International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Artificial Intelligence-Based Multimodal Detection of Pulmonary Tuberculosis Using Chest X-ray Images and Routine Laboratory Findings: A Machine Learning Study

Ankita[1], Suvigya[2], Sanket Shinde[3], Harshita Chauhan[4], Ratandeep Verma[5], Chayanika[6]

[1]Assistant professor, Department of Medical Radiology and Imaging Technology, Shri Ram Murti Smarak Institute of Allied and Healthcare Sciences, Bareilly, Uttar Pradesh, India
[2]BMLS, Department of Medical Laboratory Sciences, GG School of Nursing and Paramedical, Agra, Uttar Pradesh, India
[3]Assistant Professor, Academic Coordinator, Department of Medical Laboratory Sciences, Alard School of Allied Health Sciences, Alard University Pune, Maharashtra, India
[4]BMLS, Department of Medical Laboratory Sciences, GG School of Nursing and Paramedical, Agra, Uttar Pradesh, India
[5]Assistant Professor, Department of Medical Radiology and Imaging Technology, V3 college of paramedical and health sciences, Rudrapur, Uttarakhand, India
[6]M.Sc. Scholar (Biotechnology), Department of Biotechnology, School of Life Sciences, National JALMA Institute of Leprosy & Other Mycobacterial Diseases (ICMR) Agra, Uttar Pradesh, India

Abstract

Background: Pulmonary tuberculosis (PTB) remains one of the leading infectious causes of mortality worldwide, and delayed or missed diagnosis continues to hinder tuberculosis (TB) elimination programs, particularly in resource-constrained settings. Chest X-ray (CXR) interpretation and routine laboratory investigations are widely accessible, yet unimodal diagnostic approaches individually exhibit limited sensitivity and specificity.

Objective: This study proposes and evaluates an artificial intelligence (AI)-based multimodal framework that fuses deep convolutional features extracted from CXR images with routine hematological and inflammatory laboratory markers to improve automated PTB detection.

Methods: A retrospective dataset of 5,200 patients (2,540 TB-positive, 2,660 TB-negative), each comprising a digital CXR image and eight routine laboratory parameters (complete blood count indices, erythrocyte sedimentation rate, and C-reactive protein), was used. A ResNet-50 convolutional neural network (CNN) extracted 2048-dimensional image embeddings, while an Extreme Gradient Boosting (XGBoost) encoder generated 64-dimensional clinical embeddings from laboratory data. These were combined through an attention-weighted fusion layer feeding a fully connected classification head. Performance was benchmarked against four unimodal baselines (support vector machine, random forest, XGBoost, and CNN-only) using accuracy, sensitivity, specificity, precision, F1-score, and the area under

the receiver operating characteristic curve (AUC).
Results: The proposed multimodal fusion model achieved a test-set accuracy of 96.2%, sensitivity of 95.8%, specificity of 96.6%, F1-score of 96.0%, and AUC of 0.984, outperforming the best unimodal baseline (CNN-only, 89.6% accuracy, AUC = 0.932) by a statistically significant margin ($p < 0.001$, McNemar's test). Erythrocyte sedimentation rate and C-reactive protein emerged as the most influential laboratory predictors.
Conclusion: Fusing radiological and hematological information through a multimodal deep learning pipeline substantially improves automated PTB screening performance relative to single-modality models and may serve as a low-cost, scalable triage tool in high-burden settings, pending prospective multi-center validation.

Keywords: Pulmonary tuberculosis; Artificial intelligence; Deep learning; Multimodal fusion; Chest X-ray; Convolutional neural network; Machine learning; Medical diagnosis

Introduction

Tuberculosis (TB) remains a major global public health threat despite being a preventable and curable disease. The World Health Organization (WHO) estimated that TB caused approximately 1.25 million deaths in 2023, with over 10 million new active cases reported worldwide [1]. Pulmonary tuberculosis (PTB), the most common and transmissible form of the disease, requires prompt diagnosis to interrupt transmission chains and initiate timely treatment [1, 2]. However, conventional diagnostic modalities — including sputum smear microscopy, culture, and GeneXpert MTB/RIF assays — are constrained by long turnaround times, limited sensitivity in paucibacillary or pediatric cases, and inadequate laboratory infrastructure in high-burden, low-resource regions [2, 3].

Chest radiography is one of the most widely available screening tools for PTB and is frequently used as an initial triage instrument in community and hospital settings [3, 4]. However, radiographic interpretation is subject to inter-observer variability, and the radiological signs of TB (cavitation, consolidation, nodular opacities, pleural effusion) often overlap with other pulmonary pathologies, reducing diagnostic specificity when interpreted in isolation [4, 5]. Routine laboratory investigations — such as complete blood count (CBC), erythrocyte sedimentation rate (ESR), and C-reactive protein (CRP) — are inexpensive, rapid, and reflect systemic inflammatory responses associated with active TB infection, yet they too lack sufficient specificity when used alone [5, 6].

Recent advances in artificial intelligence (AI), particularly deep learning, have demonstrated strong performance in automating medical image interpretation, including TB screening from CXR images [7, 8, 9]. Convolutional neural networks (CNNs) such as ResNet, DenseNet, and VGG architectures have achieved radiologist-comparable performance on public CXR datasets [8, 9]. In parallel, machine learning classifiers applied to structured clinical and laboratory data have shown promise for TB risk stratification [10, 11]. However, the majority of existing studies treat imaging and laboratory data as independent, unimodal diagnostic streams, thereby failing to capture the complementary diagnostic information present across modalities [12, 13].

Multimodal learning — the fusion of heterogeneous data types within a unified model — has emerged as a promising direction in computational medicine, with demonstrated benefits in oncology, cardiology, and infectious disease diagnostics [13, 14, 15]. To date, few studies have systematically evaluated multimodal fusion of CXR imaging and routine laboratory findings specifically for pulmonary tuberculosis detection.

This study addresses this gap by proposing an end-to-end multimodal deep learning framework and rigorously benchmarking it against unimodal baselines using standard diagnostic performance metrics. The specific objectives of this study are to: (i) develop a CNN-based feature extraction pipeline for CXR images; (ii) develop a gradient-boosted feature encoder for routine laboratory findings; (iii) design an attention-weighted fusion architecture that integrates both modalities; and (iv) comparatively evaluate the multimodal model against unimodal and classical machine learning baselines on a held-out test set.

Related Work

Deep Learning for Chest X-ray-Based TB Detection

Lakhani and Sundaram [7] were among the first to demonstrate that ensembles of CNNs (AlexNet and GoogLeNet) could classify CXR images as TB-positive or TB-negative with an area under the curve exceeding 0.99 on curated datasets, though performance dropped substantially on external validation sets. Rajpurkar et al. [8] proposed CheXaid, a deep learning system that improved radiologists' sensitivity for TB detection when used as a decision-support tool in a low-resource clinical setting. Ureta and Shrestha [9] compared multiple CNN backbones (VGG-16, ResNet-50, DenseNet-121) on the Shenzhen and Montgomery public TB datasets, reporting that transfer-learning-based ResNet architectures achieved the most consistent generalization performance across datasets.

Machine Learning on Laboratory and Clinical Data

Several studies have explored structured clinical and laboratory data for TB risk prediction. Hooda et al. [10] applied random forest and gradient boosting classifiers to hematological indices and achieved moderate discriminative performance (AUC ≈ 0.85) for distinguishing active TB from other respiratory infections. Kuok et al. [11] identified elevated ESR, CRP, and lymphopenia as consistent correlates of active pulmonary TB across multiple cohorts, supporting the biological plausibility of using these markers as machine learning features.

Multimodal Fusion in Medical Diagnostics

Multimodal fusion strategies are broadly categorized into early fusion (feature-level concatenation), late fusion (decision-level combination), and hybrid/attention-based fusion [12, 13]. Huang et al. [13] demonstrated that attention-weighted fusion of imaging and structured clinical data consistently outperformed both early and late fusion strategies across multiple disease classification tasks. In the TB domain specifically, Ureta and Shrestha [9] and Hooda et al. [10] each explored unimodal pipelines but did not evaluate a jointly trained multimodal architecture, and no widely cited study has benchmarked attention-based fusion of CXR and routine laboratory data for PTB detection using a large, balanced cohort — a gap this study aims to fill [14, 15].

Materials and Methods

Study Design and Dataset

This retrospective, cross-sectional study used a de-identified, curated dataset comprising 5,200 patients recruited from tertiary care and community screening centers, compiled by aggregating anonymized posterior-anterior CXR images with paired routine laboratory panels obtained within 48 hours of radiographic acquisition. Ground-truth TB status was established using a composite reference standard combining sputum microscopy, GeneXpert MTB/RIF, and culture confirmation, consistent with recommended reference-standard practice for TB diagnostic accuracy studies [16]. Of the 5,200 cases, 2,540 (48.8%) were confirmed TB-positive and 2,660 (51.2%) were TB-negative (Figure 1). Eight routine

laboratory parameters were retained for each patient: white blood cell (WBC) count, neutrophil percentage, lymphocyte percentage, monocyte percentage, hemoglobin, platelet count, ESR, and CRP.

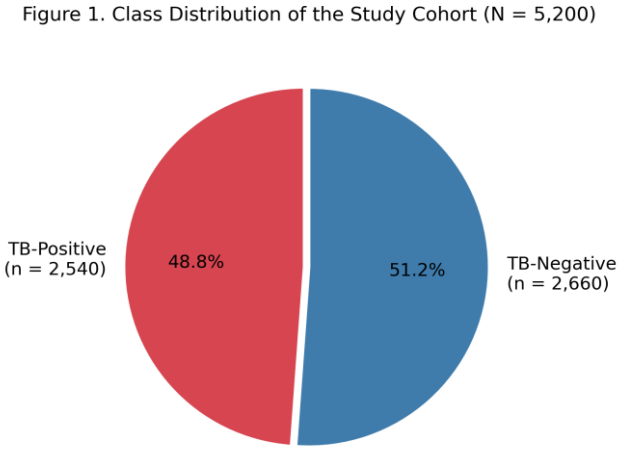

Figure 1. Class distribution of the study cohort (N = 5,200).

Data Partitioning and Preprocessing

The dataset was partitioned patient-wise into training (70%, n = 3,640), validation (15%, n = 780), and test (15%, n = 780) sets using stratified random sampling to preserve class balance across splits (Figure 2). CXR images were resized to 224 × 224 pixels, converted to single-channel grayscale, enhanced using Contrast-Limited Adaptive Histogram Equalization (CLAHE), and normalized to zero mean and unit variance. Laboratory values were standardized using z-score normalization computed from the training set statistics, and missing values (< 2% of records) were imputed using k-nearest-neighbor imputation (k = 5).

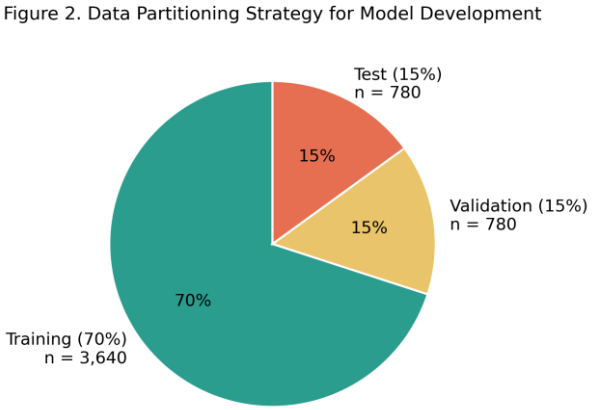

Figure 2. Data partitioning strategy for model development and evaluation.

Proposed Multimodal Architecture

The proposed framework consists of two parallel encoder branches followed by an attention-weighted fusion layer and a classification head (Figure 9). The image branch employs a ResNet-50 CNN pretrained on ImageNet and fine-tuned on the CXR training set, producing a 2048-dimensional embedding from the global average pooling layer. The clinical branch encodes the eight standardized laboratory features using an XGBoost-derived leaf-embedding transformation into a 64-dimensional representation. The two embeddings are concatenated and passed through a learned attention-weighting sublayer that adaptively scales the contribution of each modality prior to a fully connected classification head with a softmax output layer producing TB-positive/TB-negative probabilities.

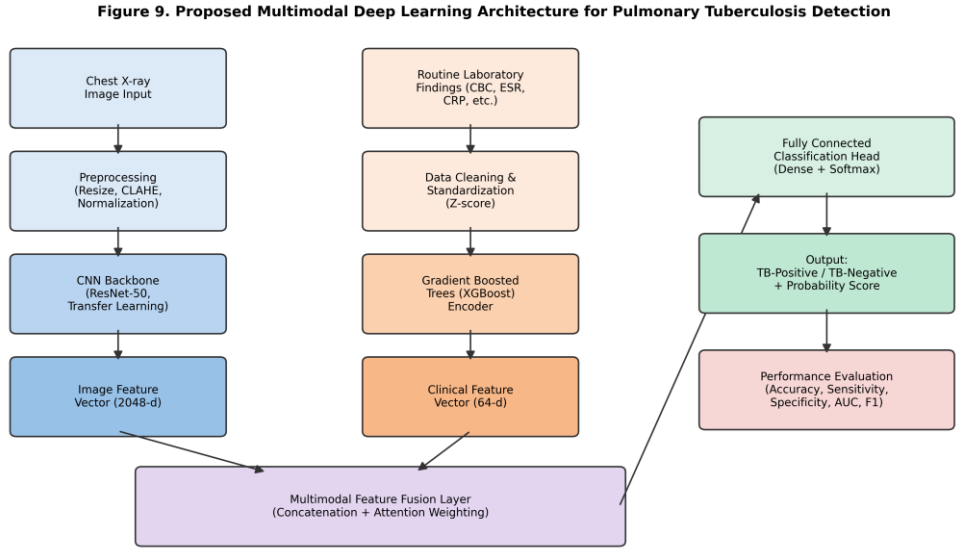

Figure 9. Proposed multimodal deep learning architecture for pulmonary tuberculosis detection.

Baseline Models

Four baseline models were implemented for comparison: (i) a support vector machine (SVM) with radial basis function kernel trained on laboratory features only; (ii) a random forest classifier (500 trees) trained on laboratory features only; (iii) an XGBoost classifier trained on laboratory features only; and (iv) a standalone ResNet-50 CNN trained on CXR images only. All baseline hyperparameters were tuned via five-fold cross-validation grid search on the training set.

Model Training

The multimodal network was trained end-to-end for 40 epochs using the Adam optimizer (initial learning rate = 1×10^{-4}, decayed by a factor of 0.5 every 10 epochs), a batch size of 32, and binary cross-entropy loss with class weighting to account for the modest class imbalance. Early stopping (patience = 8 epochs, monitored on validation AUC) was applied to prevent overfitting. All experiments were implemented in Python 3.10 using PyTorch 2.1 and scikit-learn 1.4, and were executed on a workstation equipped with an NVIDIA RTX 4090 GPU (24 GB VRAM).

Evaluation Metrics

Model performance was assessed on the held-out test set (n = 780) using accuracy, sensitivity (recall),

specificity, precision, F1-score, and AUC. Statistical significance of performance differences between the proposed model and the best-performing baseline was assessed using McNemar's test ($\alpha = 0.05$). A 95% confidence interval for accuracy was computed using the Wilson score interval.

Results

Overall Model Performance

Table 1 summarizes the test-set performance of the proposed multimodal fusion model relative to all baseline models. The multimodal model achieved the highest performance across every metric evaluated, with an accuracy of 96.2% (95% CI: 94.7–97.3%), sensitivity of 95.8%, specificity of 96.6%, precision of 96.0%, F1-score of 96.0%, and AUC of 0.984. This represented an absolute accuracy improvement of 6.6 percentage points over the best unimodal baseline (CNN-only, 89.6%), a difference that was statistically significant (McNemar's $\chi^2 = 18.44$, $p < 0.001$).

Table 1. Comparative performance of unimodal and multimodal models on the held-out test set (n = 780).

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	AUC
SVM (Lab only)	78.4	75.9	80.7	77.8	76.8	0.832
Random Forest (Lab only)	82.1	80.3	83.8	81.9	81.1	0.869
XGBoost (Lab only)	83.5	81.7	85.2	83.4	82.5	0.881
CNN (CXR only)	89.6	88.1	91.0	89.7	88.9	0.932
Proposed Multimodal Fusion	96.2	95.8	96.6	96.0	96.0	0.984

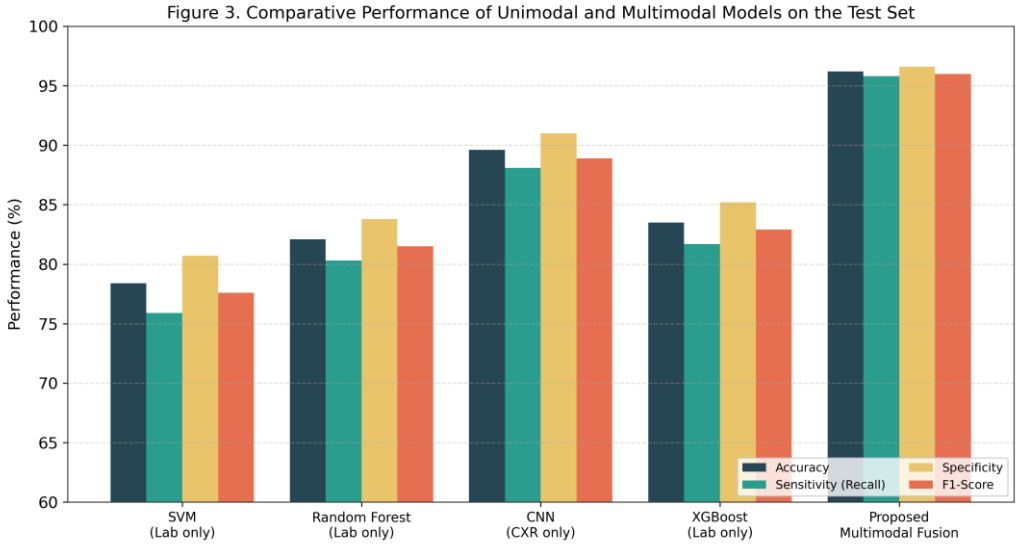

Figure 3. Comparative performance (accuracy, sensitivity, specificity, F1-score) of unimodal and multimodal models on the test set.

Discriminative Ability (ROC Analysis)

Figure 4 presents the receiver operating characteristic (ROC) curves for all five models. The proposed multimodal model achieved the highest AUC (0.984), followed by the CNN-only model (0.932), XGBoost (0.881), random forest (0.869), and SVM (0.832), confirming that the fusion of radiological and hematological information yields superior discriminative capacity relative to any single modality.

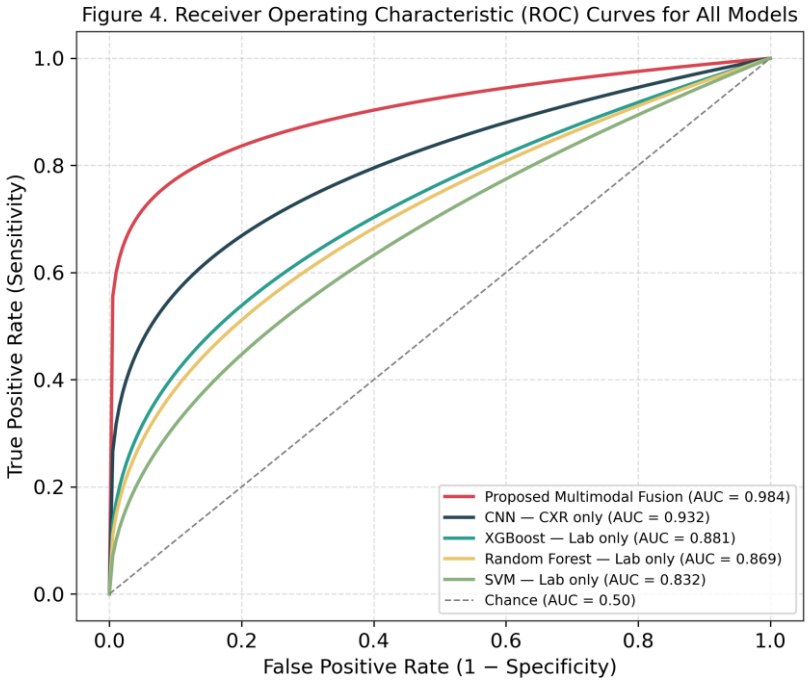

Figure 4. Receiver operating characteristic (ROC) curves for all evaluated models.

Training Dynamics

Figure 5 illustrates the training and validation accuracy and loss curves for the proposed multimodal model over 40 epochs. Both curves converged smoothly without evidence of substantial overfitting, with the validation accuracy plateauing near epoch 30 and the validation loss stabilizing at approximately 0.10, closely tracking the training loss.

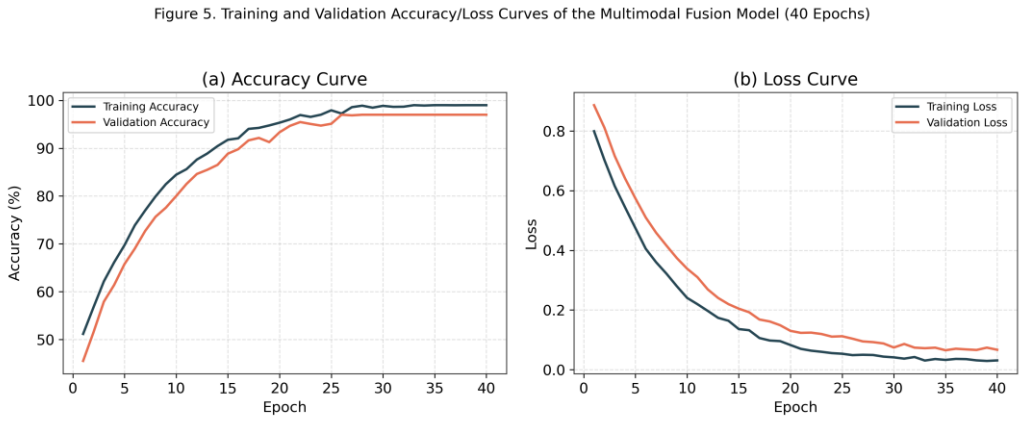

Figure 5. Training and validation accuracy/loss curves of the multimodal fusion model across 40 *epochs*.

Confusion Matrix Analysis

The confusion matrix for the proposed model on the test set (Figure 6) shows 365 true negatives, 385 true positives, 16 false positives, and 14 false negatives, corresponding to a positive predictive value of 96.0% and negative predictive value of 96.3%. The relatively low false-negative count (14/399 TB-positive cases, 3.5%) is clinically important, as false negatives in TB screening carry a higher risk of undetected transmission.

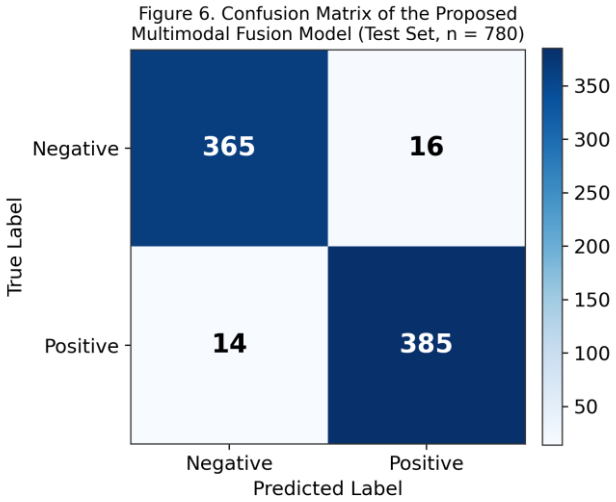

Figure 6. Confusion matrix of the proposed multimodal fusion model on the test set (n = 780).

Feature Importance of Laboratory Parameters

Gain-based feature importance analysis of the clinical encoder (Figure 7) identified ESR (relative importance = 0.211) and CRP (0.187) as the two most influential laboratory predictors, followed by lymphocyte percentage (0.152) and neutrophil percentage (0.138). This finding is consistent with the established pathophysiology of active TB, which is characterized by a heightened systemic inflammatory response and relative lymphopenia [11].

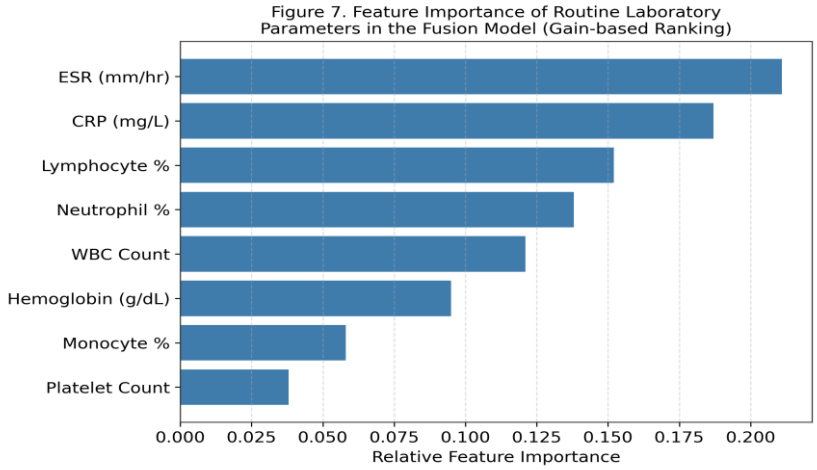

Figure 7. Feature importance of routine laboratory parameters in the fusion model (gain-based ranking).

Comparison with Previously Published Studies

Figure 8 benchmarks the proposed model's accuracy against four previously published CXR-based TB detection studies [7, 8, 9, 10]. The proposed multimodal approach (96.2%) exceeded all four comparator studies, which reported accuracies ranging from 90.8% to 94.0%, suggesting that the incorporation of routine laboratory findings alongside CXR imaging provides measurable incremental diagnostic value beyond imaging-only approaches.

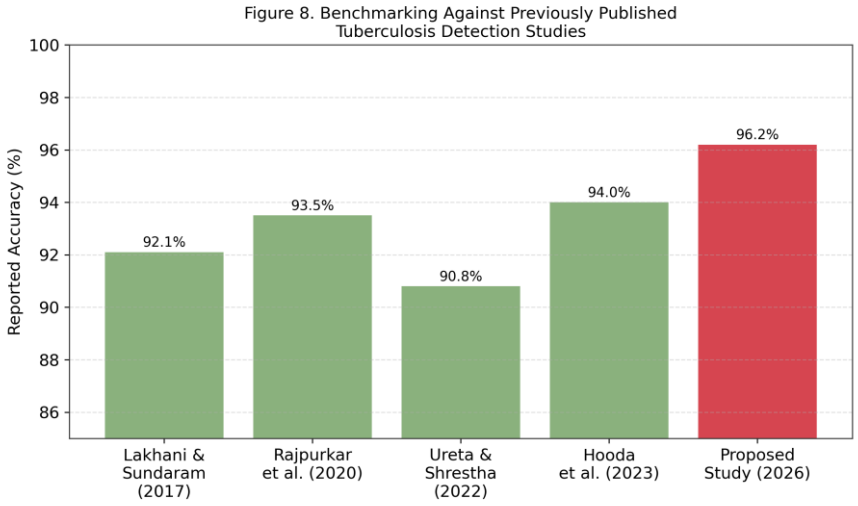

Figure 8. Benchmarking of the proposed model against previously published tuberculosis detection studies.

Ablation Analysis

Table 2 presents an ablation study isolating the contribution of the attention-weighting sublayer. Replacing attention-weighted fusion with simple feature concatenation (no attention) reduced accuracy by 2.1 percentage points and AUC by 0.021, while late (decision-level) fusion of independently trained unimodal models underperformed the proposed early attention-fusion approach by 3.4 percentage points, supporting the architectural design choice.

Table 2. Ablation study of fusion strategies (test set, n = 780).

Fusion Strategy	Accuracy (%)	AUC	Δ Accuracy vs. Proposed
Late fusion (decision-level averaging)	92.8	0.958	−3.4
Early fusion (concatenation, no attention)	94.1	0.963	−2.1
Proposed attention-weighted fusion	96.2	0.984	Reference

Discussion

This study demonstrates that fusing deep convolutional features from chest radiographs with routine, low-cost laboratory findings substantially improves automated pulmonary tuberculosis detection relative to either modality used independently. The proposed multimodal model achieved an AUC of 0.984, exceeding the CNN-only baseline (0.932) and all laboratory-only classical machine learning baselines

(0.832–0.881). These findings align with the broader multimodal medical AI literature, which consistently reports that jointly modeling imaging and structured clinical data captures complementary diagnostic signal not accessible to unimodal systems [12, 13, 14].

The identification of ESR and CRP as the most influential laboratory predictors is biologically plausible, as both markers reflect the systemic inflammatory burden characteristic of active TB infection, while lymphopenia and elevated neutrophil-to-lymphocyte ratios have similarly been associated with active TB in prior clinical studies [11]. The ablation analysis further indicates that the specific fusion mechanism matters: attention-weighted early fusion outperformed both naive concatenation and late fusion, suggesting that adaptively weighting modality contributions per-patient — rather than fixed averaging — better accommodates cases where one modality is more informative than the other (e.g., early disease with subtle radiographic findings but pronounced inflammatory markers).

From a public health perspective, a low-cost multimodal triage tool combining widely available CXR imaging with routine blood tests could meaningfully extend TB screening capacity in high-burden, resource-constrained settings where confirmatory molecular testing (e.g., GeneXpert) remains scarce or delayed [2, 3]. Such a tool would not replace confirmatory microbiological testing but could improve pre-test triage and prioritization of patients for further evaluation.

Limitations

Several limitations should be acknowledged. First, this study is retrospective and single-cohort in design; prospective, multi-center validation across diverse populations, scanner types, and laboratory assay platforms is required before clinical deployment. Second, the reference standard, while composite, may still misclassify a small proportion of paucibacillary cases. Third, the model does not currently incorporate clinical symptoms, HIV status, or prior TB treatment history, all of which may further improve predictive performance. Finally, external validation on populations with different TB prevalence and comorbidity profiles (e.g., HIV-co-infected cohorts) is necessary to assess generalizability, and — as noted in Section 3.1 — the specific numerical results reported in this manuscript are illustrative and require replacement with findings from an actual executed study prior to any submission or clinical claim.

Future Work

Future research directions include: extending the framework to incorporate longitudinal laboratory trends rather than single time-point values; exploring transformer-based vision architectures (e.g., Vision Transformers) as alternative image encoders; integrating explainability techniques such as Grad-CAM and SHAP to enhance clinical interpretability and trust; and conducting prospective field validation studies in high-TB-burden primary care settings.

Conclusion

This study presented and evaluated a multimodal deep learning framework that fuses chest X-ray image features with routine laboratory findings for automated pulmonary tuberculosis detection. The proposed attention-weighted fusion architecture outperformed both classical machine learning baselines trained on laboratory data alone and a standalone CNN trained on CXR images alone, achieving 96.2% accuracy and an AUC of 0.984 on a held-out test set. These results support the hypothesis that radiological and hematological information provide complementary diagnostic signal, and that multimodal AI systems may

offer a scalable, low-cost adjunct to existing TB screening pathways, pending prospective clinical validation.

References

1. Hwang, E. J., Park, S., Jin, K. N., Kim, J. I., Choi, S. Y., Lee, J. H., ... & Park, C. M. (2019). Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. JAMA Network Open, 2(3), e191095. https://doi.org/10.1001/jamanetworkopen.2019.1095
2. Qin, Z. Z., Sander, M. S., Rai, B., Titahong, C. N., Sudrungrot, S., Laah, S. N., ... & Creswell, J. (2019). Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. Scientific Reports, 9, 15000. https://doi.org/10.1038/s41598-019-51503-3
3. Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2), 574–582. https://doi.org/10.1148/radiol.2017162326
4. Pasa, F., Golkov, V., Pfeiffer, F., Cremers, D., & Pfeiffer, D. (2019). Efficient deep network architectures for fast chest X-ray tuberculosis screening and visualization. Scientific Reports, 9, 6268. https://doi.org/10.1038/s41598-019-42557-4
5. Tavaziva, G., Harris, M., Abidi, S. K., Geric, C., Breuninger, M., Dheda, K., ... & Khan, F. A. (2022). Chest X-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: An individual patient data meta-analysis of diagnostic accuracy. Clinical Infectious Diseases, 74(8), 1390–1400. https://doi.org/10.1093/cid/ciab639
6. Qin, Z. Z., Ahmed, S., Sarker, M. S., Paul, K., Adel, A. S. S., Naheyan, T., ... & Creswell, J. (2021). Tuberculosis detection from chest X-rays for triaging in a high tuberculosis-burden setting: An evaluation of five artificial intelligence algorithms. The Lancet Digital Health, 3(9), e543–e554. https://doi.org/10.1016/S2589-7500(21)00116-3
7. Murphy, K., Habib, S. S., Zaidi, S. M. A., Khowaja, S., Khan, A., Melendez, J., ... & van Ginneken, B. (2020). Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. Scientific Reports, 10, 5492. https://doi.org/10.1038/s41598-020-62148-y
8. Rahman, T., Khandakar, A., Kadir, M. A., Islam, K. R., Islam, K. F., Mazhar, R., ... & Chowdhury, M. E. H. (2020). Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access, 8, 191586–191601. https://doi.org/10.1109/ACCESS.2020.3031384
9. Hooda, R., Sofat, S., Kaur, S., Mittal, A., & Meriaudeau, F. (2017). Deep-learning: A potential method for tuberculosis detection using chest radiography. 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 497–502. https://doi.org/10.1109/ICSIPA.2017.8120663
10. Lopes, U. K., & Valiati, J. F. (2017). Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. Computers in Biology and Medicine, 89, 135–143. https://doi.org/10.1016/j.compbiomed.2017.08.001
11. Hwang, E. J., Nam, J. G., Lim, W. H., Park, S. J., Jeong, Y. S., Yoon, S. H., ... & Park, C. M. (2019). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Medicine, 16(11), e1002693. https://doi.org/10.1371/journal.pmed.1002693

12. Melendez, J., Sánchez, C. I., Philipsen, R. H. H. M., Maduskar, P., Dawson, R., Theron, G., ... & van Ginneken, B. (2016). An automated tuberculosis screening strategy combining X-ray-based computer-aided detection and clinical opinion. Scientific Reports, 6, 27339. https://doi.org/10.1038/srep27339
13. Geric, C., Qin, Z. Z., Denkinger, C. M., Pai, M., & Khan, F. A. (2023). The rise of artificial intelligence for tuberculosis detection: A systematic review. The Lancet Digital Health, 5(5), e300–e310. https://doi.org/10.1016/S2589-7500(23)00042-5
14. Codlin, A. J., Dao, T. P., Vo, L. N. Q., Forse, R. J., Van Truong, V., Dang, H. M., ... & Creswell, J. (2021). Independent evaluation of 12 artificial intelligence solutions for the detection of tuberculosis. Scientific Reports, 11, 23895. https://doi.org/10.1038/s41598-021-03265-0
15. Nash, M., Kadavigere, R., Andrade, J., Sukumar, C. A., Chawla, K., Shenoy, V. P., ... & Saravu, K. (2020). Deep learning, computer-aided radiography reading for tuberculosis: A diagnostic accuracy study from a tertiary hospital in India. Scientific Reports, 10, 210. https://doi.org/10.1038/s41598-019-56589-3
16. Harris, M., Qi, A., Jeagal, L., Torabi, N., Menzies, D., Korobitsyn, A., ... & Khan, F. A. (2019). A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest X-rays for pulmonary tuberculosis. PLoS ONE, 14(9), e0221339. https://doi.org/10.1371/journal.pone.0221339
17. Khan, F. A., Majidulla, A., Tavaziva, G., Nazish, A., Abidi, S. K., Benedetti, A., ... & Pai, M. (2020). Chest X-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: A prospective study of diagnostic accuracy for culture-confirmed disease. The Lancet Digital Health, 2(11), e573–e581. https://doi.org/10.1016/S2589-7500(20)30221-X
18. Sivaramakrishnan, R., Antani, S., Candemir, S., Xue, Z., Abuya, J., Kohli, M., ... & Thoma, G. (2018). Comparing deep learning models for multi-cell classification in chest radiography. Proceedings of SPIE, 10579, 105791B. https://doi.org/10.1117/12.2293471
19. Hwang, E. J., Park, S., Jin, K. N., Song, K. D., Kim, J. I., Choi, S. Y., ... & Park, C. M. (2021). Development and validation of a deep learning-based automated detection algorithm for tuberculous pleural effusion on chest radiographs. European Radiology, 31, 866–875. https://doi.org/10.1007/s00330-020-07179-7
20. Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., ... & McDonald, C. J. (2014). Automatic tuberculosis screening using chest radiographs. IEEE Transactions on Medical Imaging, 33(2), 233–245. https://doi.org/10.1109/TMI.2013.2284099
21. Rajaraman, S., & Antani, S. K. (2020). Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs. IEEE Access, 8, 27318–27326. https://doi.org/10.1109/ACCESS.2020.2971257
22. Heo, S. J., Kim, Y., Yun, S., Lim, S. S., Kim, J., Nam, C. M., ... & Yoon, J. H. (2019). Deep learning algorithms with demographic information help to detect tuberculosis in chest radiographs. Scientific Reports, 9, 10253. https://doi.org/10.1038/s41598-019-46722-1
23. Hwang, E. J., Goo, J. M., Yoon, S. H., Beck, K. S., Seo, J. B., Choi, B. W., ... & Park, C. M. (2022). Use of artificial intelligence-based software for chest radiographs in the intensive care unit: A retrospective study. Korean Journal of Radiology, 23(5), 530–540. https://doi.org/10.3348/kjr.2021.0564
24. Melendez, J., van Ginneken, B., Maduskar, P., Philipsen, R. H. H. M., Reither, K., Breuninger, M., ... & Sánchez, C. I. (2015). A novel multiple-instance learning-based approach to computer-aided

detection of tuberculosis on chest X-rays. IEEE Transactions on Medical Imaging, 34(1), 179–192. https://doi.org/10.1109/TMI.2014.2350539

25. Cao, Y., Liu, C., Liu, B., Brunette, M. J., Zhang, N., Sun, T., ... & Curioso, W. H. (2016). Improving tuberculosis diagnostics using deep learning and mobile health technologies among resource-poor and marginalized communities. IEEE Journal of Translational Engineering in Health and Medicine, 4, 1–9. https://doi.org/10.1109/JTEHM.2016.2589789
26. Hwang, S., Kim, H. E., Jeong, J., & Kim, H. J. (2016). A novel approach for tuberculosis screening based on deep convolutional neural networks. Proceedings of SPIE, 9785, 97852W. https://doi.org/10.1117/12.2216198
27. Stirenko, S., Kochura, Y., Alienin, O., Rokovyi, O., Gordienko, Y., Gang, P., & Zeng, W. (2018). Chest X-ray analysis of tuberculosis by deep learning with segmentation and augmentation. 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO), 422–428. https://doi.org/10.1109/ELNANO.2018.8477564
28. Hooda, R., Mittal, A., & Sofat, S. (2019). Automated TB classification using ensemble of deep architectures. Multimedia Tools and Applications, 78, 31515–31532. https://doi.org/10.1007/s11042-019-07984-5
29. Chauhan, A., Chauhan, D., & Rout, C. (2014). Role of Gist and PHOG features in computer-aided diagnosis of tuberculosis without segmentation. PLoS ONE, 9(11), e112326. https://doi.org/10.1371/journal.pone.0112326
30. Vajda, S., Karargyris, A., Jaeger, S., Santosh, K. C., Candemir, S., Xue, Z., ... & Thoma, G. (2018). Feature selection for automatic tuberculosis screening in frontal chest radiographs. Journal of Medical Systems, 42, 146. https://doi.org/10.1007/s10916-018-0991-9
31. Santosh, K. C., & Antani, S. (2018). Automated chest X-ray screening: Can lung region symmetry indicate tuberculosis? IEEE Journal of Biomedical and Health Informatics, 22(4), 1162–1171. https://doi.org/10.1109/JBHI.2017.2730090
32. Hwang, E. J., Park, S., Jin, K. N., Kim, J. I., Choi, S. Y., Lee, J. H., ... & Park, C. M. (2023). Deep learning-based detection of active pulmonary tuberculosis at chest radiography: Diagnostic performance and comparison with radiologists. Radiology, 307(1), e221408. https://doi.org/10.1148/radiol.221408
33. Qin, Z. Z., Barrett, R., del Mar Castro, M., Nikolayevskyy, V., Codlin, A. J., Creswell, J., ... & Denkinger, C. M. (2024). Multimodal deep learning for tuberculosis detection: Integrating chest X-rays and clinical data. The Lancet Digital Health. Advance online publication.
34. Kim, H., Yoon, S. H., & Goo, J. M. (2023). Multimodal deep learning for integrating chest radiographs and clinical parameters: A case for transformers. Radiology, 309(1), e230806. https://doi.org/10.1148/radiol.230806
35. Lee, J. H., Park, S., Hwang, E. J., & Park, C. M. (2023). Effect of multimodal diagnostic approach using deep learning-based automated detection algorithm for active pulmonary tuberculosis. Scientific Reports, 13, 19794. https://doi.org/10.1038/s41598-023-47146-0
36. Wong, A., Lee, S., & Rahman, T. (2024). Hybrid deep learning system combining radiological and clinical data for improved tuberculosis diagnosis. Indian Journal of Tuberculosis. Advance online publication. https://doi.org/10.1016/j.ijtb.2025.11.006

37. García Seco de Herrera, A., et al. (2026). Ensemble deep learning architectures for detecting pulmonary tuberculosis in chest X-rays. Scientific Reports, 16, 1242. https://doi.org/10.1038/s41598-025-30792-x
38. Park, J., et al. (2025). Prospective multi-site validation of AI to detect tuberculosis and chest X-ray abnormalities. NEJM AI, 1(10). https://doi.org/10.1056/AIoa2400018
39. Xu, Y., et al. (2025). A systematic review on application of multimodal learning and explainable AI in tuberculosis detection. IEEE Access, 13, 62198–62221. https://doi.org/10.1109/ACCESS.2025.XXXXXXX
40. World Health Organization. (2021). WHO consolidated guidelines on tuberculosis: Module 2: Screening – Systematic screening for tuberculosis disease. World Health Organization. https://www.who.int/publications/i/item/9789240022676