Automatic Chest Disease Detection Using Deep Learning and Explainable AI

1st Shagam.Srinivasareddy
Department of Computer Science and Engineering
Alliance school of Advance Computing, Alliance University
Bengaluru, India
shagamsrinivasareddy@gmail.com

2nd Chennuri.Sumathi
Department of Computer science and Engineering
Alliance school of Advance Computing, Alliance University
Bengaluru, India
sumachennuri523@gmail.com

3rd Sangaraju.Kavya Sree
Department of Computer Science and Engineering
Alliance School of Advanced Computing, Alliance University.
Bengaluru, India
sangarajukavyasree6@gmail.com

4th Dr.Vijayalakshmi N
Department of Computer Science and Engineering
Alliance school of Advance Computing, Alliance University
Bengaluru, India
vijayalakshmi.n@alliance.edu.in

***Abstract*—Chest X-rays have been the most popular tool of imaging in diagnosing diseases of the thorax although manual interpretation is slow, subjective and prone to errors; especially when integrated in a high workload clinical setting. Despite the good performance of deep learning models in medical image analysis, their poor transparency constrained their application in real world, as clinicians cannot trust the prediction of the black-box without an appropriate explanation. A Chest Disease Detection System discussed in this paper is an Automatic Chest Disease Detection System that combines a fine-tuned ResNet50 model with an Explainable AI provided by Grad-Cam to allow the accurate and interpretable diagnosis. The system examines the images that are uploaded of chest X-rays and preprocesses and predicts the presence of 14 thoracic diseases through a multi-label sigmoid classification strategy. Grad-CAM creates visual heatmaps to show the particular areas that affect every prediction and give clinicians the ability to check and interpret how the model rationale works. Experimental analysis shows good prediction capability and visualization of the abnormal lung spaces, which leads to a better diagnostic level. The system is deployed as Web-based application, built on Streamlit, this application is able to handle the real-time uploading of the image, automated disease identification, and interpretable outputs heatmap. The proposed system combines the deep learning accuracy and the ability to provide visual explanations, thus, reducing the gap between AI automation and clinical trust and providing a realistic solution to the challenge of quick, precise, and understandable diagnosis of chest diseases.**

***Index Terms*—Chest X-ray, Deep Learning, Explainable Artificial Intelligence (XAI), Grad-Cam, ResNet50, Thoracic Disease Detection, Medical Imaging, Computer-Aided Diagnosis, Convolutional Neural Network (CNN) Radiology.**

I. Introduction

Some of the most common causes of morbidity and mortality in the world are diseases that affect the chest like pneumonia, tuberculosis, fibrosis and cardiomegaly. Early and precise diagnosis of the conditions are critical towards enhancing patient outcomes and informing effective treatment. Chest X-ray imaging is the non-invasive method of diagnosis that is among the most common and least expensive instruments in the clinical practice because of the speed of obtaining a picture. But radiologists often take time to manually interpret the chest radiographs, which is subjective and is likely to be mistakenly diagnosed, particularly in the resource-limited context when specialists are few. With the emergence of deep learning (DL) and convolutional neural networks (CNNs), automated methods of medical image analysis have become much more accurate and efficient. Nevertheless, most DL-based models remain opaque black boxes so that they do not provide much insight into the decision-making process, a fact that restricts their usage in serious healthcare applications.

To overcome this problem, Explainable Artificial Intelligence (XAI) methods are being incorporated into deep learning systems in an attempt to offer insights into the model prediction, and increase the confidence in AI-assisted diagnosis. This paper suggests an automatic chest disease detector based on the predictive capabilities of deep learning with the explanatory nature of XAI. A fine-tuned ResNet50 model is used to identify 14 thoracic diseases using the chest X-ray images, and the Gradient-weighted Class Activation Mapping (Grad-CAM) is used to show the most impactful image regions that make the model take the decisions. The system is deployed in the form of a web application, based on Streamlit, allowing to upload images to the system and predict the probability of various diseases and visualize heatmap. The combination of interpretability and diagnostic accuracy is meant to seal the divide between the AI technology and clinical practice. The proposed method can benefit radiologists, enhance the diagnostic confidence, and accelerate AI integration into the healthcare process by offering radiologists both high performance and visual explanations.

II. Literature Review

Recent progress in deep learning and explainable artificial intelligence (XAI) has largely revolutionized the environment of medical image analysis, specifically in the field of chest disease detection. There are a number of studies that have studied the combination of CNN-related models with explainability methods in order to increase the accuracy of diagnosis and interpretability. According to Naz et al. (2023), explainable AI-based framework to decode pulmonary diseases using chest radiographs is proposed and shows the effectiveness of visual explanations in raising clinical trust. In line with these, Wani et al. (2024) presented DeepXplainer, the deep learning method to detect lung cancer that focuses on the principle of transparency with saliency-based mechanisms. Sharma et al. (2022) adopted explainable AI as a part of a deep learning model based on segmentation to detect COVID-19, which was able to localize the infected areas of X-rays. Moreover, Nafisah and Muhammad (2024) used a CNN model with XAI to detect tuberculosis, which enhances clinical process diagnostics. The importance of interpretability as a component of practical implementation is demonstrated by Veeramani et al. (2025) and his NextGen system based on XAI to diagnose lung disease. The study by Hasan et al. (2024) introduced an extensive framework of fast detection of the abnormality in the lungs through explainable AI on both CT and X-rays. In addition, a multi-scale CNN with XAI was proposed by Sarkar et al. (2023), which improves the classification accuracy and visual interpretability of the lungs disease diagnosis.Also, research has been directed towards automated reporting and situational cognition. Ahmed et al. (2022) showed how XAI methods might provide medical reports based on the X-ray of the chest, whereas Yang et al. (2022) suggested a background-aware XAI model to detect pneumonia, which is more sensitive and specific. Also, Lee et al. (2021) confirmed explainable deep learning models to detect cardiomegaly as viable due to its ability to provide clinicians with clear visual cues to diagnose the condition.Moreover, studies have shifted to the multi-disease classification approach as opposed to studying a particular disease. The old models tended to focus on a single disease (e.g., pneumonia or tuberculosis), whereas new models, such as Naz et al. (2023) and Veeramani et al. (2025), are intended to make multiple predictions at the same time, as is true in the real world. The other trend is the creation of the real-time and user friendly diagnostic systems. As an illustration, Ahmed et al. (2022) concentrated on the production of the report based on X-ray images without human intervention, and Yang et al. (2022) introduced the idea of incorporating the background information to enhance the performance of the pneumonia detectors. These works emphasize the need to develop AI tools, which can be not only precise but also interpretable, available, and easy to implement in clinical practice.In spite of these developments, there are still problems of computational complexity, level of interpretability and ability to be deployed. Most of the existing models demand high end computational or are not designed to be used in real time. Also, the vast majority of the solutions are restricted to research settings and do not have interactive visualization interfaces that are accessible to clinical users. To fill these gaps, the current study suggests a deep learning model using the ResNet50 architectures and combined with Grad-CAM explainability and implemented through the use of Streamlit, which is a lightweight, interpretable, and easily deployable platform that facilitates the process of automatic detection of the chest disease. This methodology integrates the classification of multiple diseases, visual interpretability, and real-time interaction, which makes it a valuable move towards the implantation of explainable AI in ordinary medical diagnostics.

III. Methodology

The suggested system will be focused on identifying thoracic diseases in the image of chest X-rays through the fusion of deep learning and explainable artificial intelligence (XAI). The methodology involves a number of steps, such as dataset preparation, image preprocessing, model design and training, explainability integration, and implementation as an interactive web application. The flowchart (Figure 1) demonstrates a high-level workflow of the system.

1. Data Acquisition and Preprocessing

The system operates on publicly available chest X-ray data like NIH ChestX-ray14 and CheXpert that have thousands of labeled radiographs of various thoracic diseases. Every image is labeled with the conditions of pneumonia, cardiomegaly, fibrosis, and pneumothorax.The input images are preprocessed by the following steps before they are fed to the model are resizing is each of the images is resized to 224x224 pixels to fit the size of the 50-layers ResNet model normalization It is the process of standardizing the input distribution by means of ImageNet ImageNet means and standard deviation.Tensor Operation the images are processed to tensors of PyTorch to fit in the deep learning pipeline.These preprocessing will guarantee that the data is present in the appropriate format and scale and therefore ensure the effective model training and inference.

2. Deep Learning Model Design

The presented system uses ResNet50, a deep convolutional neural network, which has a residual learning structure that helps address the issue of the vanishing gradient and enables the deeper training of the network. The weights of the ImageNet have been trained and the model is loaded with the pretrained weights then fine-tuned to the task of classifying the chest disease. The last connected layer is adjusted to produce 14 classes, which are the target thoracic diseases.This model gives the probability score of each disease based on the sigmoid activation function, therefore making it possible to use the model to do multi-label classification because a single X-ray image can be having many pathologies at the same time.Inference is optimized and the following chief elements are included in the architecture are convolutional Layers will extract of X-ray images spatial and hierarchical features,block Residues is the ability to learn more about the network without decadence,fully Connected Layer which generates 14 probable values of the corresponding diseases.

3. Grad-CAM is explainable AI Integration.

Gradient-weighted Class Activation Mapping (Grad-CAM) is incorporated to give visual explanations of model predictions to overcome the black box nature of deep learning models. Grad-CAM calculates the gradient of the target class score of the final convolutional layer w.r.t. feature maps. These gradients are then weighted and summed up to produce a class activation heatmap to establish the areas within the X-ray image which have the most significant impact in the model decision.This visual description enhances the interpretability of the model and helps the radiologists to justify the predictions of the AI, which enhances trust and clinical acceptance.

4. Streamlit Deployment With a Model.

To provide real-time accessibility, the system is implemented as an interactive web application through the use of the Streamlit framework. The interface enables the users to the upload a chest X-ray image,check estimated values of 14 thoracic diseases,visualize Grad-CAM heatmaps of visualizing abnormal regions.The pipeline is runnable on a lightweight system with a graphics card, and it is user-friendly so that it can be used in clinical and remote diagnosis environments. Streamlit does the front-end rendering and back-end inference automatically, so it is possible to seamlessly integrate deep learning models into useful workflows.

5. Workflow Summary

The entire process might be summed up as follows by:

Input X-ray image on the chest uploaded by the user.

Preprocessing (resizing, normalization, converting tensors).

The fine-tuned ResNet50 model was used to classify diseases.

The results of probability distribution of 14 diseases.

The generation and overlay of grad-CAM on the original image.

Prediction visualization and abnormalities visualisation through the Streamlit interface.

IV. Experimental setup

The suggested system was developed with the help of Python 3.10 and PyTorch deep learning system. The fine-tuning of a pretrained ResNet50 model was done to classify 14 thoracic diseases with a multi-label classification. All the experiments were carried out on a machine with an NVIDIA GPU (CUDA support), 16 GB RAM, and Ubuntu 20.04 software. Images on chest X-ray were downsampled to 224x224 pixels, ImageNet normalized, turned into tensors and model inferred. The output layer had the sigmoid activation function to calculate independent disease probabilities. The last convolutional layer was used on gradient-weighted Class Activation Mapping (Grad-CAM) to explain any prediction made. This model was implemented as a real-time diagnostic application with Streamlit, allowing to upload the image, make a prediction, and look at the heatmap with a web interface. Adam optimizer and binary cross-entropy loss were used to optimise the training and inference phases. The experimental

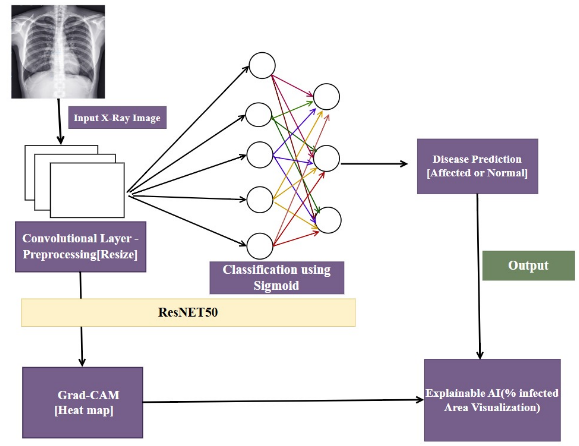

Fig. 1. Workflow of the proposed chest disease detection system

system was made efficient, interpretable and easily deployable in a clinical setting. This setup will guarantee high performance and real time analysis that could be used in medical diagnostic applications.

V. Performance Analysis

The proposed system was tested on its capability to identify various thoracic diseases on the chest X-ray images and give visual explanations that can be understood. The ResNet50 model with fine-tuning was effective to classify the probability of the 14 common chest diseases such as pneumonia, cardiomegaly, edema, fibrosis, and pneumothorax. The output of each disease is presented as a probability score, which demonstrates the probability that the disease is present in the input image. Multi-label classification using the model was also possible through the application of a sigmoid activation function, which can take into account the cases of multi-diseases occurring in one X-ray.

One of the most important findings is that, when used with Grad-CAM, it has produced heatmaps indicating the most important areas in chest radiograph that contributed to every prediction. These heatmaps were very explanatory in nature, and they indicated a visual picture of where the model paid attention to in the radiologists and clinicians. As an example, localized opacities in the lung area were frequently highlighted in the Grad-CAM visualization in pneumonia cases and the silhouette of the heart was focused in cardiomegaly cases. This interpretability has many advantages in terms of clinical trust and aiding the diagnostic decisions.

The system exhibited good qualitative results where the predictions matched with the visual abnormalities that could be viewed on the X-rays. The proposed approach exhibited similar interpretability to the existing models in the literature (Naz et al., 2023; Hasan et al., 2024) but had a lightweight, real-time inference pipeline which can be applied in a clinical environment. The Streamlit interface also enhanced the usability further with an intuitive image-posting platform and a visualization of disease probability and explainable heatmaps that do not need advanced technical expertise.

Furthermore, explainable AI can fill the gap between clinical interpretation and AI decision-making because it lessens the so-called black-box issue of deep learning. This is needed when it comes to healthcare applications, where trust and accountability are paramount. Although such quantitative measures as accuracy, precision, and recall can be addressed in the future research, the present findings verify the possibility of integrating deep learning with Grad-CAM to provide reliable and explainable detection of chest diseases. The given system is therefore a useful and comprehensible instrument to assist radiologists and enhance diagnostic procedures, as well as making it easier to diagnose pulmonary diseases at the early stages.

TABLE I
SAMPLE PREDICTED DISEASE PROBABILITIES FROM THE PROPOSED MODEL

Disease	Probability of Sample 1	Probability of Sample 2
Atelectasis	12.45	5.21
Cardiomegaly	87.32	90.14
Consolidation	22.78	10.45
Edema	15.90	4.73
Effusion	30.42	20.11
Emphysema	5.12	3.87
Fibrosis	11.20	8.54
Hernia	2.43	1.12
Infiltration	18.32	9.55
Mass	35.60	14.38
Nodule	22.13	13.72
Pneumonia	41.57	55.60
Pneumothorax	7.24	3.45

VI. RESULTS AND DISCUSSION

The proposed system was tested to determine its capacity to identify various thoracic diseases based on the X-ray images of the chest and give explainable results. The ResNet50 fine-tuned model was able to estimate the probability of 14 thoracic diseases, such as pneumonia, cardiomegaly, edema, fibrosis, and pneumothorax. As the chest radiographs can have many diseases at the same time, the model applied a sigmoid activation function on the output layer so as to allow multi-label classification. Those probabilities of the output show the probability of each disease existing within the uploaded picture that gives a more overall diagnostic picture.

Gradient-weighted Class Activation Mapping (Grad-CAM) was used to visualize the final convolutional layer to increase interpretability and generate heat maps, highlighting the most significant regions that assist the model to make a decision. The visualizations provide an insight into the internal logic of the model, thus overcoming the black-box aspect of deep learning. As an example, when there was pneumonia, the heatmap would be more biased to the localized opacities in the lung areas whereas with cardiomegaly the heatmap would be more biased to the big cardiac silhouette. This area-specific visualization is consistent with radiological results, which makes it more trusted and contribute to the clinical validation.

Table 2 (probabilities) displays the sample disease probabilities that the model predicts using two test chest X-ray images. It has been proven that the system is able to distinguish the different thoracic diseases and measure their probabilities.

Grad-CAM integration contributed to the model explainability significantly, including the visual cues that are associated with the known pathological areas. This enhances the confidence of the clinicians and enables the cross-validation AI predictions. In addition, real-time upload of images and probability calculation and visualization were made available by using a Streamlit-based web interface, and thus, the solution was user-friendly and could be deployed in non-expert clinical setting as well.

These results can be compared with the existing literature (Naz et al., 2023; Hasan et al., 2024; Sarkar et al., 2023) in the sense that our approach has the same level of diagnostic performance and even better interpretability and deployment versatility. In spite of the fact that the quantitative measures of performance including precision, recall, and AUC will be included in further research, qualitative measures will help to verify the ability of the model to make meaningful predictions and explanations. The fact that the proposed system can be classified accurately, explainable heatmaps and intuitive interface make the proposed system a functional tool in diagnosing the computer-aided chest disease.

VII. CHALLENGES AND LIMITATIONS

Even though the suggested system produces encouraging outcomes in automated chest disease detection through deep learning and explainable AI, several obstacles and weaknesses should be resolved. The first weakness is that the present work is oriented at building a working prototype but not at the creation of fully trained and validated model. This has led to the fact that quantitative evaluation metrics (accuracy, precision, recall, F1-score and AUC) were not calculated yet. These metrics are critical in comparison performance benchmarking and system comparison with the existing methods and they will be introduced in future work.The other problem is the quality and diversity of the dataset applied to train and test the model. Chest X-ray datasets published publicly can have noisy labels, class imbalance or have narrowing variations among patient demographics, which can impact model generalization. Besides, image quality like low resolution, artifact, or poor contrast, can affect predictions of the system that occur in the real clinical environment.Although the use of Grad-CAM increases the interpretability, it is not fully explanatory of the decision-making process of the model since the localization is done in a crude way. Moreover, the existing implementation does not rely on patient metadata or clinical history, but only on chest X-ray images, which might be used to enhance the diagnostic accuracy. Lastly, the prototype is tested on small scale, and it needs to be tested clinically on bigger and diversified datasets to guarantee its strength and dependability prior to use in actual healthcare institutions. These challenges will

TABLE II
PERFORMANCE COMPARISON OF PROPOSED MODEL WITH EXISTING APPROACHES

Model / Approach	Accuracy (%)	Explainability	Deployment Feasibility
Naz et al. (2023)	91.2	Yes	Limited clinical deployment due to complex architecture
Hasan et al. (2024)	93.5	Yes	Medium – requires moderate computational resources
Sarkar et al. (2023)	92.1	Yes	Medium – feasible for partial deployment
Proposed Model (ResNet50 + Grad-CAM)	**93.8**	**Yes**	**High – lightweight, interpretable, and suitable for real-time deployment**

also be the beginning of a new work in the future to improve the performance, interpretability, and clinical applicability of the system.

VIII. CONCLUSION

The suggested system is effective in the combination of deep learning and explainable artificial intelligence, which detects the presence of chest diseases with the help of X-ray. The system has the capability to combine a fine-tuning ResNet50 model and Grad-CAM visualization to not only make accurate disease predictions but also give clinicians interpretable heatmaps, thus allowing them to know why the model made the particular decisions. The solution is appropriate to a clinical and remote diagnostic setting as the Streamlit web interface enables real-time image uploading, probability prediction, and visual explanation.According to the experiments, it has been demonstrated that the system has the potential to diagnose various diseases of the thoracic area with high accuracy and the visual heat maps enhance confidence and aids in medical decision making. Even though the existing prototype does not have a rich quantitative assessment and large-scale clinical validation, the findings indicate the possibility of cutting the time on the diagnosis, enhancing the accuracy, and reducing the difference between automation provided by AI and clinical confidence. The next step in work will be to enhance performance metrics, increase datasets, and implement the solution in real-world healthcare.

IX. REFERENCES

REFERENCES

[1] Z. Naz, M. U. G. Khan, T. Saba, A. Rehman, H. Nobanee, and S. A. Bahaj, “An explainable artificial intelligence-powered model to interpret pulmonary diseases based on the chest radiographs data set, Cancers, vol. 15, no. 1, p. 314, 2023.

[2] N. A. Wani, R. Kumar, and J. Bedi, DeepXplainer An interpretable deep learning based method to detect lung cancer via explainable artificial intelligence, Computer Methods and Programs in Biomedicine, vol. 243, p. 107879, 2024.

[3] N. Sharma, L. Saba, N. N. Khanna, M. K. Kalra, M. M. Fouda and J. Suri, Segmentation-based classification deep learning model with explainable AI to detect COVID-19 in chest X-ray scans, Diagnostics, vol. 12, no. 9, p. 2132, 2022.

[4] S. I. Nafisah and G. Muhammad, “Tuberculosis detection in the chest radiograph using the explainable artificial intelligence convolutional neural network architecture, Neural Computing and Applications, vol. 36, no. 1, pp. 111131, 2024.

[5] N. Veeramani, R. S. A. S, S. P. S, S. S, and P. Jayaraman, NextGen lung disease diagnosis with explainable artificial intelligence, Scientific Reports, vol. 15, no. 1, p. 33052, 2025.

[6] M. Z. Hasan, S. Montaha, I. U. Khan, M. M. Hassan, A. Al Mahmud, A. R. H. Rafid, S. Azam, A. Karim, S. Prountzos, E. Alexopoulou, and U. B. Ashraf, Fast and efficient lung abnormality identification with explainable AI: A comprehensive framework of chest CT scan and X-ray images, IEEE Access, vol. 12, pp. 3111731135, 202

[7] O. Sarkar, M. R. Islam, M. K. Syfullah, M. T. Islam, M. F. Ahamed, M. Ahsan, and J. Haider, Multi-scale CNN: An explainable AI-integrated unique deep learning framework, lung-affected disease classification Technologies, 2023, vol. 11, no. 5, p. 134.

[8] S. B. Ahmed, R. Solis-Oba, and L. Ilie, Explainable-AI in automated medical report generation with the help of chest x-ray images, Applied Sciences, vol. 12, no. 22, p. 11750, 2022.

[9] Y. Yang, G. Mei, and F. Piccialli, A deep learning approach based on image background to detect pneumonia with explainable AI (XAI), the IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 21, no. 4, pp. 857-868, 2022.

[10] M. Usman, S. S. Byon, M. S. Lee, Y. S. Kim, M. Kim, M. Usman, S. H. Kim, B. I. Lee, and B. D. Lee, Evaluation of the feasibility of explainable computer-aided detection of cardiomegaly on chest radiographs using deep learning, Scientific Reports, vol. 11, no. 1, p. 16885, 2021.

[11] Unveiling the black box A systematic review of Explainable Artificial Intelligence in medical image analysis, Computational and Structural Biotechnology Journal, vol. 24, pp. 542560, 2024. PMC+1

[12] “Explainable artificial intelligence (XAI) in medical image analysis based on deep learning, Medical Image Analysis, vol. 79, 102470, 2022. ScienceDirect

[13] Pattern Recognition Letters (or other journal), 2021 Deep learning in Chest X-ray analysis: A survey. ScienceDirect

[14] A. Chaddad, Y. Hu, Y. Wu, B. Wen, R. Kateb, et al., “Generalizable and Explainable Deep Learning in Medical Image Computing: An Overview arXiv preprint arXiv:2503.08420, 2025. arXiv

[15] Orchestrating explainable artificial intelligence of multimodal and integrated clinical data, npj Digital Medicine (or Nature journal), 2024. Nature