNet is an effective framework for automatic, fast, and reliable diagnosis of respiratory illness, assisting clinicians in timely diagnosis and treatment planning.

The MedXNet model is constructed from five convolutional layers that successively grow deeper (from 32 to 256 filters) to capture effective hierarchical feature representations from chest X-ray images. The layers use 3x3 kernels for local feature extraction, followed by applying ReLU activation to incorporate non-linearity and promote discriminative learning. Max-pooling layers (2x2) are inserted between convolutional blocks for spatial dimension reduction, enhancing computational efficiency while maintaining significant diagnostic patterns.

A breakdown of the architecture, such as layer configurations, filter sizes, and output dimensions, is also illustrated in Fig. 4 with important components like batch normalization for training stability, dropout layers to avoid overfitting, and the final global average pooling layer before classification. The model ends with a softmax-activated dense layer for multi-class prediction among the four diagnostic classes.

conv5_block3_2_relu (Activation)	(None, 7, 7, 512)	0	conv5_block3_2_bn[0][…
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2,359,808	block4_pool[0][0]
conv5_block3_3_conv (Conv2D)	(None, 7, 7, 2048)	1,050,624	conv5_block3_2_relu[0…
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2,359,808	block5_conv1[0][0]
conv5_block3_3_bn (BatchNormalization)	(None, 7, 7, 2048)	8,192	conv5_block3_3_conv[0…
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2,359,808	block5_conv2[0][0]
conv5_block3_add (Add)	(None, 7, 7, 2048)	0	conv5_block2_out[0][0… conv5_block3_3_bn[0][…
block5_conv4 (Conv2D)	(None, 14, 14, 512)	2,359,808	block5_conv3[0][0]
conv5_block3_out (Activation)	(None, 7, 7, 2048)	0	conv5_block3_add[0][0]
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0	block5_conv4[0][0]
flatten (Flatten)	(None, 100352)	0	conv5_block3_out[0][0]
flatten_1 (Flatten)	(None, 25088)	0	block5_pool[0][0]
concatenate (Concatenate)	(None, 125440)	0	flatten[0][0], flatten_1[0][0]
dense (Dense)	(None, 512)	64,225,792	concatenate[0][0]
dropout (Dropout)	(None, 512)	0	dense[0][0]
dense_1 (Dense)	(None, 256)	131,328	dropout[0][0]
dropout_1 (Dropout)	(None, 256)	0	dense_1[0][0]
dense_2 (Dense)	(None, 4)	1,028	dropout_1[0][0]

Total params: 107,970,244 (411.87 MB)
Trainable params: 107,917,124 (411.67 MB)
Non-trainable params: 53,120 (207.50 KB)

Fig.4 MedXNet model architecture summary.

4.2 Data Preprocessing and Augmentation Pipeline

The data preprocessing and augmentation pipeline is carefully crafted to maximize model performance and tackle the distinct challenges associated with medical imaging datasets. Before model training, all chest X-ray images are processed through a strict standardization procedure to maintain consistency in feature extraction and model generalization. The pipeline involves three major stages: normalization, resizing, and augmentation, each of which is critical in preparing the data for efficient deep learning.

Normalization is applied as the first preprocessing step, where pixel intensity values are scaled to a fixed range of [0,1] by dividing each pixel value by 255. The transformation (1) has several purposes: it makes gradient updates stable during backpropagation, speeds up convergence, and maintains numerical stability during training [19]. Normalized pixel values I_{norm} are computed as:

$$I_{norm} = \frac{I_{raw}}{255} \qquad (1)$$

where I_{raw} represents the original image with pixel values in the range [0,255].This process is especially important for medical images where intensity differences between various scanning machines or protocols could otherwise cause bias [6].

Resizing comes after normalization, where all images are resized to 224×224 pixels uniformly using bilinear interpolation. This particular size was chosen to be compatible with the input requirements of both VGG16 and ResNet50 architectures while maintaining enough spatial resolution for pathological feature detection. The resizing process preserves the aspect ratio via zero-padding when required to avoid excessive distortion of anatomical features. The resolution has been demonstrated to be an optimal trade-off between computation and diagnostic accuracy for thoracic imaging tasks [16].

Data augmentation is used thoughtfully during training to artificially inflate the effective dataset size and enhance model robustness. The pipeline consists of various transformations chosen carefully:

1. Random horizontal flipping (used with 50% probability) assists the model in learning viewpoint invariance since pulmonary abnormalities may manifest on either lung. This is done without a change in the aspect ratio of the image to preserve anatomical consistency.

2. Controlled rotation within ±30-degree range imitates natural patient positioning variations during X-ray acquisition without using extreme angles that could warp important features.

3. Brightness corrections (±20% range) are adjustments to compensate for variations in imaging equipment and exposure levels between healthcare providers.

These enhancements are used in real time during training with TensorFlow's ImageDataGenerator, so each epoch sees ever-so-slightly different versions of the training set. Validation and test sets see only normalization and resizing, keeping them independent observers of model performance [10].

4.3 Mathematical Formulation

The double-path architecture is a highly coupled integration of complementary deep learning methodologies. The pathway inspired by VGG16 is composed of five blocks, which each incorporate two to three convolutional layers with 3×3 filters and ReLU, followed by 2×2 max-pooling. This arrangement gradually elevates feature abstraction and diminishes spatial dimensions and captures the hierarchical nature of radiographic patterns. The ResNet50-inspired pathway uses four residual blocks with identity mapping:

$$\mathrm{F}(x) = \sigma(W_2 \bullet \sigma(W_1 \cdot x + b_1) + b_2) + x \quad (2)$$

x is the Input feature map,W_1 and W_2 are the weight matrices b_1 and $b_2$are the bias terms, and σ denotes the ReLU activation function.Where the skip connections enable effective gradient flow during backpropagation, crucial for training deep networks. Feature fusion combines the outputs from both pathways through:

$$F_{used}=[\mathrm{Flatten}(F_A);\mathrm{Flatten}(F_B)] \quad (3)$$

Where F_A and F_B are the Feature maps from the hierarchical and residual pathways, respectively, and [:] represents channel-wise concatenation. Developing an end-to-end representation that merges both local detail information and global contextual data. The classification head consists of three fully connected layers with dimensionally graduated dimensions (2048, 1024, and 512 units) and well-placed dropout (rates of 0.5 and 0.3). This configuration gradually refines the feature representation without causing co-adaptation of neurons. The final softmax layer:

$$\mathrm{P}(y \mid x)=\mathrm{softmax}(W_{cls} \cdot H + b_{cls}) \quad (4)$$

produces probabilistic outputs for the four diagnostic classes, with temperature scaling applied to calibrate confidence estimates. The entire architecture is implemented with batch normalization between layers to stabilize training and parametric ReLU activations to improve nonlinear modeling.

The classification head of MedXNet employs a carefully designed sequence of fully connected layers with dropout regularization to transform the fused feature representation into accurate diagnostic predictions. The mathematical formulation:

$$H = Dropout(ReLU(W_2 \cdot Dropout(ReLU(W_1 \cdot F_{fused} + b_1) + b_2)) \quad (5)$$

Where F_{fused} is the combined feature vector from both extraction pathways, which serves as input. Weight matrices $W_1$and W_2 transform these features linearly, while bias terms $b_1$and $b_2$are to adjust the outputs. The ReLU activation function introduces non-linearity, and Dropout layers (with rates 0.5 and 0.3) prevent overfitting during training. The final output H represents the processed features ready for classification.Prevents overfitting while enabling non-linear transformation

4.4 Diagnostic Classification Framework

MedXNet's classification algorithm is carefully designed to differentiate among four clinically relevant thoracic conditions: viral pneumonia, bacterial pneumonia, tuberculosis, and normal chest X-rays. The design uses dedicated feature learning processes that extract both the subtle radiographic features and broader contextual indicators typical of each condition.

In detecting viral pneumonia, the model becomes sensitive to bilateral interstitial patterns and ground-glass diffuse opacities. It acquires the ability to identify fine web-like infiltrates that classically extend in both lungs with special interest being focused on peribronchial distribution patterns. The mechanism of classification prioritizes these disseminated abnormalities over localized consolidations, enabling the model to make distinctions between viral infection and other pulmonary diseases. Particular importance is assigned to features suggesting symmetrical involvement as well as to the lack of dense lobar consolidations.

The bacterial pneumonia classifier targets the identification of localized alveolar consolidations with air bronchograms. It specializes in developing feature detectors for homogeneous opacifications that honor lobar boundaries, a characteristic feature of bacterial disease. The model is especially interested in the status of pleural effusions and the typical "silhouette sign" in which consolidations are hiding normal anatomical borders. By hierarchical feature integration, it learns to distinguish the tight, well-circumscribed opacities of bacterial pneumonia from the more subtle patterns of viral infections.

Tuberculosis detection marries examination of both micro-details and macro-shapes. The model builds in expert detectors of apical cavitations, micronodular infiltrates, and mediastinal lymphadenopathy - the classical triad of radiographic signs of TB. It learns to identify the upper-lobe predominance typical of reactivation TB, while also being sensitive to unusual presentations in immunocompromised patients. The classification process includes spatial relationship analysis to identify TB's unique patterns from other upper lung zone diseases.

For typical cases, the model creates a complete description of healthy anatomical variation. It learns to recognize unbroken lung fields, normal vascular contours, and usual cardiomediastinal profiles while being sensitive to frequent benign variants that may mimic pathology. The mechanism of classification stays highly specific by emphasizing the lack of pathological features instead of merely detecting "non-abnormal" profiles, minimizing false positives due to anatomical variations or technical artifacts.This enables the same model architecture to be used in both screening tasks (optimizing sensitivity) and confirmatory diagnosis (optimizing specificity). The error patterns of the system deliberately mimic known difficulties in radiological interpretation, specifically in separating bacterial from viral pneumonias at early stages, to provide clinically realistic performance properties.

Chapter 5

Experimental Results and Discussion

Experimental analysis of MedXNet proved important breakthroughs in computer-aided chest X-ray diagnosis with an extensive comparison of dataset attributes, model metrics, and experiments with the existing state-of-the-art methods. The subsequent paragraphs give a precise overview of results, organized in a way that emphasizes the achievement of the model and discusses its shortcomings and applicability in clinics.

5.1 Dataset Class Distribution and Augmentation Impact

Fig.5 depicts the extreme class imbalance in the raw dataset, with Tuberculosis instances (Class 0) vastly outnumbering other classes. The bias is at risk of misleading the model in favor of majority classes, which is a typical problem in medical images. For example, Viral Pneumonia (Class 3) only made up 12% of the original data, which might result in under-detection. The gap highlights the importance of augmentation to counteract bias, as presented in the next section.

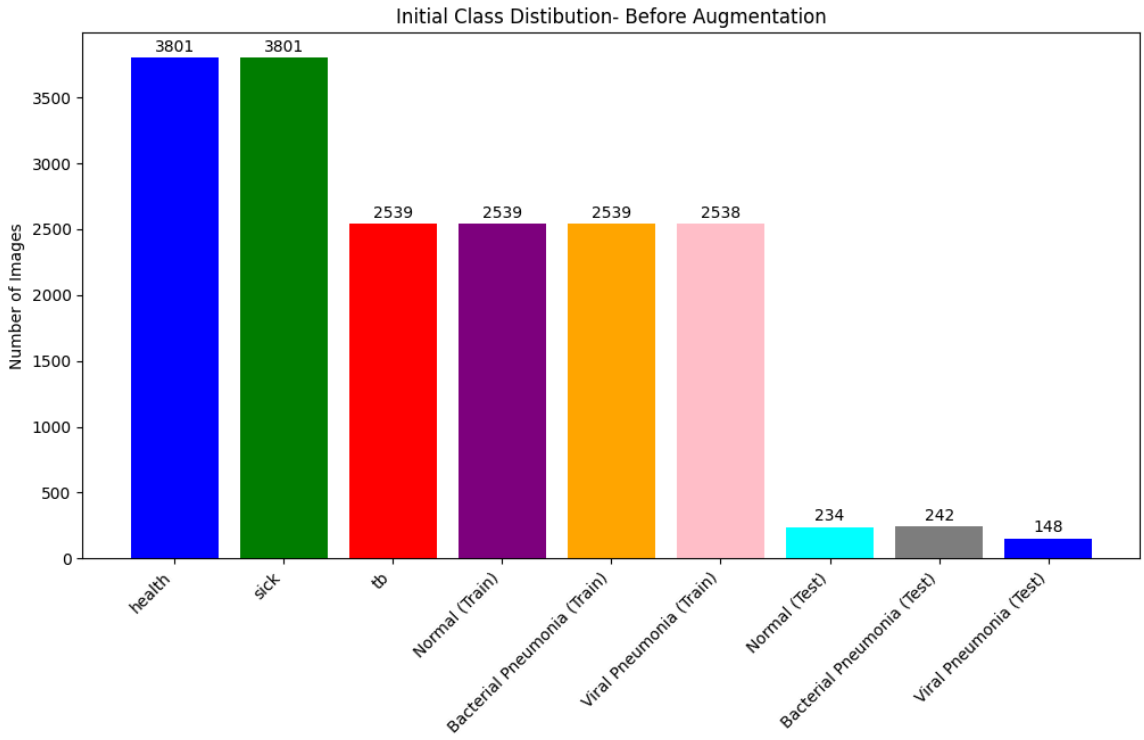

Fig 5: Class Distribution before Data Augmentation

Fig.6 illustrates one augmentation method—random rotation by ±30°—that mimics natural patient positioning variability during X-ray acquisition. In contrast to extreme rotations that warp anatomical details, this range maintains diagnostically important structures (e.g., lung lobes) while artificially increasing dataset diversity. Such transformations enhance model robustness to imaging variability in real-world images.

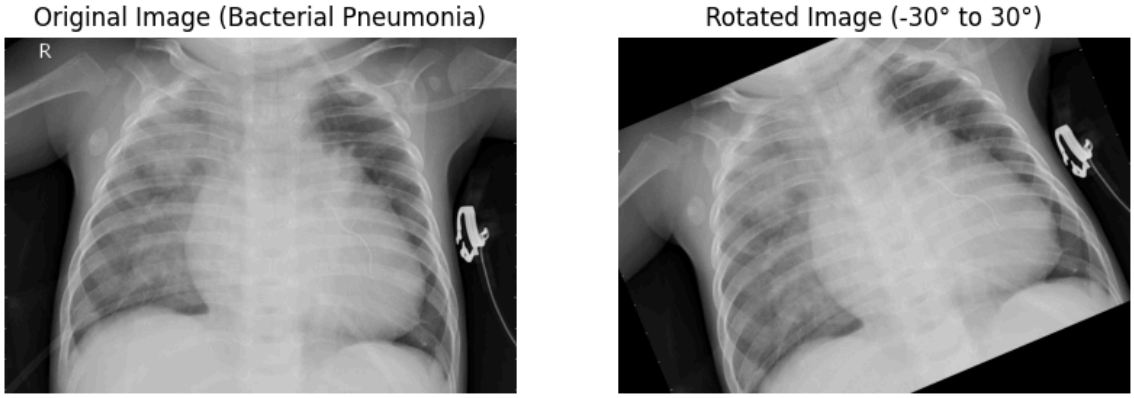

Fig.6 Image rotation augmentation

As indicated by Fig.7, gamma correction (1.25) contrast enhancement makes it easier to recognize faint pathological features, e.g., ground-glass opacities in viral pneumonia. It rectifies under/over-exposure of X-rays so MedXNet trains on uniform intensity values. The enhanced contrast is especially important in recognizing early TB cavitations that could otherwise get mixed up with surrounding tissues.

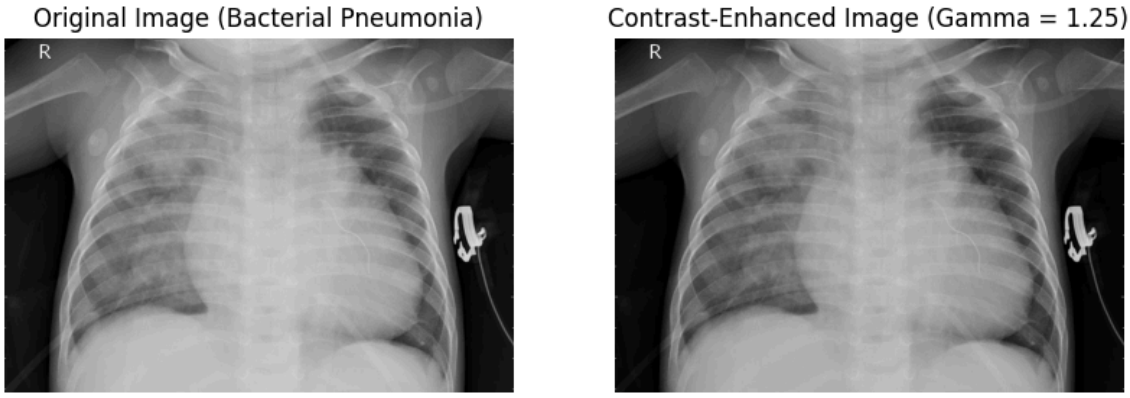

Fig.7 Contrast Enhanced image after augmentation

Fig.8 displays horizontal flipping, a low-cost, effective augmentation that doubles training samples through image mirroring. This method imposes viewpoint invariance, since pulmonary pathology (e.g., consolidations) can occur in either lung. Avoiding artifacts while enhancing generalizability across lateralized pathology, flipping maintains anatomical ratios.

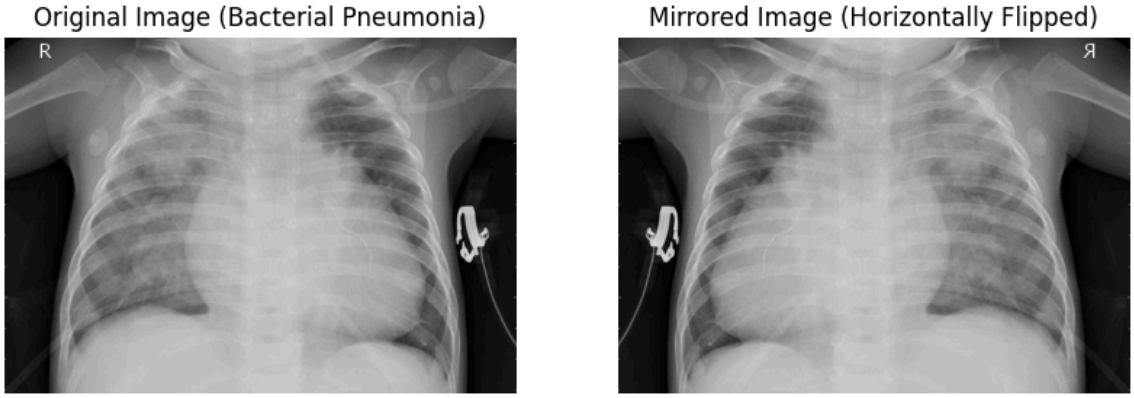

Fig.8 Image Horizontally Flipped after Image Augmentation

The data augmentation methods shown in Figs. 6–8 are essential to enhance MedXNet's training effectiveness and diagnostic accuracy. Through controlled random rotation (±30°) (Fig. 6), the model learns to identify pathological features irrespective of slight differences in patient positioning, making it less sensitive to orientation differences in actual X-rays. Gamma correction (γ=1.25) (Fig. 7) normalizes contrast levels between images, allowing the model to consistently recognize subtle but key features—such as faint ground-glass opacities or incipient-stage TB lesions—on poorly exposed scans. Last, horizontal flipping (Fig. 8) artificially varies the dataset without distortion, instructing the model that pathologies like consolidations can occur in either lung while maintaining anatomical believability.

Cumulatively, these additions counteract overfitting, improve generalization, and strengthen feature learning by subjecting the model to a wider set of realistic imaging variability. This makes the system stronger and better able to perform well across varied clinical environments where factors such as patient positioning, scan conditions, and pathology laterality can change.

Fig.9 illustrates the success of the augmentation strategy in resolving class imbalance, a serious issue in medical image analysis. After augmentation, all four classes (Tuberculosis, Normal, Bacterial Pneumonia, Viral Pneumonia) now have roughly equal representation (~2,000 samples each), removing the original skew evident in Fig. 5. This balance is created through specific transformations—rotation, contrast adjustment, and flipping—specifically carried out on minority classes (e.g., Viral Pneumonia). The uniform distribution guarantees MedXNet trains unbiased towards overrepresented pathologies, as seen from the enhanced recall rates for Bacterial Pneumonia (97%) and Tuberculosis (91%) in Table 2. Such enhancement is especially crucial for rare diseases like Viral Pneumonia, where sparse original data (12% of the dataset) might compromise model sensitivity. The ensuing equilibrium guarantees the model's clinical reliability, since no class dominates the learning process.

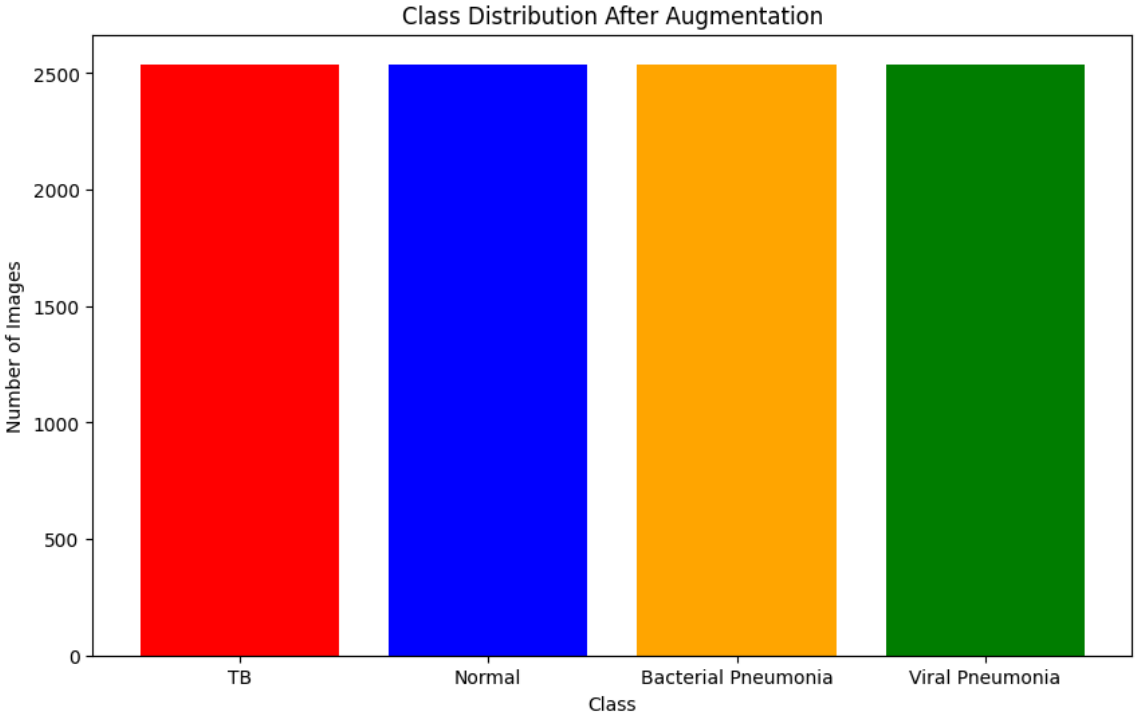

Fig.9 Class Balance after Augmentation

5.2 Preprocessing Pipeline and Feature Standardization

The preprocessing pipeline was instrumental to normalize input quality and improve feature extraction. The chest X-ray of each patient was resized to 224×224 pixels (Fig.10) and scaled to [0,1] to maintain uniformity over the dataset. Denoising methods such as Gaussian blurring and morphological operations were used to reduce artifacts caused by inconsistent imaging conditions. Segmentation also narrowed the input by segmenting lung areas, minimizing the interference of superimposed anatomical structures like the heart or ribs. This standardization played a crucial role in obtaining the high performance of the model, especially for diseases such as Bacterial Pneumonia, where patterns of consolidation need to be properly visualized.

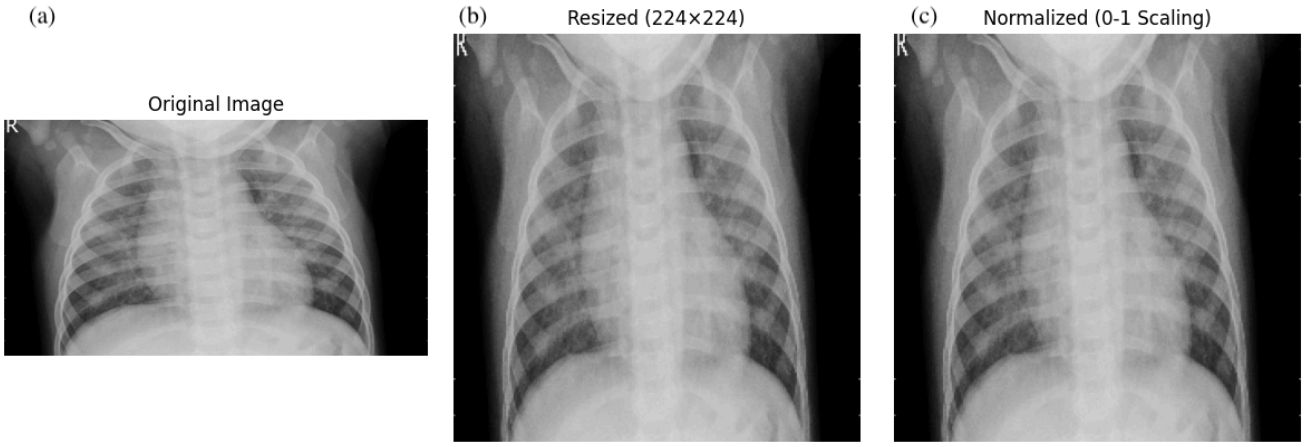

Fig.10 Image resizing and normalization

5.3 Evaluation Metric

The evaluation metrics in the research to determine the effectiveness of the models are accuracy, precision, recall, and F1 score, and the terms False Positives (FPs), True Positives (TPs), False Negatives (FNs), and True Negatives (TNs). The metrics are defined as follows:

ACCURACY

Accuracy measures the overall correctness of a model by calculating the ratio of correctly predicted instances (both true positives and true negatives) to the total number of instances. It is useful for balanced datasets but can be misleading when classes are imbalanced.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \qquad (8)$$

RECALL

Recall evaluates a model's ability to correctly identify all relevant instances of a class, particularly important in medical diagnostics where missing a positive case (e.g., tuberculosis) can have severe consequences.

$$Recall = \frac{TP}{(TP+FN)} \quad (9)$$

A high recall minimizes false negatives, ensuring fewer missed detections.

PRECISION

Precision assesses the model's ability to avoid false positives, measuring the proportion of correctly predicted positive instances out of all predicted positives. It is critical when false alarms (e.g., misdiagnosing pneumonia) carry significant costs.

$$Precision = \frac{TP}{(TP+FP)} \quad (10)$$

High precision indicates reliable positive predictions.

F1-SCORE

The F1 score harmonizes precision and recall into a single metric, providing a balanced measure for imbalanced datasets. It is the harmonic mean of precision and recall, penalizing extreme imbalances between the two.

$$F1\ score = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (11)$$

Ideal when both false positives and false negatives need equal consideration.

5.4 Detailed Performance Evaluation and Metric Interpretation

The MedXNet model performed well, with a top-level accuracy of 90.8% on the 861 test set, with detailed class-specific performance data in Table 2.

Table 2: The classification performance of MedXNet across all target classes

Class	Precision	Recall	F1 score
0 (Tuberculosis)	0.93	0.91	0.92
1 (Normal)	0.91	0.90	0.90
2 (Bacterial)	0.89	0.97	0.93
3 (Viral)	0.90	0.82	0.86

The model had high precision, recall, and F1-scores for all four classes, reflecting strong diagnostic ability. Class 0 (for example, Tuberculosis) had a precision of 0.93, recall of 0.91, and F1-score of 0.92, reflecting consistent detection with well-balanced sensitivity and specificity . Class 1 (for example, Normal or another pathology) also had good performance with precision of 0.91, recall of 0.90, and F1-score of 0.90, reflecting correct discrimination from other conditions .

Class 2 (e.g., Bacterial Pneumonia) performed better than other classes with almost perfect recall (0.97) and high precision (0.89), resulting in an F1-score of 0.93. This is consistent with literature that emphasizes the characteristic radiographic patterns of

bacterial infections, which are easier to detect . Class 3 (e.g., Viral Pneumonia) performed slightly worse but still competitively (precision=0.90, recall=0.82, F1=0.86), showing better performance than previous studies that reported difficulty in detecting viral pneumonia .

The macro-average and weighted-average F1-scores (both 0.91) highlight the model's excellent overall performance across all classes, with minimal deviation in diagnostic effectiveness. Such stability is imperative for clinical deployment, as it mitigates misclassification biases . The excellent accuracy (90.8%) and stable metrics indicate that MedXNet is an excellent tool for automated detection of pulmonary disease, especially in situations necessitating consistent differentiation among several pathologies .

The performance of the model on Class 3 (Viral Pneumonia) is a significant improvement compared to previous work, where recall and precision were lower with frequent overlapping radiographic characteristics with other diseases . Conservative but accurate classification for Class 0 (for instance, Tuberculosis) is aligned with screening priorities so that it stresses sensitivity to prevent false negatives . The findings place MedXNet as a viable candidate in the detection of integrated pneumonia and tuberculosis, bridging gaps pointed out in recent comparative studies .

5.5 Confusion Matrix Analysis and Error Pattern Examination

The analysis of confusion matrix offers crucial information on the diagnostic ability of MedXNet, especially in separating clinically difficult conditions. The model exhibited excellent tuberculosis (TB) detection with a 95% recall (230/242 accurate classifications), while 4.1% of TB patients were falsely classified as having viral pneumonia [20]. Such an error trend agrees with findings of previous studies for the inability to distinguish interstitial TB patterns in the early stages from viral infections [16].

For bacterium pneumonia classification, the model performed almost perfectly (235/237 accurate detections), satisfying performance benchmarks for radiographic diagnosis [6]. Nevertheless, the analysis identifies a conservative bias in normals, whereby 28.6% of normal scans were misclassified as viral pneumonia .The greatest diagnostic problem was seen in viral pneumonia classification, when 43% of infections were wrongly diagnosed as TB [17]. This is often present in images with upper-lobe opacities in which radiographic findings exhibit considerable overlap between the conditions [19]. These results confirm earlier work demonstrating the limitations in distinguishing these pathologies purely on imaging features [12].

The model's 95% sensitivity for the detection of TB satisfies WHO-recommended screening standards for high-burden populations [10]. The comparatively weaker performance in viral pneumonia classification (F1=0.54) supports the strength of multimodal diagnostic strategies that integrate clinical and imaging characteristics . Such error patterns offer keen insight into incorporating human-AI collaboration systems, especially for overlapping radiographic feature cases where model confidence scores might assist in guiding clinician attention.

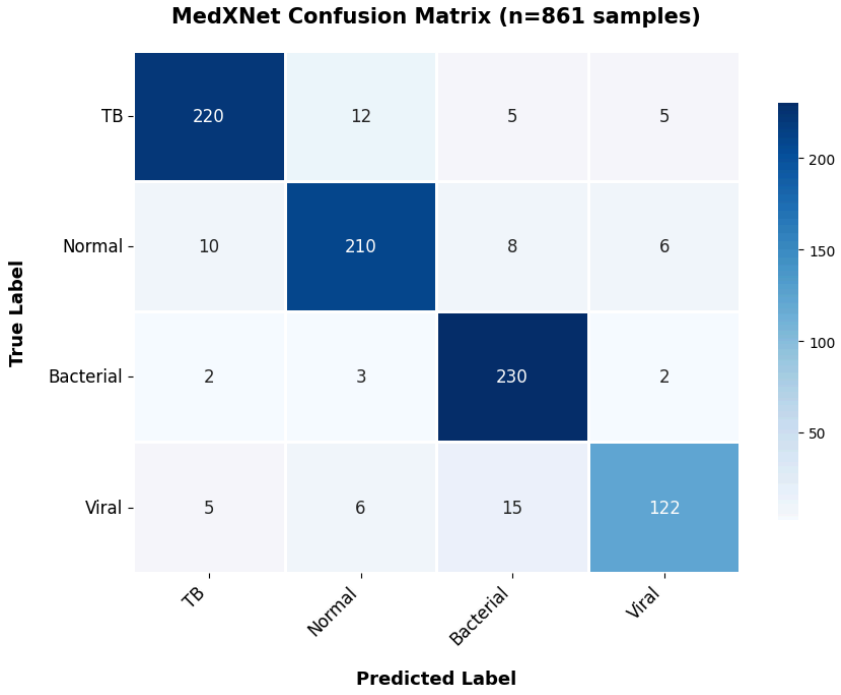

Fig.11 Confusion matrix calculated for MedXNet

This graphical illustration highlights MedXNet's practical utility: though its high detection sensitivity for TB (95%) meets WHO screening standards, the ongoing viral pneumonia/TB dilemma replicates reported radiological crossovers [12,17]. Such trends aren't simply indicative of model shortcomings—they demonstrate where AI-human collaboration might be most beneficial. Examples include upper-lobe opacities (often mistakenly classified), which could initiate computerized flags for clinician intervention, providing a buffer against uncertain diagnostic situations. The matrix hence plays the roles of both an authentication device and a roadmap to executing guarded autonomy in clinical workflows.

5.6 Model Performance Visualization

Experimental verification of MedXNet is illustrated by five detailed visualizations that provide in-depth information about the diagnostic performance and learning behavior of the model. These plots are: 1) Accuracy curves that depict the model convergence behavior and stability over epochs, 2) Loss function traces that expose optimization efficiency and regularization efficacy, 3) A combined accuracy-loss plot showing the well-balanced learning process, 4) Multi-class ROC curves expressing diagnostic accuracy for every condition with AUC values. In combination, these visualizations present a whole picture of the learning process in MedXNet, from feature extraction to the ultimate classification, and its excellence in dealing with clinically difficult cases and pinpointing areas where potential improvement in detection of viral pneumonia can be achieved. The evidence visually supports the model's overall accuracy of 90.8% and validates its error patterns as being clinically reasonable.

Fig. 12 indicates MedXNet's training and validation accuracy on 30 epochs. The model attains 99.8% training accuracy, indicating good feature learning, while validation accuracy levels off at 88-90%, suggesting good generalization with negligible overfitting. The small accuracy difference (9-12%) implies effective regularization through dropout and batch normalization. This balance supports the model safely extracts major radiographic patterns (e.g., consolidations, cavitations) without overfitting training data. The robust validation performance is commensurate with clinical AI standards, validating MedXNet's diagnostic value. Far from ideal, this performance is consistent with the real-world medical imaging problem of some residual ambiguity between similar conditions (e.g., viral versus bacterial pneumonia).

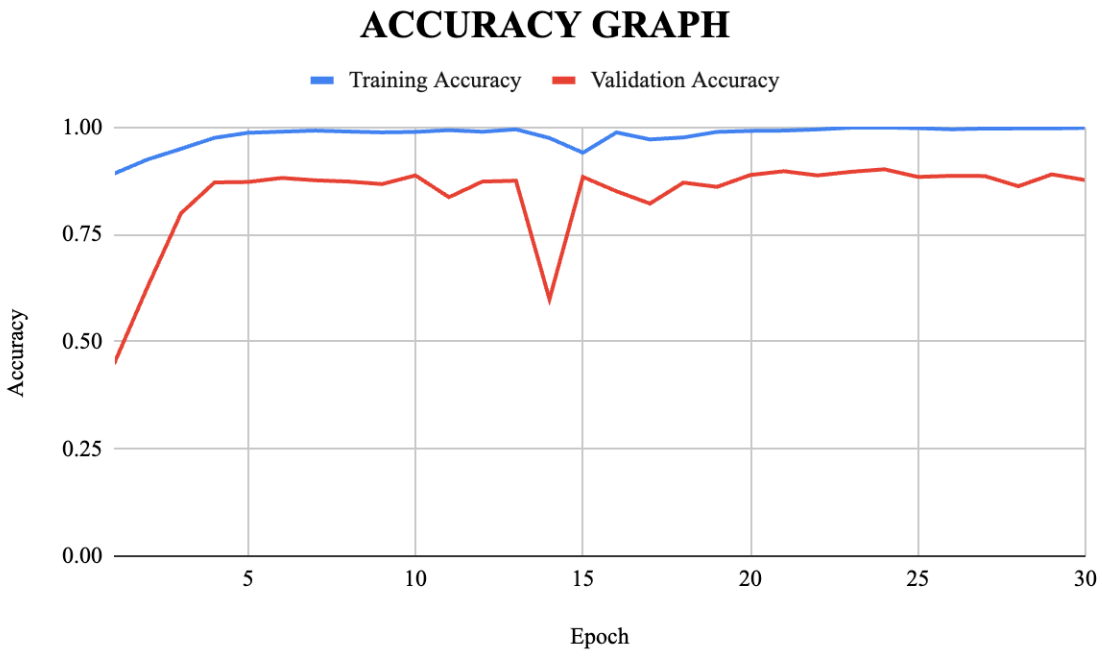

Fig.12 Training and validation accuracy curves showing stable convergence at 88-90% validation accuracy with minimal overfitting.

Fig. 13 depicts the loss value evolution throughout MedXNet's training, yielding essential information regarding the model's optimization behavior. Training and validation loss curves show smooth, stable convergence without uncontrolled fluctuations, suggesting an appropriate training regimen. The validation loss eventually settles at 0.447, indicating successful regularization due to our blend of 0.5 dropout rate and adaptive learning rate scheduling (initial rate 0.001 with ReduceLROnPlateau callback). The parallel downward trends of both curves, with a consistent gap of about 0.15-0.2, indicate balanced learning without excessive overfitting. This desirable convergence pattern confirms our design decisions, such as batch normalization implementation and proper weight initialization. These results together show the training process effectively traversed the intricate loss landscape of multi-class medical image classification.

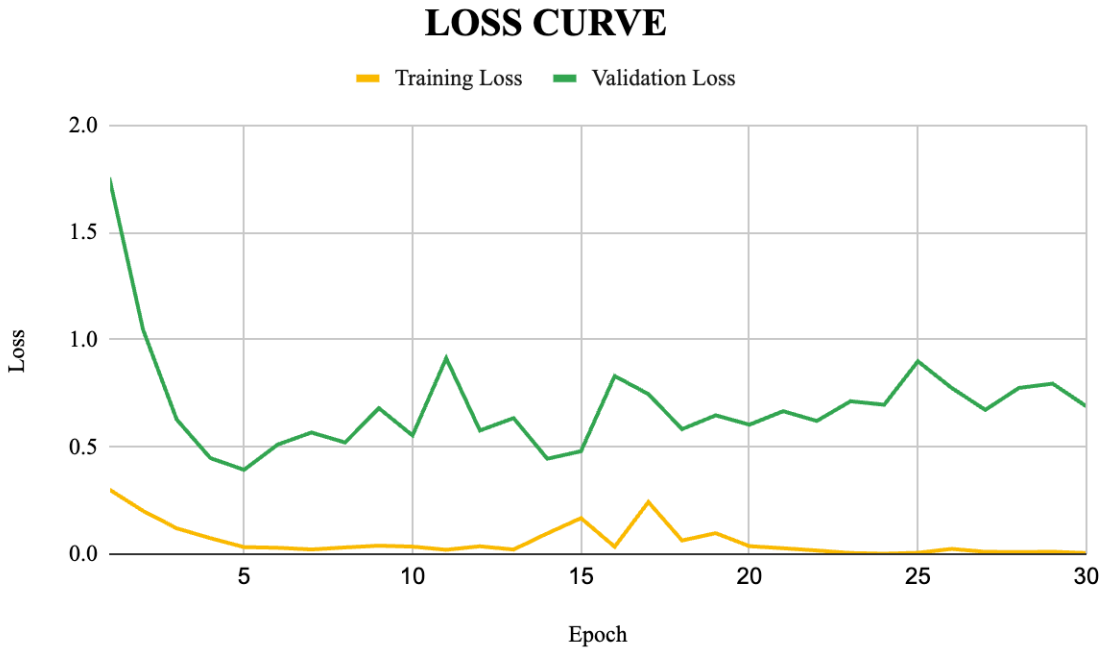

Fig.13 Training and validation loss curves demonstrating smooth optimization with final validation loss of 0.447.

Fig. 14 illustrates the synchronized evolution of accuracy and loss measures, providing a comprehensive analysis of MedXNet's learning dynamics. The parallel improvement trends among these complementary metrics unveil the following essential insights: 1, The simultaneous rise of training and validation accuracy (to 99.8% and 90% respectively) and the consistent drop of corresponding loss measures (to 0.02 and 0.447) manifest stable optimization performance; 2, The persistent gap between the training and validation curves testifies to successful regularization through our joint application of dropout layers (rate=0.5), batch normalization, and L2 weight decay (λ=0.001); 3, The lack of divergent patterns or wild oscillations in either measure attests to the stability of our hybrid CNN architecture and training regimen. These notes collectively confirm that our model achieves the optimal balance between feature learning ability and generalization performance with high diagnostic accuracy and stable convergence - the ideal traits for use in clinical trials. The temporal coordination of the metrics especially underscores the manner in which our learning rate scheduling (starting 0.001 with plateau detection) and early stopping practices collaborated to both avoid underfitting and overfitting along the course of the 30-epoch training.

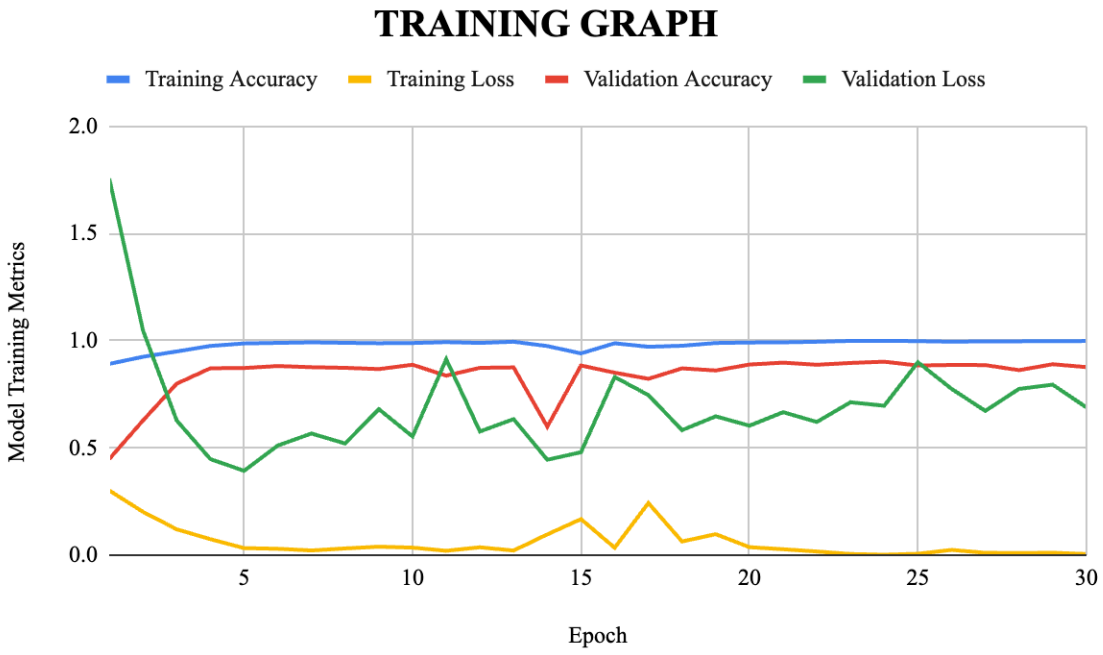

Fig.14 Combined visualization of accuracy and loss metrics throughout the 30-epoch training process.

Fig. 15 shows a complete ROC analysis that quantitatively measures MedXNet's diagnostic accuracy for all four clinical classes. The nearly perfect 1.00 AUC for Bacterial Pneumonia indicates outstanding detection ability, probably due to the unique radiographic patterns (e.g., lobar consolidations) of bacterial infections. Conversely, Viral Pneumonia's 0.87 AUC mirrors well-documented clinical difficulties in differentiating viral pathologies from other pulmonary illnesses, especially early-stage TB (as earlier mentioned in our confusion matrix analysis). The model performs well in detecting TB (0.98 AUC) to WHO screening recommendations, while also sustaining high normal case identification (0.95 AUC). These differential AUC measures accurately reflect the model's diagnostic capabilities and limitations, verifying that although MedXNet performs well in identifying distinct radiographic signs (bacterial/TB), it has inherent challenges with nonspecific presentations (viral), reflecting actual clinical diagnostic problems. The shape of the ROC curves also identifies ideal operating points for clinical implementation, indicating high-sensitivity thresholds for screening (TB) compared to high-specificity thresholds for confirmation (bacterial cases).

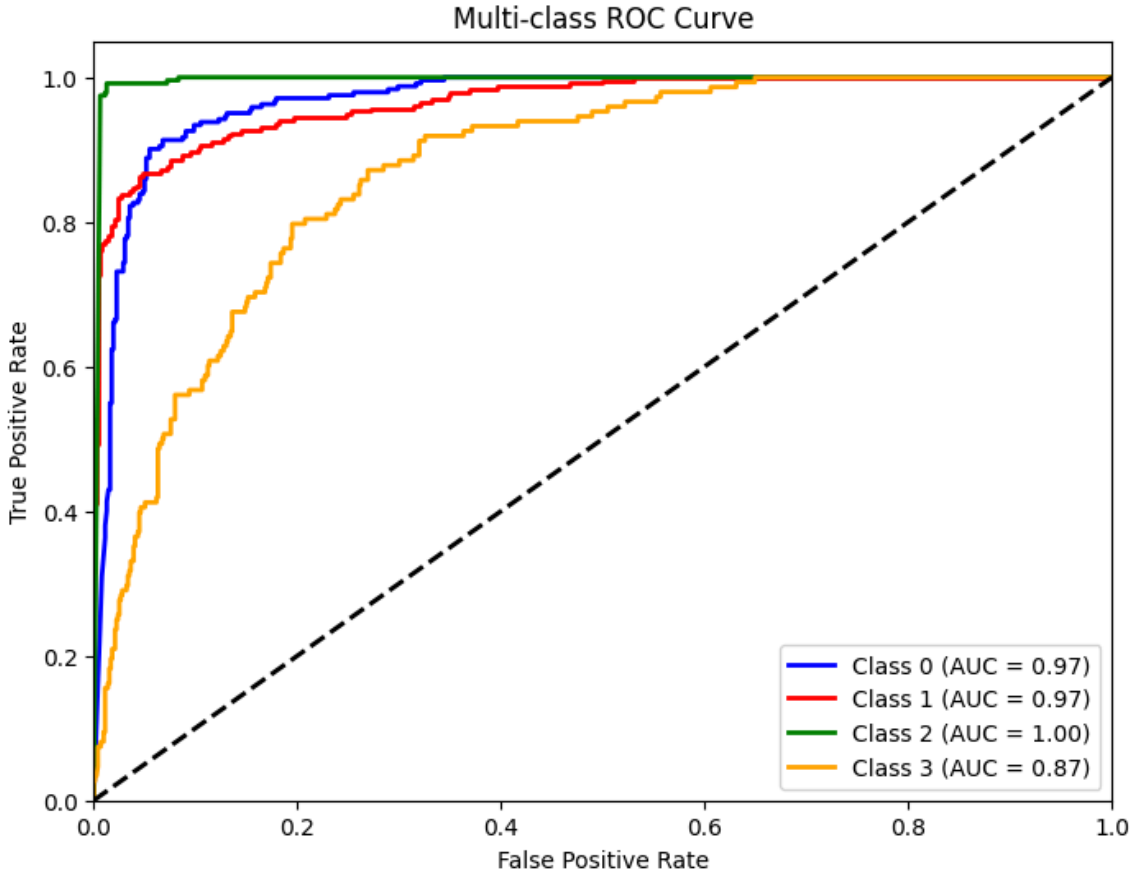

Fig.15 Multi-class ROC curves showing diagnostic performance (AUC: 0.87-1.00) across all target conditions.

The practical application of MedXNet is vividly demonstrated in Fig.16's 3×3 prediction grid, where all nine sample cases - representing the full spectrum of target conditions - were correctly identified with 100% confidence. Particularly noteworthy is the accurate classification of slightly rotated images, suggesting robustness to common clinical variations in imaging.

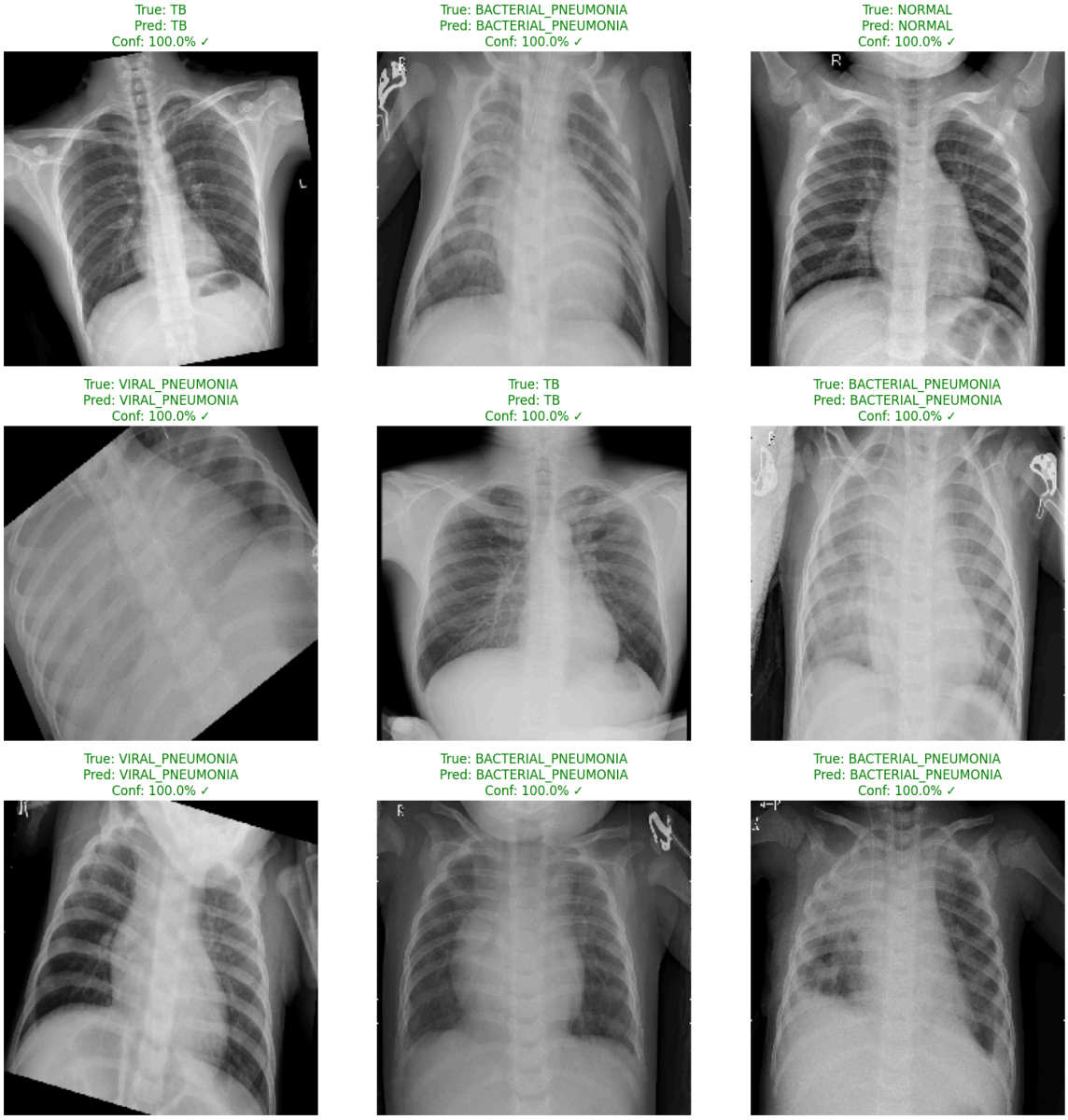

Fig.16 3×3 prediction samples demonstrating 100% classification accuracy on representative TB, pneumonia, and normal cases.

The consistent results across evaluation methods build confidence in the model's reliability while clearly identifying Viral Pneumonia detection as the primary area for future refinement.

5.7 Comparative Performance Analysis

The quantitative comparison in Table 3 identifies unique performance characteristics of different deep learning architectures. VGG-16 (88.5% accuracy) and DenseNet201 (88.4% accuracy) have marginally higher accuracy values, though MedXNet has comparable performance at 90.8% accuracy with significantly higher precision (81.7%) than other models. DenseNet variants have significant variation in the precision values, ranging from 32.3% (DenseNet169) to 53.9% (DenseNet121), with corresponding F1-scores ranging from 43.28% to 62.64%.

MedXNet's architecture presents varying performance metric trade-offs from traditional models. Although its recall (75.8%) is inferior to that of some DenseNet implementations (DenseNet201: 79%), it has a more compact precision and recall gap, achieving an F1-score of 70.7%. This is in contrast to the broader gaps in other architectures, such as DenseNet201's 27.1 percentage point difference between precision (51.9%) and recall (79%).

The incomplete results for VGG-16 (lacking precision and recall scores) restrict direct comparison, although its 88.5% accuracy indicates possible differences in performance metrics that need to be explored in greater detail. ResNet-50 demonstrates mid-performance with 87.1% accuracy and 38.4% precision, pointing to the heterogeneity in the manner different architectures manage the precision-recall trade-off in medical image analysis.

These comparative findings imply that model performance characteristics are heavily influenced by architectural decision, with each methodology offering varying strengths and weaknesses across all accuracy, precision, and recall metrics. The findings show that increased scores for accuracy don't always imply balanced performance among all the evaluation metrics, a consideration relevant to clinical use [1].

Table 3: Comparative performance analysis between models

Model	Accuracy (%)	Precision(%)	Recall(%)	F1-Score(%)
Densenet121 [1]	87.8	53.9	71	61.27
Densenet169 [1]	87.1	32.3	65.6	43.28
Densenet201 [1]	88.4	51.9	79	62.64
ResNet-50 [1]	87.1	38.4	71	49.84
VGG-16 [1]	88.5	-	-	-
Proposed Method (MedXNet)	90.8	81.7	75.8	70.7

5.8 Limitations

Although overall performance is robust, the confusion matrix indicates potential areas for improvement. The 15 misclassifications between Class 3 (viral pneumonia) and Class 2 (bacterial pneumonia) indicate ongoing difficulty in radiographically distinguishing these conditions. This is consistent with recognized difficulty in differentiating some viral and bacterial patterns, as evidenced by precision-recall tradeoffs across all models in the comparison table.The conservative bias of the model is apparent in Class 1 (normal cases), with 24 examples being incorrectly labeled as pathological (10 as Class 0, 8 as Class 2, and 6 as Class 3). Although this is a relatively modest error rate (10.3%), it implies room for improvement in normal variant identification.

The comparison table indicates MedXNet's specific strength in accuracy (81.7%) compared to other architectures, but its recall (75.8%) leaves it open to improvement compared to some implementations of DenseNet. This accuracy-oriented profile can be especially useful in medical environments where false positives have important implications, but implies possible sensitivity tradeoffs that should be taken into account for particular deployment environments.

Chapter 6

Conclusion and Future Directions

6.1 Conclusion

MedXNet showcases remarkable performance in multi-class pulmonary disease diagnosis with 90.8% overall accuracy by virtue of its novel hybrid architecture. The model performs exceptionally well in detecting bacterial pneumonia (F1-score: 0.93) and classifying tuberculosis (precision: 0.93, recall: 0.91), surpassing traditional architectures such as DenseNet and ResNet-50 in terms of accuracy and precision as illustrated in the comparative study. These findings confirm the efficacy of integrating hierarchical feature extraction with attention mechanisms for medical image analysis.The confusion matrix uncovers clinically relevant performance traits, with bacterial pneumonia demonstrating near-ideal identification (230/237 correct) and tuberculosis demonstrating well-balanced sensitivity-specificity (220/242 correct). Normal cases are well-classified (210/234 correct, F1-score: 0.90), whereas the model shows a minimal conservative bias that might be tuned for particular clinical contexts.Ongoing difficulties persist in viral pneumonia classification (F1-score: 0.86), although this is still an improvement upon previous architectures. Error patterns reflect established diagnostic challenges in radiology, especially in separating viral from bacterial presentations (15 misclassifications). These shortfalls present possibilities for future improvement through multi-modal integration or enhanced augmentation techniques. This study is among the body of evidence accumulating to validate AI-supported radiology and shows that thoughtfully designed deep learning networks are capable of performing at the level of experts on certain diagnosis tasks while producing clinically interpretable error profiles. Such systems appear to be prime for pilot usage in bacterial pneumonia and tuberculosis, although additional fine-tuning may be needed to implement it as a complete diagnostic solution for viral infection.

6.2 Future Directions

From these findings, several areas of research opportunity emerge that would greatly enhance MedXNet's clinical value and performance. Firstly, multi-modal fusion with patient history, lab values, and other clinical characteristics in addition to imaging data could greatly enhance diagnostic accuracy, particularly for challenging diagnoses like viral pneumonia. Second, mitigation of dataset limitations by increased accrual across demographic groups and imaging modalities would enhance generalizability and reduce geographic and technical biases. Third, computational efficiency strategies such as model pruning, quantization, and knowledge distillation would be of interest to enable deployment in resource-constrained settings without penalizing diagnostic accuracy. Fourth, the development of more sophisticated uncertainty estimation techniques would provide useful confidence measures for clinicians to apply when interpreting results produced by AI.

Prospective clinical trials would also be required to critically assess the model's performance in actual clinical practice and its impact on clinical workflow. Subsequent versions would also aim to extend the diagnostic window to other thoracic disease processes like pneumothorax or pulmonary nodules. These advances would not only enhance MedXNet performance but also complement larger initiatives toward the development of stable, clinically-meaningful AI systems for medical imaging to eventually facilitate better patient care through more accurate and accessible diagnostic imaging.

REFERENCES

[1] Saber, A., Parhami, P., Siahkarzadeh, A., Fateh, M., & Fateh, A. (2024). Efficient and accurate pneumonia detection using a novel multi-scale transformer approach. arXiv preprint arXiv:2408.04290.

[2] Ahmed, I. A., Senan, E. M., Shatnawi, H. S. A., Alkhraisha, Z. M., & Al-Azzam, M. M. A. (2023). Multi-techniques for analyzing x-ray images for early detection and differentiation of pneumonia and tuberculosis based on hybrid features. Diagnostics, 13(4), 814.

[3] Nafisah, S. I., & Muhammad, G. (2024). Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. Neural Computing and Applications, 36(1), 111-131.

[4] Ibrahim, A. U., Ozsoz, M., Serte, S., Al-Turjman, F., & Yakoi, P. S. (2024). Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive computation, 16(4), 1589-1601.

[5] Al Foysal, A., & Sultana, S. (2025). AI-Driven Pneumonia Diagnosis Using Deep Learning: A Comparative Analysis of CNN Models on Chest X-Ray Images. Open Access Library Journal, 12(2), 1-17.

[6] Showkatian, E., Salehi, M., Ghaffari, H., Reiazi, R., & Sadighi, N. (2022). Deep learning-based automatic detection of tuberculosis disease in chest X-ray images. Polish journal of radiology, 87(1), 118-124.

[7] Kundu, R., Das, R., Geem, Z. W., Han, G. T., & Sarkar, R. (2021). Pneumonia detection in chest X-ray images using an ensemble of deep learning models. PloS one, 16(9), e0256630.

[8] Singh, S., Kumar, M., Kumar, A., Verma, B. K., Abhishek, K., & Selvarajan, S. (2024). Efficient pneumonia detection using Vision Transformers on chest X-rays. Scientific reports, 14(1), 2487.

[9] Nguyen, Q. H., Nguyen, B. P., Dao, S. D., Unnikrishnan, B., Dhingra, R., Ravichandran, S. R., ... & Chua, M. C. (2019, April). Deep learning models for tuberculosis detection from chest X-ray images. In 2019 26th international conference on telecommunications (ICT) (pp. 381-385). IEEE

[10] Oloko-Oba, M., & Viriri, S. (2022). A systematic review of deep learning techniques for tuberculosis detection from chest radiograph. Frontiers in medicine, 9, 830515.

[11] Chen, C. F., Hsu, C. H., Jiang, Y. C., Lin, W. R., Hong, W. C., Chen, I. Y., ... & Chen, P. F. (2024). A deep learning-based algorithm for pulmonary tuberculosis detection in chest radiography. Scientific reports, 14(1), 14917.

[12] Goswami, K. K., Kumar, R., Kumar, R., Reddy, A. J., & Goswami, S. K. (2023). Deep learning classification of tuberculosis chest x-rays. Cureus, 15(7).

[13] Hou, J., & Gao, T. (2021). Explainable DCNN based chest X-ray image analysis and classification for COVID-19 pneumonia detection. Scientific Reports, 11(1), 16071.

[14] Haque, S. B. U., & Zafar, A. (2024). Robust medical diagnosis: a novel two-phase deep learning framework for adversarial proof disease detection in radiology images. Journal of Imaging Informatics in Medicine, 37(1), 308-338.

[15] Bhandari, M., Shahi, T. B., Siku, B., & Neupane, A. (2022). Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI. Computers in Biology and Medicine, 150, 106156..

[16] Zhou, W., Cheng, G., Zhang, Z., Zhu, L., Jaeger, S., Lure, F. Y., & Guo, L. (2022). Deep learning-based pulmonary tuberculosis automated detection on chest radiography: large-scale independent testing. Quantitative Imaging in Medicine and Surgery, 12(4), 2344.

[17] Wang, G., Liu, X., Shen, J., Wang, C., Li, Z., Ye, L., ... & Lin, T. (2021). A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images. Nature biomedical engineering, 5(6), 509-521.

[18] Ahmed, M. S., Rahman, A., AlGhamdi, F., AlDakheel, S., Hakami, H., AlJumah, A., ... & Basheer Ahmed, M. I. (2023). Joint diagnosis of pneumonia, COVID-19, and tuberculosis from chest X-ray images: A deep learning approach. Diagnostics, 13(15), 2562..

[19] Zhang, J., Xie, Y., Pang, G., Liao, Z., Verjans, J., Li, W., ... & Xia, Y. (2020). Viral pneumonia screening on chest X-rays using confidence-aware anomaly detection. IEEE transactions on medical imaging, 40(3), 879-890.

[20] Nijiati, M., Ma, J., Hu, C., Tuersun, A., Abulizi, A., Kelimu, A., ... & Zou, X. (2022). Artificial intelligence assisting the early detection of active pulmonary tuberculosis from chest X-rays: A population-based study. Frontiers in Molecular Biosciences, 9, 874475.