Validation Of An AI-Based Tool For Detecting Radiographic Findings Suggestive Of Tuberculosis: A Pilot Study

Yashraj Patil [a] Sushil Kachewar [b] Rahim Pathan [c]

[a]Associate Professor, Department of Radiodiagnosis, D.Y. Patil Hospital & Research Centre, Pune, India-411018 Email: med.radio@dpu.edu.in Phone Number: 0091‑ 8888867350

[b]Professor and Head, Department of Radiodiagnosis, D.Y. Patil Hospital & Research Centre, Pune, India-411018 Email: radsgk@gmail.com Phone Number: 0091-9921160357

[c]GenZ AI Labs, Pune, India email: genzailabs@gmail.com Phone Number: 0091-9923030250

Corresponding Author: Dr Sushil Kachewar

Professor and Head, Department of Radiodiagnosis, D.Y. Patil Hospital & Research Centre, Pune, India-411018 Email: radsgk@gmail.com Phone Number: 0091-9921160357

Short Title: **Validation of an AI-Based Tool for Detecting Radiographic Findings in Tuberculosis**

Keywords: Tuberculosis, Xray Chest, CXR Insightz, AI-based tool, Artificial Intelligence

Statements

Acknowledgement

Statement of Ethics

Study approval statement: *This study protocol was reviewed and* has been granted an exemption from requiring ethics approval by Institutional Ethics Subcommittee via letter bearing reference no-I.E.S.C./W/64/2025 dated 13/10/2025.

Consent to participate statement: As NO human participants were involved- only anonymized images were used hence written informed consent was not required and hence the study has been granted an exemption from requiring written informed consent by the Institutional Ethics Subcommittee.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

Author Contributions

a. YP-Data Collection, assimilation and sorting
b. SK- Data interpretation and Conclusion along with Validation
c. RP - CXR Insightz AI tool development and data churning

Data Availability Statement

The datasets generated and analysed during the study are included in this published article. Further enquiries can be directed to the corresponding authors.

1. Abstract

Introduction:

Tuberculosis (TB) remains a significant global health concern, particularly in resource-limited settings where access to expert radiological interpretation is scarce. Chest X-rays (CXR) serve as a valuable screening tool; however, their interpretation is subject to inter-reader variability. Artificial intelligence (AI)-based tools offer potential support in large-scale TB screening workflows. This pilot validation study aims to evaluate the diagnostic performance of CXR Insightz, a customized AI-based tool developed to detect radiographic findings suggestive of TB and other associated thoracic pathologies.

Objective:

To validate the performance of CXR Insightz in detecting chest radiographic abnormalities commonly associated with tuberculosis, and to assess its diagnostic agreement with expert radiologist interpretations on a curated retrospective dataset.

Materials and Methods:

A retrospective dataset of 170 chest X-rays, including 88 TB-positive and 82 normal cases (based on radiologist findings), was used for this validation study. The validation dataset was new and

independent; the AI tool, which had been pre-trained on a large proprietary dataset, was not retrained on this data. It was evaluated for its ability to detect chest radiographic pathologies commonly associated with tuberculosis, specifically: nodules, consolidation, pleural effusion, and fibrosis. The model's outputs were compared against radiologist- annotated findings considered as ground truth. Diagnostic performance was evaluated using standard metrics including accuracy, sensitivity, specificity, precision, recall, F1-score, and ROC-AUC. Statistical agreement was assessed using Cohen's Kappa and McNemar's tests.

All statistical analyses were performed using Python (version 3.10) with libraries including scikit-learn, stats models, NumPy, pandas, Matplotlib, and Seaborn

Results:

CXR Insightz achieved an overall accuracy of 93.53%, sensitivity of 100%, specificity of 86.59%, and an F1-score of 94.12% on the validation dataset. The area under the ROC curve (AUC) was 0.9329, demonstrating robust discrimination across the considered pathologies associated with TB.

A Cohen's Kappa value of 0.8698 indicated strong agreement between the AI tool and radiologist findings. McNemar's test yielded a p-value of 0.001, indicating a significant difference driven by the AI's conservative bias. All errors were false positives, reflecting the model's prioritization of sensitivity to avoid missed TB cases—an advantageous trait for a screening tool.

Conclusion:

The model demonstrated promising diagnostics performance, achieving an accuracy of 93.53% with strong agreement compared to radiologist assessments.

Its performance closely aligns with expert interpretation, supporting its utility in large-scale screening workflows, especially in settings with limited access to radiologists. Further evaluation on diverse populations and clinical integration is recommended to reinforce real-world applicability.

2. Introduction –

Tuberculosis (TB) remains a significant global health challenge, particularly in resource-limited settings with limited access to timely diagnosis. According to the World Health Organization (WHO), TB caused nearly as many deaths as COVID-19 in 2020 [1,2].

Chest X-ray (CXR) is widely used as a sensitive, cost-effective screening tool to detect radiographic abnormalities associated with TB, especially in mass screening programs. The WHO recommends CXR as a triage tool for individuals with TB symptoms or risk factors. However, its effectiveness is often limited by inter- and intra-reader variability and a shortage of trained radiologists, particularly in high-burden countries [1,3-5].

Delayed or missed recognition of TB-related radiographic findings increases TB-related mortality. Conventional diagnostic methods remain slow and resource-intensive, underlining the need for automated approaches to support the detection of TB-associated abnormalities [3-5].

The growing integration of Artificial Intelligence (AI) into medical imaging offers promising solutions to address challenges in TB screening. AI-driven computer-aided detection (CAD) tools can assist in the rapid, consistent, and large-scale interpretation of chest X-rays, thereby supporting clinical decision-making and flagging potential TB-associated findings for further evaluation [3,5,6,7].

Recognizing this potential, in 2021, the World Health Organization (WHO) updated its TB screening guidelines to recommend the use of CAD software as an alternative to human readers for analyzing digital chest X-rays in individuals older than 15 years. However, the WHO did not endorse any specific products, emphasizing the need for careful evaluation of AI tools before implementation. Rigorous validation—particularly on independent datasets reflecting real-world scenarios—remains critical for clinical adoption [4].

India's TB Burden and Study Purpose: In India, despite significant efforts and progress toward tuberculosis control, the disease burden remains substantial. As of 2022, the estimated incidence of TB was 199 cases per 100,000 population, with a mortality rate of 23 per 100,000 population [8].

The challenge in TB screening lies not only in inter-observer variability but also in accessibility to radiological expertise and the need for quick turnaround. Rapid identification and isolation of suspected active TB cases are critical, as these patients can actively spread the infection. Hence, timely triage and response are essential components in TB control.

In this pilot study, we aim to evaluate the performance of a newly developed AI-based screening tool, CXR Insightz, designed for detecting radiological findings in chest X-rays. The tool, pre-trained on a large proprietary dataset using multiple model architectures, was validated on an independent dataset comprising chest radiographs with radiologist-reported findings indicating the presence or absence of TB-associated abnormalities. This retrospective evaluation focuses on assessing key performance metrics, including sensitivity, specificity, accuracy, and the area under the receiver operating characteristic curve (ROC-AUC), to determine the tool's capability in flagging potential TB- associated findings on unseen chest radiographs [9].

Beyond screening, the CXR Insightz tool is designed with a potential feature to automatically alert

the nearest expert, enabling prompt resource allocation and patient management. This capability enhances its potential as an effective triage tool, especially in settings where time-sensitive responses are critical.

Objective -

The primary objective of this pilot study is to evaluate the diagnostic performance of CXR Insightz, an AI-based screening tool developed to flag chest X-rays with radiographic abnormalities suggestive of tuberculosis. The tool's outputs were validated against radiologist-reported findings, with performance assessed using metrics such as sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve (AUC).

Intended Use:

CXR Insightz is designed to serve as a screening aid with Portable and stationary Xray units, alerting clinicians to the presence of specific radiographic pathologies commonly associated with TB, thereby supporting timely decision-making in clinical and resource-limited settings.

Materials and Methods –

Pathologies Considered and Rationale

The AI-based screening tool evaluated in this study detects specific chest X-ray abnormalities commonly associated with pulmonary tuberculosis. The following radiographic findings were included in the analysis:

Consolidation, Pleural Effusion, Pulmonary Nodule(s) Fibrosis, and Costophrenic (CP) Angle Blunting, calcified nodules, tuberculosis.

While these pathologies are not exclusive to TB, they are frequently observed in active or healed pulmonary TB cases [1]. The AI model flags a CXR as 'positive' if one or more of these pathologies are detected. This enables the tool to serve as an early triage mechanism — prompting further clinical evaluation in settings where immediate radiologist review may be unavailable.

Study Setting and Dataset Description:

This validation study utilized a dataset of chest X-ray images sourced from an open-access repository. The dataset comprised radiographs of individuals aged 15 years and above, following international recommendations for AI-based TB screening tools [10].

Cases were selected based on the availability of radiologist-reported findings. Each case was annotated for the presence or absence of the six target pathologies described above. Additionally, normal CXRs without any pathological findings were included to evaluate the model's ability to distinguish between pathological and non-pathological scans.

No additional exclusion criteria (beyond age <15 years) were explicitly defined by the dataset source. Ethical approval documentation was not publicly available but is presumed to align with the respective repository's guidelines for research and reuse.

AI Model / Tool Description:

Brief description of the AI-based TB screening tool – model/s used.

Mention if the model uses machine learning / deep learning / CNN architecture (expand).

The tool was designed for automated analysis of posteroanterior chest radiographs.

Include information about image pre-processing (e.g., resizing, normalization) and inference pipeline.

a. Model Development and Architecture

CXR Insightz integrates deep convolutional neural networks (CNNs) trained on a large proprietary dataset of approximately 100,000 chest X-ray images. For each of the four key TB-associated

radiographic findings—consolidation, nodules, pleural effusion, and fibrosis—three independent binary classification models were developed using the following architectures:

ResNet-152 DenseNet-121 EfficientNetV2-M

Each model was optimized to classify the presence or absence of a specific pathology, enabling a focused and pathology-specific screening output.

b. Ensemble Strategy

To enhance diagnostic robustness and mitigate architecture-specific biases, an ensemble approach was employed. For each pathology, outputs from the three corresponding models were aggregated using a majority voting mechanism. A pathology was flagged as present if at least two out of the three models agreed on its presence. This ensemble design helped reduce false positives while maintaining high sensitivity and generalizability.

c. Deployment for Validation

Following training, all models were frozen and deployed in inference mode without any retraining or fine-tuning on the external validation dataset. For this study, CXR Insightz was evaluated strictly in a validation-only setting, with outputs directly compared to radiologist-annotated findings to assess diagnostic alignment.

d. Preprocessing and Inference Pipeline

Prior to inference, each chest X-ray was subjected to a standardized preprocessing pipeline. Images were resized to a consistent input dimension suitable for each architecture and normalized to ensure uniform pixel value distribution. No patient-identifiable metadata was used. The tool automatically routed each input through the ensemble models associated with each pathology, generated individual and aggregated predictions, and returned a binary output indicating the presence or absence of TB-associated abnormalities.

Validation and Testing Process:

The screening performance of the AI tool was evaluated using a multi-step validation strategy. A confusion matrix was first constructed to illustrate the distribution of true positives, true negatives, false positives, and false negatives. This helped visualize how well the tool could distinguish between CXRs with and without TB-associated radiographic abnormalities.

Subsequently, key performance metrics were calculated, including sensitivity (recall), specificity, precision, F1 score, accuracy, along with false positive and false negative rates. These metrics provided a detailed assessment of the tool's ability to flag relevant findings while minimizing incorrect alerts.

To evaluate its ability to discriminate across thresholds, Receiver Operating Characteristic (ROC) curve analysis was conducted. The Area Under the Curve (AUC) was computed with a 95% confidence interval to quantify overall screening effectiveness.

Additionally, a statistical agreement between the AI outputs and radiologist annotations was measured using Cohen's Kappa coefficient, with a p-value to assess the significance of the agreement. An error analysis was also carried out by reviewing false positive and false negative cases, highlighting areas where the tool may require further refinement for more reliable triage.

Results-

Dataset Overview

The validation dataset comprised 170 chest X-rays (CXRs) of individuals aged 15 years and above, retrospectively sourced from an open-access repository. Of these, 88 cases were annotated by expert radiologists as having one or more radiographic findings suggestive of tuberculosis. The remaining 82 cases were reported as normal, with no significant pathological abnormalities.

CXR Insightz's predictions were evaluated against these radiologist-annotated ground truth labels. The model correctly identified 88 abnormal cases (true positives) and 71 normal cases (true negatives). There were 11 false positives, where normal cases were incorrectly flagged as abnormal, and 0 false negatives, where TB-suggestive abnormalities were missed.

Despite the false positives, the tool demonstrated strong classification performance—accurately identifying all pathology cases while maintaining a manageable false positive rate. These results highlight the model's potential utility as an effective triage aid in TB screening workflows, particularly in resource-constrained settings.

Screening Performance Metrics

e. Core Metrics

CXR Insightz demonstrated strong performance on the external validation dataset, achieving an overall accuracy of 93.53%, reflecting its reliability in classifying both TB-suggestive and normal chest X-rays.

The model achieved a precision of 88.89%, indicating a low false positive rate, and a recall (sensitivity) of 100%, confirming its ability to correctly identify all cases with TB-suggestive abnormalities. The specificity—the model's capacity to accurately flag normal cases—was 86.59%.

The F1 score, which balances precision and recall, was 94.12%, underscoring the model's robustness in maintaining both sensitivity and precision. The false positive rate (FPR) and false negative rate (FNR) were 13.41% and 0.00%, respectively—both within acceptable thresholds for screening tools, with zero missed abnormal being a key highlight.

A 95% confidence interval (CI) for accuracy can be calculated as approximately 88.8%–96.53%.

f. ROC Curve Analysis:

Receiver Operating Characteristic (ROC) analysis was conducted to assess the model's classification performance. The model achieved an Area Under the Curve (AUC) of 0.9329, indicating a good balance between sensitivity and specificity. The ROC curve demonstrated that CXR Insightz could reliably distinguish between TB-suggestive and normal cases, reinforcing its potential utility as a screening tool in clinical workflows.

Model Architecture – Internal Inference

Internal assessments during development revealed architecture-specific performance nuances across the four TB-related pathologies. Each of the three convolutional neural networks (CNN) models—ResNet-152, DenseNet-121, and EfficientNetV2-M—demonstrated unique strengths:

Nodule detection showed the highest sensitivity when evaluated using EfficientNetV2-M, making it especially effective in identifying subtle radiological findings. For consolidation, DenseNet-121 consistently yielded higher specificity, contributing to reduced false positives.

All three architectures performed comparably well for pleural effusion, reflecting consistent model behavior across the board.

Fibrosis presented the greatest challenge across individual models; however, ensemble averaging significantly improved the reliability of its detection by compensating for model-specific errors.

These insights informed the final ensemble strategy, where majority voting across the three models helped leverage their complementary strengths while reducing individual biases. This contributed to the tool's overall robustness in screening for TB-suggestive abnormalities.

Statistical Agreement and Validation

To assess the AI model's concordance with radiologist interpretations, Cohen's Kappa statistic was calculated, yielding a value of 0.8698, indicating

strong agreement beyond chance. This underscores the model's reliability as a screening support tool.

Further, McNemar's test (exact version) was applied to evaluate statistical differences between the AI tool's classifications and ground truth annotations. The resulting p-value of 0.001, indicating a significant difference driven by the AI's conservative bias. All errors were false positives, reflecting the model's prioritization of sensitivity to avoid missed TB cases—an advantageous trait for a screening tool.

An analysis of misclassified cases revealed 0 false negatives and 11 false positives. While the overall misclassification rate was low, false positives may lead to additional reviews, but the absence of false negatives is particularly valuable in TB screening due to the public health risks of missed cases.

These findings highlight the importance of deploying the AI tool as an assistive triage aid, supplementing rather than replacing clinical judgment.

The ensemble predictions evaluated in this study were based on model configurations pre- optimized during internal development, incorporating architecture-specific insights to enhance screening reliability.

The evaluation findings demonstrate that the CXR Insightz exhibits strong screening performance in detecting chest X-ray pathologies suggestive of tuberculosis. The observed high agreement with radiologist annotations supports its potential utility as a triage aid in TB screening workflows.

However, the presence of false negatives emphasizes the continued need for clinical oversight, reinforcing the tool's role as a supportive, rather than standalone, solution.

Discussion

This study evaluated the diagnostic performance of CXR Insightz, to assist in the screening of pulmonary tuberculosis (TB) using chest radiographs. The model demonstrated satisfactory performance in detecting pathological findings suggestive of TB, with an overall accuracy of 93.53%.

The performance metrics and statistical agreement (Cohen's Kappa = 0.8698) indicate good consistency between the model's predictions and radiologist-confirmed findings.

McNemar's test yielded a p-value of 0.001, indicating a significant difference driven by the AI's conservative bias. All errors were false positives, reflecting the model's prioritization of sensitivity to avoid missed TB cases—an advantageous trait for a screening tool.

These results underscore the potential of CXR Insightz to function as a supporting TB screening tool, especially in settings where radiological expertise is limited or not readily available.

Internal development assessments revealed that the ensemble of three architectures—ResNet- 152, DenseNet-121, and EfficientNetV2-M—enhanced predictive performance. Each model contributed unique strengths to the ensemble: for instance, EfficientNetV2-M demonstrated higher recall, favouring sensitivity, whereas ResNet-152 yielded more conservative predictions with a lower false positive rate. The architecture-specific contributions collectively improved the model's

robustness across diverse cases. These insights were integral to the final configuration used in the present evaluation and highlight the ensemble's advantage in balancing performance trade-offs.

Several studies have explored deep learning models for directly classifying tuberculosis on chest X- rays. Lakhani and Sundaram (2017) used an ensemble of AlexNet and GoogLeNet, achieving an AUC of 0.99. Their study also highlighted that human-AI collaboration improves diagnostic accuracy, as radiologists reviewing AI-disagreement cases helped refine results [11]. Another study in 2019 proposed an efficient deep network optimized for low-resource settings, reporting an AUC of 0.973–0.991 [12]. More recently, Acharya et al. (2022) developed a normalization-free network model that achieved an AUC of 0.9938, with an accuracy of 96.91% and a sensitivity of 91.81%. Their work also included visual explanation techniques to improve model interpretability [13]. While these models focused on direct image-level classification, our tool is designed as a screening system that flags potential TB cases based on pathology findings, assisting in triage rather than final diagnosis.

Some studies introduced more complex architectures and optimization techniques to improve tuberculosis detection from chest X-rays. Rajakumar et al. (2021) combined deep features from VGG16 and VGG19 models, further optimized using the Mayfly algorithm, and reported an accuracy of 97.80% [14]. Similarly, Sathitratanacheewin et al. (2021) proposed a multi-instance learning model that analyses overlapping image patches through convolutional and complex networks, achieving an AUC of 99.00% [15]. Another study, DenResCov-19 (2021), applied a combination of DenseNet and ResNet architectures for multi-class classification, including tuberculosis, and demonstrated good performance [16]. Our study, though constrained by dataset size, leveraged an ensemble of three high-performing architectures—ResNet-152, DenseNet-121, and EfficientNetV2-M—whose pre- optimized configurations were designed to enhance reliability in real-world screening scenarios.

This design choice allowed us to test the tool's screening potential in a practical setting while keeping the workflow straightforward.

Approaches focusing on detecting multiple tuberculosis-related abnormalities are closely aligned with our pathology-based model. A study by Xie et al. (2020) developed a computer-aided system capable of identifying various TB-related lesions, such as nodules, calcifications, and pleural effusions. Their two-stage model combined lung segmentation with lesion detection, reporting good diagnostic performance [17]. While their approach provided detailed lesion-level outputs, it also involved a more complex workflow. In

contrast, our study focused on screening for the presence of any suggestive pathology, simplifying the process while still aiming to assist in identifying potential TB cases, especially in settings where quick triage is needed.

Integrating AI tools into clinical workflows has been a focus of some studies aiming to support healthcare providers, especially in resource-limited settings. CheXaid (2020) combined chest X-ray findings with clinical variables to assist physicians in diagnosing tuberculosis among HIV-positive patients. Their model not only improved physician accuracy but also performed better as a stand- alone screening tool **[18, 19].** Similar to this direction, our tool is designed with the intent to assist clinicians by flagging chest X-rays with pathologies suggestive of TB.

While this pilot study offers initial insights into the utility of an AI-based screening tool for tuberculosis, certain limitations must be acknowledged. The model was validated on a relatively small dataset of 170 chest X-rays, reduced from an initial larger set due to data quality and availability constraints, which may limit its generalizability across broader and more diverse populations. This smaller sample size could reduce statistical power, increasing the risk of overfitting to specific radiographic patterns and potentially underrepresenting rare or atypical TB presentations, demographic variations (e.g., age, ethnicity, or comorbidities), and image quality differences from real-world sources. As a result, the observed high sensitivity (100%) and overall performance metrics may not fully translate to larger, heterogeneous cohorts, where variability in TB manifestations or scanner types could affect detection accuracy.

Since the tool operates through pathology-based detection rather than direct diagnostic classification, there is a risk of missing TB cases that present atypically or share radiographic features with other conditions. Additionally, the ground truth was based solely on radiologist interpretations without microbiological confirmation, potentially introducing subjectivity. The absence of preprocessing methods such as bone suppression may also affect the model's ability to identify subtle pathological cues. Another limitation is the lack of integrated clinical data, which—if included—could enhance the screening accuracy. Demographic variability across populations could further impact generalizability and model performance. These observations provide valuable direction for future work, including expanding dataset size and diversity to thousands of cases from multi-site sources, exploring preprocessing enhancements, integrating clinical parameters, and conducting prospective real-world testing in varied demographic settings to better validate robustness and clinical utility.

While this pilot study offers initial insights into the utility of an AI-based screening tool for tuberculosis, certain limitations must be acknowledged. The model

was validated on a relatively small dataset, which may limit its generalizability across broader and more diverse populations. Since the tool operates through pathology-based detection rather than direct diagnostic classification, there is a risk of missing TB cases that present atypically or share radiographic features with other conditions. Additionally, the ground truth was based solely on radiologist interpretations without microbiological confirmation, potentially introducing subjectivity. The absence of preprocessing methods such as bone suppression may also affect the model's ability to identify subtle pathological cues. Another limitation is the lack of integrated clinical data, which—if included—could enhance the screening accuracy. Demographic variability across populations could further impact generalizability and model performance. These observations provide valuable direction for future work, including expanding dataset size and diversity, exploring preprocessing enhancements, integrating clinical parameters, and conducting real-world testing in varied demographic settings **[20].**

Despite these limitations, the study demonstrates the potential of a simplified, pathology-based AI tool to assist in tuberculosis screening through chest X-rays. Designed for quick identification of possible cases, the model offers a practical solution for settings with limited resources and radiological expertise. While preliminary, the study provides a transparent evaluation of the tool's performance and lays the groundwork for future improvements. Expanding datasets, integrating clinical information, and refining the model's capabilities could enhance its utility as a supportive triage tool in real-world screening workflows.

Conclusion

This study presents the development and evaluation of CXR Insightz, a customized AI-based tool designed to assist in tuberculosis screening using chest X-rays. The model demonstrated promising diagnostic performance, achieving an accuracy of 93.53% with strong agreement compared to radiologist assessments. These findings suggest its potential as a supportive screening aid, particularly in resource-limited settings.

However, as CXR Insightz operates by identifying radiologically visible pathologies, it is not intended to replace clinical judgment or microbiological confirmation. Given the diagnostic complexity of tuberculosis, cautious interpretation of AI-generated outputs remains essential. Nonetheless, tools like this can play a valuable role in large-scale screening efforts by flagging potentially suspicious cases for further evaluation. Future validation on larger, more diverse datasets and in real-world clinical settings will be crucial to enhance its generalizability, reliability, and practical impact.

Statements

study at the Dr D.Y. Patil Medical College, Hospital and Research center under the MOU between Gen Z AI Labs Pvt Ltd and DPU Foundation for Innovation Incubation & Entrepreneurship.

Statement of Ethics

Study approval statement: *This study protocol was reviewed and* has been granted an exemption from requiring ethics approval by Institutional Ethics Subcommittee via letter bearing reference no- I.E.S.C./W/64/2025 dated 13/10/2025.

Consent to participate statement: As NO human participants were involved- only anonymized images were used hence written informed consent was not required and hence the study has been granted an exemption from requiring written informed consent by the Institutional Ethics Subcommittee.

Conflict of Interest Statement The authors have no conflicts of interest to declare.

Data Availability Statement

The datasets generated and analysed during the study are included in this published article. Further enquiries can be directed to the corresponding authors.

References

1. The Radiology Assistant: Imaging findings in TB. Accessed March 24, 2025. https://radiologyassistant.nl/chest/tb/tuberculosis

2. Kant S, Srivastava MM. Towards Automated Tuberculosis detection using Deep Learning. In: *2018 IEEE Symposium Series on Computational Intelligence (SSCI).* ; 2018:1250-1253. doi:10.1109/SSCI.2018.8628800

3. Geric C, Qin ZZ, Denkinger CM, et al. The rise of artificial intelligence reading of chest X-rays for enhanced TB diagnosis and elimination. *Int J Tuberc Lung Dis.* 2023;27(5):367-372. doi:10.5588/ijtld.22.0687

4. Qin ZZ, Ahmed S, Sarker MS, et al. Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms. *Lancet Digit Health.* 2021;3(9):e543-e554. doi:10.1016/S2589-7500(21)00116-3

5. Jaeger S, Karargyris A, Candemir S, et al. Automatic tuberculosis screening using chest radiographs. *IEEE Trans Med Imaging.* 2014;33(2):233-245. doi:10.1109/TMI.2013.2284099

6. Hansun S, Argha A, Liaw ST, Celler BG, Marks GB. Machine and Deep Learning for Tuberculosis Detection on Chest X-Rays: Systematic Literature Review. *J Med Internet Res.* 2023;25:e43154. doi:10.2196/43154

7. Qin ZZ, Ahmed S, Sarker MS, et al. Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms. *Lancet Digit Health.* 2021;3(9):e543-e554. doi:10.1016/S2589-7500(21)00116-3

1. TB-Report_for-Web_08_10-2024-1.pdf. Accessed March 24, 2025. https://tbcindia.mohfw.gov.in/wp-content/uploads/2024/10/TB-Report_for-Web_08_10-2024-pdf

8. Zhan Y, Wang Y, Zhang W, Ying B, Wang C. Diagnostic Accuracy of the Artificial Intelligence Methods in Medical Imaging for Pulmonary Tuberculosis: A Systematic Review and Meta- Analysis. *J Clin Med.* 2022;12(1):303. doi:10.3390/jcm12010303

9. World Health Organization. *WHO Consolidated Guidelines on Tuberculosis. Module 2: Screening - Systematic Screening for Tuberculosis Disease.* World Health Organization; 2021.

10. Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology.* 2017;284(2):574-582. doi:10.1148/radiol.2017162326

11. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization | Scientific Reports. Accessed March 24, 2025. https://www.nature.com/articles/s41598-019-42557-4#Tab2

12. Vasundhara Acharya, Dhiman G, Prakasha K, et al. AI-Assisted Tuberculosis Detection and Classification from Chest X-Rays Using a Deep Learning Normalization-Free Network Model. *Comput Intell Neurosci.* 2022;2022(1):2399428. doi:10.1155/2022/2399428

13. Rajakumar MP, Sonia R, Uma Maheswari B, Karuppiah SP. Tuberculosis detection in chest X-ray using Mayfly-algorithm optimized dual-deep-learning features. *J X-Ray Sci Technol.* 2021;29(6):961-974. doi:10.3233/XST-210976

14. Sathitratanacheewin S, Sunanta P, Pongpirul K. Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. *Heliyon.* 2020;6(8):e04614. doi:10.1016/j.heliyon.2020.e04614

15. DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays - ScienceDirect. Accessed March 25, 2025. https://www.sciencedirect.com/science/article/pii/S0895611121001579?via%3Dihub

16. Xie Y, Wu Z, Han X, et al. Computer-Aided System for the Detection of Multicategory Pulmonary Tuberculosis in Radiographs. *J Healthc Eng.* 2020;2020(1):9205082. doi:10.1155/2020/9205082

17. Rajaraman S, Zamzmi G, Folio L, Alderson P, Antani S. Chest X-ray Bone Suppression for Improving

Classification of Tuberculosis-Consistent Findings. *Diagnostics.* 2021;11(5):840. doi:10.3390/diagnostics11050840

18. Rajpurkar P, O'Connell C, Schechter A, et al. CheXaid: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. *Npj Digit Med.* 2020;3(1):1-8. doi:10.1038/s41746-020-00322-2

19. Sathitratanacheewin S, Sunanta P, Pongpirul K. Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. *Heliyon.* 2020;6(8):e04614. doi: 10.1016/j.heliyon. 2020.e04614
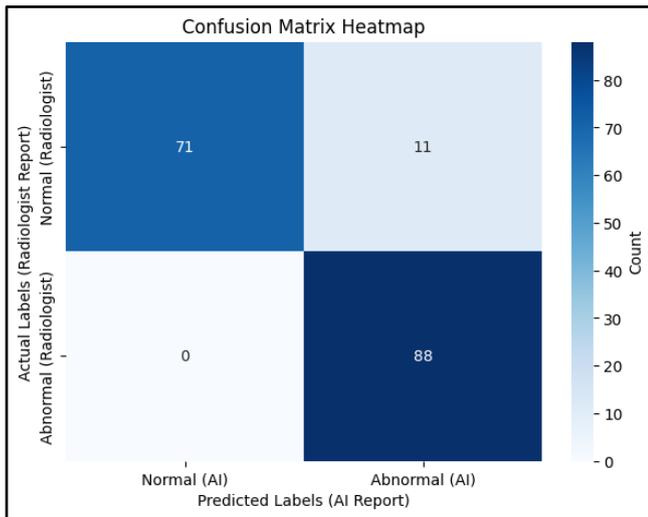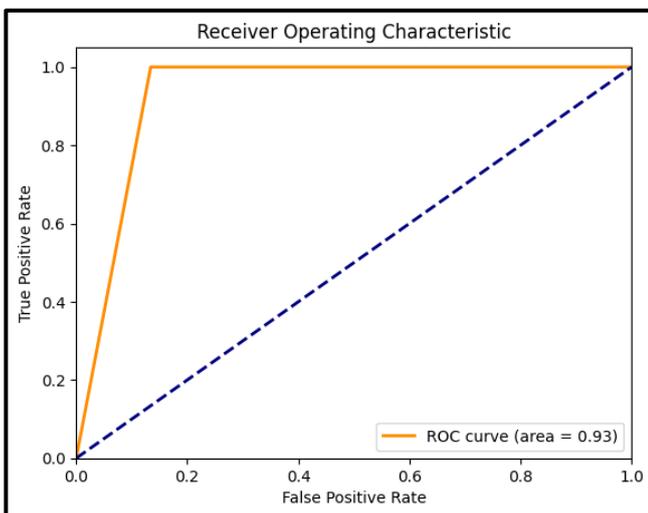
Fig. 1. Confusion Matrix Heatmap

Fig. 2: ROC Curve