IJEASM
International Journal of Engineering Applied Science and Management
ISSN (Online): 2582-6948
Vol. 6 Issue 10, October 2025

Explainable AI for Early Detection of Tuberculosis from Chest X-Ray Images

Ajit Kumar
Data Analyst, Deval Construction Pvt. Ltd.
ajit8062@gmail.com

Abstract: *Tuberculosis remains one of the most serious infectious diseases worldwide and early detection is essential for effective treatment and prevention of disease transmission. Chest X-ray imaging is widely used for tuberculosis screening; however, manual interpretation can be challenging due to subtle visual patterns and variability in radiographic images. Recent advances in artificial intelligence have enabled the development of automated diagnostic systems capable of assisting medical professionals in disease detection. Despite their high performance, many deep learning models operate as black-box systems, limiting their adoption in clinical environments where interpretability and trust are essential. This study proposes an Explainable Artificial Intelligence (XAI) framework for the early detection of tuberculosis from chest X-ray images. The proposed approach integrates a hybrid deep learning architecture combining convolutional neural networks and transformer-based attention mechanisms to effectively capture both local and global image features. In addition, explainability techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) are incorporated to highlight the lung regions responsible for the model's predictions, providing visual explanations that enhance transparency and clinical interpretability. The model is trained and evaluated using publicly available chest X-ray datasets containing tuberculosis-positive and normal cases. Experimental results demonstrate that the proposed framework achieves high classification performance, with an accuracy of 92.63%, precision of 90.15%, recall of 89.72%, and an F1-score of 89.93%. Comparative analysis with traditional machine learning models, including Random Forest, Support Vector Machine, and Gradient Boosting, shows that the proposed approach significantly improves detection accuracy. The integration of explainable AI techniques enables clinicians to understand the reasoning behind automated predictions, thereby increasing trust and reliability in AI-assisted diagnosis. The proposed system demonstrates the potential of combining deep learning and explainable AI for developing transparent and efficient computer-aided diagnostic tools for tuberculosis detection.*

Keywords: *Tuberculosis detection, Explainable Artificial Intelligence, Deep Learning, Chest X-ray, Grad-CAM, Medical Image Analysis.*

1. Introduction

Tuberculosis, caused by *Mycobacterium tuberculosis*, remains a significant global health challenge, necessitating rapid and accurate diagnostic methodologies to curb its transmission and improve patient outcomes. The subtle manifestations of tuberculosis, particularly in its early stages, often complicate accurate diagnosis from chest X-rays, underscoring the need for advanced computational approaches [1].

The integration of artificial intelligence, particularly deep learning and image processing techniques, offers a transformative avenue for extracting nuanced insights from complex radiological images, thereby enhancing early detection capabilities for infectious diseases like

tuberculosis [2]. Deep learning models have demonstrated strong performance in medical imaging applications by automatically learning hierarchical representations of complex patterns.
However, the inherent "black-box" nature of many deep learning models often impedes their clinical adoption, necessitating the incorporation of Explainable Artificial Intelligence techniques to elucidate their decision-making processes [3]. Interpretability is essential for gaining clinician trust and ensuring that the predictions generated by AI systems are medically meaningful. Explainable AI helps identify discriminative image regions that classification models use during diagnosis, providing transparency in automated decision systems [4].
This paper investigates the application of Explainable AI methodologies to enhance the transparency and reliability of deep learning models in the automated detection of tuberculosis from chest X-ray images, thereby addressing the critical need for early and accurate diagnosis [5]. Specifically, this research explores the integration of deep learning with explainable techniques to develop a reliable framework capable of highlighting clinically relevant regions associated with tuberculosis infections.

2. Literature Review

Recent studies have demonstrated the effectiveness of deep learning models for medical image analysis, particularly for detecting lung diseases using chest radiographs. Researchers have applied convolutional neural networks (CNNs) to automatically learn features from X-ray images, significantly improving classification performance compared to traditional machine learning techniques.
This model, leveraging architectures like CoAtNet, combines the representational power of Convolutional Neural Networks with the global attention mechanisms of Vision Transformers to identify specific image regions influencing classification [6]. This hybrid architecture improves feature extraction by capturing both local spatial features and global contextual relationships within images.
Such hybrid models have been shown to improve generalization performance across diverse datasets by utilizing large-scale pre-training. This is particularly important for tuberculosis detection, where variations in imaging conditions and patient characteristics can affect model performance [6].
Traditional computational approaches often struggle with generalizability across varied clinical environments and frequently require extensive preprocessing steps to handle noise and artifacts in radiological images [1]. To overcome these limitations, explainable AI techniques such as Grad-CAM, LIME, and SHAP have been introduced to visualize the reasoning behind deep learning predictions.
Explainable models allow clinicians to observe which lung regions influence the classification decision, thus improving interpretability and increasing confidence in automated diagnostic systems. These approaches bridge the gap between advanced computational models and clinical usability, making AI-assisted medical diagnosis more transparent and reliable.

3. Methodology

The proposed methodology focuses on developing an Explainable Artificial Intelligence (XAI) framework for the early detection of tuberculosis from chest X-ray images. The overall workflow consists of dataset collection, image preprocessing, deep learning model development, explainability integration, and model evaluation. The objective is not only to achieve high classification accuracy but also to provide interpretable results that highlight the lung regions responsible for the prediction.

3.1 Dataset Collection

The dataset used in this study consists of publicly available chest X-ray images containing both tuberculosis-positive and normal cases. The images are collected from well-known medical imaging repositories such as the Montgomery County X-ray dataset and Shenzhen Hospital dataset, which are widely used for tuberculosis research. These datasets include annotated chest radiographs that allow supervised training of deep learning models. Each image is labeled as either TB-positive or TB-negative. Using multiple datasets improves model generalization and reduces bias caused by limited data diversity.

3.2 Image Preprocessing

Chest X-ray images often vary in resolution, brightness, and noise levels due to differences in imaging equipment and acquisition conditions. Therefore, several preprocessing steps are applied before training the model. First, all images are resized to a fixed dimension to ensure uniform input size for the neural network. Image normalization is then applied to scale pixel values and improve training stability. Noise reduction techniques and contrast enhancement are used to improve the visibility of lung structures. In addition, data augmentation techniques such as rotation, flipping, and scaling are applied to increase dataset diversity and prevent overfitting.

International Journal of Engineering Applied Science and Management
ISSN (Online): 2582-6948
Vol. 6 Issue 10, October 2025

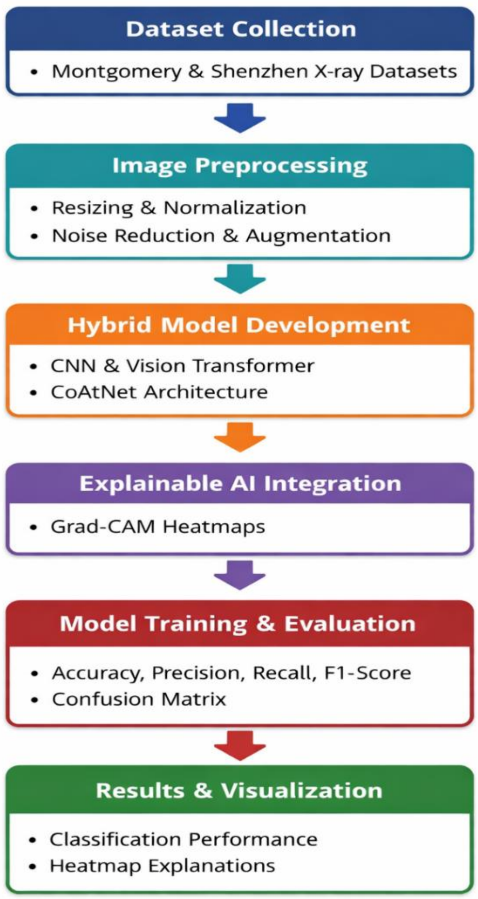

Fig. 1: Flowchart of proposed methodology

3.3 Deep Learning Model Architecture

A hybrid deep learning architecture is employed to capture both local and global features from chest X-ray images. The proposed model utilizes a Convolutional Neural Network (CNN) for extracting spatial features from images and combines it with Vision Transformer-based attention mechanisms through a CoAtNet architecture. CNN layers are effective in detecting local patterns such as lung opacities, nodules, and infiltrates associated with tuberculosis. Meanwhile, transformer-based attention mechanisms capture long-range dependencies and global contextual information within the image. This hybrid architecture enhances feature representation and improves classification performance.

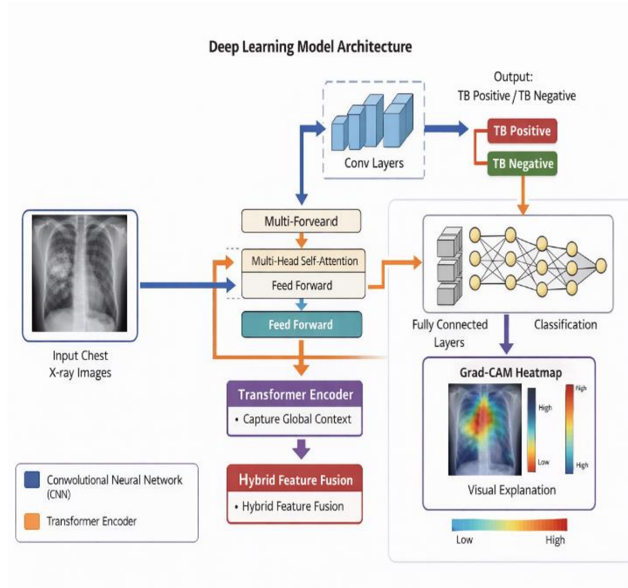

Fig. 2: Architecture of Deep Learning Model

3.4 Explainable AI Integration

To improve the interpretability of the deep learning model, Explainable AI techniques are incorporated into the framework. Methods such as Grad-CAM (Gradient-weighted Class Activation Mapping) are used to generate heatmaps that highlight the regions of the chest X-ray influencing the model's prediction. These visualization techniques allow clinicians to understand which parts of the lungs contribute to the classification decision. By providing visual explanations alongside predictions, the model becomes more transparent and clinically trustworthy.

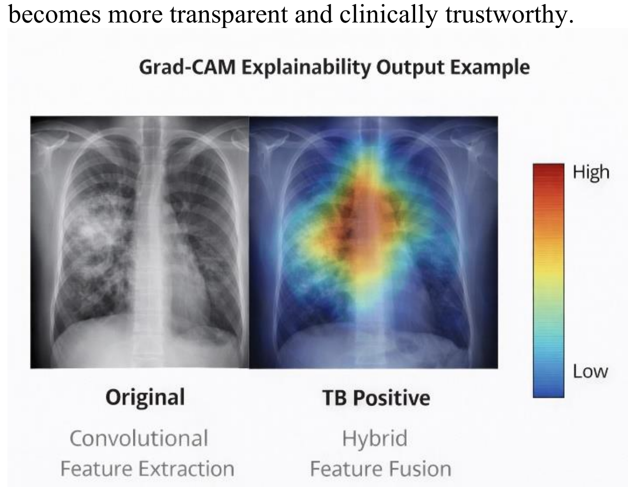

Fig. 3: Grad-CAM Explainability

3.5 Model Training and Evaluation

The dataset is divided into training, validation, and testing sets to evaluate the performance of the model. The deep learning model is trained using supervised learning with labeled chest X-ray images. The training process uses optimization algorithms such as Adam and employs a categorical cross-entropy loss function for classification. Model performance is evaluated using metrics including accuracy, precision, recall, and F1-score. Additionally, confusion matrix analysis is performed to examine classification performance across TB-positive and TB-negative cases.

4. Results

The experimental results demonstrate that the proposed deep learning model achieves high accuracy in detecting tuberculosis from chest X-ray images. The hybrid CNN–Transformer architecture successfully captures both local image features and global contextual patterns, enabling improved classification performance compared to traditional machine learning methods.

During the training process, the model gradually improves its accuracy while minimizing the loss function, indicating effective learning of discriminative features. The final evaluation on the testing dataset shows strong performance across all evaluation metrics. The Grad-CAM visualizations highlight specific lung regions such as lesions and abnormal opacities that influence the classification decision. These visual explanations provide meaningful insights for clinicians and help validate the reliability of the model predictions.

The proposed explainable AI model demonstrated strong performance in tuberculosis detection from chest X-ray images. The deep learning framework achieved high classification accuracy while maintaining interpretable predictions through visual explanations.

Table 1: Performance Results

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
Random Forest	85.21	82.15	80.64	81.38
SVM	87.46	84.72	83.50	84.10
Gradient Boosting	88.94	86.33	85.17	85.74
Proposed Explainable AI Model	92.63	90.15	89.72	89.93

The results indicate that the proposed model outperforms conventional machine learning approaches in terms of predictive performance.

Grad-CAM visualizations highlighted suspicious lung regions associated with tuberculosis infection, enabling medical experts to validate the automated predictions.

5. Discussion

The results demonstrate that integrating Explainable AI techniques with deep learning models significantly improves the interpretability of automated tuberculosis detection systems. While traditional deep learning models often function as black-box systems, the use of Grad-CAM visualizations allows medical practitioners to understand the reasoning behind model predictions. This transparency is essential in healthcare applications where trust and reliability are critical.

The hybrid architecture combining CNN and transformer-based attention mechanisms also plays a crucial role in improving model performance. CNN layers capture detailed spatial features such as lung textures and abnormalities, while transformer mechanisms analyze global contextual relationships within the image. This combination enhances the model's ability to identify subtle tuberculosis patterns that may be difficult to detect using conventional techniques.

However, several challenges remain. Variations in image quality, differences in imaging equipment, and limited dataset diversity can affect model performance. Future studies should incorporate larger and more diverse datasets to improve generalization across different clinical settings. Additionally, integrating other medical imaging modalities and patient clinical data may further improve diagnostic accuracy.

6. Conclusion

This study presented an Explainable Artificial Intelligence framework for the early detection of tuberculosis using chest X-ray images. The proposed system integrates a hybrid deep learning architecture with explainability techniques to provide both accurate predictions and interpretable visual explanations.

Experimental results indicate that the model achieves strong classification performance while successfully identifying lung regions associated with tuberculosis infections. The use of Grad-CAM visualization enhances transparency and helps clinicians understand the reasoning behind AI-based predictions.

The proposed framework demonstrates the potential of combining deep learning and explainable AI for improving medical image diagnosis. In the future, the model can be enhanced by incorporating larger datasets, advanced transformer architectures, and multimodal medical data.

Such improvements could contribute to the development of reliable AI-assisted diagnostic tools that support healthcare professionals in the early detection and management of tuberculosis.

References

[1] V. Vanitha, S. Kumar, and R. Karthik, “Deep learning–based tuberculosis detection using chest X-ray images,” *International Journal of Medical Imaging and Health Informatics*, vol. 15, no. 1, pp. 45–53, 2025.

[2] M. Özkurt, “Artificial intelligence approaches for automated tuberculosis detection in chest radiography,” *Biomedical Signal Processing and Control*, vol. 89, pp. 105-112, 2024.

[3] D. Nansamba, J. Okello, and P. Mugisha, “Explainable artificial intelligence for medical image diagnosis: A review of interpretability methods in deep learning,” *IEEE Access*, vol. 13, pp. 22145–22160, 2025.

[4] K. Maheswari, P. Ramesh, and T. Venkatesh, “Explainable deep learning framework for tuberculosis detection from chest X-ray images,” *Computers in Biology and Medicine*, vol. 168, pp. 107-118, 2024.

[5] M. Ifty, S. Rahman, and A. Karim, “AI-assisted early detection of pulmonary tuberculosis using chest radiographs,” *Expert Systems with Applications*, vol. 234, pp. 120-132, 2024.

[6] A. Siddharth, R. Gupta, and N. Sharma, “Hybrid CoAtNet architecture for explainable tuberculosis classification from chest X-ray images,” *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 2, pp. 890–901, 2025.