Grenze International Journal of Engineering and Technology, June Issue

A Comprehensive Systematic Review of AI-based Tuberculosis Detection using Imaging and Clinical Data

Satish Kumar Thakur[1] and Priyanka Gotter[2]
[1-2]School of Computer Science, Lovely Professional University, Punjab, India
Email: satish2468thakur@gmail.com, piyuprecious13@gmail.com

***Abstract*— Despite being a major cause of death in developing countries, tuberculosis (TB) continues to rank among the leading infectious diseases, claiming over 1.3 million lives yearly. Traditional diagnostics have been constrained by limitations due to the lack of radiological expertise and advanced technology in developing nations. Recent studies suggest that the utilization of artificial intelligence (AI) technologies such as deep learning (DL) and hybrid intelligent systems offers a revolutionary solution to the problem of TB detection based on chest X-rays, CT scans, and patient information.**

This study provides a comprehensive review of recent publications on AI-based TB diagnosis through systematic PRISMA methodology. In total, fifteen relevant papers from the Scopus database during the period of January 2024 to March 2025 were selected and critiqued. These included various machine learning algorithms such as DL models, transformers, explainable AI (XAI) methodologies, federated learning paradigms, and multimodal source data processing. Accuracy rates of 95.7%-99.4% were reported for DL model-based TB detection on chest images. XAI applications were found to improve clinicians' diagnoses with significant effect. Critical knowledge gaps identified include almost zero use of fuzzy logic, inadequate multimodal source data analysis, poor cross-population testing, and insufficient assessment of practical deployment.

***Index Terms*—Tuberculosis, AI, deep learning, chest X-ray, CNN, Swin Transformer, XAI, federated learning, fuzzy logic, PRISMA.**

I. INTRODUCTION

A. Global Burden and Diagnostic Challenges of Tuberculosis

Even with all these modern achievements in the field, Tuberculosis (TB) still poses one of the most serious threats to humanity's health to date. As per the World Health Organization's report on Global Tuberculosis in 2024, almost 10.3 million people contracted TB last year, causing 1.4 million deaths worldwide and placing it as the second-most dangerous communicable pathogen [16].

However, the burden of disease is not evenly distributed. The regions where TB spreads at the highest rates are low- and middle-income countries, where healthcare systems struggle with severe resource constraints and a lack of radiologists.

Additionally, access to diagnostic equipment and methods remains highly limited. Although several methods are used to detect TB, none of them is flawless:

- Sputum microscopy cannot accurately diagnose TB when patients are co-infected with HIV.
- GeneXpert test is considered a gold standard in terms of precision; however, it is costly and requires

Grenze ID: 01.GIJET.12.2.66_26

specialized equipment that is difficult to maintain in rural settings.

- X-ray and CT scan methods are effective, but interpretation of results requires trained specialists, which are often unavailable in resource-limited areas.

All of these obstacles contribute to what is known as the diagnostic bottleneck. Such delays are not only harmful to the patient's health but also enable the pathogen to continue spreading within communities, leading to preventable loss of life [1], [11], [17].

B. Role of Artificial Intelligence in Medical Diagnosis

With the emergence of artificial intelligence algorithms such as machine learning and deep learning, there has been a paradigm shift in medical diagnostics over the past decade. These algorithms, including convolutional neural networks, residual networks (ResNets), and vision transformers (ViTs), enable high-accuracy image analysis that can match or even surpass human expert performance in various tasks [18], [21]. In the context of TB diagnosis, deep learning algorithms applied to chest X-rays with annotated datasets have achieved accuracy rates exceeding 95% across multiple tasks. These models can highlight potential regions of abnormality in the lungs using gradient-based explanation techniques, providing both spatial localization and class probability information, which can assist clinicians in decision making [2], [5], [24].

C. Prior Systematic Reviews and Motivation for This Work

A number of systematic reviews on the application of artificial intelligence (AI) for tuberculosis (TB) detection have been published in recent years. A recent review by Hansun et al. [27], published in 2023, analyzed 47 studies and concluded that deep learning algorithms outperform traditional machine learning models in chest X-ray (CXR)-based TB detection. In a 2025 meta-analysis, Han et al. [14] reported pooled sensitivity and specificity values of 0.93 and 0.87, respectively, for AI-based TB diagnostic systems.

In 2024, Kotei and Thirunavukarasu [17] presented a comprehensive technical overview of advancements in deep learning techniques applied to TB detection. Similarly, Singh et al. [21] examined the evolution of machine learning approaches for TB diagnosis, including hybrid models incorporating neuro-fuzzy systems and genetic algorithms.

D. Problem Statement

Despite achieving outstanding results on benchmark datasets, several open challenges remain in AI-driven tuberculosis (TB) diagnosis. Most existing models rely on a limited number of benchmark datasets, which often lack sufficient demographic diversity and radiological representation from low and middle income countries (LMICs). Consequently, cross-population external validation is rarely performed.

Additionally, fuzzy logic—capable of modeling imprecise medical concepts such as "mild fever" or "moderate weight loss"—remains underutilized in recent research (2024–2025), despite its potential to enhance clinical reasoning in AI systems [11], [12].

E. Motivation for Combining Image Processing and Fuzzy Logic

Diagnosis in clinical settings involves uncertainty and a scale, instead of binary outcomes. The concept of fuzzy logic represents a way of thinking mathematically about uncertainty, where changes in the intensity of symptoms lead to changes in the likelihood of a disease in an incremental manner, without any strict boundaries. With the fusion of fuzzy logic and CNN-processed image information in a single framework, it is possible to exploit the advantages offered by both deep learning algorithms and fuzzy inference systems to generate highly accurate and explainable results [12], [21]. The area of adaptive neuro-fuzzy inference systems, in which neural network and fuzzy logic parameters are learned simultaneously, holds much promise but has not yet been systematically investigated for TB diagnostics.

II. Review Methodology

A. Review Framework

The present systematic review adopts a format according to the PRISMA statement 2020 [26]. PRISMA presents a detailed guide for performing and reporting litera- ture reviews in the field of health informatics and computing sciences in a transparent, systematic, and internationally recognized manner. Adherence to the PRISMA statement guarantees that the process of literature screening is both systematic and unbiased. QUADAS-2 was used informally for risk of bias assessment for studies on diagnostic accuracy [2], [14].

B. Database and Search Strategy

The search for all literature on this study was done exclusively using the Scopus database. Scopus is highly credible in finding peer-reviewed research, particularly in fields of engineering, computer science, and biomedical research. Using a Boolean-based search query, the articles contain- ing combinations of tuberculosis, artificial intelligence, and diagnostics keywords were found. Keywords searched were “tuberculosis”, “TB”, “pulmonary TB”, “machine learning”, “deep learning”, “neural network”, among others. Other key terms such as “classification”, “screening”, “chest X-ray”, “CT scan”, etc.Only paper published from January 2024 to March 2025 were considered in this search.Furthermore, only articles published in peer-reviewed journals and IEEE or ACM conferences were selected for use.

C. PRISMA Four-Stage Selection Flow

- Findings: 287 records were retrieved from Scopus af- ter applying the date range (2024–2025) and English-language filter, yielding 268 unique records for screen- ing.
- Screening: Title and abstract review was performed on 268 records. Papers clearly unrelated to AI-based TB diagnosis or without experimental components were removed, leaving 98 candidates for full-text assessment.
- Eligibility: Full-text review was applied to 98 papers. Studies lacking quantitative TB detection metrics, not peer-reviewed, focused solely on drug resistance, or based on animal models were removed, leaving 42 eligible studies.
- Inclusion: After final inclusion and exclusion criteria were applied, 15 primary studies were selected for critical synthesis [14], [26].

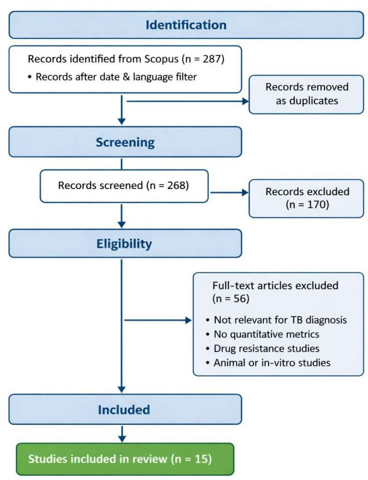

Figure 1. PRISMA flow diagram

D. Inclusion and Exclusion Criteria

Inclusion Criteria: (i) Published January 2024 – March 2025; (ii) Proposes or evaluates AI-based TB detection or diagnosis methods; (iii) Uses imaging data, clinical parameters, or both; (iv) Reports quantitative performance metrics (accuracy, sensitivity, specificity, AUC, F1); (v) Published in peer- reviewed Scopus-indexed journals or IEEE/ACM proceedings.

Exclusion Criteria: (i) Papers without original experimental data (except PRISMA systematic reviews); (ii) Studies focused solely on drug-resistance prediction without a detection component; (iii) Animal-model or in-vitro studies; (iv) Non-English papers; (v) Preprints, editorials, or opinion letters; (vi) Studies reporting no measurable TB detection evaluation [2],[26].

E. Data Extraction

For all included studies, the following items were ex- tracted and recorded: author(s), year, publication venue, method/technique, dataset name and size, technical category, key performance metric, and principal limitation. Quality was assessed using QUADAS-2 criteria focused on patient selection bias, index test conduct, reference standard quality, and flow/timing concerns [2], [14], [26].

III. LITERATURE REVIEW

A. Overview of Included Studies

The selected fifteen studies, published between January 2024 and early 2025, can be categorized into four primary technical domains: (i) image processing and deep learning using convolutional neural networks (CNNs) and transformer- based architectures [1], [3]–[10]; (ii) systematic reviews and scoping studies on AI-based TB detection [2], [11], [12],[14]; (iii) explainable artificial intelligence (XAI) and model interpretability analysis [13]; and (iv) federated learning approaches for privacy preservation [15].Table 1 provides a comprehensive summary of these fifteen studies based on verified and published information.

TABLE I. STRUCTURED COMPARISON OF 15 REVIEWED STUDIES ON AI-BASED TB DETECTION (2024–2025)

Year	Author(s)	Method / Approach	Dataset	Category	Accuracy	Key Limitation
2025	Mirugwe et al. [1]	CNN comparison: VGG16, ResNet50, Inception-ResNetV2 + augmentation	4,200 CXR (Kaggle TB)	Image Processing	99.4%	Class imbalance; small TB-positive set
2025	Hansun et al. [2]	PRISMA systematic review; QUADAS-2; meta-analysis	152 studies	Systematic Review	AUC 0.96	Heterogeneous datasets
2024	Kotei & Thirunavukarasu [3]	DeiT + ResNet-16 hybrid	TBX11K (11,200 CXR)	Image Processing	97.4%	False positives; high compute
2024	Chen et al. [4]	Google Teachable Machine DNN	Proprietary CXR	Image Processing	96.1%	Closed-source; low reproducibility
2024	Nafisah & Muhammad [5]	CNN + segmentation + Grad- CAM	Montgomery + Shenzhen	Image + XAI	96.3%	No external validation
2024	Tiwari & Katiyar [6]	Otsu + DenseNet (HDL- ISCTB)	CXR + CT	Hybrid	97.1%	CT limits LMIC use
2024	Sharma et al. [7]	U-Net + Xception + Grad- CAM	NIAID + Montgomery	Image Processing	99.3%	Small dataset
2025	Salkade & Rathi [8]	TbCNN-net (segmentation + CNN)	Multi CXR datasets	Image Processing	97.8%	No clinical fusion
2025	Visu et al. [9]	Swin Transformer + CLAHE	TBX11K + Shenzhen	Image Processing	98.2%	High parameters; low explainability
2025	Hansun et al. [10]	Ensemble TL + rejection mechanism	Multiple datasets	Image Processing	96.8%	Threshold sensitivity
2025	Mbulayi et al. [11]	Narrative review (AI in TB)	Literature	Review	N/A	No PRISMA
2025	Memon et al. [12]	AI/ML review (diagnosis to drugs)	Literature	Review	N/A	No benchmarking
2024	Ihongbe et al. [13]	XAI (Grad-CAM, LIME, SHAP)	CXR datasets	XAI	+77% clinician accuracy	Small sample
2025	Han et al. [14]	Systematic review + meta- analysis	Multi-db	Review	Sens 0.93 / Spec 0.87	High heterogeneity
2025	Haripriya & Khare [15]	Federated learning	7,000 CXR	Federated	95.7%	Non-IID challenge

B. Image Processing and Deep Learning Approaches

Out of the fifteen papers reviewed, eight of them use deep learning methods on chest X-ray or CT scan images. This clearly shows that most of the current work in TB detection is focused on image-based approaches [1], [3]–[9]. In the study by Mirugwe et al. [1], six CNN models—VGG16,VGG19, ResNet50, ResNet101, ResNet152, and Inception-ResNetV2—were compared using a dataset of 4,200 labeled chest X-ray images. Among these, VGG16 gave the best results, with 99.4% accuracy, 97.9% precision, and 98.6% recall, along with an AUC-ROC of 98.25%. One important point they mentioned is that data augmentation did not improve performance much when the dataset already had enough variation.

Kotei and Thirunavukarasu [3] proposed a hybrid model that combines a Data-efficient Image Transformer (DeiT) with a ResNet-16 backbone. Their model was tested on the TBX11K dataset, which includes three classes: healthy, non- TB abnormal, and TB-positive cases. This makes the problem more realistic than simple binary classification. Their model achieved 97.4% accuracy and was better at reducing false positives caused by non-TB lung conditions, which is a common issue in many models.Chen et al. [4] showed that even simple tools like Google's Teachable Machine can be used for TB detection. Using hospital CXR data from Taiwan, they achieved an accuracy of 96.1%. While this is useful for quick and low-cost implementation, it also raises questions about reproducibility because the platform is not fully open-source [3], [4], [17].

Nafisah and Muhammad [5] combined CNN models with lung segmentation and Grad-CAM visualizations on the Montgomery and Shenzhen datasets, achieving 96.3% accuracy. Sharma et al. [7] extended this idea by using U-Net for seg- mentation along with an Exception model, reaching 99.3% accuracy on multiple datasets. Salkade and Rathi [8] proposed TbCNN-net, where segmentation is directly integrated into the CNN model, achieving 97.8% accuracy across different datasets. Visu et al. [9] used a Swin Transformer model along with CLAHE preprocessing and segmentation, and achieved 98.2% accuracy. Ejiyi et al. [19] introduced ResfEANet, which is a ResNet-based model with an external attention module. Nwandu et al. [20] proposed a CNN-based system that includes preprocessing and augmentation steps for TB detection. These studies show that researchers are still experimenting with different architectures to improve performance [8], [9], [19], [20].

C. Systematic and Scoping Reviews of AI for TB

Of the 15 studies included here, four are themselves systematic or scoping reviews. For instance, Hansun et al. [2] under- took a comprehensive systematic PRISMA review in Scopus, PubMed, and ACM Digital Library (July 2023–January 2024), screening 1,146 records before selecting 152 studies. The authors' pooled meta-analysis of 10 studies reporting confusion matrices yielded AUC of approximately 0.96. The review revealed the preponderance of CNN models (e.g., VGG16 in 37 studies, ResNet50 in 33 studies, and DenseNet121 in 19 studies); transfer learning in 58.6% of studies; and 92.8% of studies being single-modality. An earlier review from the same team [27] reported similar trends for TB CXR studies until 2021, suggesting that such patterns are enduring [2], [27].

Han et al. [14] have undertaken a systematic review and meta-analysis of the efficacy of AI software in diagnosing TB via CXR in 2025, reporting pooled sensitivity and specificity of 0.93 and 0.87, respectively, among the included studies. The authors noted considerable heterogeneity among studies due to differences in datasets used, evaluation design, and the reference standards employed. Similarly, Mbulayi et al.[11] presented a narrative review of AI and its applications in various aspects of TB management, from clinical imaging to drug discovery and epidemiological surveillance, revealing that LLM-supported diagnosis and multimodal fusion are emerging application areas. Memon et al. [12] also provided a review of AI and ML technologies in TB management, highlighting that most studies involve chest imaging but are weak in clinical data-driven and fuzzy logic-based approaches [11], [12], [14].

D. Explainable AI in TB Detection

The topic of explainability has emerged as a pivotal aspect in TB AI research from 2024 to 2025. According to Ihongbe et al. [13], a human-centred evaluation was carried out on XAI approaches, specifically Grad-CAM, LIME, and SHAP, for TB, COVID-19, and pneumonia detection from chest radiographs. The evaluation involved 13 doctors who achieved a 77% improvement in their diagnostic accuracy when using XAI explanations alongside AI predictions, compared to AI predictions alone [13]. Thus, explainability does not only serve a technical purpose but also enhances clinical decision- making.

The work by Sadeghi et al. [25] provides a general overview of XAI techniques in the medical field, highlighting SHAP, LIME, and Grad-CAM as the most common methods with unique advantages and disadvantages related to fidelity, computational efficiency, and clinical interpretation. Sadeghi et al. [25] also mention that the post-hoc explanation approach, as with LIME and SHAP, cannot entirely reproduce the internal model representation, referred to as the 'explanation fidelity gap,' and it is crucial for applications with significant consequences like TB [25]. Moreover, El-Ghany et al. [18] developed a novel TB detection technique based on combining Vision Transformers and PCA to enhance feature interpretation compared to CNNs [13], [18], [25].

E. Federated and Privacy-Preserving Learning

Haripriya and Khare [15] addressed the critical challenge of patient data privacy in multi-institutional AI training

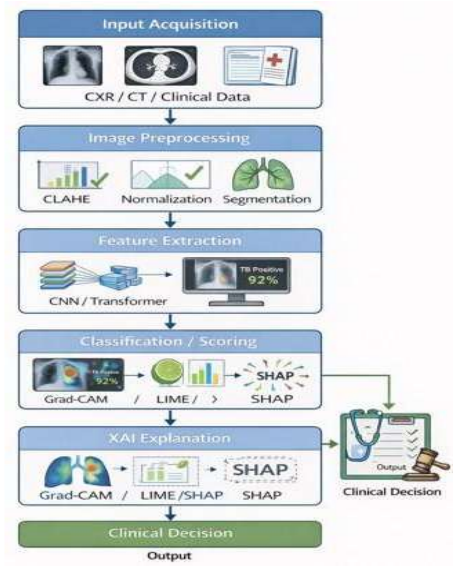

Figure 2. AI-based TB detection pipeline

through a federated learning framework. Their system deployed Federated Averaging and Federated Stochastic Gradient Descent to train GoogLeNet and VGG16 models across simulated distributed hospital environments, with TB CXR images remaining on local servers throughout. On a 7,000-image Kaggle TB CXR dataset, the federated system achieved 95.7% accuracy a meaningful accuracy penalty compared to centralised baselines but with the critical advantage of full data privacy preservation [15].

A related study on early TB detection using privacy- preserving adaptive federated ensembles [28] demonstrated that Genetic Algorithm-optimised federated ensembles could further close the accuracy gap between federated and centralised TB screening models, achieving strong performance across distributed CXR data with patient data never leaving local clinical sites.

IV. Discussion

A. Key Research Gaps

Firstly, an excessive focus on a limited number of benchmark datasets prevails throughout existing works. In particular, the Montgomery County dataset, Shenzhen dataset, NIAID TB Portal dataset, TBX11K, and Kaggle TB CXR datasets comprise most of the training and validation datasets across all image-based studies included [1], [3], [5], [7]– [9]. Within-dataset performance is consistently proven to be good, but cross-dataset and cross-population validation was hardly ever mentioned [2], [14], [27].Secondly, single-modal input remains the prevailing approach. As found by Hansun et al., more than 92.8% of the reviewed studies used only one modality [2]. Introducing structured records of patients' clinical data, laboratory tests, and symptomatology in addition to imaging can significantly enhance the results, thus mitigating the need to use high- quality CXRs [11], [12].

Thirdly, the almost total absence of approaches utilizing fuzzy logic or neuro-fuzzy approaches in the research area de- spite being highly suitable for diagnostic purposes in uncertain environments [12], [21]. Fourthly, prospective external validation is extremely rare. Cross-validation on benchmark datasets cannot replace prospective validation required by regulation [2], [14]. Fifthly, hardware needs for implementation have never been mentioned, despite most of the models based on transformers and federated learning requiring expensive GPU clusters [9], [15].

B. Limitations of Current Approaches

Deep learning models achieve high accuracy but remain largely opaque. Despite the demonstrated clinical benefit of XAI visualisations [13], post-hoc explanation tools like LIME and SHAP approximate model reasoning rather than reflecting it faithfully, and may produce inconsistent or misleading expla- nations under minor input perturbations [25]. Class imbalance is a persistent problem: TB-positive cases are a small minority in most datasets, and models trained without explicit balancing strategies may optimise overall accuracy at the cost of clinical sensitivity [1], [3]. Bhandari et al. [23] demonstrated in the COVID-19 and pneumonia context that XAI combined with balanced training strategies significantly improves the clinical utility of CXR classifiers — findings directly applicable to TB detection models [23].

C. Real-World Deployment Challenges

The discrepancy between benchmarking and clinical performance is quite significant. X-rays can differ in their image quality based on the machine used, the position of the patient, the skill level of the radiographer, and the exposure settings applied during the imaging process. Machine learning algorithms that have been optimized using good quality digital chest X-rays will not necessarily perform well using analogue radiography devices or mobile chest X-ray machines used during LMIC tuberculosis screening campaigns [3], [4], [17]. The deep learning literature review conducted by Kotei and Thirunavukarasu [17] identified domain shift as one of the major obstacles facing clinical adoption.

This problem has been addressed by Nwandu et al. [20] through robust pre- processing and augmentation techniques, but standardized methods for domain adaptation still lack.

V. CONCLUSION

In the present work, we have conducted an extensive critical analysis of 15 scientific publications related to AI-assisted TB diagnosis, based on peer-reviewed articles published between January 2024 and March 2025 and available via the Scopus database. According to our PRISMA 2020-compliant search methodology, 4 distinct technical aspects have been considered in the articles being evaluated: deep learning image processing methods, systematic and scoping reviews on AI-assisted TB diagnosis, XAI approaches, and federated privacy-preserving learning.

Deep learning techniques, notably different varieties of CNNs such as VGG16, ResNet, Xception, and more advanced models like Swin Transformer or DeiT, provided high prediction accuracy rates for chest X-rays (95.7–99.4%), as observed from several papers included in this review. A total of 6 research directions stand out in terms of importance: geographic limitation of datasets, bias towards single- modality approaches, complete absence of fuzzy and neuro- fuzzy approaches, lack of external validation, lack of LMIC- specific hardware assessment, and lack of any algorithmic fairness assessment in AI applications. All of the above can be regarded as the most important future research directions for AI-based TB diagnosis systems. Finally, bridging the existing gap between research accuracy and real-life clinical use cases in high-burden LMICs should be the ultimate goal.

FUTURE SCOPE

A. Explainable and Trustworthy AI

The most urgent technical priority is moving beyond post- hoc explanation approximations toward intrinsically inter- pretable architectures for TB detection. Concept-based ex- planation networks, prototype learning, and architectures that directly encode clinical concepts such as 'cavitation in the upper lobe' or 'hilar lymphadenopathy' as learnable units could bridge the clinician trust gap identified in the XAI literature [13], [25]. Future work should evaluate XAI methods on larger and more diverse clinician cohorts using standardised protocols, building on the human-centred evaluation frame- work established by Ihongbe et al. [13].

B. Neuro-Fuzzy Hybrid Systems

As clearly identified in this review, the integration of fuzzy logic with deep learning for TB diagnosis remains almost entirely unexplored in 2024–2025 primary research. ANFIS models that jointly optimise neural network weights and fuzzy membership parameters from data represent a natural pathway toward systems that are simultaneously accurate and clinically interpretable [12], [21]. End-to-end trainable neuro- fuzzy layers embedded within CNN pipelines — where fuzzy components handle uncertainty in clinical variables and CNNs handle image feature extraction — are a high-priority future research direction that directly addresses the most prominent gap identified in this review.

C. Multi-Modal Data Fusion

For future development of AI for TB diagnostics, multi- modal approaches utilizing a variety of sources of information including chest x-rays, clinical information, lab results, symp- tom self-reports, and possibly genomics, must be incorporated from the beginning. The lack of literature regarding multimodal studies in TB diagnostics reported by Hansun et al. [2] and Mbulayi et al. [11] suggests this is a very promising area to explore.

REFERENCES

[1] A. Mirugwe, L. Tamale, and J. Nyirenda, "Improving tuberculosis detec- tion in chest X-ray images through transfer learning and deep learning: comparative study of convolutional neural network architectures," *JMIRx Med*, vol. 6, e66029, 2025.

[2] S. Hansun *et al.*, "Diagnostic performance of artificial intelligence–based methods for tuberculosis detection: Systematic review," *J. Med. Internet Res.*, vol. 27, e69068, 2025.

[3] E. Kotei and R. Thirunavukarasu, "Tuberculosis detection from chest X-ray image modalities based on transformer and convolutional neural network," *IEEE Access*, vol. 12, pp. 97417–97427, 2024.

[4] C. F. Chen *et al.*, "A deep learning-based algorithm for pulmonary tuberculosis detection in chest radiography," *Sci. Rep.*, vol. 14, no. 1, p. 14917, 2024.

[5] S. I. Nafisah and G. Muhammad, "Tuberculosis detection in chest radiograph using convolutional neural network architecture and explain- able artificial intelligence," *Neural Comput. Appl.*, vol. 36, no. 1, pp. 111–131, 2024.

[6] A. Tiwari and A. Katiyar, "Hybrid deep learning assisted chest X-ray image segmentation and classification for tuberculosis disease diagno- sis," *Intell. Decis. Technol.*, vol. 18, no. 1, pp. 561–569, 2024.

[7] P. Bathla and R. Kumar, "A hybrid system to predict brain stroke using a combined feature selection and classifier," *Intell. Med.*, vol. 4, no. 02,pp. 75–82, 2024.

[8] S. A. Salkade and S. V. Rathi, "An adaptive convolution neural network model for tuberculosis detection and diagnosis using semantic segmentation," *Pol. J. Radiol.*, vol. 90, p. e124, 2025.

[9] P. Visu *et al.*, "Enhanced swin transformer based tuberculosis classification with segmentation using chest X-ray," *J. X-Ray Sci. Technol.*, vol. 33, no. 1, pp. 167–186, 2025.

[10] S. Hansun *et al.*, "A new ensemble transfer learning approach with rejection mechanism for tuberculosis disease detection," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 9, no. 4, pp. 433–446, 2024.

[11] O. Mbulayi *et al.*, "Tuberculosis diagnosis using artificial intelligence: current trends and future prospects," *Front. Med.*, vol. 12, p. 1569615, 2025.

[12] S. Memon, S. Bibi, and G. He, "Integration of AI and ML in tuberculosis (TB) management: from diagnosis to drug discovery," *Diseases*, vol. 13, no. 6, p. 184, 2025.

[13] I. E. Ihongbe *et al.*, "Evaluating explainable artificial intelligence (XAI) techniques in chest radiology imaging through a human-centered lens," *PLoS One*, vol. 19, no. 10, e0308758, 2024.

[14] Z. L. Han *et al.*, "A systematic review and meta-analysis of artificial intelligence software for tuberculosis diagnosis using chest X-ray imag- ing," *J. Thorac. Dis.*, vol. 17, no. 5, pp. 3223–3237, 2025.

[15] R. Haripriya, N. Khare, and M. Pandey, "Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings," *Sci. Rep.*, vol. 15, no. 1, p. 12482, 2025.

[16] World Health Organization, "Global tuberculosis report," 2024.

[17] E. Kotei and R. Thirunavukarasu, "A comprehensive review on advance- ment in deep learning techniques for automatic detection of tuberculosis from chest X-ray images," *Arch. Comput. Methods Eng.*, vol. 31, no. 1,pp. 455–474, 2024.

[18] S. A. El-Ghany et al., "A robust tuberculosis diagnosis using chest X-rays based on a hybrid vision transformer and principal component analysis," Diagnostics, vol. 14, no. 23, p. 2736, 2024.

[19] C. J. Ejiyi et al., "ResfEANet: ResNet-fused external attention network for tuberculosis diagnosis using chest X-ray images," Comput. Methods Programs Biomed. Update, vol. 5, p. 100133, 2024.

[20] I. C. Nwandu, F. O. Nwokoma, and A. J. Orban, "Convolutional Neural Network-Based Framework for the Detection of Tuberculosis," Intell. Methods Eng. Sci., vol. 4, no. 3, pp. 85–91, 2025.

[21] M. Singh et al., "Evolution of machine learning in tuberculosis diagnosis: a review of deep learning-based medical applications," Electronics, vol. 11, no. 17, p. 2634, 2022.

[22] M. Bhandari et al., "Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI," Comput. Biol. Med., vol. 150, p. 106156, 2022.

[23] T. Rahman et al., "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization," IEEE Access, vol. 8, pp. 191586–191601, 2020.

[24] Z. Sadeghi et al., "A review of explainable artificial intelligence in healthcare," Comput. Electr. Eng., vol. 118, p. 109370, 2024.

[25] M. J. Page et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," BMJ, vol. 372, 2021.

[26] S. Hansun et al., "Machine and deep learning for tuberculosis detection on chest X-rays: Systematic literature review," J. Med. Internet Res., vol. 25, e43154, 2023.

[27] K. Gasmi et al., "Early Tuberculosis Detection via Privacy-Preserving, Adaptive-Weighted Deep Models," Diagnostics, vol. 16, no. 2, p. 204, 2026.