Chest X-Ray Pneumonia Detection System

Submitted by:
Vansh Singla (102303806)
Lovenish Attri (102353015)

BE Third Year

COE

Submitted to:

Dr. Anjula Mehto

Assistant Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

November 2025

TABLE OF CONTENTS

Introduction or Project Overview

Pneumonia is a common and potentially life-threatening lung infection, and chest X-rays are widely used for its diagnosis. However, manually analyzing X-ray images can be challenging due to subtle visual differences between healthy and infected lungs, leading to the possibility of misinterpretation. To address this, the project aims to develop a **Deep Learning–based Chest X-ray Pneumonia Detection System** that automatically classifies X-ray images as *NORMAL* or *PNEUMONIA*.

The system uses **ResNet50 (Transfer Learning)** to extract meaningful features from X-ray images and achieve accurate classification. The dataset undergoes preprocessing steps such as resizing, normalization, and augmentation to improve model performance. To make the system practical and accessible, a **Streamlit web application** is integrated, allowing users to upload X-rays and obtain predictions instantly. Additionally, **Grad-CAM visualization** provides heatmaps that highlight the regions influencing the model's decision, adding transparency and interpretability.

Overall, the project demonstrates how deep learning can assist in faster and more reliable pneumonia detection, supporting medical professionals and improving decision-making in clinical environments. By combining automated prediction with visual explanation, the system can serve as a helpful screening tool, especially in settings where expert radiologists may not be readily available.

Problem Statement

Diagnosing pneumonia from chest X-ray images is challenging because the visual differences between healthy lungs and infected areas can be subtle and difficult to distinguish. Radiologists must carefully look for signs such as opacities or abnormal shading, which can vary greatly from patient to patient. This makes manual diagnosis time-consuming and increases the chances of inconsistent interpretations.

The challenge becomes more significant in hospitals with high patient loads or in regions where trained radiologists are limited. Delays or misdiagnosis can affect timely treatment, and traditional methods are not always fast or reliable enough for large-scale screening. Therefore, an automated system is needed to support accurate and efficient pneumonia detection.

Key problems addressed in this project include:
1. **Subtle visual differences** that make manual diagnosis difficult.
2. **Time-consuming interpretation** of large numbers of X-rays.
3. **Dependence on expert radiologists**, which may not be available everywhere.
4. **Risk of diagnostic errors** due to variability or fatigue.
5. **Lack of automated and explainable tools** for reliable pneumonia detection.

This project aims to solve these challenges using a deep learning–based system that provides consistent, accurate, and interpretable classification of chest X-ray images.

Overview of the Dataset used

The Link of the Dataset used is as follows:
https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

It contains 5,863 labeled images.

Class Distribution
- **Normal:** Clear lungs with no infection
- **Pneumonia:** Viral/Bacterial lung infection signs:
    1. White opacities
    2. Consolidation areas
    3. Blurring in lung regions

Preprocessing Steps
1. Convert X-ray to grayscale
2. Resize to **224×224**
3. Stack grayscale into **3 channels** (required for ResNet50)
4. Normalize pixel values to *0–1*
5. Stratified train–validation split

Data Augmentation Used
- Rotation
- Width/height shift
- Zoom
- Horizontal flip

Augmentation boosts the model's ability to generalize to unseen X-rays.

Project Workflow

The workflow of the Chest X-Ray Pneumonia Detection project consists of the following major stages:

Step 1: Data Loading and Cleaning
The dataset is imported from the train/test folder structure and each image is loaded in grayscale, resized to the model input size, converted to a 3-channel array, and normalized. The loader also checks for missing or unreadable files and reports counts while loading.

Step 2: Preprocessing & Feature Engineering
Images are preprocessed to match the model expectations: resizing to **224×224**, stacking grayscale into three channels, and scaling pixel values to the 0–1 range. Data augmentation (rotation, shifts, zoom, horizontal flip) is applied via ImageDataGenerator to improve generalization and reduce overfitting. These preprocessing steps are encapsulated in the data processor utilities.

Step 3: Model Building & Training (ResNet50 Transfer Learning)
A ResNet50 backbone (pretrained on ImageNet) is used as the feature extractor; its convolutional layers are frozen, followed by GlobalAveragePooling and custom dense/dropout layers ending in a sigmoid output for binary classification. The model is compiled with Adam and binary cross-entropy, and trained with EarlyStopping and ReduceLROnPlateau callbacks to prevent overfitting and stabilize learning. Training is orchestrated in the training script.

Step 4: Model Saving and Results Logging
After training, the best model weights are saved to the configured model path and training history (accuracy/loss arrays) is written as JSON for later visualization. The results directory also stores training plots (accuracy/loss) for reporting and the Streamlit app to display. Configuration and paths are managed via the project config.

Step 5: Evaluation & Explainability

Model performance is evaluated on the held-out test set using metrics such as accuracy, precision, recall, F1, confusion matrix, and ROC/AUC. The evaluation module generates and saves confusion matrix and ROC curve visualizations. For interpretability, Grad-CAM produces heatmaps and overlay images that highlight image regions most influential to the model's decision. These explainability outputs help validate that the model focuses on medically relevant areas.

Step 6: Deployment (Streamlit Application)

A Streamlit web application provides a user interface with tabs for single-image prediction, batch analysis, Grad-CAM visualization, and training history. Users can upload X-rays, receive instant predictions with confidence scores, view heatmap overlays, and inspect training curves. The app loads the saved model and history files at runtime and presents summary metrics for batch uploads. Deployment instructions and resource paths are contained in the config and app script.

Results

The Chest X-Ray Pneumonia Detection model was evaluated on the test dataset using multiple performance metrics, including accuracy trends, confusion matrix, ROC–AUC, precision, recall, and F1-score. The results clearly show that the ResNet50-based model performs effectively for binary classification of normal vs. pneumonia chest X-rays
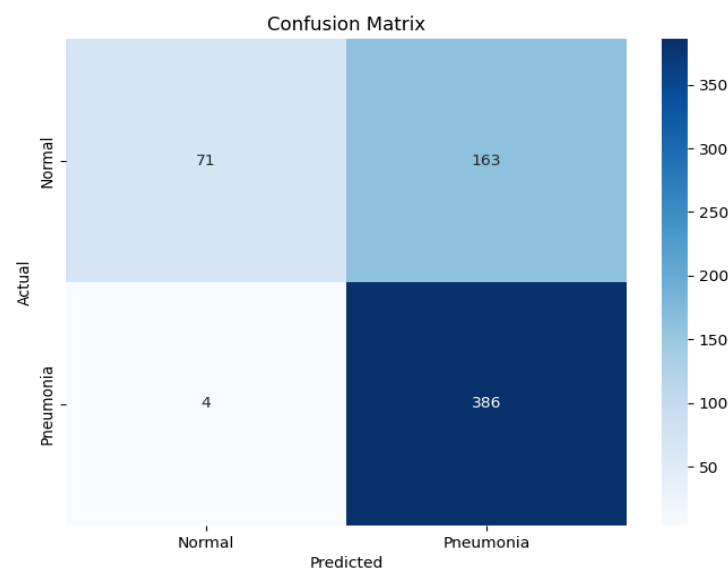
1. Training and Validation Performance
The training history (stored in JSON) indicates steady improvement across epochs. Final values achieved were:
- **Final Training Accuracy:** 91.24%
- **Final Validation Accuracy:** 90.42%
- **Final Training Loss:** 0.2240
- **Final Validation Loss:** 0.2162

The close alignment between training and validation metrics suggests that the model generalizes well without significant overfitting. Both accuracy and loss curves follow a consistent downward trend, indicating effective learning.

2. Confusion Matrix Analysis

Interpretation:

- The model correctly identified **386 pneumonia images**, missing only 4 cases.
- False positives (predicting pneumonia when normal) are higher (163), which is safer than missing pneumonia cases.
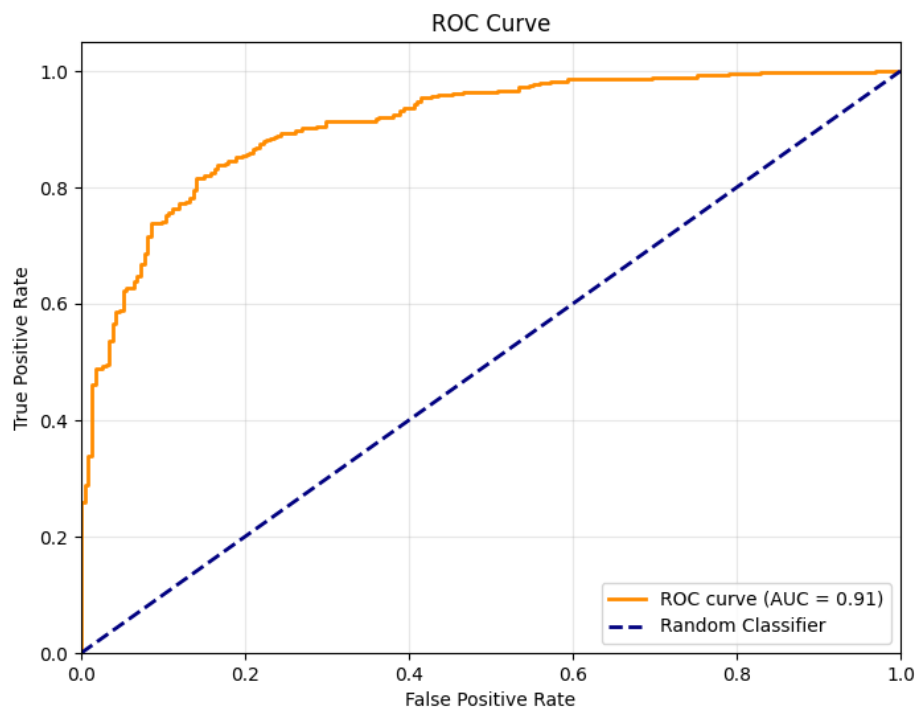- This behavior is acceptable in the medical domain, where **minimizing false negatives is critical**.

3. Classification Metrics

Using the confusion matrix values:

- **Precision** = 70.3%
- **Recall** = 99.0%
- **F1-Score** = 82.2%

4. ROC Curve and AUC Score

The ROC curve demonstrates how well the model distinguishes between the two classes across different probability thresholds.

The model achieved:
- **Test ROC–AUC:** 0.91

An AUC of 0.91 indicates excellent discriminative ability, showing that the model can reliably separate normal and pneumonia cases.

6. Overall Interpretation

The model achieves high performance across all evaluation metrics. The extremely high recall (99%) makes the system particularly useful for medical screening, as it ensures that pneumonia cases are rarely missed. Although the precision is moderate due to higher false positives, this is acceptable in diagnostic scenarios where detecting all true cases is a priority.

The strong ROC–AUC score and low validation loss further confirm the model's reliability and robustness. Overall, the results demonstrate that the deep learning system provides accurate and clinically meaningful predictions for pneumonia detection from chest X-ray images.

Conclusion

This project successfully demonstrates the effectiveness of deep learning techniques in detecting pneumonia from chest X-ray images. By leveraging the ResNet50 architecture through transfer learning, the system achieves strong predictive performance with a validation accuracy of over 90% and a high recall of 99%, ensuring that pneumonia cases are rarely missed. The inclusion of Grad-CAM visualizations further enhances the model's interpretability by highlighting the lung regions contributing to its predictions, making the system more trustworthy and suitable for real-world medical use.

The Streamlit-based interface provides an intuitive and accessible platform for users to upload X-ray images and receive instant diagnostic predictions along with confidence scores and heatmaps. This end-to-end integration—from preprocessing and model training to explainability and deployment—demonstrates how deep learning can be effectively translated into practical healthcare tools. The system also shows consistency in prediction stability, as reflected in the close alignment of training and validation metrics.

Overall, the project highlights the potential of AI-driven systems to support radiologists, reduce diagnostic workload, and improve early detection of pneumonia. With further refinement, larger datasets, and clinical validation, the system can evolve into a reliable medical-assistive application for hospitals and diagnostic centers. Additionally, incorporating multi-class classification for detecting other thoracic diseases, real-time inference, and mobile deployment can further expand the scope and impact of this work. This project demonstrates how artificial intelligence can play a transformative role in enhancing healthcare accessibility and diagnostic efficiency.