International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

ClinXAI: An Explainable AI-Powered Dual Medical Diagnosis System

Y. Dasaratha Rami Reddy[1], S.D. Soha Tabassum[2], L. Naga Aparna[3], M. Susmitha[4], R. Ambika [5], V. Reddy Lalitha [6]

[1]Assoc.Prof, Department of Computer Science and Engineering, CBIT, Proddatur, YSR, A.P
[2]UG Student, Department of Computer Science and Engineering(DS), CBIT, Proddatur, YSR, A.P
[3]UG Student, Department of Computer Science and Engineering(DS), CBIT, Proddatur, YSR, A.P
[4]UG Student, Department of Computer Science and Engineering(DS), CBIT, Proddatur, YSR, A.P
[5]UG Student, Department of Computer Science and Engineering(DS), CBIT, Proddatur, YSR, A.P
[6]UG Student, Department of Computer Science and Engineering(DS), CBIT, Proddatur, YSR, A.P
*Corresponding Author E-mail: 222p1a3241@gmail.com

Abstract

ClinXAI: An Explainable AI-Based X-Ray Analysis and Clinical Decision Support System is a smart healthcare application developed to assist medical professionals in interpreting X-ray images accurately and efficiently. The system applies deep learning techniques to examine medical X-rays and identify possible abnormalities such as lung infections, pneumonia, fractures, and other visible conditions.A key feature of this system is explainability. Instead of only producing a prediction, the model highlights the specific regions in the X-ray that influenced its decision. This helps doctors understand the reasoning behind the output, increasing confidence and transparency in AI-assisted diagnosis.

In addition to image analysis, the system includes a clinical decision support component that provides basic guidance related to the detected condition. By combining automated image interpretation with understandable explanations, ClinXAI aims to reduce diagnostic time, support medical practitioners in high-workload environments, and improve overall patient care. The system is especially useful in areas where specialist radiologists are limited, enabling faster and more reliable preliminary screening. Designed to assist medical professionals in diagnosing pneumonia from chest X-rays combined with patient clinical symptoms. The main goal of this project is to build a trustworthy and easy-to-use medical diagnosis support system. By providing clear explanations, the system improves confidence among doctors, supports better decision-making, and helps in faster and safer diagnosis in healthcare environments.

Keywords: Explainable Artificial Intelligence (XAI); X-Ray Image Analysis; Clinical Decision Support System (CDSS); Deep Learning, Convolutional Neural Networks (CNN); Medical Image Processing; Disease Detection; Heatmap Visualization.

1 Introduction

Medical imaging is an important tool for diagnosing many health conditions at an early stage. Among different imaging methods, X-ray imaging is widely used because it is quick, affordable, and available in most hospitals. It helps in identifying problems such as lung infections, pneumonia, tuberculosis, and bone fractures. However, interpreting X-ray images accurately requires experience and careful observation. In many healthcare centers, especially in rural or resource-limited areas, there is a shortage of skilled radiologists, which can lead to delayed or inaccurate diagnoses.Artificial Intelligence (AI) has recently shown great potential in supporting medical image analysis. Deep learning models, particularly Convolutional Neural Networks (CNNs), can automatically learn patterns from X-ray images and detect abnormalities with good accuracy. Although these models perform well, many of them do not explain how their decisions are made. This lack of transparency can reduce trust among healthcare professionals.ClinXAI is proposed as an explainable AI-based X-ray analysis and clinical decision support system. It not only predicts diseases from X-ray images but also highlights important regions that influence the decision. By providing both results and explanations, the system aims to improve trust, support doctors, and enhance diagnostic efficiency.

1.1Existing System

Existing AI-based medical diagnosis systems mainly focus on image classification using deep learning models such as Convolutional Neural Networks (CNNs). These systems analyze medical images like X-rays or CT scans to detect diseases and provide automated predictions. While many of these models achieve high accuracy, they typically rely on a single type of data and do not incorporate patient clinical information. As a result, they may miss important contextual factors that doctors consider during diagnosis. Another major limitation of existing systems is the lack of explainability. Many deep learning models operate as black boxes, providing only a final prediction without showing how the decision was made. This reduces trust among healthcare professionals and makes it difficult to validate AI-based recommendations. Additionally, some systems do not provide risk levels or confidence scores, which are important for understanding disease severity. Therefore, although existing approaches demonstrate strong performance, they often lack transparency, multimodal integration, and practical clinical usability.

1.2 Proposed System

To overcome the limitations of existing systems, ClinXAI is proposed as a multimodal and explainable AI-based medical diagnosis framework. Unlike traditional models that rely only on image data, ClinXAI integrates both medical images and structured clinical parameters to improve diagnostic accuracy. This combined approach better reflects real-world clinical decision-making. ClinXAI also incorporates explainable AI techniques to enhance transparency. Visual explanation methods highlight important regions in medical images, while feature-based analysis explains the contribution of clinical parameters to the final prediction. In addition, the system provides a confidence-based risk score to assist doctors in understanding the severity of the condition. By combining multimodal learning, explainability, and real-time performance, ClinXAI offers a more reliable and clinically meaningful diagnostic support system

2 Literature Review

Recent advancements in Artificial Intelligence (AI) and Deep Learning have significantly improved medical image analysis, particularly in X-ray diagnosis. Convolutional Neural Networks (CNNs) are widely used for detecting diseases such as pneumonia, lung infections, and fractures from chest X-ray images. Studies published in journals like Nature have shown that deep learning models can achieve performance comparable to radiologists in specific diagnostic tasks. Large public datasets provided by organizations such as the National Institutes of Health (NIH) have further supported research in automated disease detection systems.

However, many existing AI-based diagnostic systems operate as “black-box” models, providing predictions without explaining how decisions are made. This lack of transparency creates trust issues among healthcare professionals. To address this problem, researchers introduced Explainable AI (XAI) techniques such as Grad-CAM, which visually highlights important regions in X-ray images responsible for the model’s prediction. These visualization methods improve interpretability and help doctors better understand and validate AI decisions.

Recent research also emphasizes multimodal approaches that combine medical images with clinical data such as patient symptoms, age, and medical history. These integrated systems show improved diagnostic accuracy compared to image-only models. Despite these advancements, challenges remain in ensuring reliability, transparency, and real-time clinical deployment. Therefore, modern systems focus on developing accurate, explainable, and user-friendly decision support tools for healthcare applications.

3 Methodology

The proposed framework, ClinXAI, is built upon a multimodal explainable artificial intelligence architecture that combines medical image interpretation with the analysis of structured clinical information to produce clear and interpretable diagnostic outcomes. The overall approach is organized into five key phases: collecting the required data, preparing and cleaning the inputs, developing and training the predictive models, integrating outputs from multiple data sources, and generating explanations to support transparent decision-making.

3.1 Data Acquisition

Chest X-ray scans obtained from publicly accessible medical imaging repositories, which are standardized and prepared before being used for analysis. Structured Clinical Data: Individual patient characteristics, such as age, sex, oxygen saturation levels (SpO_2), presence of fever, and presence of cough. By combining visual findings from radiological images with key physiological and demographic patient details, the multimodal architecture enhances the reliability and context awareness of the final diagnostic prediction.

3.2 Data Preprocessing

3.2.1 Image Preparation

To ensure uniformity in the input data, every medical image was subjected to a series of preprocessing operations:

- Adjustment to a predefined image size (for example, 224 × 224 pixels)
- Central cropping to eliminate unnecessary background portions
- Normalization of pixel intensity values
- Transformation into tensor representation to enable compatibility with deep learning models

3.2.2 Structured Data Preparation

The structured clinical information was refined through the following procedures:

- Encoding gender into numerical format
- Converting fever and cough indicators into binary values
- Applying scaling techniques to continuous variables such as age and oxygen saturation
- Combining all processed attributes into a standardized fixed-length

3.3 Model Development

3.3.1 Image-Based Deep Learning Model

A convolutional neural network (CNN) architecture was employed for image classification. The network extracts hierarchical spatial features from X-ray images and produces a probability score representing the likelihood of disease occurrence.

The model output is defined as:

$$Pv \in [0,1]$$

where Pv represents the probability derived from the image-based analysis.

3.3.2 Structured Data Machine Learning Model

A tree-based classifier is used to analyse clinical parameters. Tree-based models are selected due to their ability to:

- Capture nonlinear relationships
- Handle mixed feature types
- Provide feature importance measures The structured data model produces:

$$Pt \in [0,1]$$

where Pt represents the probability derived from patient-level features.

3.4 Multimodal Fusion

To improve robustness and reduce reliance on a single modality, a late fusion strategy is implemented.

The final diagnostic probability was computed as follows:

$$Pfinal = (\alpha \times Pv) + (\beta \times Pt)$$

where:

$\alpha + \beta = 1$

$\alpha = 0.6$ (image weight)

$\beta = 0.4$ (clinical weight).

The classification decision is obtained using a predefined threshold τ:

$$\text{If } Pfinal \geq \tau \rightarrow \text{Positive}$$

$$\text{If } Pfinal < \tau \rightarrow \text{Negative}$$

Risk stratification, which is further categorized into Low, Medium, and High risk based on probability intervals.

3.4.1 Visual Explanation

Gradient-based localization is applied to the final convolutional layers of the image model to identify influential regions. A heatmap was generated and overlaid on the original image to highlight the areas contributing to the prediction.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

3.4.2 Feature Contribution Analysis

For structured data predictions, feature contribution values are computed using model-based interpretability techniques. Each input parameter was assigned an importance score indicating its positive or negative influence on the final outcome. This dual explainability framework enables clinicians to verify the reasoning of AI.

4. System Architecture

The ClinXAI architecture illustrates a layered framework where medical images and clinical data are processed through separate analytical models and combined using a fusion mechanism to produce the final diagnosis. It also includes an explainability module that generates visual heatmaps and feature contribution scores to ensure transparency and easy interpretation of results.

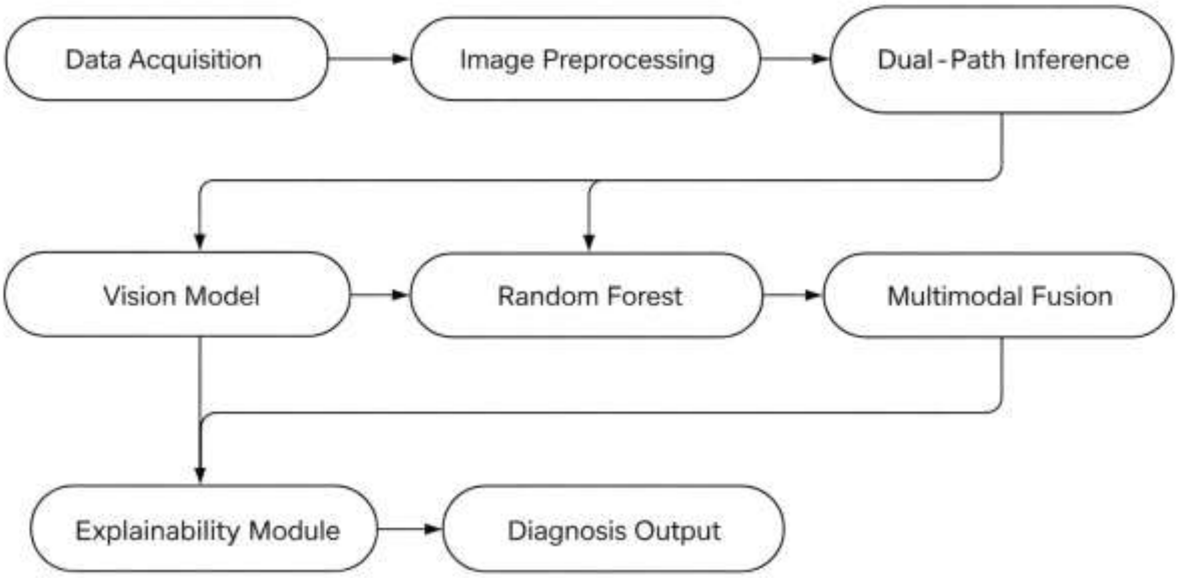


Fig. 1. ClinXAI System Architecture: An illustration of the ClinXAI framework integrating medical image and clinical data streams with multimodal fusion

5. SYSTEM MODULES

The ClinXAI framework is an interpretable multimodal diagnostic system that integrates radiological evidence and clinical indicators to estimate disease risk.

Instead of relying on a single modality, the system imitates clinical reasoning where physicians consider imaging findings together with patient symptoms. The architecture processes heterogeneous inputs through parallel learning streams and later combines their evidential confidence into a unified diagnostic decision.

5.1 Multimodal Input Module

The system receives patient observations from two complementary sources:

- Radiographic modality: chest X-ray scan
- Clinical modality: structured physiological and symptomatic attributes

Each record is represented as a tuple:

$\mathbf{X = (I, C)}$

where

I denotes the radiographic image and $\mathbf{C = \{c_1, c_2, ..., c_n\}}$ denotes clinical parameters.

Before inference, the module performs syntactic validation and semantic range verification to ensure physiologically feasible values. Invalid entries are rejected to avoid uncertainty propagation in later stages.

5.2 Image Standardisation Module

Medical images differ in acquisition resolution and exposure intensity. Therefore a deterministic transformation is applied:

$$\mathbf{I' = T(I)}$$

The transformation $\mathbf{T(\cdot)}$ performs spatial alignment, scale normalisation, and intensity stabilisation.The objective is not augmentation but distribution harmonisation so that the feature extractor learns pathology instead of device characteristics.

5.3 Radiological Inference Module

The standardised image $\mathbf{I'}$ is processed by a deep convolutional representation learner that maps visual patterns to diagnostic likelihood:

$$\mathbf{P_v = F_v(I')}$$

The function F_v extracts hierarchical features corresponding to anatomical structures and abnormal opacity patterns.

5.4 Decision Generation Module

The final probability **Pdiag** is mapped to a clinically interpretable category:

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

- Low if $P_{diag} < \tau_1$
- Moderate if $\tau_1 \leq P_{diag} < \tau_2$
- High if $P_{diag} \geq \tau_2$

The module outputs:

- predicted condition
- confidence score
- risk category
- visual and clinical explanations

This converts statistical inference into actionable decision support.

5.5 System Integration Layer

All modules operate sequentially in a service pipeline where acquisition, inference, fusion, and explanation remain logically independent.

This modular separation allows model replacement or retraining without changing the external diagnostic interface, supporting long-term deployment.

6 Result and Discussion

The ClinXAI system was assessed using real-time multimodal inputs that combined chest radiographic images with structured clinical information. The purpose of this evaluation was to determine whether the framework could deliver dependable diagnostic predictions while simultaneously offering clear and interpretable explanations suitable for clinical use. During the evaluation phase, a chest X-ray image was processed together with relevant patient information, including age, gender, oxygen saturation percentage, and symptom indicators such as fever and cough. After submission of these inputs, the inference engine computed a disease probability score and concurrently produced both visual and feature-level interpretability outputs.

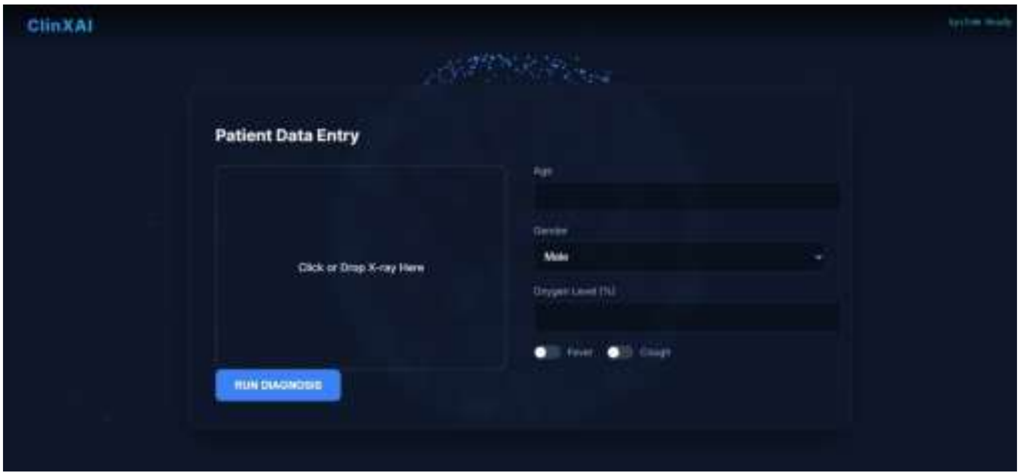


Figure 2 displays the patient input interface, emphasizing the modular design of the system. Radiographic uploads and structured clinical data are handled through separate input channels before being combined at the fusion stage. This separation ensures that the model maintains functional independence between modalities and increases robustness in scenarios where partial clinical data may be unavailable.

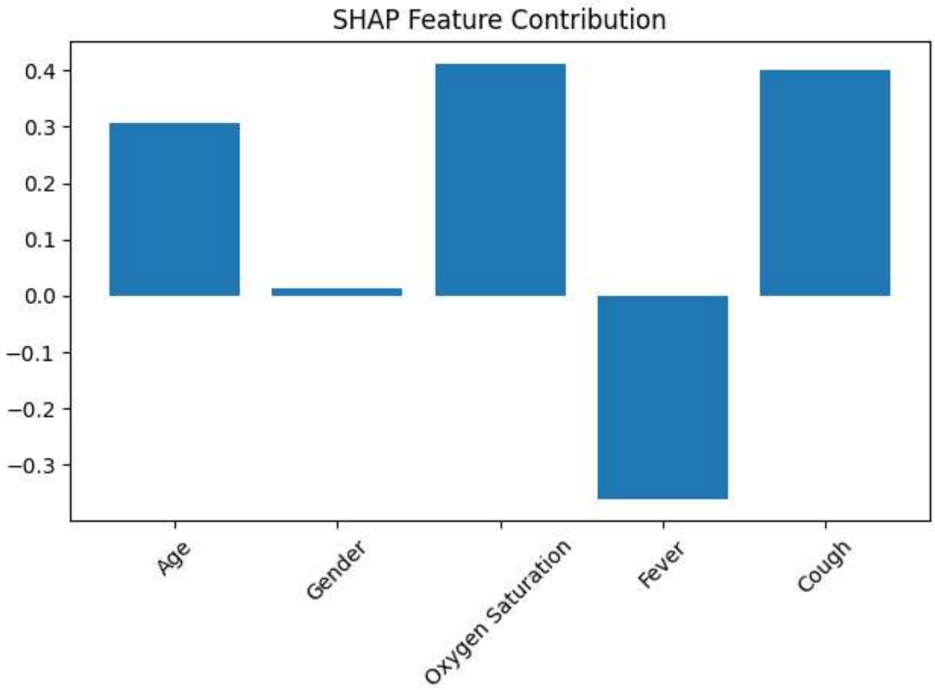
Figure 3 represents the system's state prior to diagnostic execution. The progression from this initial interface to the final high-risk prediction demonstrates smooth integration of preprocessing, multi-modal inference, feature fusion, and explanation components within a unified pipeline.

The system generated a confidence score of 85.0%, classifying the case as High Risk. The predefined decision threshold was established at 50%, meaning that any prediction exceeding this value is marked for clinical attention. The high probability value indicates that the fusion layer effectively integrated imaging features and clinical attributes to estimate the likelihood of pathology.

In addition to spatial visualization, the SHAP-based analysis quantifies the contribution of individual clinical parameters to the final prediction score. Variables such as oxygen saturation, cough, and age show measurable influence on the output probability. This feature-level interpretability enables clinicians to understand not only the anatomical basis of the decision but also the relative impact of patient-specific clinical indicators.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Performance Evaluation



Conclusion

ClinXAI represents a promising step toward explainable multimodal medical diagnosis, yet its current scope highlights important challenges that must be addressed before widespread adoption. The reliance on high-quality, diverse datasets and the sensitivity to class imbalance reveal potential risks in generalization, particularly for rare conditions or underrepresented populations. Technical constraints such as noisy imaging, missing clinical values, and computational demands further emphasize the need for refinement. While interpretability tools like heatmaps and feature importance provide valuable insights, they still fall short of fully revealing the complex reasoning within deep neural networks. Moreover, the system's operation in controlled environments limits its immediate applicability in real-world healthcare settings.

References

[1] J. P. Cohen et al., "TorchXRayVision: A library of chest X-ray datasets and models," *Journal of Open Source Software*, vol. 5, no. 54, p. 2733, 2020.

[2] P. Rajpurkar et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," 2017. from: https://arxiv.org/abs/1711.05225

[3] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.from: https://arxiv.org/abs/1610.02391

[4] A. Ngiam et al., "Multimodal deep learning," in *Proceedings of the 28th International Conference on Machine Learning (ICML)*, 2011.from:https://dl.acm.org/doi/10.5555/3104482.3104569

[5] G. Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.from: https://www.sciencedirect.com/science/article/pii/S1361841517301135

[6] E. S. Berner, "Clinical decision support systems: State of the art," *AHRQ Publication*, 2009.from:https://www.ncbi.nlm.nih.gov/books/NBK470264/