XAI-LENS: An Explainable AI–Based Classification of Lung Disorders from X-ray Images

By
Juiria Humayan
213-15-4394

Md. Najmus Sakib Nahid
213-15-4575

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science in Computer Science and Engineering

Supervised by

Mr. Amir Sohel

Assistant Professor
Department of Computer Science and
Engineering Daffodil International University

Co-Supervised by

Dr. Arif Mahmud

Associate Professor
& Associate Head
Department of Computer Science and
Engineering Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

September 16, 2025

APPROVAL

This Project titled **"XAI-LENS: An Explainable AI–Based Classification of Lung Disorders from X-ray Images"**, submitted by Juiria Humayan, ID No: **213-15-4394** and Md. Najmus Sakib Nahid, ID No: **213-15-4575** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **16 September, 2025.**

BOARD OF EXAMINERS

Dr. S.M Aminul Haque — **Chairman**
Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md Alamgir Kabir — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Mr. Md Assaduzzaman — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Zulfiker Mahmud — **External Examiner**
Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Mr. Amir Sohel, Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Mr. Amir Sohel
Assistant Professor
Department of Computer Science & Engineering
Daffodil International University

Co-Supervised by:

Dr. Arif Mahmud
Associate Professor & Associate Head
Department of Computer Science & Engineering
Daffodil International University

Submitted by:

Juiria

Juiria Humayan
Student ID: 213-15-4394
Department of Computer Science & Engineering
Daffodil International University

Nahid

Md. Najmus Sakib Nahid
Student ID: 213-15-4575
Department of Computer Science & Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project(FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Mr. Amir Sohel**, Senior Lecturer of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Machine Learning, and Deep Learning** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Accurate and efficient classification of lung diseases from chest X-ray (CXR) images is vital for timely diagnosis and treatment, particularly in healthcare environments with limited resources. This study introduces LightXrayNet, a lightweight convolutional neural network designed for multi-class classification of CXR images into nine categories: Normal, Pneumonia, Higher Density, Lower Density, Obstructive Pulmonary Diseases, Degenerative Infectious Diseases, Encapsulated Lesions, Mediastinal Changes, and Chest Changes. The dataset was sourced from the publicly available *X-ray Lung Diseases Images (9 Classes)* repository on Kaggle and subjected to a comprehensive preprocessing pipeline, including adaptive CLAHE-based contrast enhancement, resizing, normalization, light augmentation (horizontal flip, ±5° rotation), and data splitting. LightXrayNet's performance was benchmarked against three pretrained CNNs—DenseNet201, ResNet50V2, and InceptionV3—using metrics such as accuracy, precision, recall, F1-score, confusion matrices, and training efficiency. Experimental results show that LightXrayNet achieved a test accuracy of 99.22% and near-perfect values across all classes, while requiring substantially less training time compared to deeper pretrained architectures. These findings demonstrate the potential of LightXrayNet as a practical and deployable solution for automated lung disease detection, with strong applicability in resource-constrained healthcare settings.

Table of Contents

Approval i

Declaration ii

Acknowledgements iii

Abstract iv

List of Figures vii

List of Tables viii

1 Introduction 1

1.1 Introduction……………………………………………………………..1
1.2 Motivation………………………………………………………………2
1.3 Objectives……………………………………………………………….2
1.4 Methodology…………………………………………………………….3
1.5 Project Outcome………………………………………………………...3
1.6 Organization of the Report……………………………………………...4

2 Background 5

2.1 Introduction…………………………………………………………...…...5
2.2 Literature Review………………………………………………………….5
2.3 Gap Analysis……………………………………………………………..10
2.4 Summary…………………………………………………………………11

3 Research Methodology 12

3.1 Methodology/Requirement Analysis & Design Specification………………...12
3.1.1 Overview…………………………………………………………….12
3.1.2 Proposed Methodology/ System Design……………………………….12
3.2 Detailed Methodology and Design……………………………………………13
3.3 Project Plan……………………………………………………………………18
3.4 Task Allocation……………………………………………………………...19
3.5 Summary……………………………………………………………………..20

4 Implementation and Results 21

4.1 Environment Setup…………………………………………………………..21
4.2 Testing and Evaluation/Performance/ Comparative Analysis…………….........22
4.3 Results and Discussion………………………………………………………27
4.4 Summary……………………………………………………………………..30

5 Engineering Standards and Design Challenges **32**
5.1 Compliance with the Standards……………………………………………………32
5.1.1 Software Standards……………………………………………………...32
5.1.2 Hardware Standards……………………………………………………..32
5.1.3 Communication Standards………………………………………………32
5.2 Impact on Society, Environment and Sustainability……………………………32
5.2.1 Impact on Life……………………………………………………………...32
5.2.2 Impact on Society & Environment………………………………………33
5.2.3 Ethical Aspects…………………………………………………………..33
5.2.4 Sustainability Plan……………………………………………………….33
5.3 Project Management and Financial Analysis……………………………………33
5.4 Complex Engineering Problem…………………………………………………33
5.4.1 Complex Problem Solving………………………………………………33
5.4.2 Engineering Activities…………………………………………………...34
5.5 Summary……………………………………………………………………...35

6 Conclusion **36**
6.1 Summary…………………………………………………………………..36
6.2 Limitation……………………………………………………………….…36
6.3 Future Work…………………………………………………………….….36

References **37**

List of Figures

3.1 Proposed Methodology. 12

3.2.1 Working Flowchart. 13

3.2.2 Sample x-ray images…………………………..…... 14

3.2.3 CLAHE enhancement results for normal lung disease class. 15

4.2.1 Training and Validation Accuracy/Loss Analysis – InceptionV3. 22

4.2.2 Confusion Matrix for InceptionV3.. 23

4.2.3 Training and Validation Accuracy/Loss Analysis – ResNet50V2..………….......... 23

4.2.4 Confusion Matrix for ResNet50V2.. 24

4.2.5 Training and Validation Accuracy/Loss Analysis – DenseNet201. ……….......... 24

4.2.6 Confusion Matrix for DenseNet201. 25

4.2.7 Training and Validation Accuracy/Loss Analysis – LightXrayNet. ………......... 25

4.2.8 Confusion Matrix for LightXrayNet. 26

4.3 Implementation of Grad-CAM, LIME,

Score-CAM among the nine classes…... 28

List of Tables

2.1 Summary of Literature Reviewed. 07

2.3 Gap Analysis Table. 10

3.2.3 Training Arguments. …...16

3.4 Task Allocation. 19

4.1 System Environment Configuration. 21

4.2 Time Complexity Analysis. 26

4.3 Performance comparison…………………………………………………………..27

4.3.1 Cross Validation. 27

4.3.2 Statistical Analysis. ..../..28

5.1 Mapping with Complex Engineering Problem………………….................................33

5.2 Mapping with Knowledge Profile. 34

5.3 Mapping with Complex Engineering Activities…………………...............................34

Chapter 1
Introduction

1.1 Introduction

Lung diseases cause many health problems around the world and lead to a lot of sickness and deaths. From the data of World Health Organization, tuberculosis kills more than 1.3 million people every year [1]. Other lung infections like pneumonia and COVID-19 also put a heavy strain on hospitals and healthcare systems globally. These problems are even worse in countries like Bangladesh, where tuberculosis is common, and outbreaks of viral pneumonia and COVID-19 often happen, making it hard for their healthcare system to cope. Finding lung diseases early is very important to reduce their spread, start treatment quickly, and help patients get better [2].

Chest X-rays are one of the easiest and cheapest ways to check for lung diseases. But reading these X-rays can be hard because some lung problems look very similar, and doctors may disagree on what they see. This has led to more interest in using deep learning, a type of artificial intelligence, to analyze chest X-rays automatically. Existing pre-trained models like ResNet50V2, DenseNet201 also InceptionV3 performs well for this task but due to the model complexity they required very large datasets for model training and accurate classification, also the dataset must be balanced else model sometimes overfits to a single class. Besides, high computational power is required to train these models. On the other hand, we have models like EfficientNet that require less computational power but it also underperforms this task.

Keeping all the above issues in mind we developed a small, lightweight custom convolutional neural network model named as LightXrayNet that is specially designed for this task. It uses special images preprocessing technique before feeding the data to the neural networks. Our proposed model was compared with others pretrained models mentioned before to check its performance top notch while maintaining the efficiency. Also to make the model more trustworthy, explainable AI techniques like Grad-CAM, LIME and Score-CAM, were implemented to show which parts of the X-rays the model used to make its decisions. Because of its speed, accuracy, and more reliability, LightXrayNet could be very useful for diagnosing lung diseases automatically in countries with limited healthcare resources.

1.2 Motivation

Our goal is to create a lightweight CNN model that can train fast and also give superior results in classifying lung diseases. Many deep learning models perform well but need powerful computers, which makes them hard to use in places with limited resources. Also, many existing models are difficult to interpret, which lowers doctors' trust in them. To solve this, LightXrayNet, a small convolutional neural network, is introduced. It provides high accuracy while requiring little computing power. It also uses explainable AI methods to make its decisions clearer to testers and build more confidence in its predictions.

1.3 Objectives

All the objectives of this project is listed below;

1. Create LightXrayNet, a small and efficient convolutional neural network (CNN) that can classify multiple lung diseases using chest X-ray images.

2. Improve quality of the chest X-ray images before analysis by using an image preprocessing method called (CLAHE). This technique makes the images clearer and helps the CNN detect important features more easily.

3. Benchmark the performance of LightXrayNet against established pretrained convolutional neural networks, including ResNet50V2, DenseNet201, and InceptionV3, under standardized experimental conditions.

4. Add Explainable Artificial Intelligence (XAI) techniques, specifically Gradient-weighted Class Activation Mapping (Grad-CAM), Local Interpretable Model-agnostic Explanations (LIME), and Score-CAM, to improve clinical interpretability of model predictions.

5. Evaluate the efficiency, robustness, and deployability of LightXrayNet in practical healthcare environments.

1.4 Methodology

In the very first step, we collected the dataset from a publicly available resource; kaggle which contains nine different categories of chest x-ray images. After gathering all the images, we started the image preprocessing step where each image was processed through an instructed pipeline. The image preprocessing pipeline includes several steps likes; Contrast Limited Adaptive Histogram Equalization (CLAHE), then resizing to 224 by 224 pixel, after this the pixel values were scaled between 0 to 1 range, also called as normalization step. In the mid step a light augmentation technique was performed to ensure all the class has the same number of images in it. After completing all these image preprocessing we splitted the whole dataset into three parts, namely train, test and validation set.

After completing the first step, we started designing our custom CNN model using the core layers like convolutional, pooling and batch normalization. While designing the model layers we tested several combinations of kernel size, number or filters, strides, paddings and so on. We altered the model layers like adding a batch normalization layer or dropout layer in between the pooling and convolutional layer. After designing the core architecture we trained the model each time and noted the results. We also conduct ablation study like changing the classification head, batch size, activation function, optimizer and learning rate to get the best performance from our proposed model while training it in the shortest time possible

In the third step, the pretrained models were also trained and their performance were noted to compare the performance against our custom model. The evaluation metrics that were used to compare the performance are (accuracy, precision, recall and F1-score). Also the confusion matrix was analyzed for each model. Then we performed stratified cross validation of 5 folds in both our custom model and the pretrained ones, to find the significance of our model.

Lastly, we implemented some Explainable AI techniques (Grad-CAM, LIME and Score cam) to highlight the lung regions where our model is focusing to make the predictions. This technique increases the trustworthiness of our model predictions.

1.5 Project Outcome

We received an accuracy of 99.22% recorded as top performer compared to other pretrained models, after training our model with the optimized hyperparameters. Besides, the integration of Explainable artificial intelligence (AI) techniques enhances the transparency of LightXrayNet's predictions. By identifying specific regions of X-ray images that influence classification outcomes, the model enables clinicians to assess and validate its outputs. The combined attributes of accuracy, computational efficiency, and interpretability support the model LightXrayNet as a reliable tool for automated lung disease detection in clinical practice.

1.6 Organization of the Report

The rest of this report is organized as follows:

Chapter 1 introduces the project and sets the context. It outlines the global and local (Bangladesh) burden of lung diseases, explains why automated analysis of chest X-rays is needed, and states the motivation for building a lightweight yet explainable model. It then lists the objectives, gives a brief overview of the methodology (dataset, CLAHE-based preprocessing, the LightXrayNet architecture, and baseline comparisons), summarizes the expected outcomes, and finally describes how the rest of the report is organized.

Chapter 2 provides the background study. It reviews prior research on CXR-based disease classification, highlights similar applications of deep learning in medical imaging, and distills the key limitations of existing work—such as heavy computation and limited interpretability—into a clear gap analysis that motivates LightXrayNet.

Chapter 3 details the research methodology and system design. It explains how the dataset was prepared (CLAHE, resizing, grayscale conversion, normalization, stratified splits), describes the architecture of LightXrayNet and the rationale behind its lightweight design, and lays out the benchmarking setup against DenseNet201, ResNet50V2, and InceptionV3. It also summarizes functional/non-functional requirements, the workflow/DFD, UI considerations, and the project plan and task allocation.

Chapter 4 presents implementation and results. It documents the experimental environment, evaluation metrics, and comparative findings across all models, including confusion matrices, ROC-AUC curves, learning curves, and a time-complexity analysis. It also reports the effect of design choices via ablation studies and illustrates explainability outcomes with Grad-CAM, Score-CAM, and LIME.

Chapter 5 discusses engineering standards and design challenges. It maps the project to relevant software/hardware/communication standards, considers ethical and sustainability aspects, and reflects on societal impact. It further aligns the work with complex engineering problem categories and activities, as required by the FYDP guidelines.

Chapter 6 concludes the study by summarizing the main findings, acknowledging limitations (e.g., single-dataset dependency, lack of severity grading), and outlining future directions such as cross-dataset validation, multi-modal integration, mobile/edge deployment, and prospective clinical studies.

Each chapter is designed to provide a clear progression from problem identification to solution development, ensuring that the research objectives are met and the project's contributions are thoroughly discussed.

Chapter 2

Background

2.1 Introduction

This chapter is all about the theoretical and practical context for this research problem. Detection of lung diseases using chest X-ray (CXR) images has received considerable attention because X-ray imaging is more affordable and accessible than other medical imaging modalities. Over the past decade, researchers have explored computational methods to improve diagnostic accuracy, evolving from traditional machine learning algorithms to advanced deep learning frameworks.

2.2 Literature Review

Present times, chest X-ray (CXR) based disease identification has gained more attention because it is affordable, accessible, and valuable for diagnosis. As deep learning (DL) has advanced, researchers have created automated systems that make lung disease classification faster and more accurate. Early methods used traditional ML classifiers, but convolutional neural networks (CNNs) and transfer learning with pretrained models like ResNet, DenseNet, and MobileNet have shown better results [6,7,8]. More recent work has focused on customized lightweight CNNs [9,11], hybrid deep learning models [15], and explainable AI frameworks [13] to tackle issues such as computational cost, class imbalance, and making results easier to interpret in clinical settings. The next section reviews related studies that show how these methods have developed and why they matter for our approach.

Initial efforts in automated chest X-ray (CXR) analysis utilized traditional machine learning algorithms and shallow neural networks. These investigations demonstrated that automated classification systems can provide substantial support to radiologists. However, they also identified significant challenges in managing heterogeneous data sources. For example, Mohan et al. (2024) developed convolutional neural network (CNN), image retrieval (IR), and VGG16 models using two public datasets, achieving classification accuracies of 97%, 96%, and 93%, respectively. Although these models enhanced manual classification, they exhibited limited generalizability to different image acquisition conditions.

With the rise of deep learning, many works shifted toward large pretrained CNNs. [6] Alshmrani et al. (2023) proposed a multi-class deep learning system to classify TB, lung opacity, lung cancer, pneumonia, normal, and COVID-19. Although the model improved accuracy, it suffered from high computational costs. In another study, [7] Hao et al. (2021) applied EfficientNet-B7, achieving 85.32% accuracy on the NIH dataset and 96.1% on a hospital dataset, but highlighted that class imbalance weakened robustness. Similarly, [8] Metwally et al. (2022) experimented with EfficientNetB0, Xception, and NasNetLarge on 7,135 CXR images. While NasNetLarge achieved 91.3% accuracy, the study reported execution-time inefficiencies, suggesting that deep transfer learning often comes at the cost of practicality.

To overcome the heavy resource demands of large CNNs, several researchers designed lightweight or custom architectures. [9] Shamrat et al. (2023) proposed MobileLungNetV2, a modified MobileNetV2, which achieved 96.97% accuracy. However, the training process required considerable hardware, limiting deployment in low-resource settings. [10] Kim et al. (2022) fine-tuned EfficientNet v2-M to detect pneumonia and pneumothorax, scoring 91.65% accuracy on the NIH dataset but struggling with interpretability. Likewise, [11] Bhandari et al. (2023) developed a lightweight CNN that achieved 94.31% accuracy but admitted that its “black-box” nature reduced trust in clinical environments.

As model interpretability became critical for clinical acceptance, several studies incorporated explainable AI (XAI). [12] Al-Sheikh et al. (2023) combined custom CNNs with ensemble techniques, obtaining 98.6% accuracy on X-rays and 98.8% on CT scans. Still, they noted inconsistencies in image quality across datasets. Building on this, [13] Ifty et al. (2024) employed transfer learning with Xception alongside XAI methods, reporting 96.21% accuracy while improving clinical trust. A more recent contribution came from [14] Aldamani et al. (2024), who introduced LungVision, an edge-compatible system trained on a nine-class dataset. Their EfficientNetB2 model achieved an F1-score of 98.58% and, through quantization-aware training, was successfully deployed on mobile devices, highlighting progress toward portable and real-time diagnostic support.

Other researchers have explored hybrid or multi-label frameworks to improve classification efficiency and flexibility. [15] Farhan and Yang (2023) proposed a Hybrid Deep Learning Algorithm (HDLA) that combined ResNet50 feature extraction with classifiers such as DNN, SVM, and AdaBoost. Their approach reached 98.99% accuracy but lacked the ability to assess disease severity or deliver real-time predictions. Similarly, [16] Irtaza et al. (2024) addressed multi-label classification by testing multiple CNNs—MobileNetV1, EfficientNet, DenseNet121, ResNet50, InceptionV3, and Xception—on the NIH Chest X-ray14 dataset. MobileNetV1 achieved the best results, with 93.4% accuracy, but synthetic images generated through GAN augmentation proved only marginally effective.

While significant advancements have been made in lung disease classification, several limitations remain unaddressed. High computational overhead [6,9], reliance on balanced datasets [7,16], and the inefficacy of synthetic data generation [16] hinder the deployment and scalability of existing models. Additionally, sensitivity to varying image quality [13,12] reduces clinical trust and real-world applicability. Limited integration of traditional ML techniques and interpretability methods [14] restricts broader diagnostic insight despite achieving mobile deployment. To address these gaps, this study aims to develop a framework where a hybrid approach combining both traditional ML and DL methods could provide a balanced solution for lung disease classification from CXR images. Researchers are now focusing on developing frameworks that leverage both methodologies' strengths while addressing their respective limitations regarding robustness against varying image quality.

Table 2.1: Summary of Literature Reviewed.

Author (s)	Year	Title	Methodology	Key Findings
Mohan et al.[5]	2024	A multiclass deep learning algorithm for healthy lung, Covid-19 and pneumonia disease detection from chest X-ray images.	CNN, IR, VGG16 on two open-source CXR datasets.	Accuracy up to 97%; limited generalization due to dataset diversity.
Alshmrani et al.[6]	2023	A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images.	Deep CNN on TB, lung opacity, cancer, pneumonia, normal, COVID-19.	Improved classification, but high computational cost.
Hao et al.[7]	2021	Multi-Class Classification of Lung Diseases Using CNN Models.	EfficientNet B7 with preprocessing and fine-tuning.	Accuracy: 85.32% (NIH), 96.1% (SCH); limited robustness.
Metwally [8]	2022	Automatic Detection and Multi-Class Classification of COVID-19, Pneumonia, and Tuberculosis Diseases in Chest X-ray Images Using Deep Learning Techniques.	NasNetLarge, EfficientNetB0, Xception on 7,135 X-rays.	Accuracy up to 91.3%; slow execution and high cost.

Shamrat et al.[9]	2023	High-precision multiclass classification of lung disease through customized MobileNetV2 from chest X-ray images.	Customized MobileNetV2 (“MobileLungNet V2”) with augmentation.	Accuracy: 96.97%; heavy training load and poor scalability
Kim et al.[10]	2022	Deep learning in multi-class lung diseases’ classification on chest X-ray images.	EfficientNet v2-M on NIH and SCH datasets.	Accuracy: 91.65% (NIH); interpretability issues.
Bhandari et al.[11]	2023	Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI.	Custom lightweight CNN.	Accuracy: 94.31%; interpretability gap.
Al-Sheikh et al.[12]	2023	Multi-class deep learning architecture for classifying lung diseases from chest X-Ray and CT images.	Custom CNN + Ensemble learning.	Custom CNN + Ensemble learning.
Ifty et al.[13]	2024	Explainable AI framework utilizing advanced deep learning for lung disease classification.	Xception with transfer learning + XAI.	Accuracy: 96.21%; improved interpretability.
Aldamani et al.[14]	2024	LungVision: X-ray Imagery Classification for On-Edge	EfficientNetB2 with quantization for mobile	F1-score: 98.58%; lightweight and mobile deployable.

		Diagnosis Applications.	deployment.	
Farhan & Yang [15]	2023	Automatic lung disease classification from the chest X-ray images using hybrid deep learning algorithm.	Hybrid DL (ResNet50 feature extraction + SVM, AdaBoost, DNN).	Accuracy: 98.99%; lacked severity and real-time capacity.
Irtaza et al.[16]	2024	Multi-Label Classification of Lung Diseases Using Deep Learning.	Multiple pretrained models (MobileNetV1, DenseNet, ResNet50, InceptionV3, Xception).	Accuracy: 93.4%; struggled with class imbalance.

2.3 Gap Analysis

Many researchers have explored deep learning techniques for lung disease detection, but there are still some limitations that they haven't overcome yet. Problems like imbalance of the dataset, model overfitting to data, and more computational cost still persist in many research as given below in the table. Moreover the development of custom cnn's has some issues like, missing Explainable AI implementation, no time complexity analysis or significance testing.

Table 2.3 : Gap Analysis Table

Paper Name	Limitation Addressed	How Our Work Minimizes the Limitation
Mohan et al. (2024)	Overfitting, data disparity.	LightXrayNet uses robust preprocessing, SpatialDropout, and light augmentations (horizontal flip, small rotations of up to 5°) to reduce overfitting, improve generalization, and handle data disparity.
Alshmrani et al. (2023)	High computational cost.	Our lightweight CNN design minimizes resource requirements and reduces computational demands.
Hao et al. (2021)	Dependence on balanced datasets.	Stratified data splitting and adaptive preprocessing enhance robustness on imbalanced datasets.
Metwally (2022)	Computationally expensive models.	LightXrayNet achieves comparable accuracy with far fewer parameters and faster training.
Shamrat et al. (2023)	Resource-intensive training.	Compact architecture reduces resource usage while maintaining high accuracy.
Kim et al. (2022)	Low interpretability	Integration of explainable AI (Grad-CAM, LIME, Score-CAM) provides transparency in predictions.

Bhandari et al. (2023)	Limited explainability	LightXrayNet incorporates multiple XAI methods to enhance clinical trust.
Al-Sheikh et al. (2023)	Reduced clinical trust due to dataset inconsistencies.	Robust preprocessing ensures consistent performance across varying image quality.
Ifty et al. (2024)	Sensitivity to image quality variations.	Our approach strengthens robustness against inconsistent image quality.
Farhan & Yang (2023)	No severity estimation, limited multi-class performance.	LightXrayNet focuses on accurate multi-class classification as a foundation for future severity grading.
Irtaza et al. (2024)	Ineffectiveness of GAN-generated synthetic images.	Our study avoids synthetic data, relying instead on carefully balanced real datasets.
Aldamani et al. (2024)	Limited interpretability in edge-deployable models.	LightXrayNet combines efficiency with XAI techniques to maintain transparency while remaining lightweight.

2.4 Summary

This chapter is all about the related works that have been done in this specific field. This includes the traditional machine learning approach then the more advanced deep learning approach, using the transfer learning approach and so on. The literature review shows some promising results but the challenges and limitations with the work need to be fixed too. The gap analysis showed how these limitations restrict real-world adoption and outlined how LightXrayNet is designed to address them by combining efficiency, robustness, and transparency, setting the stage for the methodology described in the next chapter.

Chapter 3

Research Methodology

3.1 Methodology

3.1.1 Overview

This section describes the actual working of this project. It begins from the raw dataset that was collected from the kaggle repository, then the preprocessing stage where all the images were processed through an image preprocessing pipeline. Then the custom and pretrained models were trained with these images. Then in the evaluation part models were compared with one another using the classification report. Moreover advanced evaluation methods like stratified 5-fold cross validation, significance testing and Explainable AI were also implemented.

3.1.2 Proposed Methodology

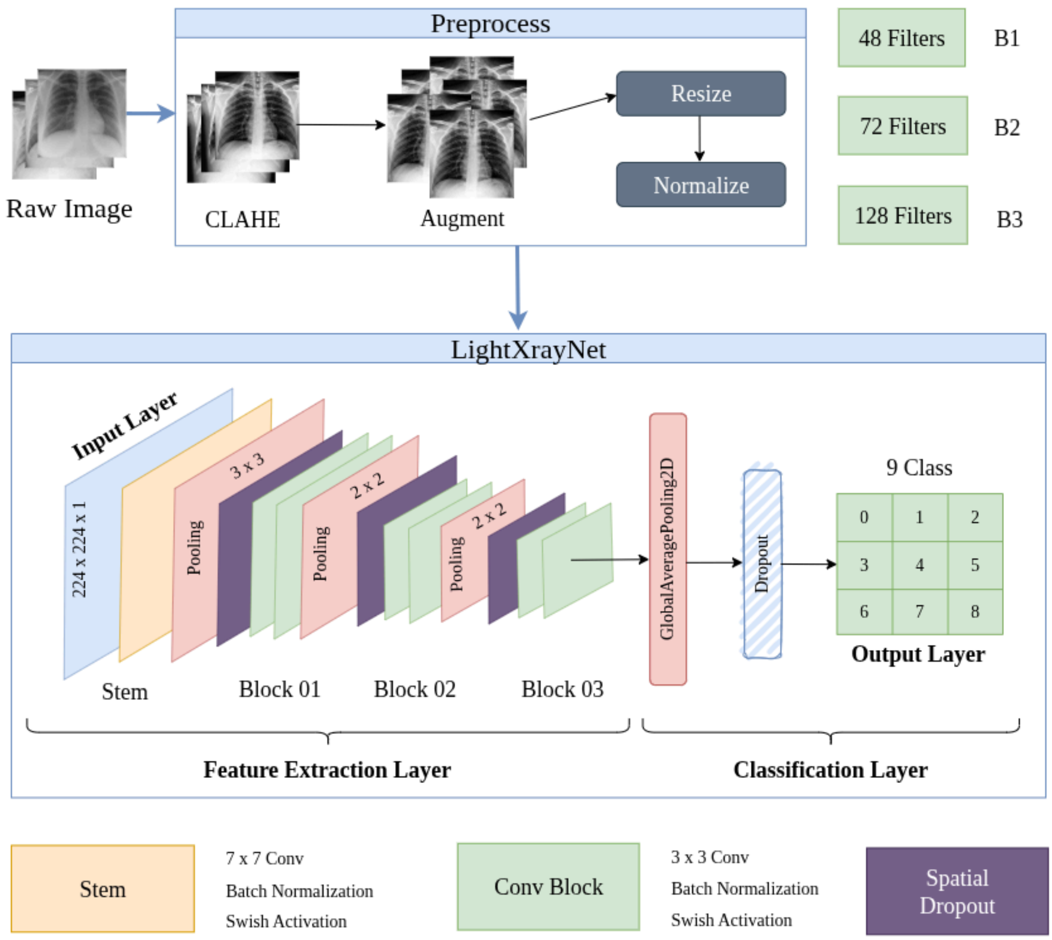


Figure 3.1: Proposed Methodology

3.2 Detailed Methodology and Design

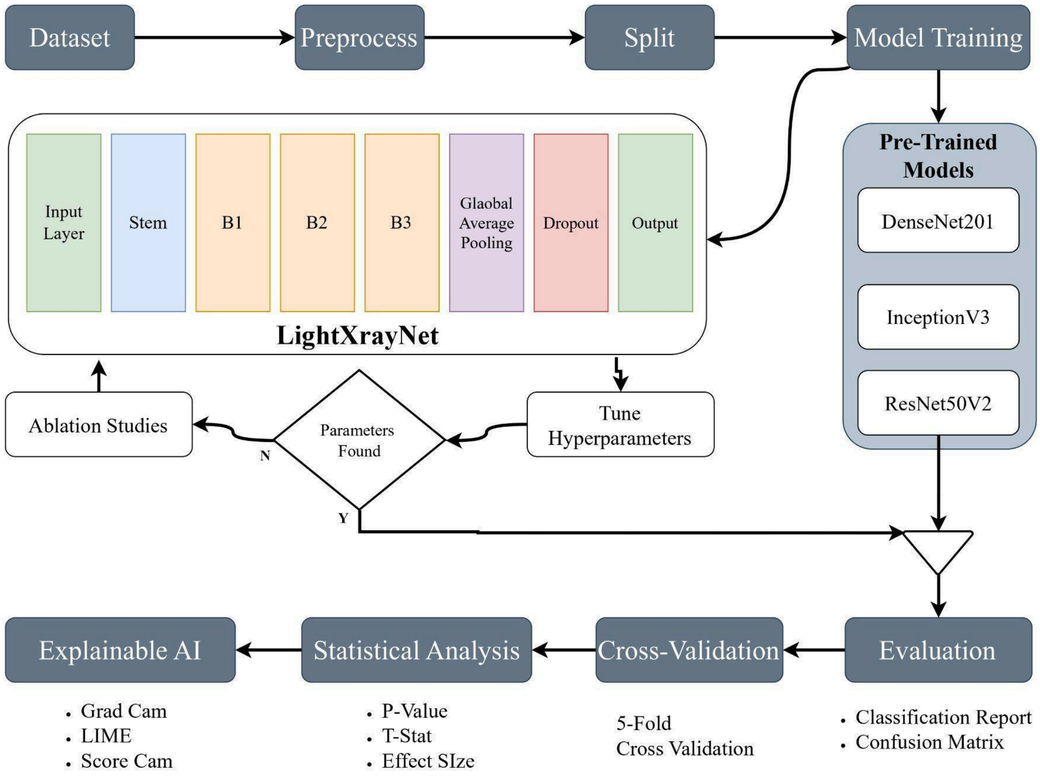


Figure 3.2.1 : Working Flowchart

1. Dataset Description

We used the publicly available X-ray Lung Diseases Images (9 Classes) dataset [3]. The largest group is Normal, with 1,340 images, followed by Pneumonia with 1,060 cases. Other groups are Higher Density (678 images), Lower Density (629 images), and Obstructive Pulmonary Diseases (644 images). The dataset also includes 594 images of Degenerative Infectious Diseases and 658 of Encapsulated Lesions. There are 596 cases of Mediastinal Changes and 544 images in the Chest Changes category. This range of cases gives the model enough examples of different thoracic conditions to help it recognize subtle differences among various lung disorders.

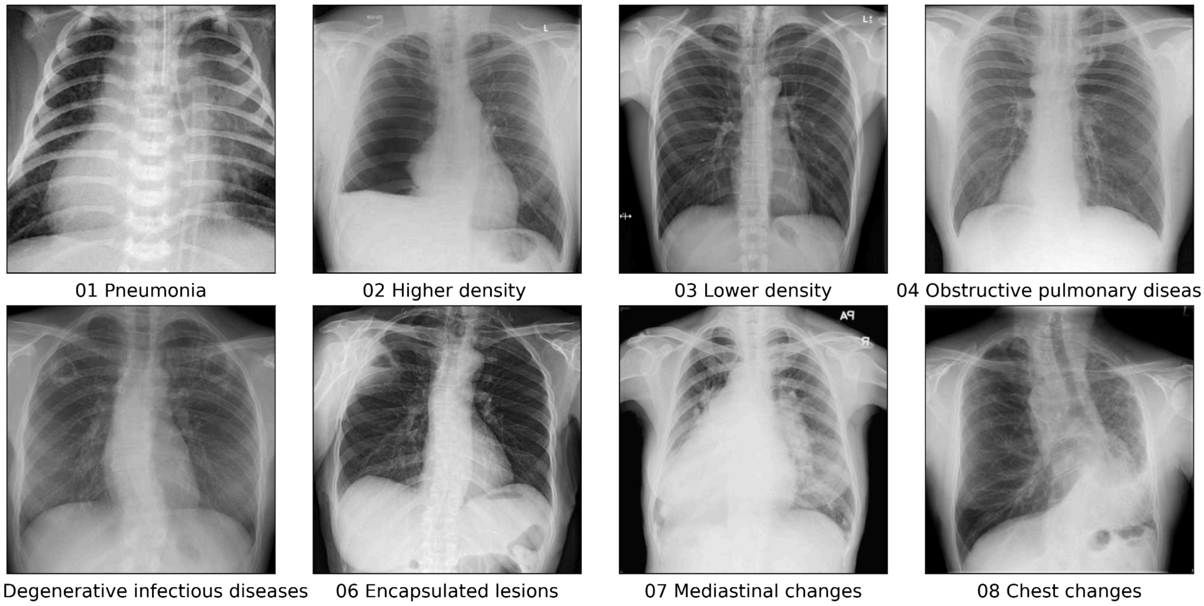


Figure 3.2.2 : Sample x-ray images

2. Data Preprocessing

- **Contrast Enhancement with CLAHE**

Contrast Limited Adaptive Histogram Equalization (CLAHE) enhances the visibility of local details in medical images by dividing the image into small tiles and applying histogram equalization to each. The “clip limit” parameter prevents over-amplification of noise, while the “tile grid size” determines the number of regions processed. We employed CLAHE to improve local contrast, particularly in regions containing subtle pathological cues. Unlike traditional CLAHE implementations that rely on fixed parameter values, our approach applied a grid search strategy to systematically explore a predefined range of clip limit and tile grid size combinations. The configuration that produced the highest local contrast enhancement was selected for each image, ensuring improved feature visibility without introducing excessive noise.

$$s = (L - 1) \times CDF(r) \tag{1}$$

where L is the number of gray levels, and CDF(r) is the clipped and redistributed cumulative distribution function.

In **Figure 3.2.2**, for each class, the first column presents the original X-ray, the second column displays the CLAHE-enhanced X-ray using the optimally selected parameters, and the third and fourth columns show the corresponding intensity histograms before and after enhancement. The grid search–based CLAHE approach resulted in clear improvements in local contrast, particularly in regions containing subtle pathological features.

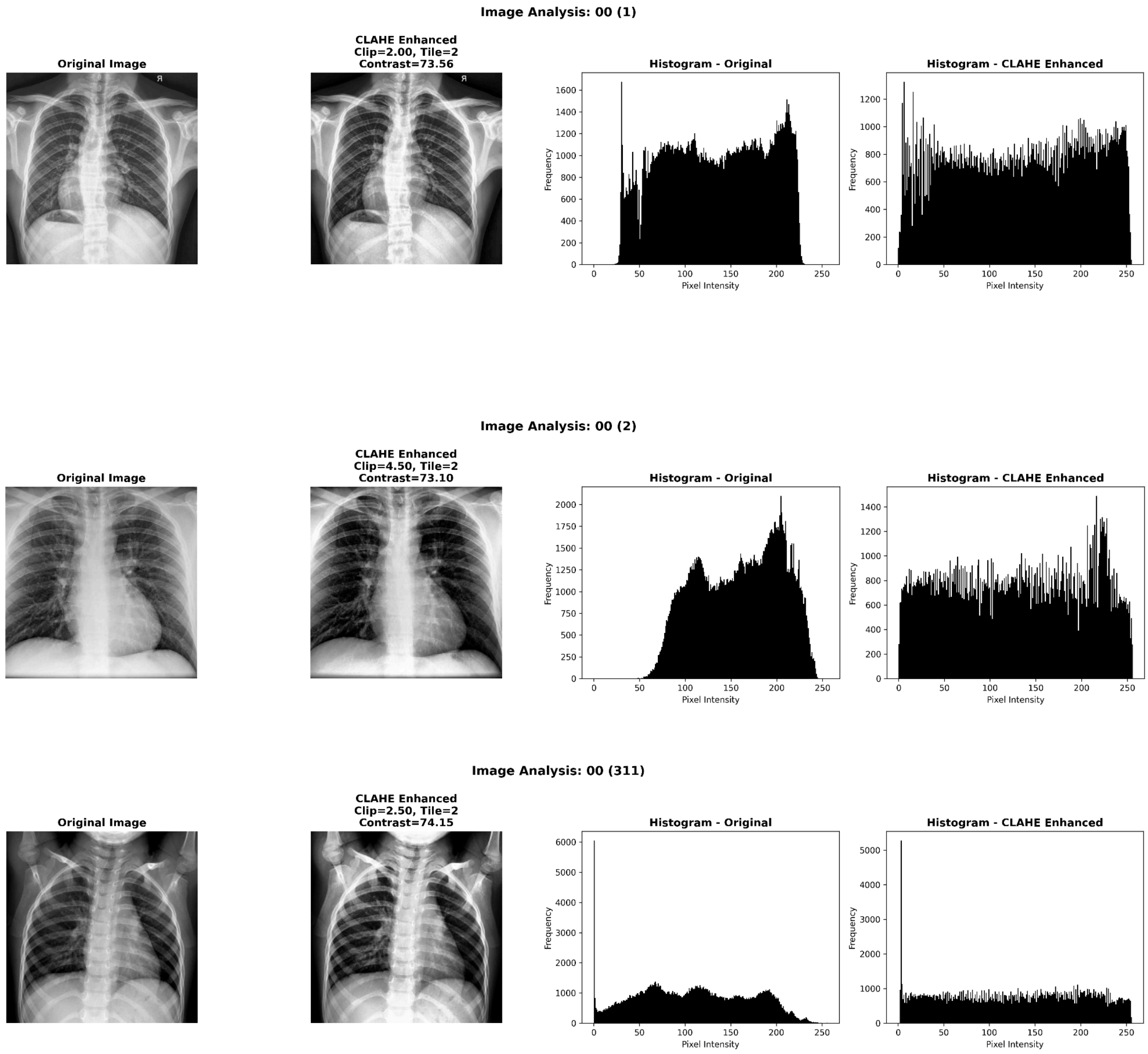


Figure 3.2.3 : CLAHE enhancement results for normal lung disease class

- **Data Augmentation**

To address class imbalance and improve model generalization, we applied light augmentation techniques. Specifically, horizontal flipping and small rotations (±5°) were used.

- **Resizing**

Deep learning models require inputs of fixed dimensions. Since the original X-ray images vary in size, all images were resized to 224 × 224 pixels to match the input size expected by the pretrained networks. Resizing was performed using bilinear interpolation, which calculates the intensity of each new pixel I (x,y) as a weighted average of the four nearest pixels in the original image:

$$I(x, y) = \sum_{i=0}^{1} \sum_{j=0}^{1} wij \ . \ I(xi, yj) \qquad (2)$$

where w ij are weights based on the relative distances from the new pixel to the original pixels. This method preserves image smoothness and reduces distortion.

- **Normalization**

Neural networks train more effectively when input features are scaled to a consistent range. Pixel intensities originally range from 0 to 255; these were normalized to the range [0, dividing each pixel value p by 255]:

$$P' = \frac{P}{255} \tag{3}$$

Normalization helps stabilize gradients during training and accelerates convergence.

3. Model Selection & Training

Before training, the dataset was split into 80% training, 10% validation, and 10% testing using stratified sampling to preserve class balance across all nine categories.

First we trained our proposed model LightXrayNet with this dataset. While training we conducted an ablation study to find the optimal parameters for our model that gives both, highest accuracy and fastest model training. The ablation study begins with altering the classifier heads like GlobalAveragePooling2D, GlobalMaxPooling2D, Flatten and Dual Pooling (combination of average pooling and max pooling). The classifier head with best accuracy was taken further for the next ablation study. After selecting the classifier head we tested several optimizers like Adam, Nadam, Adamax, RMSprop and noted results for all of them. Similarly learning rate, batch size, activation function, epoch was also noted with the help of ablation study. The hyper parameters that provide stable training with maximum accuracy are shown in Table 3.2.3.

After training our model, we trained the remaining pretrained models (DenseNet201, ResNet50V2 and InceptionV3). While training a custom callback function was used to record the training time (first epoch, average epoch and total training time).

Table - 3.2.3 : Training Arguments

Parameter	Value
Model	LightXrayNet
Classifier Head	GlobalAveragePooling2D
Loss Function	Cross-entropy loss

Epochs	40
Batch Size	48
Optimizer	Nadam
Learning Rate	0.0003
Activation Function	Swish
Dropout Rate	0.3
Input Image Size	224 × 224 (grayscale)
Evaluation Strategy	Stratified 5-fold CV

4. Model Evaluation

To evaluate how well our models perform in classifying lung diseases, we used the classification report metrics. These help us understand different aspects of the model's predictions. Accuracy tells us the overall percentage of correct predictions out of all predictions made. Precision measures how many of the predicted positive cases were actually positive. Recall (also called sensitivity) shows how many of the actual positive cases the model correctly identified. F1-score is a balance between precision and recall, giving a single score that considers both.

- **Accuracy :** Shows how accurate the model is by finding out what fraction of all its predictions were right, whether they were positive or negative.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4)$$

- **Recall :** It's also called sensitivity, and it shows how well the model finds real positive cases.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

- **Precision :** Shows how many of the results the model says are positive are actually positive, which tells us how well the model avoids making mistakes by saying something is positive when it is not.

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

- **F1- Score :** A way to combine precision and recall into a single score that balances both.

$$\text{F1-score} = \frac{2 \times (\mathit{Precision} \times \mathit{Recall})}{\mathit{Precision} + \mathit{Recall}} \quad (7)$$

3.3 Project Plan

The project will be executed in **four main phases** over a period of 36 weeks (from Week 12 to Week 48). Each phase consists of specific activities designed to achieve the research objectives efficiently:

1. **Data Collection Phase (Week 12–18)**
   Gather multi-class chest X-ray datasets from public repositories.
   Perform dataset verification and quality checks.

2. **Data Preprocessing Phase (Week 19–24)**
   Apply CLAHE for contrast enhancement.
   Resize all x-ray images to 224 by 224 pixels.
   Normalize the pixel values.

3. **Model Training & Evaluation Phase (Week 25–36)**
   Train the proposed LightXrayNet model and compare with pretrained models (ResNet50V2, DenseNet201, InceptionV3).
   Conduct ablation studies to evaluate the impact of architectural and preprocessing choices.
   Perform 5-fold cross-validation and statistical significance testing.
   Generate performance metrics: Accuracy, F1-score, AUC, Confusion Matrix.

4. **Demonstration Phase (Week 37–48)**
   Integrate Explainable AI methods (Grad-CAM, LIME, Score-CAM) for interpretability. Prepare documentation, final report, and presentation.

3.4 Task Allocation

Juiria Humayan (213-15-4394)

1. Data Collection
2. Data preprocessing (CLAHE, resizing, normalization, augmentation – horizontal flip & ±5° rotation)
3. Literature review and related works analysis

Md. Najmus Sakib Nahid (213-15-4575)

1. Model architecture design (LightXrayNet)
2. Model training, evaluation and hyperparameter tuning
3. Explainable AI integration (Grad-CAM, LIME, Score-CAM)

Joint Contribution

1. Documentation and report writing
2. Preparation of final presentation

This table depicts the timeline of the principal activities in each period of the project, from week 12 to week 48.

Table - 3.4 : Task Allocation

Task	Weeks 12–18	Weeks 19–24	Weeks 25–30	Weeks 31–36	Weeks 37–42	Weeks 43–48
Data collection phase	■					
Preprocess all the data		■				
Model training			■	■		
Ablation studies			■	■		
Evaluation & Analysis			■	■		
Explainable AI integration					■	
Documentation & Report						■

3.5 Summary

This chapter described the methodology, system design, requirements, workflow, project plan, and task allocation. By employing LightXrayNet, the study ensures high diagnostic accuracy while maintaining computational efficiency and transparency, making it deployable in real-world healthcare scenarios.

Chapter 4

Implementation and Results

4.1 Environment Setup

All the image preprocessing work was completed in a HP laptop running Fedora Linux operating system, having a total of 8 GB of RAM and Intel Core i5 11th gen processor. Beside image pre-processing, data splitting, model training and evaluation, cross validation, significance testing and implementation of Explainable AI was done in Kaggle Notebook, with the help of NVIDIA CUDA-enabled GPU and 16 GB of RAM. We used python as our main programming language, then the Tensorflow framework for model design, training and evaluation. Also the pandas and numpy library was used for data handling, moreover for visualization related tasks we used the matplotlib and seaborn library.

Table - 4.1 : System Environment Configuration

Component	Details
Programming Language	Python 3.12
Development Platforms	Kaggle Notebook (NVIDIA Tesla T4 GPU) (16 GB RAM)
Deep Learning Stack	TensorFlow v: 2.18.0, Keras v: 3.8.0
Dataset Source	Kaggle – X-ray Lung Diseases Images (9 Classes) [3]
Evaluation Libraries	scikit-learn, NumPy, SciPy
Data Handling	Pandas, NumPy
Visualization Tools	Matplotlib, Seaborn
Utility	Time, Random, Shutil

4.2 Comparative Analysis

We compared each model to show where they performed well or had limitations, focusing on predictive accuracy, how quickly they learned, and how efficiently they used computing resources.

InceptionV3 showed the weakest performance among the evaluated models. Its training and validation curves (Figure 4.2.1) indicated slow and unstable convergence, with training accuracy climbing close to 97% while validation accuracy plateaued around 92–93% and fluctuated heavily. Training loss steadily declined, but validation loss remained higher (0.20–0.25), showing weaker generalization.

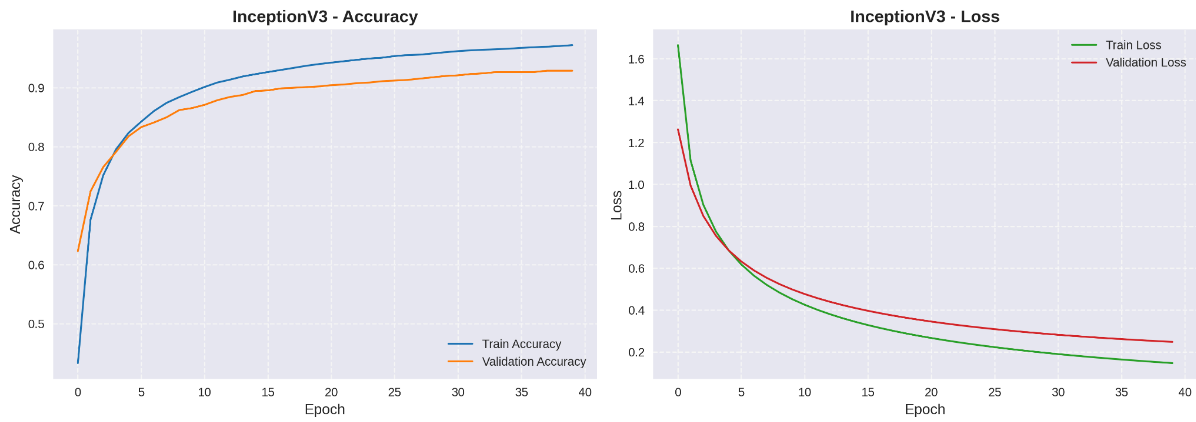


Figure 4.2.1 : Training and Validation Accuracy/Loss Analysis InceptionV3

The confusion matrix (Figure 4.2.2) confirmed this trend: pneumonia was often misclassified as normal or higher density, and errors were more widely spread across categories. Despite taking 380 seconds to train, InceptionV3 only achieved 92.67% test accuracy, making it the least effective model.

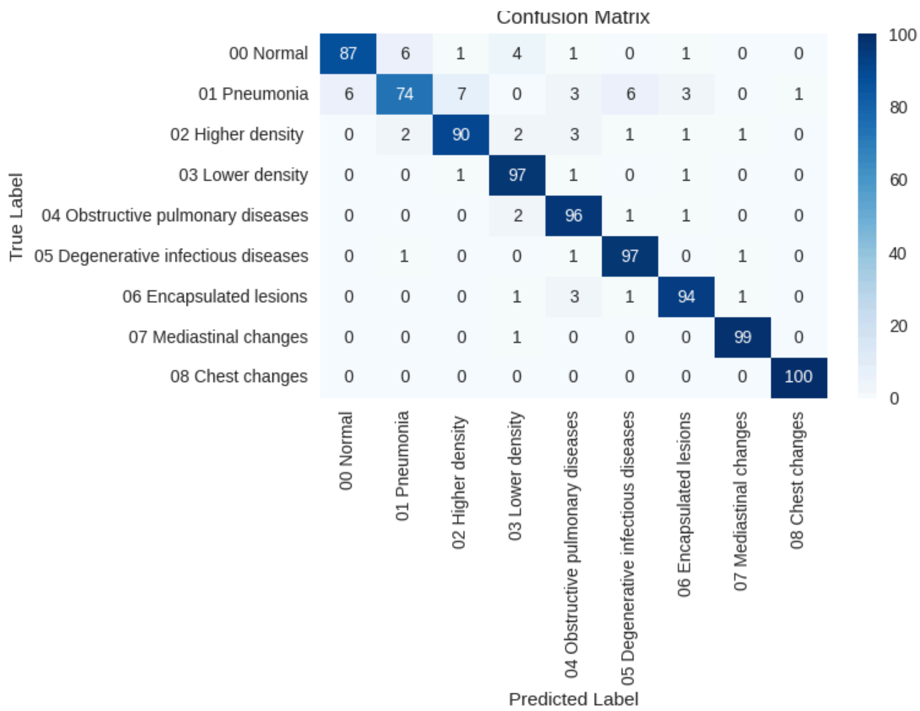


Figure 4.2.2 : Confusion Matrix for InceptionV3

ResNet50V2 recorded the strongest results among the pretrained architectures. Its training curves (Figure 4.2.3) showed consistent convergence, with accuracy approaching 98% and validation stabilizing at ~97%. A mild divergence after epoch 20 hinted at slight overfitting, but generalization remained solid.

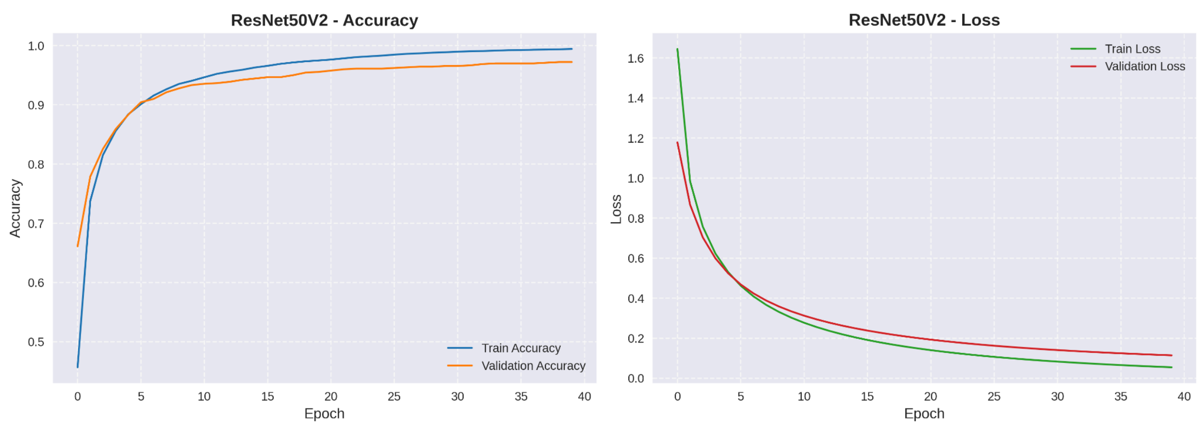


Figure 4.2.3 : Training and Validation Accuracy/Loss Analysis ResNet50V2

The confusion matrix (Figure 4.2.4) highlighted improved class separation, with only limited overlap in pneumonia predictions, while most other conditions—such as encapsulated lesions and mediastinal changes—were identified almost perfectly. ResNet50V2 trained faster than InceptionV3 (375 seconds total, ~9s/epoch) and achieved 98.22% test accuracy.

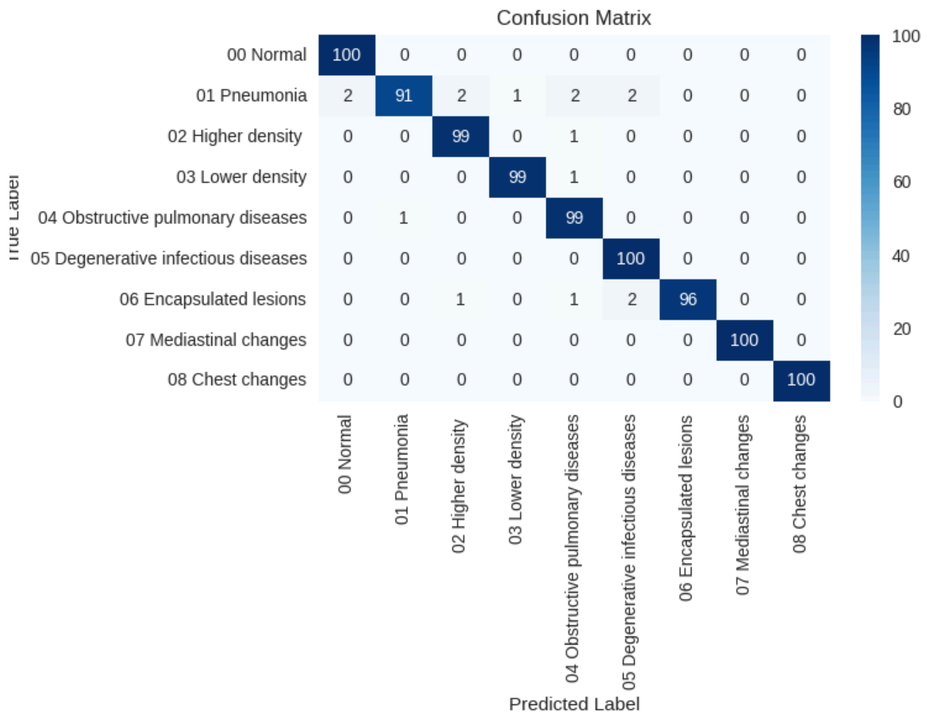


Figure 4.2.4 : Confusion Matrix for ResNet50V2

DenseNet201 achieved more stable and reliable performance. Its training curves (Figure 4.2.5) illustrated gradual but steady convergence, with training accuracy exceeding 97% after 35 epochs and validation accuracy holding at ~95%.

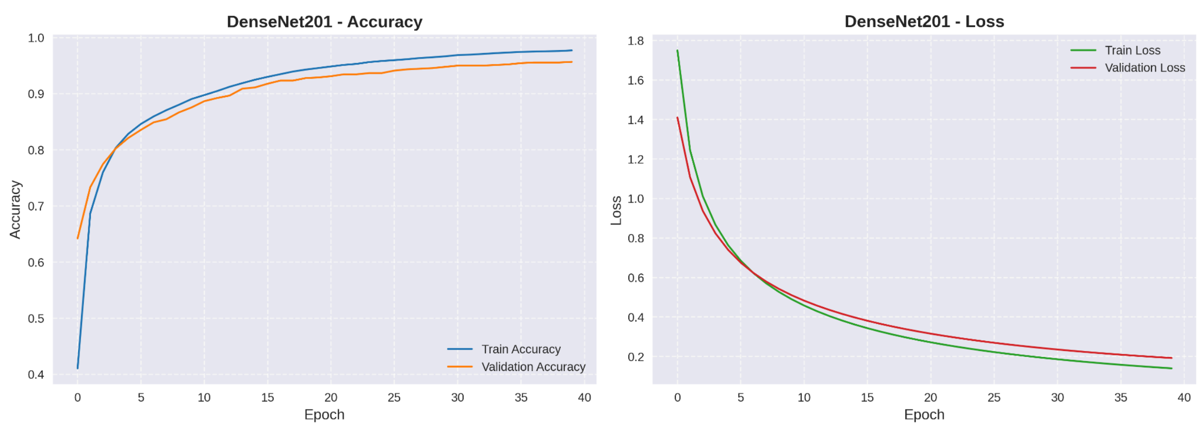


Figure 4.2.5 : Training and Validation Accuracy/Loss Analysis DenseNet201

The confusion matrix (Figure 4.2.6) showed clear diagonal dominance, meaning almost all classes were correctly predicted, with only minor confusion between pneumonia and normal or encapsulated lesions. However, this accuracy came at a cost: DenseNet201 required 706 seconds total training time (106s first epoch), making it the slowest despite delivering 96.56% accuracy.

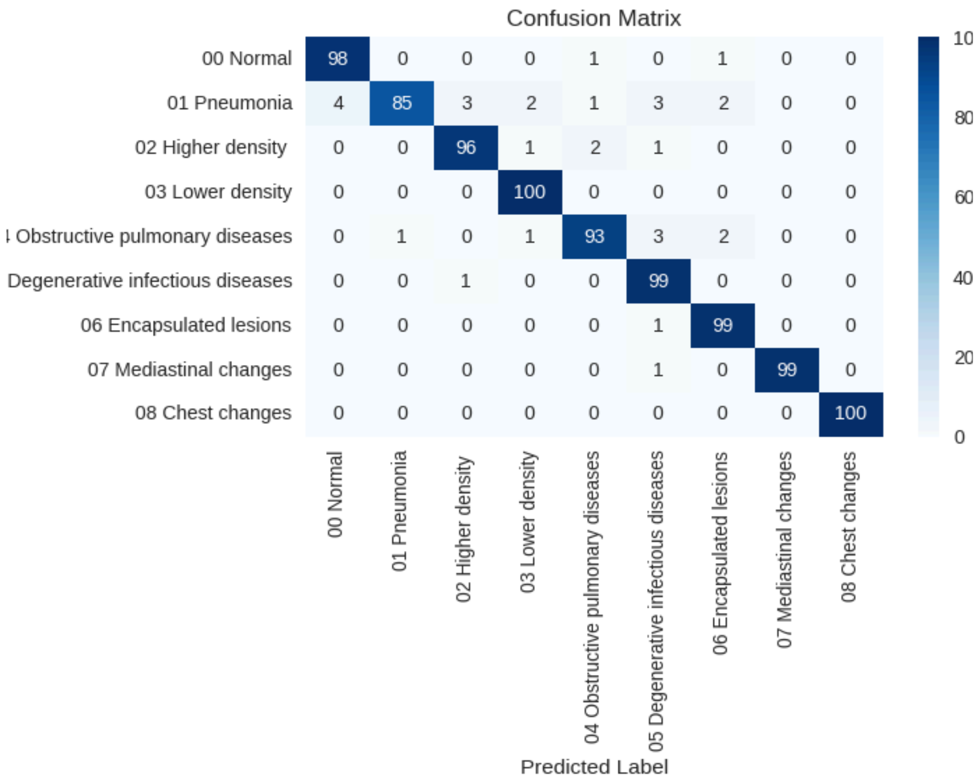


Figure 4.2.6 : Confusion Matrix for DenseNet201

LightXrayNet, the proposed model, achieved accuracy comparable to ResNet50V2 but with far greater efficiency. Its training curves (Figure 4.2.7) showed the fastest and most stable convergence. Training accuracy rose from 27% in the first epoch to above 90% by epoch 11, while validation accuracy tracked closely at ~99% without divergence. Loss values dropped steeply (training <0.02, validation ~0.12), indicating excellent generalization and minimal overfitting.

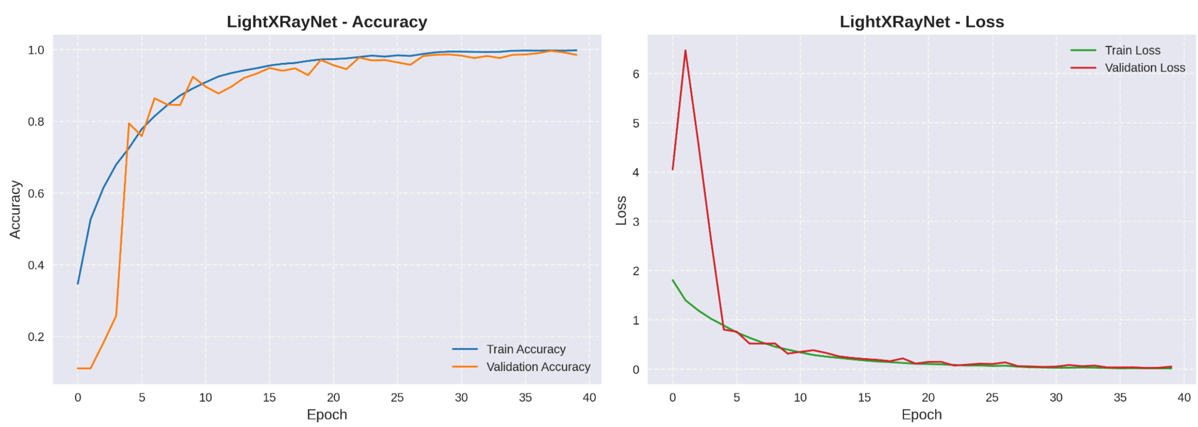


Figure 4.2.7 : Training and Validation Accuracy/Loss Analysis LightXrayNet

The confusion matrix (Figure 4.2.8) demonstrated near-perfect predictions across all categories, with 100% accuracy for mediastinal changes, encapsulated lesions, and degenerative infectious diseases, and only minor misclassifications for pneumonia. Most importantly, LightXrayNet trained in just 138 seconds total (~3s/epoch), proving that lightweight models can deliver state-of-the-art results.

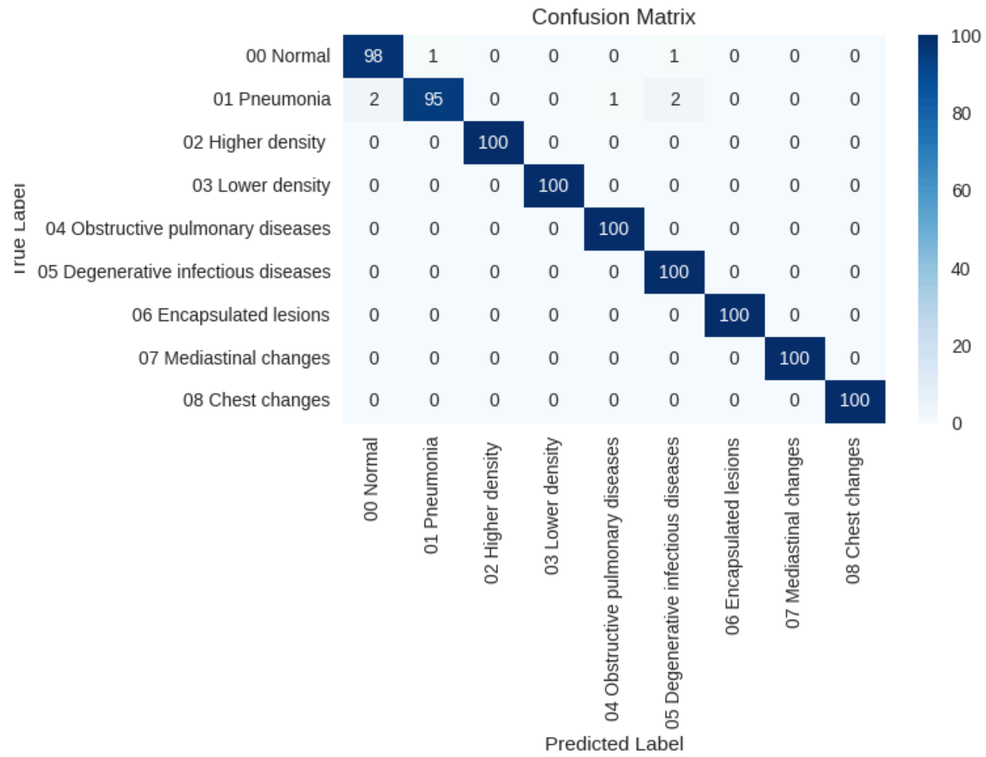


Figure 4.2.8 : Confusion Matrix for LightXrayNet

Overall, the results highlight a clear trade-off. InceptionV3 was the weakest due to instability and low accuracy. DenseNet201 balanced speed and accuracy but still showed mild overfitting. ResNet50V2 was the most accurate pretrained model but prohibitively slow. LightXrayNet, in contrast, delivered high accuracy (99.22%) with unmatched efficiency, making it the most practical choice for real-world deployment.

Table - 4.2 : Time Complexity Analysis

Model Name	**First Epoch** (Sec)	**Average Epoch** (Sec)	**Total Time** (Sec)
InceptionV3	44	9	380
ResNet50V2	31	9	375
DenseNet201	106	18	706
LightXrayNet	18	3	138

4.3 Results and Discussion

Table 4.3. outlines the summarized information of all the models. Among the pretrained ones, InceptionV3 achieved the weakest results, with an accuracy, precision, recall, and F1-score of 0.93. Both ResNet50V2 and DenseNet201 performed better, reaching scores of 0.98 and 0.97 respectively across all metrics, confirming their strength in medical image classification tasks. However, the custom-designed MedNetMini (LightXrayNet) surpassed all pretrained models, achieving 0.99 across accuracy, precision, recall, and F1-score. This demonstrates that a carefully designed lightweight model can outperform larger pretrained networks while maintaining computational efficiency, making it more practical for real-world healthcare use.

Table - 4.3 : Performance comparison

Model Name	**Accuracy**	**Precision**	**Recall**	**F1**
InceptionV3	0.93	0.93	0.93	0.93
ResNet50V2	0.98	0.98	0.98	0.98
DenseNet201	0.97	0.97	0.97	0.97
LightXrayNet	0.99	0.99	0.99	0.99

To evaluate model stability and generalization, a 5-fold cross-validation was conducted, with results shown in Table 4.3.1. LightXrayNet consistently maintained outstanding performance, achieving an average accuracy of 0.985 ± 0.001 across folds. In contrast, pretrained models showed sharp declines when evaluated under the same setup. ResNet50V2 dropped to 0.600 ± 0.011, DenseNet201 to 0.511 ± 0.016, and InceptionV3 to 0.706 ± 0.011. These findings highlight a key weakness of pretrained models: while they may perform strongly in a single train–test split, they fail to generalize well across folds. LightXrayNet, on the other hand, demonstrated both high accuracy and stability, reinforcing its suitability for clinical deployment where reliability is critical.

Table - 4.3.1 : Cross Validation

Model	**Accuracy (mean±SD)**
InceptionV3	0.706 ± 0.011
ResNet50V2	0.600 ± 0.011
DenseNet201	0.511 ± 0.016
LightXrayNet	0.985 ± 0.001

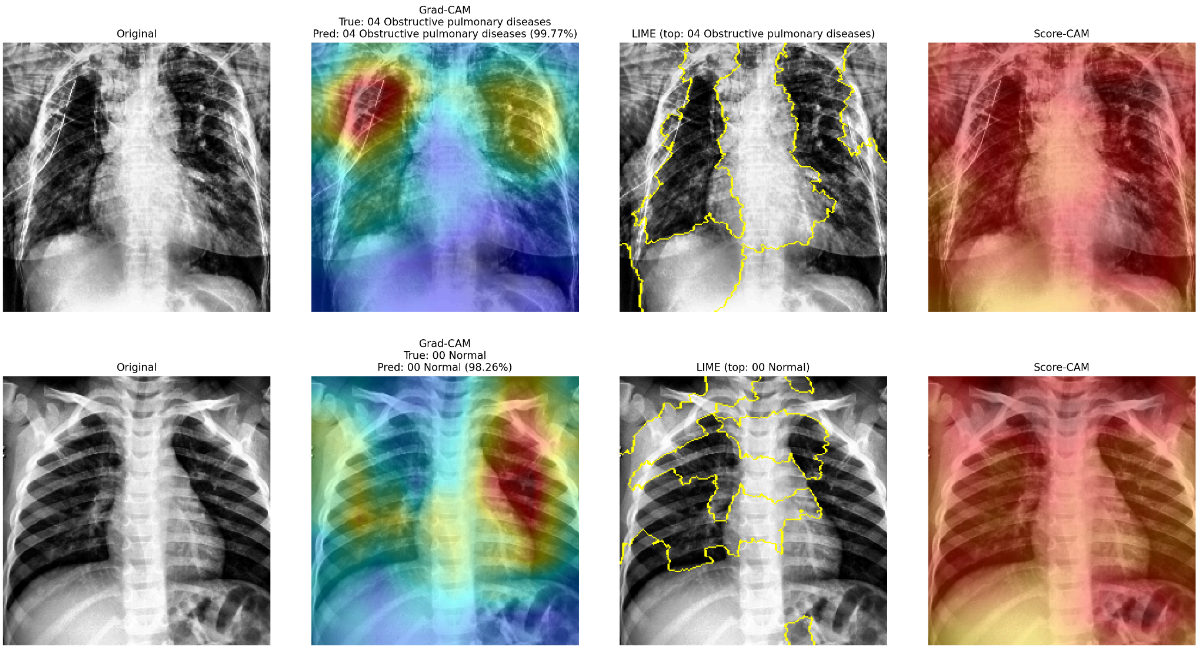
The performance differences were further validated using paired t-tests, with results summarized in Table 4.3.2. When compared against InceptionV3, ResNet50V2, and DenseNet201, LightXrayNet achieved extremely high **t-statistics** (52.76, 76.38, and 65.50

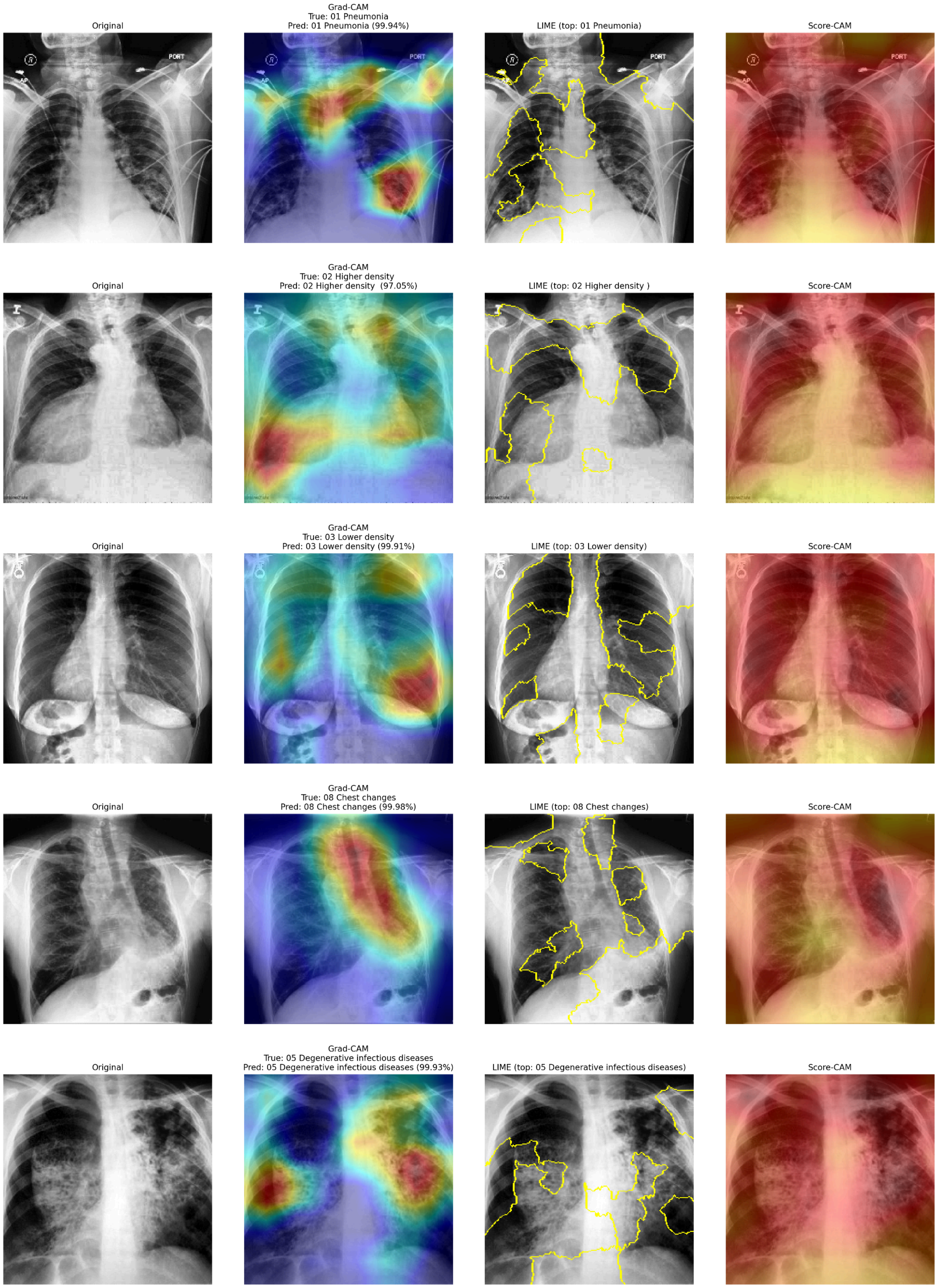
respectively), all with p-values far below 0.000001. The effect sizes (Cohen's d) were also exceptionally large, ranging from 23.59 to 34.16, which indicates not only statistical significance but also substantial practical importance. These results confirm that the improvements delivered by LightXrayNet are both real and highly meaningful, leaving little doubt about its superiority over the pretrained models.

Table - 4.3.2 : Statistical Analysis

Model A	Model B	T-Statistic	P-Value	Effect Size (cohens_d_paired)
LightXrayNet	InceptionV3	52.76	0.00000077	23.59
LightXrayNet	ResNet50V2	76.38	0.00000018	34.16
LightXrayNet	DenseNet201	65.50	0.00000032	29.29

Beyond raw performance, LightXrayNet was evaluated for interpretability using explainable AI techniques, including Grad-CAM, LIME, and Score-CAM. These methods allowed visualization of the regions in chest X-rays that influenced the model's predictions. As shown in Figure 4.3, LightXrayNet consistently focused on clinically meaningful areas such as pulmonary opacities in pneumonia cases or encapsulated regions in lesion cases. This transparency helps clinicians understand *why* a prediction was made, building confidence in the system's output. In contrast, pretrained models functioned largely as "black boxes," offering no built-in interpretability. By integrating XAI, LightXrayNet bridges the gap between high performance and clinical trust, which is essential for adoption in real-world healthcare environments.




Original
Grad-CAM
True: 01 Pneumonia
Pred: 01 Pneumonia (99.94%)
LIME (top: 01 Pneumonia)
Score-CAM
PORT
AP
Original
Grad-CAM
True: 02 Higher density
Pred: 02 Higher density (97.05%)
LIME (top: 02 Higher density)
Score-CAM
Original
Grad-CAM
True: 03 Lower density
Pred: 03 Lower density (99.91%)
LIME (top: 03 Lower density)
Score-CAM
Original
Grad-CAM
True: 08 Chest changes
Pred: 08 Chest changes (99.98%)
LIME (top: 08 Chest changes)
Score-CAM
Original
Grad-CAM
True: 05 Degenerative infectious diseases
Pred: 05 Degenerative infectious diseases (99.93%)
LIME (top: 05 Degenerative infectious diseases)
Score-CAM

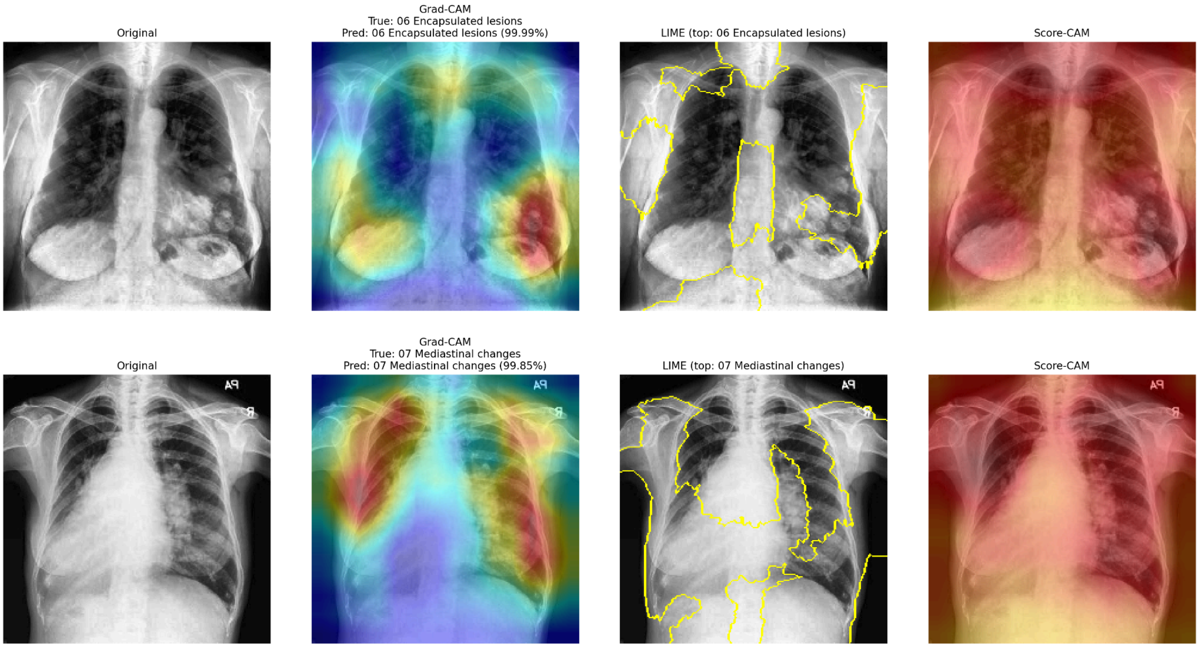


Figure - 4.3 (implementation of Grad-CAM, LIME, Score-CAM among the nine classes)

In summary, the comparative evidence highlights that LightXrayNet offers the best overall balance among all models. While pretrained architectures like InceptionV3, ResNet50V2, and DenseNet201 achieved good accuracy, they required longer training times and heavier computational resources. In contrast, LightXrayNet delivered high accuracy with significantly faster and more efficient training, while also demonstrating strong generalization and interpretability. These advantages make it a more practical and trustworthy choice for real-world healthcare deployment, especially in environments where computational resources are limited. Building on this analysis, the next section (4.3) provides a deeper evaluation of performance, cross-validation, statistical analysis, and explainability to further validate these findings.

4.4 Summary

This chapter described the design, implementation, and experimental evaluation of LightXrayNet in comparison with pretrained models. The experimental results indicate that pretrained networks such as InceptionV3, ResNet50V2, and DenseNet201 achieved high accuracy but required substantial computational resources and exhibited slower training times. LightXrayNet achieved comparable or superior performance, reaching up to 99% accuracy, precision, recall, and F1-score, while maintaining a lightweight and efficient architecture. Cross-validation demonstrated the stability of LightXrayNet, which consistently maintained high accuracy across folds. In contrast, pretrained models exhibited a marked reduction in generalization. Statistical analyses confirmed that the observed performance improvements were highly significant, with large effect sizes across all comparisons.

Beyond performance metrics, LightXrayNet integrates explainability through Grad-CAM, LIME, and Score-CAM visualizations, which support clinical interpretability and trust in model predictions. The combination of high accuracy, computational efficiency, robust generalization, and transparent decision-making demonstrates the practical advantages of

LightXrayNet compared to larger pretrained architectures. Consequently, LightXrayNet is suitable for deployment in healthcare settings, especially in resource-constrained environments where efficiency and reliability are critical.

Chapter 5

Engineering Standards and Design Challenges

5.1 Compliance with the Standards

Following the standards were key to making LightXrayNet reliable, reproducible, and deployable. Aligning with software, hardware, and communication standards helped us achieve technical robustness and long-term sustainability.

5.1.1 Software Standards

All software for LightXrayNet was selected according to international standards to make sure not only readability, reproducibility, but also sustainability. The software that was used are : Python (following PEP8 guidelines), TensorFlow, Keras. IEEE software testing standards were applied while doing unit tests and performance checks across preprocessing, training, and evaluation modules, ensuring the system is not only reliable but also trustworthy.

5.1.2 Hardware Standards

The hardware for LightXrayNet was selected to meet computing standards. Training used NVIDIA GPUs which provide the parallel processing, besides it is supported by CUDA to ensure consistent performance. For deployment, the model was optimized with TensorFlow. Combining these makes the work more productive and efficient.

5.1.3 Communication Standards

The project followed clear communication standards. The dataset was sourced from Kaggle, and All experiments, from preprocessing to training and evaluation, were reported according to IEEE standards.

5.2 Impact on Society, Environment and Sustainability

LightXrayNet's impact goes beyond technology. By tackling healthcare challenges in both well-resourced and limited-resource settings, it contributes to patient care, societal benefit, ethical practice, and long-term sustainability.

5.2.1 Impact on Life

LightXrayNet offers fast, efficient and automated diagnosis of lung diseases. Besides, it offers a much quicker diagnosis with less cost and time which can help speed up the whole process in an efficient way.

5.2.2 Impact on Society & Environment

As discussed in 5.2.1, it is the ultimate efficient solution for diagnosing many lung diseases. Its efficient design reduces energy use, making it more environmentally friendly.

5.2.3 Ethical Aspects

Ethics were central to the project. The dataset was sourced from Kaggle, and All experiments, from preprocessing to training and evaluation, were reported according to IEEE standards.

5.2.4 Sustainability Plan

The project was developed to support sustainability in areas such as 1. Technical sustainability, 2. Economical sustainability, 3.Research sustainability.

5.3 Project Management and Financial Analysis

This project offers a minimal costing plan as most of the work was done in kaggle environment (GPU-based training), and the potential system that can be developed for this will be able to run on phones and standard devices making it cost effective.

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

Table 5.1: Mapping with Complex Engineering Problem.

EP1 Dept of Knowledge	EP2 Range Of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent Of Stake-holder Involvement	EP7 Interdependence
✔	✔	✔	✔			✔

EP1 – Depth of Knowledge

This project applied some of the core engineering principles—like image preprocessing, convolutional neural networks, and evaluation metrics along with advanced techniques like: transfer learning, lightweight CNN design, and explainable AI.

EP2 – Range of Conflicting Requirements

The project -”LightXrayNet” is designed in a way that can work on a less powerful computer. It also helps to balance between being correct and being fast.

EP3 – Depth of Analysis

The analysis for this project was done in many levels like simple accuracy measurement, cross-validation, ablation studies, statistical analysis and so on.

EP4 – Familiarity of Issues

The LightXrayNet addresses issues like imbalance dataset, overfitting etc showing it handles common problems with useful solutions.

EP7 – Interdependence

LightXrayNet performs effectively because its hardware and software are connected and work together seamlessly.

Mapping with Knowledge Profile

Table 5.2: Mapping with knowledge Profile.

K1 Natural Science	K2 Mathematics	K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K7 Comprehension	K8 Research Literature
✔	✔	✔	✔	✔	✔		✔

K1 – Natural Science: This project uses chest X-ray imaging, which relies on physics and biology.

K2 – Mathematics: Mathematical foundations such as statistics, linear algebra, and probability are central.

K3 – Engineering Fundamentals: The Core engineering methods are applied via image preprocessing, system design, and performance evaluation, providing the technical backbone of the work.

K4 – Specialist Knowledge: The project applies advanced methods in computational modeling, optimization, and efficient system design.

K5 – Engineering Design: The system is built to be both accurate and efficient, so it can run not only on regular computers but also mobile devices.

K6 – Engineering Practice: The project shows practical use such as : training on GPUs.

K8 – Research Literature: integrating 12+ papers and articles about lung disease classification.

5.4.2 Engineering Activities

Mapping with Complex Engineering Activities

Table 5.3: Mapping with Complex Engineering Activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity

✔	✔	✔	✔	✔

EA1 – Range of Resources: The project used a set of tools such as 1.Kaggle datasets, 2.TensorFlow/Keras, 3.GPU training.

EA2 – Interaction: It brings together a team of computer science, medical imaging, and AI expertise.

EA3 – Innovation: LightXrayNet offered an efficient way to diagnose lung disease in less time , money and space .

EA4 – Impact: The model provides quick diagnosis, it has low-resource settings and uses less energy.

EA5 – Familiarity: It tackles challenges like class imbalance and overfitting.

5.5 Summary

Chapter 5 covers the engineering standards, design challenges, and impact of LightXrayNet. The project followed software, hardware, and communication standards as mentioned. Moreover, It applied core engineering principles and advanced techniques like transfer learning, lightweight CNN design, and explainable AI. The LightXrayNet not only delivers fast, automated lung disease diagnosis, but also improves diagnosis in a cost effective and energy-efficient way.

Chapter 6

Conclusion

6.1 Summary

This work introduced a custom CNN model: LightXrayNet, a lightweight neural network that helps to classify multiple lung diseases using chest X-ray images. The whole process began with data preprocessing(CLAHE, resizing, normalizing), model building, to compare with well-known pretrained models like DenseNet201, ResNet50V2, and InceptionV3 to provide a fair comparison. Performance was recorded using standard metrics (accuracy, precision, recall, F1-score, confusion matrices, time complexity). In this project AI tools like Grad-CAM, Score-CAM, and LIME were also applied.

6.2 Limitation

Although the project performed well, it has some limitations. The model is built using only one dataset, so it may not work on other images from outside our used dataset. LightXrayNet runs efficiently, but it still needs more optimization before it can be used in real time on mobile or devices in clinical settings.

6.3 Future Work

To overcome these limitations and make LightXrayNet more useful, several future steps can be taken like : training the model using different dataset both from online and offline resources. The framework could also be improved.

References

[1] World Health Organization. (2025). Tuberculosis. WHO. https://www.who.int/news-room/fact-sheets/detail/tuberculosis

[2] Bangladesh intensifies fight against Tuberculosis: Key stakeholders convene to mitigate challenges. (2024). Brac.net. https://www.brac.net/latest-news/item/1423-bangladesh-intensifies-fight-against-tuberculosis-key-stakeholders-convene-to-mitigate-challenges

[3] Dataset: X-ray Lung Diseases Images (9 Classes). Kaggle. https://www.kaggle.com/datasets/fernando2rad/x-ray-lung-diseases-images-9-classes

[4] Tyagi, S., Srivastava, S., & Sahana, B.C. (2025). Design of optimized fourth order PDE filter for restoration and enhancement of Microbiopsy images of breast Cancer. Multimed Tools Appl, 84, 13223–13253. https://doi.org/10.1007/s11042-024-19527-8

[5] Mohan, G., Subashini, M. M., Balan, S., & Singh, S. (2024). A multiclass deep learning algorithm for healthy lung, Covid-19 and pneumonia disease detection from chest X-ray images. Discover Artificial Intelligence, 4(1), 20.

[6] Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. Alexandria Engineering Journal, 64, 923-935.

[7] Hao, R., Qiang, Y., Liao, X., Yan, X., & Ji, G. (2021). Multi-Class Classification of Lung Diseases Using CNN Models. Applied Sciences, 11(19), 9289. https://doi.org/10.3390/app11199289

[8] Metwally, M. A. H. (2022). Automatic Detection and Multi-Class Classification of COVID-19, Pneumonia, and Tuberculosis Diseases in Chest X-ray Images Using Deep Learning Techniques.

[9] Shamrat, F. J. M., Azam, S., Karim, A., Ahmed, K., Bui, F. M., & De Boer, F. (2023). High-precision multiclass classification of lung disease through customized MobileNetV2 from chest X-ray images. Computers in Biology and Medicine, 155, 106646.

[10] Kim, S., Rim, B., Choi, S., Lee, A., Min, S., & Hong, M. (2022). Deep learning in multi-class lung diseases' classification on chest X-ray images. Diagnostics, 12(4), 915.

[11] Bhandari, M., Shahi, T.B., & Siku, B. (2023). Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI. PMC.

[12] Al-Sheikh, M. H., Al Dandan, O., Al-Shamayleh, A. S., Jalab, H. A., & Ibrahim, R. W. (2023). Multi-class deep learning architecture for classifying lung diseases from chest X-Ray and CT images. Scientific Reports, 13(1), 19373.

[13] Ifty, S., et al. (2024). Explainable AI framework utilizing advanced deep learning for lung disease classification. arXiv preprint. https://arxiv.org/abs/2404.11428

[14] Aldamani, R., Abuhani, D. A., & Shanableh, T. (2024). LungVision: X-ray Imagery Classification for On-Edge Diagnosis Applications. Algorithms, 17(7), 280.

[15] Farhan, A. M. Q., & Yang, S. (2023). Automatic lung disease classification from chest X-ray images using hybrid deep learning algorithm. Multimedia Tools and Applications, 1.

[16] Irtaza, M., Ali, A., Gulzar, M., & Wali, A. (2024). Multi-Label Classification of Lung Diseases Using Deep Learning. IEEE Access.

[17] Hong, M., Rim, B., Lee, H., Jang, H., Oh, J., & Choi, S. (2021). Multi-Class Classification of Lung Diseases Using CNN Models. Applied Sciences, 11(19), 9289. https://doi.org/10.3390/app11199289

[18] Ismael, A. M., & Şengür, A. (2021). Deep learning approaches for COVID-19 detection based on chest X-ray images. Expert Systems with Applications, 164, 114054. https://doi.org/10.1016/j.eswa.2020.114054

[19] Kim, C., et al. (2023). Deep learning to estimate lung disease mortality from chest radiographs. Radiology, 308(2), e222881.

[20] Chouat, I., Echtioui, A., Khemakhem, R., Zouch, W., Ghorbel, M., & Hamida, A. B. (2022). COVID-19 detection in CT and CXR images using deep learning models. Biogerontology, 23(1), 65–84. https://doi.org/10.1007/s10522-021-09946-7

[21] Huang, T., Yang, R., Shen, L., Feng, A., Li, L., He, N., ... & Lyu, J. (2022). Deep transfer learning to quantify pleural effusion severity in chest X-rays. BMC Medical Imaging, 22(1), 100.

[22] Huynh, J., Masoudi, S., Noorbakhsh, A., Mahmoodi, A., Kligerman, S., Yen, A., ... & Hsiao, A. (2022). Deep Learning Radiographic Assessment of Pulmonary Edema: Optimizing Clinical Performance, Training With Serum Biomarkers. IEEE Access, 10, 48577-48588.

[23] Miles, S. E., Sandrini, A., Johnson, A. R., & Yates, D. H. (2008). Clinical consequences of asbestos-related diffuse pleural thickening: a review. Journal of Occupational Medicine and Toxicology, 3(1), 20.

[24] Rugg, A. L., & Lee, J. J. (2022). Incidental finding of lung hernia in a patient with a remote history of empyema status post video-assisted thoracoscopic surgery. Radiology Case Reports, 17(3), 798-801.

[25] Ayalew, A. M., Bezabh, Y. A., Abuhayi, B. M., & Ayalew, A. Y. (2024). Atelectasis detection in chest X-ray images using convolutional neural networks and transfer learning with anisotropic diffusion filter. Informatics in Medicine Unlocked, 45, 101448.

[26] MDPI Diagnostics. (2024). Diagnostics 14(12), 1243. https://www.mdpi.com/2075-4418/14/12/1243

213-15-4575

ORIGINALITY REPORT

22%	18%	16%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	3%
2	iieta.org Internet Source	1%
3	ouci.dntb.gov.ua Internet Source	1%
4	trilogi.ac.id Internet Source	1%
5	www.mdpi.com Internet Source	1%
6	"Data Mining and Information Security", Springer Science and Business Media LLC, 2025 Publication	1%
7	J. Glory Precious, Rajkumar S, Sarayu Pratika B, Vaisnav R R, Syed Sheik Mohamed M, V Sapthagirivasan. "A Lightweight Hybrid DL Model for Multi-Class Chest X-ray Classification for Pulmonary Diseases", Biomedical Physics & Engineering Express, 2025 Publication	1%
8	Submitted to Liverpool John Moores University Student Paper	<1%
9	Qi Peng, Yi Cai, Jiankun Liu, Xing Chen, Zefeng Wang, Jiayuan Xie, Qing Li. "CKE-Former: A	<1%