



AI-Powered Multi-Disease Chest X-Ray Analysis And Explainable AI System Using Densenet121 and Grad-Cam

Jatavath Asha^{a*}, T. Malathi^b

^aMCA (Data Science), School of Informatics, Aurora Higher Education and Research Academy, Uppal, Hyderabad, Telangana 500098, India

^bAssistant Professor and Head of Department,,Dept. of Computer Application, Aurora Higher Education and Research Academy, Uppal, Hyderabad, Telangana 500098, India

DOI : <https://doi.org/10.55248/gengpi.07.0926.3025>

ABSTRACT :

Chest X-ray imaging is widely used for examining abnormalities associated with the lungs and respiratory system. The increasing availability of medical image datasets has created opportunities for applying deep learning techniques to assist in the preliminary analysis of chest radiographs. However, classification of different chest conditions within a single system and interpretation of the resulting predictions remain important challenges.

This paper presents an AI-Powered Multi-Disease Chest X-Ray Analysis System for classifying chest X-ray images into four categories: COVID-19, Normal, Pneumonia, and Tuberculosis. The implemented system uses a DenseNet121-based transfer learning model for multi-class image classification. The dataset consists of approximately 7,135 chest X-ray images organized into training, validation, and testing sets. The input images are resized to 224×224 pixels before classification. The model produces class-wise probability values, and the class with the highest probability is presented as the predicted category.

To improve the interpretability of the model output, Gradient-weighted Class Activation Mapping (Grad-CAM) is integrated into the system. Grad-CAM uses gradient information associated with a target class to generate a visual heatmap highlighting image regions that contributed to the model's prediction. A basic image-validation mechanism is also included to reject obviously unsuitable inputs before the prediction process.

The trained DenseNet121 model achieved approximately 90.93% training accuracy and 84.21% validation accuracy after 10 epochs. The trained model was integrated into a Flask-based web application, allowing users to upload chest X-ray images and view the predicted class, model confidence, class-wise probabilities, original X-ray image, and corresponding Grad-CAM visualization. The implemented prototype demonstrates the feasibility of combining multi-class chest X-ray classification with explainable artificial intelligence within a single web-based system.

Keywords: Chest X-Ray Analysis, Multi-Disease Classification, Deep Learning, DenseNet121, Transfer Learning, Explainable AI, Grad-CAM, COVID-19, Pneumonia, Tuberculosis, Flask

Introduction

Chest X-ray (CXR) imaging is one of the commonly used imaging techniques for examining abnormalities involving the lungs and other structures within the chest. Because X-ray examinations are relatively accessible and can provide useful visual information about pulmonary conditions, they are frequently considered in the assessment of respiratory diseases. The growing availability of medical image datasets has also created opportunities for applying artificial intelligence and deep learning techniques to computer-assisted chest X-ray analysis.

Convolutional Neural Networks (CNNs) have been widely applied to medical image classification because they can learn visual representations directly from image data. However, training a deep neural network from the beginning generally requires substantial labelled data and computational resources. Transfer learning provides an alternative approach in which knowledge obtained from a previously trained model is adapted to a new classification task. Research on chest X-ray classification has demonstrated the use of transfer learning with architectures such as ResNet50 and DenseNet121, showing its potential for extracting useful image representations in medical imaging tasks.

Several studies have investigated automated detection of individual or multiple respiratory diseases from chest X-ray images. In particular, research has explored the classification of conditions such as COVID-19, pneumonia, and tuberculosis using deep learning techniques. One study investigated a multi-category chest X-ray classification approach involving COVID-19, pneumonia, and tuberculosis and additionally applied explainable artificial intelligence techniques including Grad-CAM, LIME, and SHAP. These studies demonstrate the potential of combining image classification with explanatory techniques, while also indicating the importance of developing systems that can handle multiple disease categories within a common workflow.

An important challenge associated with deep learning-based medical image classification is the interpretability of model predictions. A classification model can produce a predicted class and an associated probability value, but these outputs alone do not indicate which regions of an input image

contributed to the prediction. This can make it difficult for users to understand the visual basis of an individual model output. Gradient-weighted Class Activation Mapping (Grad-CAM) addresses this problem by using gradients associated with a target class to generate a localization map over the input image. The original Grad-CAM work demonstrated that these visualizations can highlight image regions that are important for a CNN's prediction without requiring architectural changes or retraining of the model.

The present work develops an AI-Powered Multi-Disease Chest X-Ray Analysis System that combines multi-class image classification and explainable visualization in a single web-based application. The implemented system classifies chest X-ray images into four categories: COVID-19, Normal, Pneumonia, and Tuberculosis. A DenseNet121-based transfer learning model is used for classification, with input images processed at a resolution of 224×224 pixels. The system produces class-wise probability values and identifies the class having the highest probability as the predicted category.

In addition to classification, the system incorporates a basic image-validation mechanism before prediction. The validation checks characteristics such as image dimensions, aspect ratio, and colour-channel variation to reject obviously unsuitable inputs. This validation is intended as a preliminary input-screening mechanism and is not a medical-grade chest X-ray detector. After a valid image is processed, the system generates the predicted class, model confidence, probability distribution across the four classes, and a Grad-CAM visualization.

The trained model was integrated into a Flask-based web application to provide an accessible interface for uploading images and viewing the analysis results. The implemented prototype achieved approximately 90.93% training accuracy and 84.21% validation accuracy after 10 training epochs. The application therefore demonstrates an end-to-end workflow in which a chest X-ray image is uploaded, validated, processed by the trained DenseNet121 model, classified into one of four categories, and accompanied by an explanatory Grad-CAM visualization.

The primary contribution of this work is the integration of four-class chest X-ray classification, probability-based prediction, basic input validation, and Grad-CAM-based explainability within a single web application. The system is intended as a research prototype for preliminary computational analysis and is not presented as a replacement for professional medical examination or clinical diagnosis.

Literature Survey

1.1. Multi-Disease Chest X-Ray Classification

Deep learning has increasingly been investigated for automated analysis of chest X-ray images because convolutional models can learn discriminative visual patterns directly from radiographic images. Early research in this area often concentrated on individual diseases or a small number of categories. With the availability of combined chest X-ray datasets, researchers have also explored multi-class classification approaches capable of distinguishing between different pulmonary conditions.

A relevant study by Rahman et al. investigated a deep transfer learning approach for distinguishing COVID-19, pneumonia, tuberculosis, and healthy cases from chest X-ray images. The proposed DenResCov-19 framework combined DenseNet-121 and ResNet-50 through an additional convolutional component. The study evaluated two-class, three-class, and four-class classification settings, including the combination of COVID-19, pneumonia, tuberculosis, and healthy categories. This work demonstrates the feasibility of using transfer learning for multiple respiratory disease categories within a common classification framework.

Another important study by Bhandari et al. (2022) developed a deep learning framework for classifying chest X-ray images associated with COVID-19, pneumonia, and tuberculosis. The researchers used a publicly available dataset containing 7,132 chest X-ray images and incorporated multiple explainable AI techniques, including Grad-CAM, LIME, and SHAP. Their study reported classification results together with visual explanations intended to make the model outputs easier to interpret.

These studies indicate that multi-disease classification can provide a unified computational workflow instead of developing a separate model for every disease. However, differences in datasets, preprocessing methods, model architectures, class distributions, and evaluation procedures mean that the results of one study cannot be directly treated as the expected performance of another system.

1.2. Deep Learning and Transfer Learning for Chest X-Ray Analysis

Training a deep neural network entirely from the beginning can require substantial labelled data and computational resources. Transfer learning provides an alternative by adapting representations learned from a previously trained network to a new image classification task.

DenseNet121 has been investigated for chest radiograph classification because of its ability to reuse features through dense connections. In a study focused on COVID-19 classification from chest radiographs, researchers trained DenseNet121 under different training strategies and examined whether additional training using related chest radiography datasets could improve classification performance. The study used 1,060 radiographs from 475 patients, including laboratory-confirmed COVID-19 cases.

Further research has also applied DenseNet121 as a transfer-learning backbone for COVID-19, pneumonia, and normal chest X-ray classification. For example, a more recent study developed a DenseNet121-based architecture with multi-scale feature extraction and attention mechanisms for three-category chest X-ray classification.

The existing literature therefore supports the use of transfer learning as a practical approach for medical image classification. In the present work, DenseNet121 is used as the main feature-extraction and classification architecture and is adapted to the four categories COVID-19, Normal, Pneumonia, and Tuberculosis.

1.3. Explainable AI and Grad-CAM

Although deep learning models can produce useful classification outputs, understanding the visual basis of a prediction is an important consideration when applying such models to medical images. A model may provide a predicted category and probability without directly indicating which regions of the image influenced the prediction.

Grad-CAM, introduced by Selvaraju et al., provides a method for producing class-specific visual explanations from convolutional neural networks. It uses gradients associated with a target class to generate a localization heatmap over the input image.

Bhandari et al. applied Grad-CAM together with LIME and SHAP to chest X-ray classification involving COVID-19, pneumonia, and tuberculosis. Their work demonstrated how visual explanation techniques can be incorporated alongside disease classification to provide additional information about model decisions.

More recent research has continued to investigate explainable AI for chest X-ray analysis. For example, a 2025 study combined a Vision Transformer with Grad-CAM for tuberculosis detection, using the visualization to highlight regions associated with model predictions.

In the proposed system, Grad-CAM is used specifically to generate a visual representation of the regions contributing to the DenseNet121 prediction. The visualization is presented alongside the original X-ray image so that the model output is not limited to a numerical class and probability.

1.4. Research Gap

The reviewed studies demonstrate the potential of deep learning, transfer learning, and explainable AI for chest X-ray analysis. However, several areas remain relevant when developing an integrated research prototype.

First, existing studies differ considerably in their disease categories and datasets. Some focus on individual diseases, while others combine a limited number of respiratory conditions. Although multi-disease classification has been investigated, there remains scope for implementing a unified four-class workflow containing COVID-19, Normal, Pneumonia, and Tuberculosis using a single DenseNet121-based model.

Second, many research studies primarily concentrate on model development and classification performance. The complete workflow from image upload and basic input validation to prediction, probability display, and visual explanation is not always provided as a simple web-based prototype.

Third, explainability is addressed by several studies using techniques such as Grad-CAM, LIME, or SHAP, but the choice and integration of these methods vary between systems. A lightweight implementation using Grad-CAM can provide a practical visual explanation without requiring multiple explanation frameworks.

Based on these observations, the present work focuses on integrating multi-class chest X-ray classification, basic input validation, probability-based prediction, DenseNet121 transfer learning, Grad-CAM visualization, and Flask-based web deployment within a single prototype.

1.5. Problem Statement

Chest X-ray images are widely used for examining abnormalities associated with the lungs and respiratory system. However, identifying different respiratory conditions from X-ray images can be challenging because several diseases may produce visually similar patterns. Conventional interpretation depends on trained medical professionals and can become time-consuming when a large number of images need to be examined. Deep learning-based image classification has therefore been investigated as a computational approach for assisting with preliminary analysis of chest radiographs.

Existing research has demonstrated the use of deep learning and transfer learning for detecting individual diseases or a limited number of disease categories from chest X-ray images. Several approaches have focused on conditions such as COVID-19, pneumonia, or tuberculosis, while some studies have investigated multi-class classification. However, the available approaches differ in their datasets, model architectures, preprocessing procedures, disease categories, and deployment methods. In addition, a classification model may provide only a predicted category and probability without presenting a visual indication of the regions that influenced the prediction.

Another important issue is the handling of unsuitable input images. A system designed specifically for chest X-ray analysis should provide at least a preliminary mechanism for identifying obviously unsuitable inputs before applying the trained classification model. Without such validation, unrelated or inappropriate images may be processed by the model and produce meaningless predictions.

Therefore, there is a need for an integrated chest X-ray analysis system that can classify multiple disease categories within a common workflow, provide class-wise probability information, perform preliminary input validation, and offer an interpretable visualization of the model prediction. The proposed system addresses these requirements through a DenseNet121-based transfer learning model, basic image validation, probability-based multi-class classification, Grad-CAM explainability, and a Flask-based web interface. The system focuses on four categories: COVID-19, Normal, Pneumonia, and Tuberculosis.

The developed system is intended as a research prototype for preliminary computational analysis and not as a replacement for professional medical examination or clinical diagnosis.

1.6. Objectives

The major objectives of the implemented system are:

- To develop a multi-class chest X-ray classification system capable of categorizing chest X-ray images into four classes: COVID-19, Normal, Pneumonia, and Tuberculosis.
- To implement a DenseNet121-based transfer learning model for extracting relevant image features and performing multi-class classification of chest X-ray images.
- To prepare and preprocess chest X-ray images by organizing the dataset into appropriate training, validation, and testing sets and resizing input images to 224×224 pixels for model processing.
- To generate class-wise probability values and prediction confidence so that the system presents the model's estimated probability for each of the four categories along with the predicted class.
- To implement basic input validation for identifying obviously unsuitable images before they are passed to the classification model.
- To integrate Grad-CAM-based explainability to generate a visual heatmap showing image regions that contributed to the selected model prediction.
- To develop a Flask-based web interface through which users can upload a chest X-ray image and view the corresponding classification and explainability results.
- To integrate the complete analysis workflow including image upload, validation, preprocessing, DenseNet121 prediction, probability calculation, confidence display, and Grad-CAM visualization within a single application.
- To evaluate the implemented system through functional testing and model performance measurements, including training and validation accuracy and verification of the complete prediction workflow.

Proposed Methodology

The proposed methodology follows an integrated, modular approach for multi-disease chest X-ray classification and explainable image analysis. The system combines dataset preparation, image preprocessing, DenseNet121-based transfer learning, multi-class prediction, probability analysis, preliminary image validation, Grad-CAM-based explainability, and Flask-based web integration within a unified workflow.

1.7. Startup Readiness Assessment Framework

The system uses a publicly available chest X-ray dataset containing images belonging to four categories: COVID-19, Normal, Pneumonia, and Tuberculosis.

The dataset contains approximately 7,135 chest X-ray images, organized into training, validation, and testing subsets.

Table 1 – Dataset Distribution

Dataset Category	Number of Images
Training	6,326
Validation	38
Testing	771
Total	7,135

The four disease categories used by the implemented classification model are:

Class	Description
COVID-19	Chest X-ray images associated with COVID-19
Normal	Chest X-ray images without the target disease categories
Pneumonia	Chest X-ray images associated with pneumonia
Tuberculosis	Chest X-ray images associated with tuberculosis

The dataset is used to train and evaluate the DenseNet121-based classification model. The same class ordering is maintained during model training and deployment to ensure consistency between the trained model and the Flask application.

1.8. Image Preprocessing

Before classification, the input chest X-ray image is loaded and resized to 224×224 pixels, corresponding to the input resolution used by the trained DenseNet121 model.

The preprocessing workflow includes image loading, resizing, conversion into a numerical image array, and preparation of the input tensor for model prediction.

During web-based prediction, the uploaded image is processed using the same input preparation procedure used for the trained model. Maintaining a consistent preprocessing pipeline helps ensure that images submitted through the Flask application are handled in the same manner as the images used during model development.

1.9. DenseNet121-Based Classification

The core classification component uses DenseNet121 with transfer learning. DenseNet121 is a convolutional neural network architecture in which feature representations are propagated through densely connected layers.

In the implemented system, the trained DenseNet121 model performs four-class classification of chest X-ray images. The model produces an output probability for each of the four categories.

The predicted class is selected using the class having the highest output probability:

$$\text{Predicted Class} = \text{argmax}(P_1, P_2, P_3, P_4)$$

where the four probability values correspond to COVID-19, Normal, Pneumonia, and Tuberculosis.

The model was trained for 10 epochs. The final training accuracy obtained during training was approximately 90.93%, while the validation accuracy was approximately 84.21%.

1.10. Prediction and Confidence Analysis

After the model processes an input image, the system obtains probability values for all four classes. These values provide the model's estimated distribution across the available categories.

The system identifies the category with the highest probability as the predicted class. The probability associated with this selected class is displayed as the **model confidence**.

For example, for a verified sample prediction, the system produced the following probability distribution:

Class	Probability
COVID-19	75.23%
Normal	4.11%
Pneumonia	13.25%
Tuberculosis	7.40%

In this example, COVID-19 has the highest model probability and is therefore displayed as the predicted category.

These probability values represent the model output and should not be interpreted as clinical certainty or a medical diagnosis.

1.11. Preliminary Image Validation

A preliminary validation mechanism is incorporated before performing disease classification. The purpose of this step is to identify obviously unsuitable inputs before they are passed to the trained model.

The implemented validation checks basic image characteristics including:

- Minimum image dimensions
- Image aspect ratio
- Variation between colour channels

Images that do not satisfy the basic validation conditions are rejected and are not passed through the prediction and Grad-CAM workflow.

This mechanism is intended only as a *basic input-screening method*. It is not a medical-grade X-ray detector and cannot guarantee that every accepted image is a valid chest X-ray.

1.12. Grad-CAM-Based Explainability

To improve the interpretability of the classification output, the system incorporates Gradient-weighted Class Activation Mapping (Grad-CAM).

Grad-CAM uses gradient information associated with the selected target class to generate a heatmap representing image regions that contributed to the model's prediction. The heatmap is overlaid on the original chest X-ray to produce a visual explanation of the model output.

In the implemented DenseNet121 model, the Grad-CAM process uses the convolutional feature representation associated with the final DenseNet feature layer. The generated visualization is saved and displayed on the result page together with the original X-ray image.

The Grad-CAM output does not independently determine the disease category. Instead, it provides an interpretability visualization associated with the classification produced by the trained model.

1.13. Flask Based Application

The trained model is integrated into a Flask-based web application to provide a user-accessible interface for performing the analysis.

The web application follows the workflow:

- User opens the X-ray upload page.
- User uploads an image.
- The system performs preliminary image validation.
- A valid image is resized and prepared for prediction.
- DenseNet121 generates the four-class probability distribution.
- The class with the highest probability is selected.
- Model confidence and class-wise probabilities are calculated.
- Grad-CAM visualization is generated.
- The result page displays the original X-ray, predicted category, confidence, probability distribution, and Grad-CAM visualization.

This provides an end-to-end workflow from image upload to classification and visual explanation.

1.14. Overall System Workflow

The overall methodology can therefore be represented as:

Chest X-Ray Dataset → Image Preprocessing → DenseNet121 Transfer Learning → Four-Class Prediction → Probability & Confidence Analysis → Grad-CAM Explainability → Flask Web Interface → Result Display

The methodology integrates the individual components into a single analysis pipeline. Dataset preparation and model training provide the classification model, while the Flask application provides the deployment interface. Preliminary validation is performed before classification, and Grad-CAM is applied after prediction to provide a visual explanation.

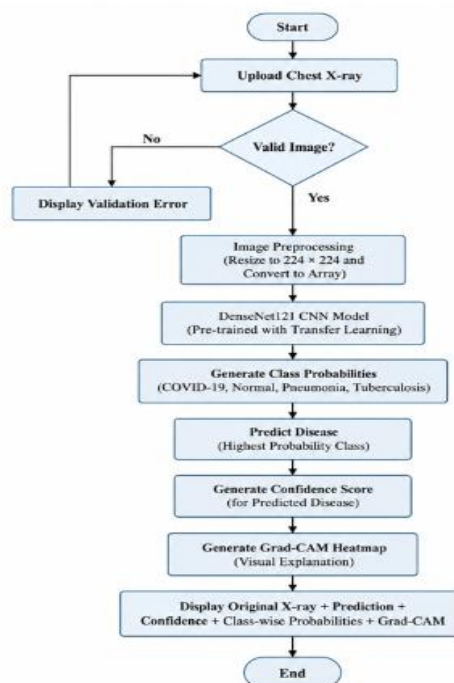


Fig. 1 – Proposed Methodology Workflow

System Architecture and Implementation

1.15. System Architecture

The proposed system follows a sequential architecture in which the uploaded chest X-ray image passes through a series of processing modules before the final result is presented to the user. The architecture integrates image upload, preliminary image validation, image preprocessing, DenseNet121-based classification, probability analysis, Grad-CAM visualization, and result presentation through a Flask web application.

The system begins when the user uploads a chest X-ray image through the web interface. The uploaded image is first passed to the validation module, which performs basic checks on the image before allowing it to enter the prediction pipeline. If the image satisfies the validation conditions, it is resized to 224×224 pixels and converted into the required numerical representation.

The processed image is then passed to the trained DenseNet121 transfer learning model. The model produces probability values for the four target categories: COVID-19, Normal, Pneumonia, and Tuberculosis. The category with the highest probability is selected as the predicted class, and its probability is presented as the model confidence.

After prediction, the Grad-CAM module generates a heatmap based on the selected class. The heatmap is combined with the original X-ray image to provide a visual representation of the regions associated with the model prediction.

Finally, the Flask application renders the results through the web interface. The result page displays the original X-ray, predicted category, confidence percentage, class-wise probabilities, and Grad-CAM visualization. The implementation report describes this integration as the connection of the upload, validation, AI prediction, probability analysis, Grad-CAM, and result presentation modules through Flask.

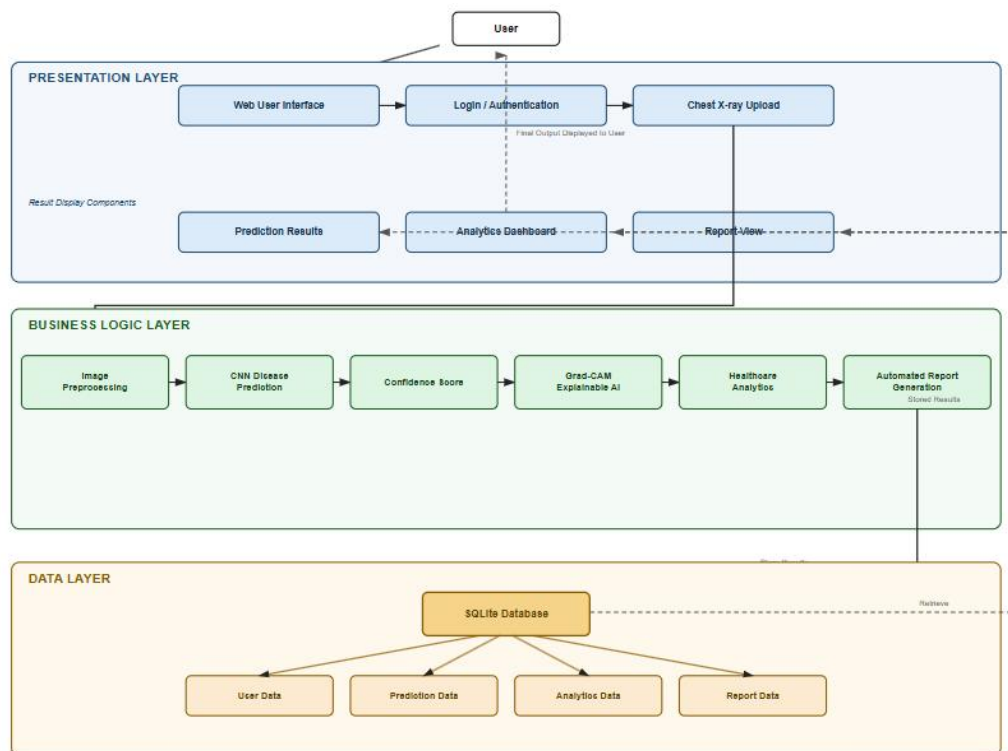


Fig. 2 – System Architecture of the Platform

1.16. Software and Development Environment

The system was developed using Python-based artificial intelligence and web development technologies. Model development and training were performed using Google Colab, while the trained model was integrated and tested locally through a Flask application.

The major software components used in the implementation are shown below.

Table 2 – Software and Development Environment

Software / Technology	Purpose
Python	Main programming language
TensorFlow	Deep learning framework
Keras	Neural network and model handling
DenseNet121	Image classification architecture
Flask	Web application development
NumPy	Numerical and array operations
Pillow	Image processing
Google Colab	Model training and experimentation
Kaggle	Dataset source
PyCharm / VS Code	Source-code development
Web Browser	Application testing

A separate Python virtual environment was used to isolate the project dependencies. The implementation report specifies that the environment contains the packages required for TensorFlow, Keras, Flask, NumPy, Pillow, and other supporting libraries.

1.17. Dataset and Model Implementation

The implemented system uses a chest X-ray dataset containing approximately 7,135 images belonging to four categories:

- COVID-19
- Normal
- Pneumonia
- Tuberculosis

The images were organized into training, validation, and testing directories before model development. The trained classification model uses *DenseNet121 with transfer learning* and is configured for four-class classification.

The model was trained for 10 epochs and achieved approximately 90.93% training accuracy and 84.21% validation accuracy. The resulting model was saved in Keras format as: CNN_MultiDisease_XRay_4Class.keras. The trained model was subsequently transferred to the local Flask application for inference rather than retraining the model whenever a new image was uploaded.

1.18. DenseNet121 Classification

DenseNet121 is used as the core deep learning architecture for the classification task. The implementation uses transfer learning so that the model can utilize learned visual feature representations instead of developing a complete convolutional network from the beginning.

For every valid input image, the model produces four output probabilities corresponding to the four classes. The prediction module identifies the class with the highest probability.

The model therefore performs the following classification:

Input: Chest X-ray image

Output: Four class probabilities

$$P=[P_{\text{COVID}}, P_{\text{Normal}}, P_{\text{Pneumonia}}, P_{\text{TB}}]$$

The predicted class is determined by:

$$\text{Class}_{\text{predicted}} = \text{argmax}(P)$$

The probability corresponding to the selected class is displayed as the model confidence.

1.19. Image Validation and Preprocessing

Before prediction, the uploaded image passes through a preliminary validation stage. The implementation checks characteristics including image dimensions, aspect ratio, and basic RGB-channel variation. The purpose is to reject obviously unsuitable images before they reach the trained classification model.

For a valid image, the preprocessing pipeline performs the following operations:

- Load the uploaded image.
- Check the image dimensions and basic characteristics.

- Resize the image to 224×224 pixels.
- Convert the image into a numerical array.
- Prepare the input tensor for the trained model.
- Pass the processed image to DenseNet121.

The validation mechanism is intentionally described as *basic or preliminary validation*. It does not constitute a medical-grade system for determining whether an image is a chest X-ray.

1.20. Grad-CAM Implementation

Grad-CAM is integrated into the system to provide an interpretable visualization of the model prediction. After the DenseNet121 model generates the prediction, the selected class is used as the target for the Grad-CAM calculation.

The implementation uses the DenseNet121 convolutional layer conv5_block16_concat for generating the activation information required for the visualization. The generated activation map is converted into a heatmap and overlaid on the original chest X-ray.

The resulting visualization highlights regions associated with the model's prediction. The Grad-CAM output is therefore used as an *explanation of the model prediction*, rather than as an independent disease-detection mechanism.

Results and Evaluation

The implemented system was evaluated based on its ability to classify chest X-ray images into four categories, generate class-wise prediction probabilities, calculate model confidence, produce Grad-CAM visualizations, and complete the integrated prediction workflow through the Flask web application. Functional testing was also performed to verify the individual modules and their integration.

1.21. Model Training and Validation Results

The DenseNet121-based transfer learning model was trained using the prepared chest X-ray dataset containing approximately 7,135 images across four classes: COVID-19, Normal, Pneumonia, and Tuberculosis.

The model was trained for 10 epochs. The final observed training accuracy was approximately 90.93%, while the validation accuracy reached approximately 84.21%.

Table 3 – Model Training Results

Parameter	Result
Model Architecture	DenseNet121
Learning Approach	Transfer Learning
Number of Classes	4
Dataset Size	7,135 images
Training Epochs	10
Training Accuracy	90.93%
Validation Accuracy	84.21%
Input Image Size	224×224 pixels

The difference between training and validation accuracy indicates that the model performed differently on the training and validation data. Therefore, these values are reported separately rather than treating the training accuracy as the overall model accuracy.

1.22. Multi-Disease Prediction Results

The developed system successfully generates predictions for the four target categories:

- COVID-19
- Normal
- Pneumonia
- Tuberculosis

For every valid input image, the model generates a probability value for each class. The category having the highest probability is selected as the predicted class.

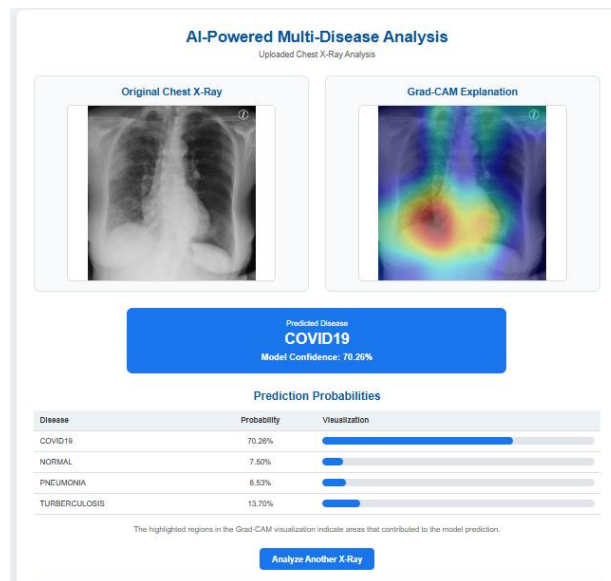
For example, a verified COVID-19 sample produced the following output:

Table 4 – Sample Prediction Result

Class	Probability
COVID-19	75.23%
Normal	4.11%
Pneumonia	13.25%
Tuberculosis	7.40%
Predicted Class	COVID-19
Model Confidence	75.23%

The result demonstrates that the system does not display only the selected class. It also provides the probability distribution across all four categories, allowing the complete model output to be viewed.

The reported confidence represents the model's output probability for the selected class and should not be interpreted as clinical certainty.

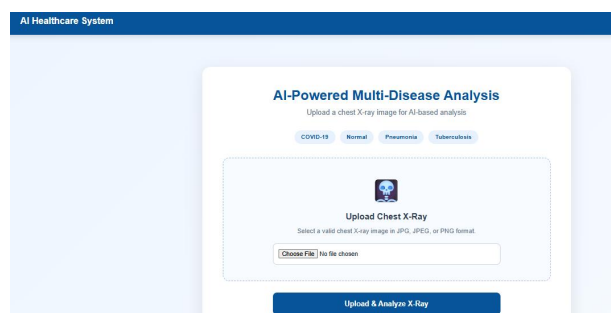
**Fig. 3 – Multi-Disease Prediction Result**

1.23. Image Validation Results

The system includes a preliminary image-validation stage before disease prediction. The validation mechanism checks basic characteristics of the uploaded image, including image dimensions, aspect ratio, and basic RGB-channel variation.

During functional testing, valid supported image files were accepted for processing, while obviously unsuitable inputs were rejected before the prediction stage.

This validation mechanism provides an additional input-screening step and prevents obviously inappropriate images from being directly passed to the trained model. However, it is a *basic heuristic validation mechanism* and should not be considered a medical-grade method for determining whether an image is a chest X-ray.

**Fig. 4 – Image Validation and Upload Interface**

1.24. Grad-CAM Explainability Results

Grad-CAM was successfully integrated with the DenseNet121 classification model to provide a visual explanation associated with the

predicted class.

After the model generates a prediction, Grad-CAM produces a heatmap representing image regions associated with the selected prediction. The generated visualization is displayed together with the original chest X-ray on the result page.

The implementation uses the DenseNet121 convolutional layer conv5_block16_concat for generating the Grad-CAM visualization.

The Grad-CAM output provides additional information beyond the numerical prediction by allowing the user to visually inspect the regions highlighted by the model.

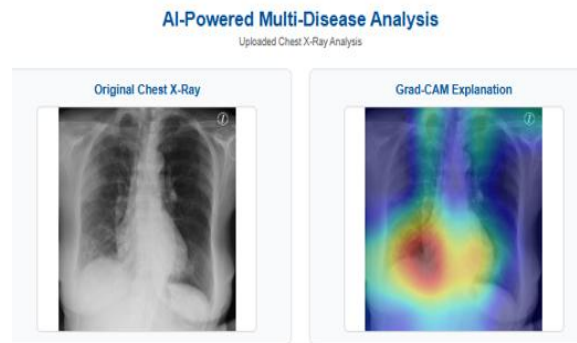


Fig. 5 – Original Chest X-Ray and Grad-CAM Visualization

Discussion

The implemented system demonstrates an integrated approach to multi-disease chest X-ray image classification by combining deep learning-based prediction with explainable visualization. Unlike a workflow that produces only a disease label, the developed prototype provides the predicted category, class-wise probability values, model confidence, and a Grad-CAM visualization within the same application.

The use of DenseNet121 with transfer learning provides a practical approach for extracting image features and performing classification across four categories: COVID-19, Normal, Pneumonia, and Tuberculosis. Previous research has also investigated DenseNet121 and transfer-learning strategies for chest X-ray classification, supporting the suitability of transfer learning as a practical approach for this type of image-analysis task.

The model achieved approximately 90.93% training accuracy and 84.21% validation accuracy after 10 epochs in the implemented experiment. These values indicate the observed performance of the developed model on the training and validation data used during development. They should not be interpreted as a clinical diagnostic accuracy or as directly comparable with results from studies using different datasets, data splits, preprocessing methods, and evaluation protocols.

The inclusion of class-wise probability values provides additional information beyond the selected prediction. Instead of displaying only the highest-probability category, the system presents the model's estimated probabilities for all four classes. This allows the output distribution to be examined for each individual image.

Explainability is another important component of the developed system. Grad-CAM provides a heatmap associated with the selected prediction, allowing the user to visually inspect regions that contributed to the model output. Previous chest X-ray research has similarly investigated Grad-CAM and other XAI techniques for explaining predictions involving COVID-19, pneumonia, and tuberculosis.

The preliminary image-validation mechanism adds an input-screening step before classification. It checks basic image characteristics such as dimensions, aspect ratio, and RGB-channel variation. However, this mechanism is heuristic rather than medical-grade and therefore cannot guarantee that every accepted image is a valid diagnostic chest X-ray.

The Flask-based implementation demonstrates how the trained model can be integrated into an accessible web interface. The resulting workflow connects image upload, validation, preprocessing, classification, probability calculation, Grad-CAM generation, and result presentation. This provides a complete research prototype rather than an isolated machine-learning model.

Overall, the developed system demonstrates the feasibility of combining multi-class chest X-ray classification, probability-based prediction, preliminary image validation, explainable visualization, and web-based deployment within one application. The system should nevertheless be considered a research prototype, and its outputs should complement rather than replace professional medical assessment.

Limitations and Future Work

Although the implemented system provides an integrated chest X-ray classification and explainability workflow, several limitations remain.

The first limitation is related to the dataset. The model was developed using a specific publicly available dataset containing approximately 7,135 images distributed across four categories. Therefore, the observed model performance may depend on the characteristics, class distribution, image quality, and labeling of this dataset. Performance on other datasets or real-world clinical images may differ.

The second limitation concerns the evaluation results. The current implementation reports training and validation accuracy, while the verified sample prediction is used to demonstrate the deployed workflow. A more comprehensive evaluation would require additional metrics such as precision, recall, F1-score, specificity, sensitivity, confusion matrix analysis, and class-wise performance on an independently maintained test set.

The third limitation is the basic image-validation mechanism. The current validation checks image dimensions, aspect ratio, and basic colour-channel characteristics. It does not use a dedicated medical-image classifier to determine whether an uploaded image is actually a chest X-ray. Therefore, it should be regarded only as a preliminary screening mechanism.

Another limitation is related to model interpretability. Although Grad-CAM provides a useful visual representation of regions associated with a prediction, the heatmap should not be interpreted as a definitive identification of a disease lesion. It represents information derived from the model's internal feature representations.

The system is also currently deployed as a *local Flask prototype*. It has not been evaluated as a production hospital system or integrated with clinical information systems. Consequently, issues such as large-scale deployment, security, patient-data privacy, authentication, and interoperability with medical information systems are outside the scope of the current implementation.

Future development can focus on expanding the dataset with larger and more diverse chest X-ray collections from different sources and populations. This could provide a broader basis for evaluating model generalization.

The classification model can also be evaluated using additional performance measures, including precision, recall, F1-score, sensitivity, specificity, confusion matrices, and class-wise ROC-AUC where appropriate.

A dedicated *X-ray-versus-non-X-ray validation model* could replace the current heuristic validation mechanism. Such a classifier could provide a more systematic approach to screening uploaded images before disease classification.

Further work can investigate alternative transfer-learning architectures and compare their performance under the same experimental conditions. Ensemble models and carefully controlled fine-tuning strategies could also be explored.

The explainability component can be extended by comparing Grad-CAM with other XAI techniques such as *LIME* and *SHAP*. Previous research has already explored combinations of these methods for chest X-ray analysis.

The current prototype can also be extended with the *healthcare analytics dashboard*, *persistent database storage*, *automated report generation*, and *additional analysis features* originally planned for the broader project scope. These extensions should be implemented and separately evaluated before being described as completed system functionality.

Finally, future versions could investigate secure deployment, user authentication, privacy protection, larger-scale testing, and controlled evaluation using clinically curated datasets.

Conclusion

This paper presented an AI-Powered Multi-Disease Chest X-Ray Analysis System that combines deep learning-based classification with explainable artificial intelligence in a web-based prototype. The implemented system classifies chest X-ray images into four categories: COVID-19, Normal, Pneumonia, and Tuberculosis.

A DenseNet121-based transfer learning model was developed and trained using approximately 7,135 chest X-ray images. After 10 epochs, the model achieved approximately 90.93% training accuracy and 84.21% validation accuracy. The trained model was subsequently integrated into a Flask application for local inference.

The developed workflow includes preliminary image validation, image preprocessing, four-class prediction, class-wise probability calculation, model confidence display, and Grad-CAM-based visualization. The integration of these components allows the user to view both the numerical prediction and an associated visual explanation through a single interface.

The Grad-CAM component adds an interpretability layer to the classification workflow by producing a heatmap associated with the selected prediction. This is particularly relevant to medical-image research because classification results can otherwise be difficult to interpret from numerical outputs alone. Related research has also demonstrated the use of Grad-CAM and other XAI methods for multi-disease chest X-ray classification.

The functional evaluation further demonstrated that the implemented modules operate together as an end-to-end application. The system successfully connects image upload, validation, preprocessing, DenseNet121 inference, probability analysis, Grad-CAM generation, and result presentation.

The main contribution of the work is therefore the integration of multi-disease chest X-ray classification and explainable AI within a lightweight Flask-based research prototype. However, the system is not intended to replace professional medical interpretation or serve as an autonomous clinical diagnostic system. Further evaluation using larger and more diverse datasets, additional performance metrics, stronger image validation, and controlled clinical datasets would be required before considering broader real-world application.

Acknowledgements

The author would like to express sincere gratitude to the management and faculty of Aurora Higher Education and Research Academy, Hyderabad, for providing the academic environment, facilities, and resources required to carry out this project.

The author would like to express special gratitude to Mrs. T. Malathi, Department of Computer Applications, for her valuable guidance, constructive suggestions, continuous encouragement, and support throughout the development of the project and preparation of this research paper.

The author also thanks the Department of Computer Applications for providing the technical and academic support necessary for completing the work. Finally, the author expresses sincere gratitude to family members, friends, and everyone who provided support and encouragement directly or indirectly during the development and completion of this project.

REFERENCES

- Wang, L., Lin, Z. Q., & Wong, A. (2020). COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, 10, 19549. <https://doi.org/10.1038/s41598-020-76550-z>
- Hemdan, E. E.-D., Shouman, M. A., & Karar, M. E. (2020). COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images. *arXiv preprint arXiv:2003.11055*.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4700–4708.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- 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. <https://doi.org/10.1016/j.combiomed.2022.106156>
- Yao, J., Wang, X., Li, S., et al. (2022). Deep transfer learning for the multilabel classification of chest X-ray images. *Diagnostics*, 12(6), 1457. <https://doi.org/10.3390/diagnostics12061457>
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., et al. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., et al. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5), 1122–1131.e9.
- Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Illcus, S., Chute, C., et al. (2019). CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1), 590–597.