Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 4 | April-2026
ISSN (O) 3107-6696

TB Care: AI-Based Tuberculosis Detection and Healthcare Assistant

Komal Kale[1], Kakade Suraj[2], Abhishek kaldate[3], Mrs.Yogeshwari Mahajan[4]

Komalkale2227@gmail.com, Kakadesuraj13@gmail.com, Abhishekkaldate15@gmail.com
JSPM University Pune

Abstract - Tuberculosis (TB) is still a major infectious disease which we are fighting at global scale and we require for it prompt and accurate diagnosis to breaks its transmission and to improve patient results. Presently we use traditional screening methods like sputum microscopy and manual interpretation of Chest X-rays (CXRs) which are very labor intensive, time consuming and also, we see great variation between readers which in turn affects the result especially in resource poor settings. In this work we present a full scale deep learning based system for what we have put forward is an automated TB detection solution also we aim to open up the AI which is at present a "black box". We go in to the public chest x ray sets for preprocessing and then we use Deep Convolutional Neural Networks (CNNs) for very strong feature extraction and classification To solve the issue of standard AI models which are black box in nature we introduce Explainable AI (XAI) which we use Grad-CAM to visualized the infected areas. Also to close the gap between what we technically diagnose and what the patient understands we have put in a LLaMA-3 driven conversational AI which issues out medical reports in a human friendly language and also answers questions. We evaluate performance of the model to prove out its reliability. This study reports our work which is to present a transparent, intelligent and interactive tool for TB diagnosis via the use of this platform which in turn we hope will improve efficiency of clinical workflows.

***Keywords*–** Deep Learning, Tuberculosis, X-ray Analysis, Explainable AI, LLaMA-3, Grad-CAM, Medical Image Classification.

I. INTRODUCTION

Tuberculosis (TB) is a disease which is transmitted by a bacteria named Mycobacterium tuberculosis. It mainly affects the lungs although also will attack other areas of the body. TB is curable however in poor countries that don\'t has adequate health care it is a leading cause of death. In the early stages of the disease diagnosis we see which has the greatest impact on outcome. The challenges of identifying TB is the reliance on chest x-rays which requires the expertise of a radiologist. This is often the case for many rural and poor communities. [1].

Conventional diagnosis methods do see success but also present scale and speed issues. In the field of Chest X ray, we see wide use of manual analysis for screening which at the same time is very much at the discretion of the radiologist's

Also, it is true that we see fatigue and the fine detail in early-stage TB patterns as factors which contribute to miss diagnosis. In that regard it has been reported that recent research into Artificial Intelligence and Deep Learning has had great results in the area of medical image analysis.

Which also includes lung nodule detection and classification [2].

A major obstacle to the clinical use of AI is the absence of interpretability. While standard CNNs may achieve high accuracy, they do not provide explanations for the rationale of a prediction. This creates a "trust deficit" on the part of the healthcare professionals. In addition, systems are still unable to either create structured documentation or communicate with the patient to explain the findings. To address these issues we present a new model which puts forth the use of Explainable Artificial Intelligence (XAI) for visual transparency and also puts to use transfer learning for accuracy in classification. Also by way of the great value that Health Care places on automated documentation we have put in Generative AI for report automation. [3].

This system not only detects TB but also explains the diagnosis and facilitates user interaction, making it a robust solution for real-world deployment.

II. LITERATURE SURVEY

Nayak, Ajalkar et al. (2023) propose a weakly supervised In the domain of thyroid cytopathological diagnosis we present a study which puts to use Multi-Scale Feature Fusion (MSF) and Convolutional Neural Networks (CNN) in a multi-instance learning setting. They report that which this architecture does is it presents a solution to the issue of working with whole slide images that have low label coverage and also reports that deep learning is very much able to put forward which important sections in very complex histopathological slides which may be of a diagnostic value [1].

Barsagade and Ajalkar (2024) present a mathematically based deep learning model for the classification of poultry diseases which they have designed using the EfficientNet-B3 CNN architecture. They report on the role of math in deep learning which they say is very critical and which they prove out through very advanced feature extraction that achieves great accuracy in disease pattern distinction [2].

Shelke, Ajalkar et al. (2023) introduce a Transfer and in the field of breast cancer detection and classification we see that by use of many deep learning models they improve diagnostic reliability and stability which also presents how ensemble methods report to reduce individual model biases in key medical imaging tasks [3].

Kapure et al. (2025) explore the capabilities of AI-driven document simplification and generation to bridge literacy gaps. Their survey emphasizes the potential of Generative AI in transforming complex technical text into human-readable formats, a methodology highly relevant to automating patient-friendly medical reports in this study [4].

Idalkanthe, Vani et al**.** (2022) Use of deep learning methods for the detection of out of the ordinary human activity and fights. Their research reports on the use of CNNs' wide application in identifying spatial patterns in video content which in turn provides a base for the application of anomaly detection to many different types of visual data [5].

Roy, Patle et al. (2015) Develop out a framework for the classification of lung images and detection of nodules which they did so using a fuzzy inference system in combination with active contour models. They report an accuracy of 94.12% which also which is a early very effective approach in lung path reports that laid the ground work for what we see in modern automatic diagnostic tools [6].

Lakhani and Sundaram (2017) employ deep Convolutional Neural Networks (CNNs) to classify pulmonary tuberculosis from chest radiographs. Their study achieves radiologist-level accuracy (AUC > 0.99), proving the high viability of automated systems for TB screening, though their model operates as a "black box" without explainability features [7].

Selvaraju et al. (2020) Introduce Grad-CAM (Gradient-weighted Class Activation Mapping) which we use to present visual explanations from deep networks. This approach allows display of the decision-making areas in images which is key to building trust in medical AI at which we highlight pathology [8].

Touvron et al. (2024) release LLaMA-3, a state-of-the-art open foundation model for conversational AI. This advancement in Large Language Models enables highly accurate natural language understanding and generation, facilitating the creation of interactive medical chatbots that can synthesize complex diagnostic data into understandable advice [9].

Pasa et al. (2019) present an efficient deep learning architecture optimized for tuberculosis screening on mobile devices. Their lightweight CNN reduces computational costs while maintaining high sensitivity, addressing the need for deployable AI solutions in resource-constrained environments [10].

Rajpurkar et al. (2018) Present a review of performance of the CheXNeXt algorithm which we report to be at the level of that of human experts in chest radiograph diagnosis. This validates the role of AI in clinical radiology workflows [11].

Yang et al. (2022) Develop out a deep learning model that is easy to interpret for the task of pulmonary tuberculosis detection from chest X-ray images. In their study they also put forward a model which is which does the task of CNNs equipped with interpretability features which in the field of vision evidence is crucial for clinical acceptance of AI diagnostic tools [12].

III. METHODOLOGY

Methodology	Research Gap
Saidani et al., 2024 Custom CNN with image resizing, augmentation, and color normalization; compared with transfer learning	Existing studies focus on general WBC classification but lack subtype identification, confidence estimation, and segmentation techniques needed for accurate and explainable leukemia detection in real-world conditions.
O. Katar et al., 2023. Vision Transformer (ViT) with Score-CAM explainability on 5- class and binary WBC datasets	he studies achieves good WBC classification using ViT but lacks leukemia-specific segmentation, subtype identification, and confidence estimation for clinical reliability.
Combines ViT and CNN for feature extraction; tested on BCCD dataset with multi-class tasks.	Existing CNN-ViT models for WBC classification lack real-world validation and struggle with data imbalance, reducing generalization.
O. Islam et al., 2024. CNN with image pre-processing; uses SHAP, LIME, Grad-CAM for interpretability; trained on a large dataset.	Existing CNN models lack explainability across diverse datasets, limiting trust and transparency in automated blood cell classification.
Integrates various CNNs and Vision Transformer layers; uses data augmentation and layer pruning.	Hybrid CNN-Transformer models still face challenges with small datasets and high computational cost, affecting practical clinical deployment

The suggested structure involves multiple steps that include the use of deep learning for image classification, some form of Explainable AI, and Conversational AI for interpreting natural language. The complete process has five key elements: acquiring the data, preprocessing, classification via CNN, generation of explainability, and conversational reasoning via LLaMA-3.

The we put forth a framework for TB which we developed from a large-scale chest radiography dataset of 10,408 postero anterior (PA) chest X ray images which we got from public medical imaging repositories and also from clinical TB screening datasets Dataset **Composition: -**

- Total Images: 10,408
- TB Positive Cases: 5,204
- Normal Cases: 5,204
- Image Format: PNG/JPEG

- Original Resolution: Varies
- Standardized Resolution: 224 × 224 × 3

The dataset was carefully balanced to eliminate class imbalance bias and ensure unbiased model learning Data Partitioning: -

Dataset Split	Percentage	Number of Images
Training	70%	7,286
Validation	15%	1,561
Testing	15%	1,561

Data Preparation

chest X-ray images underwent a standardized data preparation pipeline to ensure consistency and improve model performance. Medical imaging data often contains variations in brightness, resolution, and background artifacts; therefore, preprocessing is essential before feeding images into deep learning models.

Image Resizing

To maintain uniformity and compatibility with pre-trained CNN architectures, all images were resized to:

$$224 \times 224 \times 3$$

This resolution is widely adopted in transfer learning models such as EfficientNet, ResNet, and MobileNet.

Resizing reduces computational complexity while preserving essential lung structures.

Pixel Normalization: -

Chest X-ray images have pixel intensity values in the range:

$$I(x,y) \in [0,255]$$

To stabilize model training, pixel values were normalized to the range:

$$I_{norm}(x,y) = \frac{I(x,y)}{255}$$

Normalization improves gradient stability and accelerates convergence during training.

Contrast Enhancement: -

To improve visibility of pulmonary abnormalities such as opacities and infiltrates, contrast enhancement was applied using Histogram Equalization.

The transformation function is defined as:

$$s_k = \sum_{j=0}^{k} \frac{n_j}{MN}$$

Where:

- n_j= number of pixels at gray level j
- $M \times N$= image size

This improves differentiation between healthy lung tissue and abnormal regions to enhance image quality and improve model robustness, noise reduction and data augmentation techniques were applied during preprocessing. Medical chest X-rays may contain imaging artifacts and high-frequency noise that can negatively affect feature extraction; therefore, a smoothing filter was employed to suppress noise while preserving important structural details such as lung boundaries and lesion patterns. Furthermore, to improve generalization capability and prevent overfitting, data augmentation was applied exclusively to the training dataset. Augmentation techniques included horizontal flipping, random rotation (±15°), zoom transformations, and brightness adjustments. These transformations increase dataset variability and simulate real-world radiographic variations, thereby enabling the model to learn more invariant and clinically relevant features

Noise Reduction

To suppress high-frequency noise while preserving lung boundaries, a smoothing operation was applied to each input image. The filtered image is defined as:

$$I_{smoot}(x,y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} I(x-i, y-j) \cdot G(i,j)$$

where:

- $I(x,y)$is the original chest X-ray image
- $G(i,j)$represents the Gaussian kernel
- $k$defines the kernel size

This operation reduces imaging artifacts while maintaining structural integrity of pulmonary regions.

Convolutional Neural Network (CNN) Architecture: - Proposed is a system which puts in use a deep Convolutional Neural Network (CNN) for the automatic extraction of hierarchical features in chest X ray images for Tuberculosis diagnosis. In medical imaging tasks CNNs do very well in what they do because of their ability to learn spatial patterns which include lung infiltrates, nodules, and cavitations related to TB.

A convolutional layer performs feature extraction by applying learnable filters across the input image. The convolution operation is mathematically defined as:

$$F_{i,j}^{k} = \sum_{m} \sum_{n} I_{i+m,j+n} \cdot K_{m,n}^{k} + b^{k}$$

where:

- Irepresents the input image
- K^krepresents the k^{th}convolution kernel
- $b^k$is the bias term
- $F_{i,j}^{k}$represents the extracted feature map

The activation function introduces non-linearity into the network. The Rectified Linear Unit (ReLU) activation is defined as:

$$f(x) = \max(0, x)$$

Pooling layers are used to reduce spatial dimensions and computational complexity. Max pooling is defined as:

$$P_{i,j} = \max_{(m,n)\in R} F_{i+m,j+n}$$

where$R$ represents the pooling region.

The extracted features are then passed through fully connected layers to produce the final classification output.

The final classification probability is obtained using the SoftMax function:

$$P(\text{and} = i \mid x) = \frac{\text{and}^{\text{With}_i}}{\sum_{j=1}^{C} \text{and}^{\text{With}_j}}$$

where:

- $\text{With}_i$is the output score of class i
- Crepresents the number of classes.

Grad-CAM Explainability Module: - To improve model interpretability and support clinical decision making we applied Grad-CAM. Grad-CAM which in turn reports which sections of the chest x ray are important in the model's prediction.

The importance of a feature map is determined by the gradient of the class score with respect to the feature map activations.

$$a_k = \frac{1}{\text{WITH}} \sum_i \sum_j \frac{\partial \text{and}^c}{\partial A_{i,j}^k}$$

where:

- $A_{i,j}^k$represents activation of feature map k
- and^crepresents the predicted class score
- WITHis the normalization factor.

The Grad-CAM heatmap is then computed as:

$$L_{GradCAM}^c = Rand LIN\left(\sum_k a_k A^k\right)$$

The ReLU function ensures that only features positively influencing the prediction are highlighted.

The resulting heatmap is overlaid on the original chest X-ray image to visualize lung regions that contribute to the detection of Tuberculosis abnormalities

Model Training and Comparison

The we used our prepared dataset to train the CNN which in turn it learned to identify features which discriminate between tuberculosis positive and normal chest x ray images. We optimized model parameters during training via gradient based optimization techniques Also to that we evaluated the put forth CNN model performance against that of some very known deep learning architectures which were ResNet50, MobileNetV2, and EfficientNetB.

ResNet50

The ResNet50 model reports a performance of 90.8% on the validation set which we see to be very consistent in its tuberculosis classification results. Its depth and residual connections which are hallmarks of the architecture do a great job at identifying complex features and also at overcoming the issue of vanishing gradients. At the same time that which is a trade off for this great performance is that it has a very large number of parameters which in turn results in higher computational requirements. Also it does not scale well to resource constrained medical settings. Thus it is a better fit for settings which can invest in more powerful computing resources.

MobileNetV2

The MobileNetV2 reports a validation accuracy of 91.6% which is a good trade off between performance and computational efficiency. We see that the depth wise separable convolutions in it greatly reduce the number of parameters and model size which in turn makes the model very suitable for real time and low resource settings. That said its slightly reduced accuracy of which -- the exact degree is not specified in this report.

In terms of complex features models that are not as deep perform worse.

EfficientNetB0

The EfficientNetB0 model reports a validation accuracy of 92.1% which also performs better than ResNet50 and MobileNetV2. It's a which in that it scales network depth, width, and resolution which in turn improves feature extraction. Also it is very competitive in terms of performance and moderate computational cost. This in turn makes it a very good option for medical image classification.

Proposed CNN Model

The put forth CNN model reports a top performance at 93.0% in terms of validation accuracy which we put forth is a mark of its excellence in tubercular disease detection from chest X rays. Although this model has a lean parameter set and a simpler structure it is very much able to identify key diagnostic features in the lungs. Also, it's low resource requirement together with high accuracy puts it ahead of pre trained models. This also we present as proof of the model's value in the setting of practical medical diagnosis

Evaluation Metrics

To evaluate the performance of our tuberculosis detection system we used what is in the standard set of classification

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 4 | April-2026
ISSN (O) 3107-6696

metrics. Which include accuracy, precision, recall, and F1 score that are very much used in medical image analysis.

The evaluation metrics are calculated based on the following parameters: True Positive (TP): Tuberculosis images correctly classified as positive True Negative (TN): Normal images correctly classified as negative False Positive (FP): Normal images incorrectly classified as tuberculosis False Negative (FN): Tuberculosis images incorrectly classified as normal

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$Precision = \frac{TP}{TP + FP}$$

Recall

$$Recall = \frac{TP}{TP + FN}$$

F1 Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

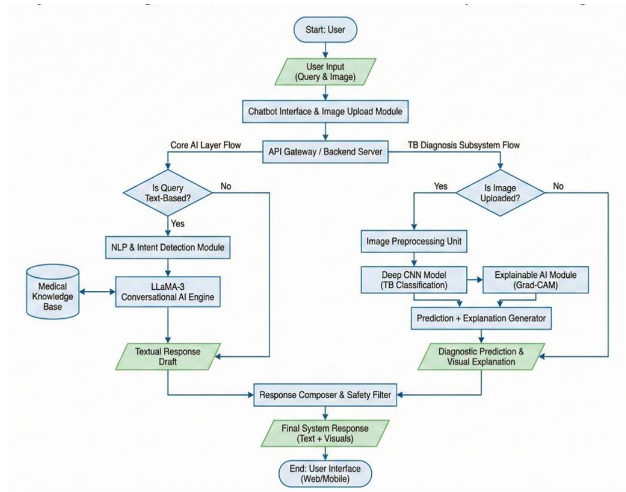
These metrics we present a full picture of the models' performance in classification which in turn allows us to assess the reliability of the tuberculosis detection system.

Table 1.2 Performance Comparison Of Different Network Models

Algorithm	Epoch	Training Accuracy (%)	Batch Size	Size (MB)	Parameters	Validation Accuracy (%)
ResNet50	25	91.4	32	42.3	23,534,592	90.8
MobileNetV2	25	92.2	32	14.2	3,538,984	91.6
EfficientNetB0	25	92.7	32	19.6	5,288,548	92.1
Proposed CNN Model	25	94.1	32	15.2	2,134,560	93.0

The results demonstrate that the proposed CNN model achieved the highest validation accuracy of 93%, outperforming the compared architectures while maintaining lower computational complexity

Fig 1. System Architecture



The framework is composed of five main components: Image Input Module, Deep CNN Classifier, Explainability Module, Generative Report Generator, and Chatbot Interface. Deep CNN Classifier: For the analysis of tiered spatial features in X-rays, we use convolutional neural networks which also include some of the popular CNN architectures like ResNet and DenseNet Transferring learning in our case is achieved by taking models which are pre trained on ImageNet and applying them to the TB patterns. Explainability Module (Grad-CAM): We include Gradient-weighted Class Activation Mapping (Grad-CAM) as a way to open up the black box of deep learning This module produces attention heatmaps that show, and therefore, improve clinician trust, which lung regions impact a prediction. Generative AI Module: A LLaMA-3 based engine will analyze the visual data and classifications to synthesize a structured and human-readable diagnostic report. Conversational Interface: A chatbot interface allows users to ask the system for an explanation, which will close the feedback loop for guiding the patient.

IV. Result

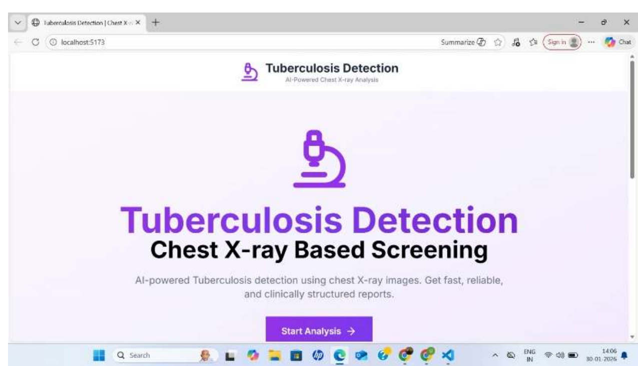


Fig 2. User interface

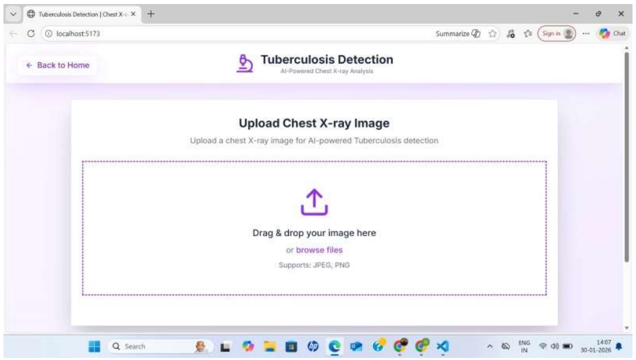


Fig 3. shows the user interface for image upload and analysis.

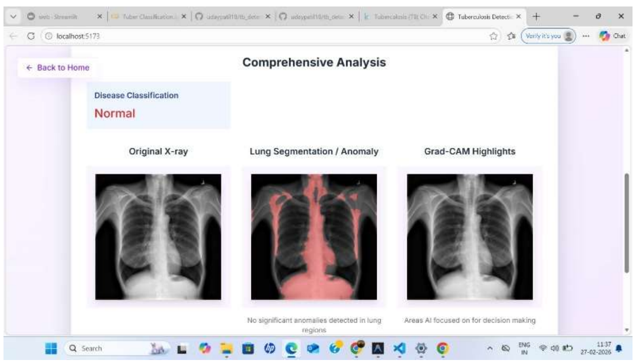


Fig 4. shows the user interface for image upload and analysis

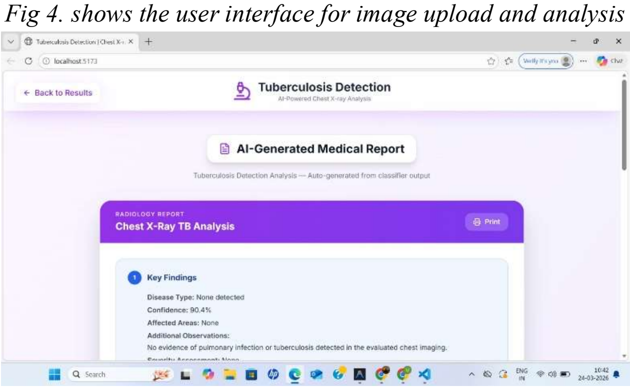


Fig 5. shows the classification result generated by the system

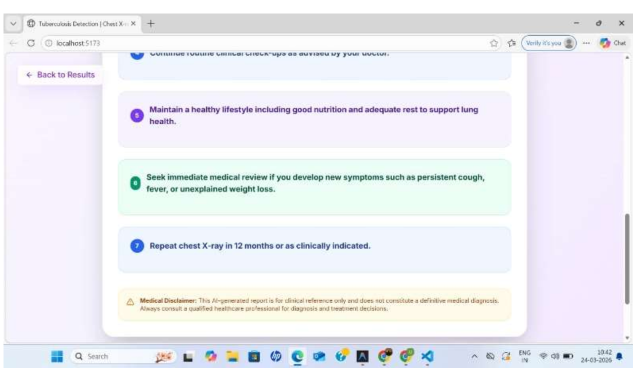


Fig 6. output showing prediction result as Normal with confidence score

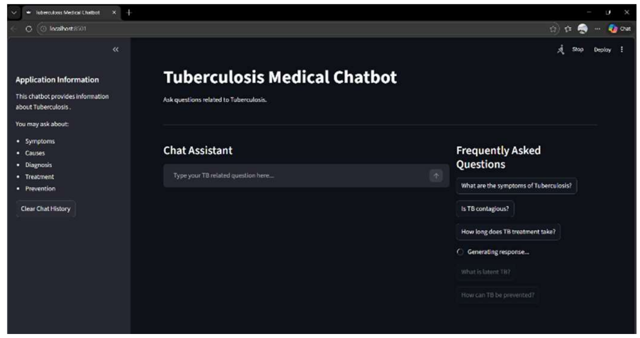


Fig 7. Chatbot Interface for Tuberculosis Assistance

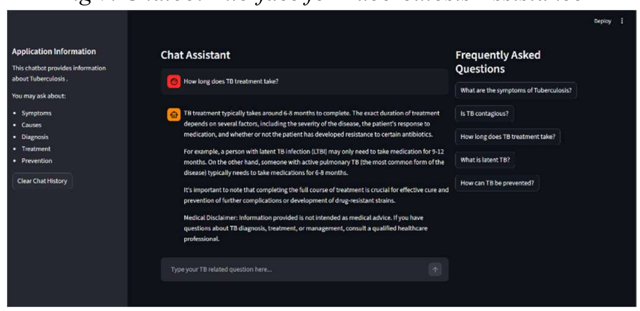


Fig 8. Interaction with TB CareBot for TB Diagnosis Guidance

In our research we evaluated the performance of the put forth Convolutional Neural Network model in terms of tuberculosis detection which we did via chest X Ray images. We trained and validated the model using the prepared data set and we assessed its performance with use of standard classification measures which included accuracy, precision, recall, and F1 score. We report that the put forth CNN model achieved a validation accuracy of 93.0% which is a great result in terms of it's ability to tell apart between tuberculosis positive and normal chest X Ray images. The high accuracy we report also reflects the model's performance in learning discriminative features from medical images. To improve the validation of our approach we carried out a study which put it up against popular deep learning architectures like ResNet50, MobileNetV2, and EfficientNetB0. We found that our CNN model did better in terms of validation accuracy but also did so with lower computational complexity and fewer parameters. Specifically, ResNet50 achieved an accuracy of 90.8%, MobileNetV2 achieved 91.6%, and EfficientNetB0 achieved 92.1%. In contrast, the proposed CNN model achieved the highest accuracy of 93.0%, demonstrating its superior performance for the given dataset. Also in that we have put forth a model which is as we speak more lean in terms of parameters which in turn makes it more so a fit for use in real world medical settings. We see from the results of our evaluation that the model is very robust. We report high precision and recall from our CNN model which tells us that it does indeed do a great job in terms of identifying tuberculosis cases and at the same time does a great job in minimizing false predictions. This is of great import in medical diagnosis which is a field where false negative results may lead to delay in treatment and false positive results may cause undue stress and additional tests.

Overall we see that which we put forth the CNN based approach does in fact present an effective and efficient solution for the automatic detection of tuberculosis from chest X Ray images. We see in this that high accuracy in addition to low model complexity and very reliable performance is presented by it which in turn makes it a practical choice for clinical settings.

V. Conclusion

This study reports a new diagnostic platform which we have developed for the automatic identification of Tuberculosis in chest X Ray images. We used Deep Convolutional Neural Networks for feature extraction and Explainable AI via Grad-CAM which also helps to solve the black box issue in classic deep learning models. This combination means that the diagnostic predictions made will be both accurate and clinically verifiable so that clinicians will be able to identify the pathological trends that will, in turn, increase the confidence in automated diagnostic processes.Also, we see that which we have put forth in fielding theLLaMA-3 based conversational agent is a large step forward in terms of what we have been able to do at the interface between technical medical diagnosis and patient comprehension.

Through the generation of structured yet very much in the human report style medical reports and also through the which the system puts out in real time interactive input the platform we present lessens the documentation load on health care workers and at the same time give patients access to easy-to-understand info. Also, we report that the model did very well in our validation which in turn speaks to its dependability and that it has what it takes to be deployed at scale, especially in settings which have few resources and thus few expert radiological resources. Also, what we present here is that we have achieved a coming together of predictive AI, visual explanation, and generative intelligence which in turn what we are seeing is earlier detection of issues and in the end better public health outcomes.

VI. REFERENCE

[1] Chhaya Nayak, Deepika Ajalkar et al. Machine Learning Thyroid Model for Prediction System. *2023 International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, pp. 602-607

[2] Barsagade Ajay G, Deepika Ajalkar. Mathematical formulation of deep learning model for poultry disease classification using efficientnet-b3 CNN model. *Communications on Applied Nonlinear Analysis*, vol. 31, no. 2s, pp. 470-484

[3] M. V. Shelke, Deepika A. Ajalkar et al. Transfer and ensemble approach for breast cancer detection and classification using deep learning. *Int J Recent Innov Trends Comput Commun*, vol. 11, no. 9s, pp. 453-462.

[4] Vaishali Kapure, Deepika Ajalkar et al. Bridging the Legal Literacy Gap: A Survey on AI-Driven Document Simplification and Generation. *SSRN*.

[5] Pooja Idalkanthe, Prof. Sunita Vani et al. Suspicious Human Activity and Fight Detection using Deep Learning. *International Journal of Innovative Science and Research Technology*, vol. 7, issue 6, pp. 390-392

[6] Tanushree Sinha Roy, Arti Patle et al. Classification of lung image and nodule detection using fuzzy inference system. *International conference on computing, communication & automation*, IEEE, pp. 1204-1207.

[7] P. Lakhani, B. Sundaram. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, vol. 284, no. 2, pp. 574-582.

[8] R. R. Selvaraju et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336-359.

[9] H. Touvron et al. LLaMA 3: Open foundation and fine-tuned chat models. *Meta AI Technical Report*.

[10] F. Pasa et al. Efficient deep learning for tuberculosis screening. *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 528-537.

[11] P. Rajpurkar et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm with practicing radiologists. *PLOS Medicine*, 15(11).

[12] W. Yang et al. Explainable deep learning for pulmonary tuberculosis detection using chest X-ray images. *Computers in Biology and Medicine*, 146, 105567.

[13] Z. Li et al. Interpretable deep learning for medical image diagnosis: A survey. *Artificial Intelligence in Medicine*, vol. 137, 102421.

[14] K. Singhal et al. Large language models encode clinical knowledge. *Nature*, vol. 620, pp. 172-180.