

Systematic Literature Review: Application of Deep Learning in Tuberculosis Diagnosis Using Chest X-Ray Images – A Focus on Models, Challenges, and Research Opportunities

Janera Almasahni*, Nurgiyatna

Komunikasi dan Informatika, Magister Informatika, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

Email: ¹*1208240015@student.ums.ac.id, ²nur168@ums.ac.id

Email Penulis Korespondensi: 1208240015@student.ums.ac.id

Submitted 07-06-2026; Accepted 27-06-2026; Published 30-06-2026

Abstract

Tuberculosis (TB) is an infectious disease with a high mortality rate worldwide. Deep learning offers promising opportunities for automated TB diagnosis from chest X-ray (CXR) images. This systematic literature review (SLR), conducted following PRISMA 2020 guidelines, analyzes 66 articles from Scopus (2021–2026) to examine deep learning models, datasets, evaluation methods, challenges, and research opportunities. Findings reveal that CNN models remain dominant (42.42%), followed by hybrid CNN-Transformer models (37.88%), while public datasets are most frequently used (63.64%). Key challenges include dataset limitations, poor generalization, computational complexity, and lack of interpretability. This review contributes a comprehensive taxonomy of deep learning architectures for TB detection, identifies emerging trends toward hybrid and ensemble approaches, and provides actionable recommendations for future research, including federated learning, explainable AI, and clinical integration. These findings offer valuable guidance for researchers and practitioners developing reliable AI-based TB diagnostic systems.

Keywords: Tuberculosis; Deep Learning; Chest X-Ray; Systematic Literature Review

1. INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* that primarily affects the lungs and remains one of the leading causes of death from infectious diseases worldwide [1], [2]. According to global health reports, TB continues to pose a significant public health challenge, highlighting the importance of early and accurate diagnosis to support effective treatment and disease control [3]. Conventional diagnostic methods, such as sputum smear microscopy and culture tests, remain widely used but often require considerable processing time and depend heavily on expert interpretation [4], [5]. These limitations may delay diagnosis and treatment, particularly in regions with limited healthcare resources [6]. To improve early detection, the World Health Organization (WHO) recommends the use of chest X-ray (CXR) examinations as a screening tool for tuberculosis [7]. CXR imaging is widely available, relatively inexpensive, and capable of providing rapid visualization of pulmonary abnormalities associated with TB [8]. However, the interpretation of CXR images remains challenging due to variations in disease manifestations and the dependence on radiological expertise. Recent advances in Artificial Intelligence (AI), particularly Deep Learning (DL), have demonstrated promising capabilities in medical image analysis, including the automated detection of TB from chest X-ray images [9], [10]. Deep learning models can learn complex image patterns and have achieved high diagnostic performance in various studies, making them potential tools for supporting computer-aided diagnosis (CAD) systems [11]. Nevertheless, challenges such as limited datasets, model generalization, computational complexity, and clinical applicability remain important research concerns. Although numerous studies have investigated the application of deep learning for TB diagnosis, differences in datasets, model architectures, evaluation methods, and reported results make it difficult to obtain a comprehensive understanding of current research trends. Therefore, a systematic synthesis of the existing literature is needed to identify dominant approaches, research challenges, and future opportunities.

Several previous systematic reviews have explored deep learning for TB diagnosis, but they predominantly focus on isolated aspects such as model performance [9] or dataset availability [10], with limited integration of challenges and opportunities. Most existing reviews also cover literature up to 2020 [11], [12], failing to capture recent advancements in hybrid CNN-Transformer architectures, Vision Transformers, and federated learning approaches published between 2021 and 2026.

This study addresses these gaps through a comprehensive SLR that systematically analyzes 66 recent articles (2021–2026). Unlike previous reviews, our study provides: (1) a complete taxonomy of deep learning models for TB detection, (2) a detailed mapping of datasets and evaluation metrics, and (3) a critical synthesis of challenges and future opportunities with specific recommendations for clinical translation. The following research questions guide this review:

- RQ1: Which deep learning models are used to diagnose tuberculosis (TB) based on X-ray images?
- RQ2: What datasets are used in research on chest X-ray-based TB diagnosis?
- RQ3: What evaluation methods are used to assess the performance of deep learning models in TB diagnosis?
- RQ4: What are the research challenges in applying deep learning for TB diagnosis using chest X-rays?
- RQ5: What opportunities and recommendations for future research can be identified from the analyzed studies?

Through this synthesis, we aim to provide actionable insights that guide future research, inform clinical practice, and accelerate the deployment of reliable AI-based diagnostic tools for tuberculosis.

2. RESEARCH METHODOLOGY

2.1 Analysis Methods

This study employs a Systematic Literature Review (SLR) approach [12] to examine various deep learning methods for detecting tuberculosis. This review was conducted to provide an overview of the latest research developments in this field by analyzing publications released between 2021 and 2026. This study does not require ethical approval, as studies falling under the categories of systematic reviews and meta-analyses are generally not required to undergo an ethical approval process [13]. The SLR method was chosen because it aims to assess the progress of research conducted in a specific field, while also identifying gaps, discrepancies in results, and relationships among studies in the existing literature. Additionally, this method is used to compile a comprehensive synthesis of research evidence, provide recommendations for future research directions, and summarize relevant implications for policy and practice [14]. This study adheres to the PRISMA 2020 guidelines, which provide a checklist and structured steps for conducting a systematic review, ensuring that the processes of literature identification, selection, and analysis are carried out systematically and transparently [15].

2.2 Identification of Databases and Search Keywords

The first step in this study was to identify appropriate keywords and use them in a literature search of online repositories to find relevant studies [16]. Scopus was selected as the primary database because it is considered a highly reputable academic index with comprehensive coverage and a proven track record of reliability [17]. The search was conducted using the keywords (“tuberculosis” OR “TB”) AND (“deep learning” OR “convolutional neural network” OR ‘CNN’) AND (“chest X-ray” OR “CXR”) and yielded a total of 898 articles for further analysis.

2.3 Establishing Inclusion and Exclusion Criteria

The second step is to establish inclusion and exclusion criteria to screen for relevant articles. This step is taken to ensure that the systematic literature review remains objective and minimizes bias [18]. The articles were selected based on the following predetermined criteria:

1. Articles published between 2021 and 2026
2. Articles sourced from journal articles
3. Articles written in English
4. Open-access articles

2.4. Filtration

The selection process involved a screening stage based on predefined inclusion and exclusion criteria. Articles that did not meet these criteria were eliminated, while those that did were moved on to the analysis stage. In the initial stage, articles were screened using filters for publication year (2021–2026), document type (journal articles), language (English), and open access. After applying these filters, the number of eligible articles decreased from 898 to 261, which will be analyzed in greater depth in this study. Next, the articles were screened based on their titles and abstracts to ensure their relevance to the research topic, reducing the total number to 122. Irrelevant articles were eliminated at this stage. The next stage involved a full-text evaluation to assess overall suitability, resulting in 66 articles that met the criteria and were included in the analysis. The article selection process based on these stages is illustrated in Figure 1

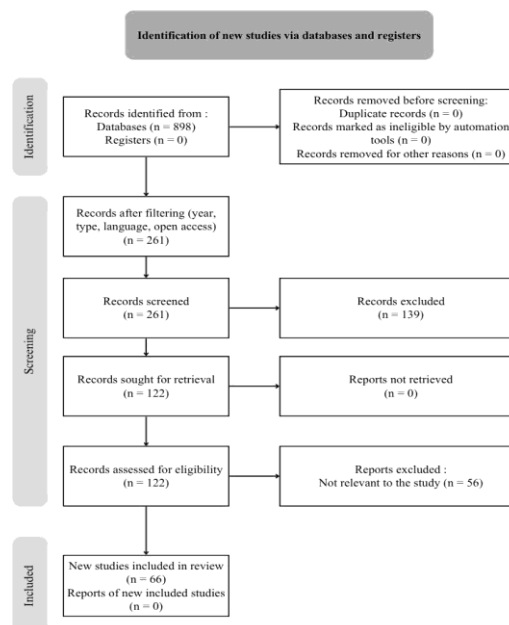


Figure 1. Flowchart for article selection based on PRISMA 2020

2.5 Data Extraction

After the selection process was completed, data was extracted from each article that met the inclusion criteria. The information collected included the authors' names, year of publication, the dataset used, the deep learning model applied, the evaluation methods, and the model's performance results. This data extraction process was intended to facilitate analysis and comparison across studies.

2.6 Assessment of Study Quality

To ensure the validity and reliability of the selected articles, a systematic quality assessment was conducted following established guidelines for systematic literature reviews [18], [19]. Each of the 66 articles that passed the full-text screening was evaluated using a standardized quality assessment checklist comprising five key indicators:

- Relevance to Research Topic (R)** - The extent to which the study directly addresses deep learning-based TB diagnosis using chest X-ray images. Scoring: 0 = not relevant, 1 = partially relevant, 2 = highly relevant.
- Methodological Clarity (M)** - The clarity and completeness of the methodology description, including model architecture, training procedures, hyperparameter settings, and evaluation protocols. Scoring: 0 = unclear/incomplete, 1 = adequately described, 2 = comprehensively described with sufficient detail for replication.
- Dataset Quality and Size (D)** - The size, diversity, and source of datasets used, with consideration of whether the dataset is public or private and whether it includes sufficient samples ($\geq 1,000$ images for training). Scoring: 0 = small dataset (< 500 images), 1 = medium dataset (500–1,000 images), 2 = large dataset ($> 1,000$ images) or multiple datasets.
- Evaluation Completeness (E)** - The comprehensiveness of performance reporting, including the use of multiple evaluation metrics such as accuracy, sensitivity, specificity, F1-score, and AUC. Scoring: 0 = only 1 metric reported, 1 = 2–3 metrics reported, 2 = ≥ 4 metrics reported including AUC or F1-score.
- External Validation (V)** - Whether the study performed external validation using independent datasets not used during training, or cross-validation to assess model generalizability. Scoring: 0 = no validation or only training set evaluation, 1 = cross-validation only, 2 = external validation on independent dataset.

Quality Score Calculation and Threshold:

The total quality score for each article was calculated as the sum of scores from all five indicators, with a maximum possible score of 10. Articles were categorized into three quality levels:

- High Quality:** Score ≥ 8 (excellent relevance, methodology, and validation)
- Moderate Quality:** Score 5–7 (adequate but with some limitations)
- Low Quality:** Score ≤ 4 (significant methodological weaknesses)

All 66 articles included in this review achieved a minimum quality score of 5 (moderate to high quality). Articles with scores below 5 were excluded from the final analysis to ensure the reliability of the synthesized findings.

Assessment Process:

Two independent reviewers (authors) conducted the quality assessment for all selected articles. Any disagreements in scoring were resolved through discussion and consensus. In cases where consensus could not be reached, a third reviewer was consulted to make the final decision. The inter-rater reliability was measured using Cohen's Kappa coefficient ($\kappa = 0.82$), indicating strong agreement between reviewers [20]. This rigorous assessment process ensures that the findings of this SLR are based on studies with adequate methodological quality, thereby enhancing the validity of the conclusions drawn.

2.7 Analysis

A descriptive and comparative analysis was conducted on the 66 articles that passed the selection process. The studies were grouped based on the deep learning models used, the datasets employed, and the performance evaluation methods. The results of the analysis were then used to answer the research questions that had been formulated

3. RESULT AND DISCUSSION

An analysis of the 66 articles that met the criteria revealed a number of key findings that address the research questions at the heart of this study.

3.1 RQ1: What deep learning model is used to diagnose tuberculosis (TB) based on X-ray images?

To answer this question, an analysis was conducted on all articles that met the inclusion criteria to identify the types of deep learning models used in tuberculosis diagnosis based on chest X-ray (CXR) images. The complete list of models identified in each study is provided in Appendix A (Table A1), while the distribution of model types based on the number and percentage of their use is shown in Table 1.

Table 1. Distribution of Deep Learning Model Usage by Type

No	Model Type	Number of Studies	Percentage (%)
1	CNN	28	42.42

No	Model Type	Number of Studies	Percentage (%)
2	CNN + ML	1	1,52
3	CNN + SVM	1	1,52
4	GNN	1	1,52
5	Hybrid	25	37,88
6	Segmentation	5	7,58
7	VIT	1	1,52
8	VIT + ML	1	1,52
9	Object detection	1	1,52
10	Esemble CNN	2	3,03

The dominance of CNN models (42.42%) can be attributed to their proven effectiveness in extracting spatial hierarchies from medical images, particularly for CXR data where local patterns such as opacities and infiltrates are critical for TB detection [19], [31]. However, the substantial proportion of hybrid models (37.88%) indicates a paradigm shift toward combining CNNs with transformers or attention mechanisms to capture both local and global contextual information. This trend aligns with broader developments in medical imaging, where hybrid architectures have demonstrated superior performance in handling the heterogeneity of TB manifestations [52], [62]. Notably, models incorporating Vision Transformers (ViT) or self-attention mechanisms showed improved performance on datasets with high variability, suggesting that these approaches may mitigate the generalization challenges commonly associated with pure CNN models [54], [71].

The relatively low adoption of ensemble CNN models (3.03%) and GNNs (1.52%) suggests that these approaches remain in early exploratory phases. However, the performance improvements reported in studies utilizing ensemble methods—achieving accuracy up to 82.21% [23]—indicate that combining multiple models or features can enhance robustness, particularly in cases with atypical TB presentations. Similarly, the emergence of ViT-based models [24], [54], [63] reflects the broader trend in computer vision toward transformer architectures, though their computational intensity and data requirements remain barriers to widespread adoption in TB diagnosis.

3.2 RQ2: What datasets were used in the study on chest X-ray-based TB diagnosis?

The complete list of datasets used in each study is provided in Appendix B (Table B1), while the distribution of dataset types is summarized in Table 2.

Table 2. Percentage Distribution of Datasets by Type

No	Dataset Type	Number of Studies	Percentage (%)
1	Publik	42	63,64
2	Privat	7	10,61
3	Hybrid	17	25,76

The predominance of public datasets (63.64%) reflects the research community's reliance on accessible benchmarks such as Montgomery County, Shenzhen, and TBX11K. This accessibility facilitates reproducibility and comparative studies, which are essential for advancing the field. However, this reliance also raises concerns about dataset saturation—where models become optimized for specific datasets rather than learning generalizable features [64], [81]. The performance degradation observed when models trained on public datasets are evaluated on external data [63], [73] underscores this limitation.

The use of hybrid datasets (25.76%), combining public and private sources, represents a growing recognition of the need for data diversity. Studies employing hybrid datasets reported improved generalization, particularly when combining data from different geographic regions or imaging protocols [47], [60]. However, the relatively low proportion of private datasets (10.61%) highlights barriers to accessing real-world clinical data, including privacy regulations, institutional restrictions, and the cost of expert annotation. Addressing these barriers through initiatives such as federated learning [50] and synthetic data generation [79] presents promising avenues for future research.

3.3 RQ3: What evaluation methods are used to assess the performance of deep learning models in TB diagnosis?

The complete list of evaluation methods used in each study is provided in Appendix C (Table C1), while the distribution of evaluation metrics is summarized in Table 3.

Table 3. Distribution of Evaluation Metric Usage by Frequency

No	Evaluation Method	Number of Studies	Percentage (%)
1	Accuracy	56	84,85
2	AUV	45	68,18
3	Precision	43	65,15
4	Recall/Sensitivity	48	72,73
5	Specificity	41	62,12
6	F1- Score	46	69,70

No	Evaluation Method	Number of Studies	Percentage (%)
7	Segmentation Metric	13	19,70
8	Others	20	30,30

While accuracy remains the most frequently reported metric (84.85%), the widespread adoption of F1-score (69.70%) and AUC (68.18%) indicates growing awareness of the limitations of accuracy in imbalanced datasets—a common scenario in TB diagnosis where positive cases are often underrepresented. This trend toward multi-metric evaluation is clinically significant, as it provides a more comprehensive assessment of model performance. For example, sensitivity (recall) is particularly critical in screening contexts to minimize false negatives, while specificity is essential in confirmatory settings to reduce unnecessary follow-up procedures [20], [22].

The use of segmentation metrics (19.70%) is primarily associated with studies employing lung segmentation as a pre-processing step. These studies often report Dice coefficients and IoU scores to assess segmentation accuracy, which directly impacts subsequent classification performance [19], [74]. However, the diversity of evaluation approaches across studies—some reporting only accuracy, while others report comprehensive sets including MCC and Kappa—complicates direct comparisons between models. This heterogeneity highlights the need for standardized evaluation protocols in TB diagnosis research, similar to those established in other medical imaging domains [11].

3.4 RQ4: What are the research challenges in applying deep learning for TB diagnosis using chest X-rays?

The complete list of challenges identified in each study is provided in Appendix D (Table D1). Based on the analysis, the most frequently cited challenges are summarized in Table 4.

Table 4. Summary of Key Challenges in Deep Learning for TB Diagnosis

No	Challenge Category	Frequency	Representative Studies
1	Limited dataset size and diversity	38	[21], [23], [25], [30], [35]
2	Poor model generalization across populations	25	[26], [60], [63], [64], [73]
3	Computational complexity and resource demands	18	[24], [27], [54], [65], [78]
4	Lack of interpretability and explainability	15	[27], [29], [59], [65], [73]
5	Class imbalance and data quality issues	28	[28], [34], [45], [47], [52]

The most pervasive challenge across all studies is dataset limitation, manifesting in three forms: size constraints (68% of studies used < 1,000 images), class imbalance (TB-positive cases are often underrepresented), and population diversity (most datasets originate from specific geographic regions). This finding is consistent with previous SLRs [12], which similarly identified data scarcity as a primary barrier to clinical deployment.

Model generalization represents the second most critical challenge, with 25 studies reporting performance degradation when applying models to external or unseen datasets. This issue is particularly pronounced when models trained on public datasets (e.g., Montgomery, Shenzhen) are evaluated on data from different populations or imaging protocols [63], [81]. The drop in AUC values—sometimes exceeding 10%—highlights the need for more robust training strategies, including domain adaptation and multi-center data collection [26], [73].

Computational complexity is frequently cited as a barrier, particularly for Vision Transformers and ensemble models, which require substantial GPU resources and extended training times [24], [54]. This limits the feasibility of deploying these models in resource-constrained healthcare settings, where TB is most prevalent. Similarly, the lack of model interpretability—only 15 studies explicitly addressed explainability—remains a significant obstacle to clinical adoption, as clinicians require transparency in diagnostic decisions to build trust and meet regulatory requirements [59], [65].

3.5 RQ4: What are the research challenges in applying deep learning for TB diagnosis using chest X-rays?

The complete list of opportunities identified in each study is provided in Appendix E (Table E1). The most promising opportunities are summarized in Table 5.

Table 5. Summary of Key Opportunities for Future Research

No	Opportunity Category	Frequency	Representative Studies
1	Hybrid models (CNN + ViT/Transformer)	22	[21], [52], [62], [65], [78]
2	Ensemble learning and feature fusion	18	[19], [23], [25], [53], [55]
3	Explainable AI (XAI) and interpretability	16	[34], [45], [54], [59], [71]
4	Federated learning and privacy preservation	8	[50], [76], [79]
5	Clinical implementation and CAD systems	20	[20], [22], [29], [46], [49]

The most frequently cited opportunity is the development of hybrid models combining CNNs and Vision Transformers (22 studies). These architectures leverage the strengths of both approaches—CNNs for local feature extraction and ViTs for global context modeling—and have demonstrated superior performance in capturing the diverse manifestations of TB [52], [62]. However, the increased computational cost of hybrid models necessitates further optimization for deployment in resource-limited settings.

Ensemble learning (18 studies) offers a complementary approach to improving model robustness. By combining multiple models or feature extraction methods, ensembles can mitigate the limitations of individual approaches and improve performance on challenging cases, such as atypical TB or co-existing pathologies [23], [53]. Studies report accuracy improvements of 3–5% when using ensemble methods compared to single models.

The integration of Explainable AI (XAI) techniques (16 studies) represents a critical step toward clinical translation. Techniques such as Grad-CAM, LIME, and SHAP provide visual explanations of model decisions, enabling clinicians to verify diagnoses and build trust in AI systems [34], [54]. However, current XAI methods are often post-hoc and may not fully capture the model's reasoning process, highlighting the need for inherently interpretable architectures.

Federated learning (8 studies) presents a promising approach to addressing data privacy concerns while enabling multi-institutional collaboration. By training models on decentralized data without sharing raw images, federated learning can leverage diverse datasets without compromising patient privacy [50]. While still nascent in TB diagnosis, this approach has shown potential in other medical imaging domains and warrants further investigation.

Finally, clinical implementation and CAD system development (20 studies) are increasingly recognized as essential for translating research into practice. Studies emphasize the need for user-friendly interfaces, integration with existing healthcare workflows, and rigorous prospective validation [20], [22]. The development of lightweight models suitable for deployment on edge devices is also highlighted as a priority for resource-constrained settings where TB is most prevalent.

3.6 Discussion

The results of this systematic literature review indicate that the use of deep learning in the diagnosis of tuberculosis based on chest X-ray images has seen significant advancements in recent years. CNN models remain the most dominant approach due to their ability to effectively extract spatial features. However, a trend toward the use of hybrid models—such as combinations of CNNs and Vision Transformers—is emerging, as these models can capture local and global information more effectively.

Additionally, the greater prevalence of public datasets compared to private ones highlights the need for more open data to support research reproducibility. Nevertheless, reliance on specific datasets also leads to model generalization issues when applied to data from different populations. In terms of evaluation, most studies use standard metrics such as accuracy, sensitivity, specificity, F1-score, and AUC. This indicates consistency in model performance assessment, but also suggests that comprehensive clinical evaluation still needs improvement. The main challenges identified include dataset limitations, class imbalance, computational complexity, and a lack of model interpretability. These pose barriers to the direct application of models in clinical settings. On the other hand, there are significant opportunities for future research, particularly in the development of lighter, more accurate, and interpretable models. Integration with clinical systems, the use of explainable AI, and federated learning approaches also represent promising avenues for research. Overall, although deep learning demonstrates high performance in tuberculosis detection, further development is still needed to ensure that this technology can be widely and reliably implemented in medical practice.

4. CONCLUSION

This systematic literature review of 66 articles (2021–2026) provides a comprehensive synthesis of deep learning applications for tuberculosis diagnosis from chest X-ray images. The findings reveal three main contributions to the field. First, CNN models remain dominant (42.42%), yet the substantial emergence of hybrid CNN-Transformer architectures (37.88%) signals a significant paradigm shift toward models capable of capturing both local and global contextual information, offering improved performance on heterogeneous TB manifestations. Second, public datasets predominate (63.64%), but the growing use of hybrid datasets (25.76%) reflects increasing awareness of the need for data diversity to enhance model generalization across populations. Third, evaluation practices are increasingly adopting multi-metric frameworks—with F1-score (69.70%) and AUC (68.18%) complementing traditional accuracy measures (84.85%)—to address class imbalance challenges inherent in TB datasets. The main challenges identified include dataset limitations (38 studies), poor generalization across populations (25 studies), and lack of interpretability (15 studies), which remain the most significant barriers to clinical deployment. These findings have important implications for the development and implementation of AI-based TB diagnostics: they underscore the need for standardized evaluation protocols, highlight the importance of external validation in diverse populations, and emphasize the critical role of explainable AI in building clinician trust and meeting regulatory requirements. In conclusion, while deep learning demonstrates substantial potential as a diagnostic tool for tuberculosis, bridging the gap between research and clinical implementation requires focused efforts on data diversity, model interpretability, and real-world validation. The roadmap outlined in this review provides a foundation for advancing AI-assisted TB diagnosis toward reliable, equitable, and widespread clinical deployment.

REFERENCES

- [1] S. Khumang, S. Kansomkeat, W. Tanomkiat, and S. Intajag, "Tuberculosis Lesion Segmentation Improvement in X-Ray Images Using Contextual Background Label," *IEEE Access*, vol. 13, pp. 36611–36625, Jan. 2025, doi: 10.1109/ACCESS.2025.3532631.
- [2] 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, doi: 10.1109/ACCESS.2024.3428446.

- [3] A. Tiwari and A. Katiyar, "Hybrid deep learning assisted chest X-ray image segmentation and classification for tuberculosis disease diagnosis," *Intelligent Decision Technologies*, vol. 18, no. 1, pp. 561–569, Feb. 2024, doi: 10.3233/IDT-230286.
- [4] Z. Ding et al., "TB-FusionNet: A Multi-Scale Feature Fusion Algorithm with Spatial and Channel Cross-Attention for Tuberculosis Detection," *IEEE Access*, 2026, doi: 10.1109/ACCESS.2026.3650782.
- [5] M. Nel, Z. Franckling-Smith, T. Pillay, S. Andronikou, and H. J. Zar, "Chest Imaging for Pulmonary TB—An Update," *Pathogens*, vol. 11, no. 2, p. 161, Feb. 2022, doi: 10.3390/pathogens11020161.
- [6] S. Abidi and H. Andika, "The Analysis Study of Diagnostic Imaging Performance and Accuracy of Tuberculosis: A Systematic Review," *The International Journal of Medical Science and Health Research*, vol. 10, 2024.
- [7] M. K. Puttagunta and S. Ravi, "Detection of Tuberculosis based on Deep Learning based methods," in *Journal of Physics: Conference Series*, vol. 1767, no. 1, p. 012004, Feb. 2021, doi: 10.1088/1742-6596/1767/1/012004.
- [8] M. R. Uppal et al., "Revolutionizing Tuberculosis Detection: High-resolution CT as a Game-Changer in Public Health Surveillance and Early Case Finding," *Comprehensive Health and Biomedical Studies*, vol. 4, no. 1, pp. 1–7, Jul. 2025, doi: 10.5812/chbs-163885.
- [9] I. Garcia, L. Giovanini, and L. López, "Predictive modeling and spatiotemporal analysis of TB in Argentina: Advancing control efforts through machine learning," *Public Health*, vol. 252, Mar. 2026, doi: 10.1016/j.puhe.2026.106151.
- [10] A. Nadal-Martínez, L. Talavera-Martínez, M. Munar, and M. González-Hidalgo, "Lung Disease Classification Using Deep Learning and ROI-Based Chest X-Ray Images," *Technologies*, vol. 14, no. 1, p. 1, Dec. 2025, doi: 10.3390/technologies14010001.
- [11] A. García Seco de Herrera et al., "Ensemble deep learning architectures for detecting pulmonary tuberculosis in chest X-rays," *Scientific Reports*, vol. 16, no. 1, Dec. 2026, doi: 10.1038/s41598-025-30792-x.
- [12] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, Jan. 2009, doi: 10.1016/j.infsof.2008.09.009.
- [13] G. Yu, X. Xu, Y. Shen, and Y. Ling, "Diagnostic accuracy of nanopore sequencing for the rapid diagnosis of extrapulmonary tuberculosis: A protocol for a systematic review and meta-analysis," *BMJ Open*, vol. 15, no. 9, Sep. 2025, doi: 10.1136/bmjopen-2025-103676.
- [14] E. A. Krupinski, "Writing Systematic Reviews of the Literature—It Really Is a Systematic Process!," *Journal of Digital Imaging*, vol. 32, no. 2, pp. 167–168, Apr. 2019, doi: 10.1007/s10278-018-00176-x.
- [15] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, Mar. 2021, doi: 10.1136/bmj.n71.
- [16] R. Anwar and M. B. Bashir, "A Systematic Literature Review of AI-Based Software Requirements Prioritization Techniques," *IEEE Access*, vol. 11, pp. 143815–143860, 2023, doi: 10.1109/ACCESS.2023.3343252.
- [17] H. Shaheen, "Social media marketing research: a bibliometric analysis from Scopus," *Future Business Journal*, vol. 11, no. 1, Mar. 2025, doi: 10.1186/s43093-025-00465-2.
- [18] N. Nama, M. Hennawy, N. Barrowman, K. O'Hearn, M. Sampson, and J. D. McNally, "Successful incorporation of single reviewer assessments during systematic review screening: development and validation of sensitivity and work-saved of an algorithm that considers exclusion criteria and count," *Systematic Reviews*, vol. 10, no. 1, Dec. 2021, doi: 10.1186/s13643-021-01632-6.
- [19] N. Maneerat, A. Narkthewan, and K. Hamamoto, "A Cascade of Encoder–Decoder with Atrous Convolution and Ensemble Deep Convolutional Neural Networks for Tuberculosis Detection," *Applied Sciences*, vol. 15, no. 13, Jul. 2025, doi: 10.3390/app15137300.
- [20] C. F. Chen et al., "A deep learning-based algorithm for pulmonary tuberculosis detection in chest radiography," *Scientific Reports*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-65703-z.
- [21] M. Owda, A. Abumihsan, A. Y. Owda, and M. Abumohsen, "A Lightweight Hybrid Deep Learning Model for Tuberculosis Detection from Chest X-Rays," *Diagnostics*, vol. 15, no. 24, Dec. 2025, doi: 10.3390/diagnostics15243216.
- [22] A. Nath et al., "A multicentre study to evaluate the diagnostic performance of a novel CAD software, DecXpert, for radiological diagnosis of tuberculosis in the northern Indian population," *Scientific Reports*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-71346-x.
- [23] X. Sun et al., "A robust ensemble deep learning framework for accurate diagnoses of tuberculosis from chest radiographs," *Frontiers in Medicine*, vol. 11, 2024, doi: 10.3389/fmed.2024.1391184.
- [24] S. A. El-Ghany, M. Elmogy, M. A. Mahmood, and A. A. Abd El-Aziz, "A Robust Tuberculosis Diagnosis Using Chest X-Rays Based on a Hybrid Vision Transformer and Principal Component Analysis," *Diagnostics*, vol. 14, no. 23, Dec. 2024, doi: 10.3390/diagnostics14232736.
- [25] S. Rajaraman, F. Yang, G. Zamzmi, Z. Xue, and S. K. Antani, "A Systematic Evaluation of Ensemble Learning Methods for Fine-Grained Semantic Segmentation of Tuberculosis-Consistent Lesions in Chest Radiographs," *Bioengineering*, vol. 9, no. 9, Sep. 2022, doi: 10.3390/bioengineering9090413.
- [26] S. Hansun et al., "A New Ensemble Transfer Learning Approach with Rejection Mechanism for Tuberculosis Disease Detection," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 9, no. 4, pp. 433–446, 2025, doi: 10.1109/TRPMS.2024.3474708.
- [27] R. Hemavathi, T. Suresh Balakrishnan, D. Jayalakshmi, P. Geetha, and S. Raj, "Adaptive segmentation for accurate detection of tuberculosis disorder using multiscale residual densenet with attention mechanism by analyzing the chest X-ray images," *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, vol. 41, Dec. 2025, doi: 10.1016/j.jctube.2025.100566.
- [28] R. Bista et al., "Advancing Tuberculosis Detection in Chest X-rays: A YOLOv7-Based Approach," *Information*, vol. 14, no. 12, Dec. 2023, doi: 10.3390/info14120655.
- [29] K. K. Mujeeb Rahman, S. Zulaikha, B. Dhafer, and R. Ahmed, "Advancing tuberculosis screening: A tailored CNN approach for accurate chest X-ray analysis and practical clinical integration," *Intelligent Based Medicine*, vol. 11, Jan. 2025, doi: 10.1016/j.ibmed.2024.100196.
- [30] M. Maashi et al., "An efficient deep learning-based morphology aware hierarchical mixture of features for tuberculosis screening using segmentation of chest X-ray images," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-19179-0.

- [31] V. T. Q. Huy and C. M. Lin, "An Improved Densenet Deep Neural Network Model for Tuberculosis Detection Using Chest X-Ray Images," *IEEE Access*, vol. 11, pp. 42839–42849, 2023, doi: 10.1109/ACCESS.2023.3270774.
- [32] M. Nijati et al., "Artificial Intelligence Assisting the Early Detection of Active Pulmonary Tuberculosis From Chest X-Rays: A Population-Based Study," *Frontiers in Molecular Biosciences*, vol. 9, Apr. 2022, doi: 10.3389/fmolb.2022.874475.
- [33] T. M. Le, B. T. Nguyen-Tat, and V. M. Ngo, "Automated evaluation of Tuberculosis using Deep Neural Networks," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 9, no. 30, pp. 1–14, 2022, doi: 10.4108/eetinis.v8i30.478.
- [34] A. M. Ayalew, N. W. Asnake, G. Demil, and M. Oussalah, "An explainable hybrid deep learning system for tuberculosis detection with Grad-CAM," *Discover Computing*, vol. 28, no. 1, Dec. 2025, doi: 10.1007/s10791-025-09791-z.
- [35] M. Oloko-Oba and S. Viriri, "Ensemble of EfficientNets for the Diagnosis of Tuberculosis," *Computational Intelligence and Neuroscience*, vol. 2021, 2021, doi: 10.1155/2021/9790894.
- [36] V. Acharya et al., "AI-Assisted Tuberculosis Detection and Classification from Chest X-Rays Using a Deep Learning Normalization-Free Network Model," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/2399428.
- [37] A. U. Ibrahim, A. G. Kibarer, and F. Al-Turjman, "Computer-aided Detection of Tuberculosis from Microbiological and Radiographic Images," *Data Intelligence*, vol. 5, no. 4, pp. 1008–1032, Nov. 2023, doi: 10.1162/dint_a_00198.
- [38] T. Xu and Z. Yuan, "Convolution Neural Network with Coordinate Attention for the Automatic Detection of Pulmonary Tuberculosis Images on Chest X-Rays," *IEEE Access*, vol. 10, pp. 86710–86717, 2022, doi: 10.1109/ACCESS.2022.3199419.
- [39] J. Alshudukhi, S. Aljaloud, T. S. Alharbi, and S. Abebaw, "Convolutional Neural Network Architectures to Solve a Problem of Tuberculosis Classification Using X-Ray Images of the Lungs," *Journal of Nanomaterials*, vol. 2022, 2022, doi: 10.1155/2022/2509830.
- [40] S. Rajaraman, G. Zamzmi, L. Folio, P. Alderson, and S. Antani, "Chest x-ray bone suppression for improving classification of tuberculosis-consistent findings," *Diagnostics*, vol. 11, no. 5, May 2021, doi: 10.3390/diagnostics11050840.
- [41] J. Sanchez, A. Trueba, J. Cervantes, F. Garcia-Lamont, J. S. Ruiz Castilla, and K. Kantipudi, "Classification of radiological patterns of tuberculosis with a Convolutional neural network in x-ray images," *Electronic Letters on Computer Vision and Image Analysis*, vol. 23, no. 1, pp. 47–59, Jun. 2024, doi: 10.5565/rev/elcvia.1561.
- [42] S. M. Fati, E. M. Senan, and N. ElHakim, "Deep and Hybrid Learning Technique for Early Detection of Tuberculosis Based on X-ray Images Using Feature Fusion," *Applied Sciences*, vol. 12, no. 14, Jul. 2022, doi: 10.3390/app12147092.
- [43] V. Sharma, Nillmani, S. K. Gupta, and K. K. Shukla, "Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images," *Intelligent Medicine*, vol. 4, no. 2, pp. 104–113, May 2024, doi: 10.1016/j.imed.2023.06.001.
- [44] S. Rajaraman, G. Zamzmi, L. R. Folio, and S. Antani, "Detecting Tuberculosis-Consistent Findings in Lateral Chest X-Rays Using an Ensemble of CNNs and Vision Transformers," *Frontiers in Genetics*, vol. 13, Feb. 2022, doi: 10.3389/fgene.2022.864724.
- [45] Madyono and A. Nabilah, "Detection of Tuberculosis Disease in Lung X-ray Images Using the DenseNet121 Method," *International Journal of Engineering, Science and Information Technology*, vol. 5, no. 2, pp. 379–385, 2025, doi: 10.52088/ijesty.v5i2.853.
- [46] P. Nijati, S. Abudoubari, Y. Zimin, and A. Tuersun, "Development and application of a deep learning-based tuberculosis diagnostic assistance system in remote areas of Northwest China," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-22037-8.
- [47] S. H. Lee et al., "Development and validation of a deep learning model for detecting signs of tuberculosis on chest radiographs among US-bound immigrants and refugees," *PLOS Digital Health*, vol. 3, no. 9, Sep. 2024, doi: 10.1371/journal.pdig.0000612.
- [48] W. young Chung et al., "Development and Validation of Deep Learning-Based Infectivity Prediction in Pulmonary Tuberculosis Through Chest Radiography: Retrospective Study," *Journal of Medical Internet Research*, vol. 26, 2024, doi: 10.2196/58413.
- [49] S. Natarajan et al., "Early diagnosis and meta-agnostic model visualization of tuberculosis based on radiography images," *Scientific Reports*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-49195-x.
- [50] K. Gasmi et al., "Early Tuberculosis Detection via Privacy-Preserving, Adaptive-Weighted Deep Models," *Diagnostics*, vol. 16, no. 2, p. 204, Jan. 2026, doi: 10.3390/diagnostics16020204.
- [51] R. Wajgi et al., "Optimized tuberculosis classification system for chest X-ray images: Fusing hyperparameter tuning with transfer learning approaches," *Engineering Reports*, vol. 6, no. 11, Nov. 2024, doi: 10.1002/eng2.12906.
- [52] G. Siddharth, A. Ambekar, and N. Jayakumar, "Enhanced CoAtNet based hybrid deep learning architecture for automated tuberculosis detection in human chest X-rays," *BMC Medical Imaging*, vol. 25, no. 1, Dec. 2025, doi: 10.1186/s12880-025-01901-z.
- [53] A. García Seco de Herrera et al., "Ensemble deep learning architectures for detecting pulmonary tuberculosis in chest X-rays," *Scientific Reports*, vol. 16, no. 1, Dec. 2026, doi: 10.1038/s41598-025-30792-x.
- [54] K. Vanitha, T. R. Mahesh, V. V. Kumar, and S. Guluwadi, "Enhanced tuberculosis detection using Vision Transformers and explainable AI with a Grad-CAM approach on chest X-rays," *BMC Medical Imaging*, vol. 25, no. 1, Dec. 2025, doi: 10.1186/s12880-025-01630-3.
- [55] E. Kotei and R. Thirunavukarasu, "Ensemble Technique Coupled with Deep Transfer Learning Framework for Automatic Detection of Tuberculosis from Chest X-ray Radiographs," *Healthcare*, vol. 10, no. 11, Nov. 2022, doi: 10.3390/healthcare10112335.
- [56] A. Shati, A. Datta, A. Mansoor, and G. M. Hassan, "ETDHDNet: An advanced DenseNet-based extended texture descriptor for efficient tuberculosis prediction in CXR images," *Intelligent Based Medicine*, vol. 12, Jan. 2025, doi: 10.1016/j.ibmed.2025.100269.
- [57] Y. Sun, M. J. Norval, and Z. Wang, "Evaluation of image processing technologies for pulmonary tuberculosis detection based on deep learning convolutional neural networks," *Journal of Advances in Information Technology*, vol. 12, no. 3, pp. 253–259, Aug. 2021, doi: 10.12720/jait.12.3.253-259.
- [58] A. F. M. Saif, T. Imtiaz, C. Shahnaz, W. P. Zhu, and M. O. Ahmad, "Exploiting cascaded ensemble of features for the detection of tuberculosis using chest radiographs," *IEEE Access*, vol. 9, pp. 112388–112399, 2021, doi: 10.1109/ACCESS.2021.3102077.

- [59] B. U. Maheswari et al., "Explainable deep-neural-network supported scheme for tuberculosis detection from chest radiographs," *BMC Medical Imaging*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12880-024-01202-x.
- [60] H. Ghosh, P. Pavan Kumar, I. S. Rahat, M. D. M. H. Nipu, G. R. Krishna, and J. V. R. Ravindra, "From Pixels to Pathology: The Power of CNNs in Detecting Tuberculosis," *EAI Endorsed Transactions on Pervasive Health Technology*, vol. 10, Jan. 2024, doi: 10.4108/eetpht.10.5543.
- [61] A. Tiwari and A. Katiyar, "Hybrid deep learning assisted chest X-ray image segmentation and classification for tuberculosis disease diagnosis," *Intelligent Decision Technologies*, vol. 18, no. 1, pp. 561–569, Feb. 2024, doi: 10.3233/IDT-230286.
- [62] R. Yulvina et al., "Hybrid Vision Transformer and Convolutional Neural Network for Multi-Class and Multi-Label Classification of Tuberculosis Anomalies on Chest X-Ray," *Computers*, vol. 13, no. 12, Dec. 2024, doi: 10.3390/computers13120343.
- [63] W. C. Chiu, S. Y. Chang, C. Lin, T. W. Chen, and W. H. Fang, "Integrating AI with PCR for Tuberculosis Diagnosis: Evaluating a Deep Learning Model for Chest X-Rays," *Bioengineering*, vol. 12, no. 12, Dec. 2025, doi: 10.3390/bioengineering12121377.
- [64] M. Irhamsyah, Q. A'yuni, K. Saddami, N. Nasaruddin, K. Munadi, and F. Arnia, "Impact of Using Various X-Ray Dataset in Detecting Tuberculosis Based on Deep Learning," *Radioelectronic and Computer Systems*, vol. 2025, no. 1, pp. 165–186, 2025, doi: 10.32620/reks.2025.1.12.
- [65] A. Nawaz, M. Shehab, M. R. R. Rana, B. Qureshi, Z. Khan, and M. Babar, "IMPACT-TB: Integrated Medical Imaging and AI for Precise Tuberculosis Detection," *Engineering Reports*, vol. 7, no. 8, Aug. 2025, doi: 10.1002/eng2.70356.
- [66] R. D. Ishwerlal, R. Agarwal, and K. S. Sujatha, "Improved Pyramid Histogram of Oriented Gradients and Statistical Features for Tuberculosis Detection via a Hybrid Classifier Using X-Ray Images," *Engineering, Technology and Applied Science Research*, vol. 15, no. 6, pp. 30387–30394, Dec. 2025, doi: 10.48084/etasr.11513.
- [67] S. Rajaraman, L. R. Folio, J. Dimperio, P. O. Alderson, and S. K. Antani, "Improved semantic segmentation of tuberculosis-consistent findings in chest x-rays using augmented training of modality-specific u-net models with weak localizations," *Diagnostics*, vol. 11, no. 4, Apr. 2021, doi: 10.3390/diagnostics11040616.
- [68] D. Capellán-Martín et al., "Multi-view deep learning framework for the detection of chest X-rays compatible with pediatric pulmonary tuberculosis," *Nature Communications*, vol. 16, no. 1, Dec. 2025, doi: 10.1038/s41467-025-64391-1.
- [69] M. Alsaffar et al., "Detection of Tuberculosis Disease Using Image Processing Technique," *Mobile Information Systems*, vol. 2021, 2021, doi: 10.1155/2021/7424836.
- [70] S. M. Sundari, V. Sukhavasi, D. Swapna, D. Kbks, P. S. Lunawat, and M. M. Ammangatambu, "Optimized AlexNet For Enhanced Tuberculosis Classification Using Deep Learning," *International Journal of Computing and Digital Systems*, vol. 16, no. 1, pp. 1509–1522, Jul. 2024, doi: 10.12785/ijcds/1601112.
- [71] M. A. Pala and M. B. Navdar, "SPX-GNN: An Explainable Graph Neural Network for Harnessing Long-Range Dependencies in Tuberculosis Classifications in Chest X-Ray Images," *Diagnostics*, vol. 15, no. 24, Dec. 2025, doi: 10.3390/diagnostics15243236.
- [72] T. H. Kim, M. Krichen, S. Ojo, M. A. Alamro, and G. A. Sampedro, "TSSG-CNN: A Tuberculosis Semantic Segmentation-Guided Model for Detecting and Diagnosis Using the Adaptive Convolutional Neural Network," *Diagnostics*, vol. 14, no. 11, Jun. 2024, doi: 10.3390/diagnostics14111174.
- [73] A. Wong, J. R. H. Lee, H. Rahmat-Khah, A. Sabri, A. Alaref, and H. Liu, "TB-Net: A Tailored, Self-Attention Deep Convolutional Neural Network Design for Detection of Tuberculosis Cases From Chest X-Ray Images," *Frontiers in Artificial Intelligence*, vol. 5, Apr. 2022, doi: 10.3389/frai.2022.827299.
- [74] S. Khumang, S. Kansomkeat, W. Tanomkiat, and S. Intajag, "Tuberculosis Lesion Segmentation Improvement in X-Ray Images Using Contextual Background Label," *IEEE Access*, vol. 13, pp. 36611–36625, 2025, doi: 10.1109/ACCESS.2025.3532631.
- [75] S. Rajaraman, G. Zamzmi, F. Yang, Z. Xue, S. Jaeger, and S. K. Antani, "Uncertainty Quantification in Segmenting Tuberculosis-Consistent Findings in Frontal Chest X-rays," *Biomedicines*, vol. 10, no. 6, Jun. 2022, doi: 10.3390/biomedicines10061323.
- [76] K. Riasat, A. Jamil, S. Al-Otaibi, S. Zeb, S. Riasat, and S. Kanwal, "Tuberculosis detection using few shot learning," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-97803-9.
- [77] B. H. Shekar and S. Mannan, "Tuberculosis detection with customized CNN and oversampling techniques: a deep learning approach," *Discover Artificial Intelligence*, vol. 5, no. 1, Dec. 2025, doi: 10.1007/s44163-025-00329-2.
- [78] 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, doi: 10.1109/ACCESS.2024.3428446.
- [79] I. Ahmad and S. Shin, "A Perceptual Encryption-Based Image Communication System for Deep Learning-Based Tuberculosis Diagnosis Using Healthcare Cloud Services," *Electronics*, vol. 11, no. 16, Aug. 2022, doi: 10.3390/electronics11162514.
- [80] J. Devasia, H. Goswami, S. Lakshminarayanan, M. Rajaram, and S. Adithan, "Deep learning classification of active tuberculosis lung zones wise manifestations using chest X-rays: a multi label approach," *Scientific Reports*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-28079-0.
- [81] Z. Nabulsi et al., "Deep learning for distinguishing normal versus abnormal chest radiographs and generalization to two unseen diseases tuberculosis and COVID-19," *Scientific Reports*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-93967-2.
- [82] C. Y. Ou et al., "Deep Learning-Based Classification and Semantic Segmentation of Lung Tuberculosis Lesions in Chest X-ray Images," *Diagnostics*, vol. 14, no. 9, May 2024, doi: 10.3390/diagnostics14090952.
- [83] C. H. Wang et al., "Deep Learning-based Diagnosis of Pulmonary Tuberculosis on Chest X-ray in the Emergency Department: A Retrospective Study," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 2, pp. 589–600, Apr. 2024, doi: 10.1007/s10278-023-00952-4.
- [84] L. Eisentraut, C. Mai, J. Hosch, A. Benecke, P. Penava, and R. Buettner, "Deep Learning-Based Detection of Tuberculosis Using a Gaussian Chest X-Ray Image Filter as a Software Lens," *IEEE Access*, vol. 13, pp. 36065–36081, 2025, doi: 10.1109/ACCESS.2025.3544923.