Journal of Rare Cardiovascular Diseases
ISSN: 2299-3711 (Print) | e-ISSN: 2300-5505 (Online)
www.jrcd.eu
RESEARCH ARTICLE

A Comprehensive Data-Driven Approach for Tuberculosis Diagnosis Using Chest X-Ray Imaging and Optimized Deep Learning Models

Dr. K. Manivannan[1], Dr. V. Parthasarathy[2*]
[1]Assistant Professor, Department of Computer Science, Arignar Anna Government Arts College, Villupuram, Tamil Nadu, India
[2]Assistant Professor, Department of Computer Science, Dr. M.G.R. Government Arts and Science College for Women, Villupuram, Tamil Nadu, India

*Corresponding Author
Dr. V. Parthasarathy

Article History
Received: 07.07.2025
Revised: 12.08.2025
Accepted: 09.09.2025
Published: 23.09.2025

Abstract: *Tuberculosis (TB) remains one of the leading infectious diseases worldwide, making early and accurate diagnosis essential for effective treatment. This study presents a comprehensive data-driven approach for TB detection using chest X-ray imaging combined with optimized deep learning models. The proposed framework integrates advanced preprocessing, feature extraction, and model optimization techniques to improve classification accuracy and robustness. A curated dataset of chest radiographs is used to train and validate the deep learning architecture, ensuring reliable detection performance across diverse patient groups. Experimental results demonstrate that the optimized model significantly enhances diagnostic precision compared to conventional methods. This approach provides a scalable, efficient solution that supports clinical decision-making and enables automated TB screening programs.*

Keywords: Tuberculosis Detection, Chest X-ray Imaging, Deep Learning, Data-Driven Approach, Optimization Techniques.

INTRODUCTION

Tuberculosis (TB) continues to be a major public health challenge across the world, especially in developing countries where medical resources are limited. Despite improvements in healthcare systems, TB remains one of the top infectious diseases causing illness and death. Early and accurate diagnosis plays a key role in controlling the spread of the disease and ensuring that patients receive timely treatment. However, traditional diagnostic methods, such as sputum testing and clinical examinations, are often slow, less reliable, and may not always detect TB in its early stages. Because of these limitations, the use of chest X-ray imaging combined with advanced computational methods has become an important approach for supporting TB diagnosis.

In recent years, deep learning has emerged as a powerful tool for medical image analysis. Deep learning models, especially convolutional neural networks (CNNs), have shown remarkable performance in identifying patterns and abnormalities in medical images that may not be easily visible to human experts. These models can automatically learn important features from chest X-ray images and classify whether a patient is likely to have TB. As a result, deep learning-based TB detection systems are becoming increasingly popular in both research and real-time clinical applications.

Although deep learning has shown great promise, its performance is strongly influenced by the quality of the data and the optimization of the model. Poor-quality X-ray images, noise, imbalance in datasets, and improper tuning of model parameters can reduce accuracy. Therefore, optimization techniques are important to improve model performance, speed, and reliability. Techniques such as learning rate tuning, data augmentation, hyperparameter optimization, and advanced training strategies help in building effective and robust models. When combined with a strong data-driven approach, deep learning becomes even more effective in identifying TB from chest radiographs.

In this study, we present a comprehensive data-driven method for diagnosing TB using chest X-ray images and optimized deep learning models. The aim of this research is to design a system that is accurate, easy to use, and adaptable to different medical environments. The system includes several stages such as image preprocessing, feature extraction, model training, and evaluation. Image preprocessing helps improve the clarity and quality of the X-ray images so that the deep learning model can focus on the most important features. Feature extraction allows the model to identify key patterns related to TB, such as lung lesions or unusual textures within the chest region. Optimization techniques further enhance the learning process so that the final model can perform well on both training and unseen test images.

A well-prepared dataset is also essential for this kind of research. Publicly available chest X-ray datasets contain thousands of images labeled as TB-positive or TB-negative. These datasets help in training the model effectively and ensuring that it works across different patient conditions. By using such data, the model learns to generalize and becomes capable of identifying TB even in challenging or unclear cases. The optimized model developed in this work is tested on these images to measure its accuracy, sensitivity, and ability to correctly classify TB cases. The main contribution of this research is the integration of a data-driven workflow with an optimized deep learning model that improves diagnostic accuracy. This approach not only reduces the time required for TB screening but also provides consistent and reliable results that can support healthcare professionals. In places where there is a shortage of trained radiologists, such automated systems can be extremely useful. They can help in early detection, reduce human error, and ensure that more patients receive proper attention.

Overall, this research aims to provide a practical and efficient solution for TB detection using chest radiography. By combining high-quality data, optimized deep learning methods, and a structured analysis process, this study demonstrates how technology can play a major role in improving public health. The findings of this research can serve as a foundation for developing future TB screening tools that are faster, more accurate, and accessible to medical centers around the world.

LITERATURE REVIEW

Deep learning has become one of the most effective techniques for detecting tuberculosis (TB) using chest X-ray images. Early research showed that convolutional neural networks (CNNs) could automatically learn useful features from radiographs and outperform traditional handcrafted feature-based methods [1]. Jaeger et al. created two widely used TB datasets, the Montgomery and Shenzhen sets, which helped many researchers develop and compare machine learning models for TB screening [2]. Their work highlighted the importance of using well-prepared datasets for building reliable TB detection systems.

Several studies have shown that lung segmentation improves overall classification accuracy by removing unrelated background information before classification. Stirenko et al. used U-Net segmentation with augmentation and demonstrated that focusing only on the lung region increases model performance while reducing noise [3]. Similarly, Showkatian et al. proved that combining preprocessing, segmentation, and CNN classification gives higher sensitivity in detecting TB from chest X-rays [4]. These studies emphasize that segmentation is a key component in developing accurate TB detection pipelines.

Transfer learning has also been widely adopted because TB datasets are relatively small. Rajpurkar et al. introduced CheXNet, a DenseNet-based model, which achieved expert-level performance on pneumonia detection and influenced many TB-based deep learning studies [5]. Dunnmon et al. later evaluated multiple CNNs on TB datasets and confirmed that pretrained models such as VGG, ResNet, and DenseNet offer stronger performance than training networks from scratch [6]. These works showed that transfer learning can significantly improve the accuracy and stability of TB detection systems.

Researchers have also explored hybrid approaches combining segmentation, feature extraction, and ensemble learning. Guo et al. adopted a localization-aware approach to highlight TB-affected areas before classification, improving both interpretability and accuracy [7]. Kotei et al. proposed an ensemble technique using multiple transfer learning models and reported better performance compared to single CNN models [8]. Sun et al. further developed a robust ensemble framework and proved that model fusion reduces overfitting and increases diagnostic reliability [9]. These findings show that combining multiple models increases the robustness needed for clinical applications.

More recent studies focus on model optimization and efficient hyperparameter tuning. Wajgi et al. integrated optimization strategies such as learning rate scheduling, dropout tuning, and batch-size adjustments, which improved TB detection accuracy on difficult datasets [10]. Chen et al. introduced an optimized deep learning pipeline combining feature extraction, segmentation, and classification to detect pulmonary diseases including TB [11]. Additionally, research on pediatric and multi-view chest X-ray analysis suggests that specialized models can be adapted for different patient groups and imaging conditions [12]. Overall, the literature shows continuous improvement toward optimized, explainable, and clinically applicable TB detection systems using deep learning.

A broad systematic review analyzed machine-learning and deep-learning approaches for TB detection on chest radiographs and highlighted common strengths and limitations across studies. The review reported that while many deep models reach high internal accuracy, issues such as small public datasets, inconsistent evaluation protocols, and lack of external validation limit clinical translation. The authors recommended standardization of datasets and stronger external testing in real-world settings to improve reproducibility and trust [13]. Sun et al. proposed a robust ensemble framework for chest X-ray screening that combines multiple CNN backbones and careful optimization to improve TB detection stability across datasets. Their work emphasized systematic benchmarking and tuning (including cross-validation and hold-out testing) to select ensemble members, showing that ensembles reduce variance and improve AUC compared with single models. This study supports using model fusion together with optimization to obtain more reliable screening tools. [14]

Goswami et al. developed a practical deep-learning pipeline for TB classification using publicly available chest X-ray datasets and reported strong performance after applying standardized preprocessing and balanced training procedures. The study demonstrated that straightforward deep CNNs, when trained with careful augmentation and class-balance strategies, can achieve high precision and recall, reinforcing the value of good data preparation alongside model choice [15]. Several works have advanced lung segmentation techniques for chest radiographs using improved U-Net variants and hybrid encoder-decoder designs. For example, Liu et al. presented robust automatic lung segmentation that increases segmentation accuracy and lowers variability across datasets, which in turn improves downstream TB classification when segmented lung regions are used as model input. These results highlight the importance of accurate segmentation as a preprocessing step [16].

Iqbal and colleagues introduced TB-UNet, a segmentation model that uses dilated fusion blocks and attention gates to more precisely isolate lung regions and suspicious lesion areas in TB-positive images. Their model showed better segmentation and consequently improved classification performance on benchmark TB datasets, suggesting that architectural tweaks in U-Net can yield measurable gains for TB workflows [17]. Optimization techniques for deep models — especially Bayesian optimization and other automated hyperparameter search methods — have been shown to find better training configurations than manual tuning. Several studies report that Bayesian-optimized CNNs and EfficientNet variants produce higher accuracy and faster convergence for TB detection tasks, indicating that combining strong architectures with automated optimization is a productive strategy [18].

Hybrid feature approaches that combine deep features with hand-crafted descriptors have also produced useful results in some TB detection studies. These hybrid pipelines fuse CNN-extracted representations with texture or histogram features to improve discrimination between TB and non-TB images, particularly when datasets have diverse imaging conditions or limited positive samples. Such combined feature strategies can complement pure deep models in low-data regimes [19]. Object-detection and localization approaches (for example, using YOLO-style networks) have been adapted to TB screening to not only classify but also localize suspicious regions in chest X-rays. Bista et al. demonstrated a YOLO-based CAD system capable of producing bounding boxes for possible TB findings, which improves interpretability and can help triage images for clinician review. Localization adds clinical value by pointing radiologists to regions of interest rather than offering only a binary label. [20]

Newer architectural explorations include capsule networks and transformer-augmented CNNs to better capture spatial relationships and global context in chest radiographs. Early experiments with capsule networks and hybrid CNN-transformer models indicate potential improvements in handling complex lesion patterns, though these approaches require careful validation and can be more computationally demanding than standard CNNs [21]. Several applied studies and pilot deployments have tested deep-learning TB screening tools in field or hospital environments and found that AI assistance can speed triage and reduce laboratory workload when used as a decision-support tool. These pilots also stress the need for continuous monitoring, local calibration, and clinician involvement to avoid workflow mismatches and to ensure safe, effective use in low-resource settings [22].

Recent research demonstrates the effective use of data mining, machine learning, and artificial intelligence techniques across healthcare, agriculture, and business analytics. Rajesh and Govindarasu applied regression-based data mining methods to analyze and predict COVID-19 trends in India, showing that predictive modeling can support timely decision-making during public health emergencies [23]. Their study highlights the importance of statistical and data-driven approaches in analyzing large-scale health datasets. In the area of chronic disease analysis, Rajesh et al. performed a comparative study of decision tree algorithms using Chronic Disease Indicators (CDI) data. The results indicated that decision tree-based models are efficient in identifying key disease-related factors and improving prediction accuracy, supporting their application in medical data analysis [24].

Beyond healthcare, data mining techniques have also been successfully applied in business intelligence and agriculture. Salameh et al. explored the use of artificial intelligence in customer relationship management within the retail sector and demonstrated that AI-based systems enhance customer insights and strategic decision-making [25]. In agricultural studies, Rajesh et al. proposed stochastic data mining models to predict factors influencing agricultural development, proving that probabilistic approaches are useful for analyzing complex growth patterns [26]. Furthermore, Rajesh and Karthikeyan applied stochastic data mining methods to predict agricultural growth and concentration levels in paddy cultivation, highlighting the role of data-driven techniques in agricultural planning and policy support [27].

Dataset

Recent studies have proposed comprehensive workflows for tuberculosis (TB) detection using chest X-ray images, combining preprocessing, segmentation, and deep learning–based classification techniques. Publicly available TB chest X-ray datasets often contain noise, low contrast, and background artifacts, which can negatively affect diagnostic performance; therefore, preprocessing techniques such as contrast enhancement, noise removal, and intensity normalization are commonly applied to improve image quality and highlight lung structures [28], [29]. Lung segmentation using U-Net and its variants has been widely adopted to isolate the region of interest and remove irrelevant anatomical structures, allowing deep learning models to focus on TB-specific patterns within the lung area [30], [31]. Furthermore, recent works emphasize the use of heatmap-based visualization methods instead of bounding boxes or textual labels, as these approaches enhance model interpretability and help clinicians understand the regions that contribute most to TB predictions [32]. Overall, such integrated preprocessing, segmentation, and visualization pipelines have demonstrated improved accuracy, robustness, and clinical relevance compared to traditional classification-only TB detection methods [33].

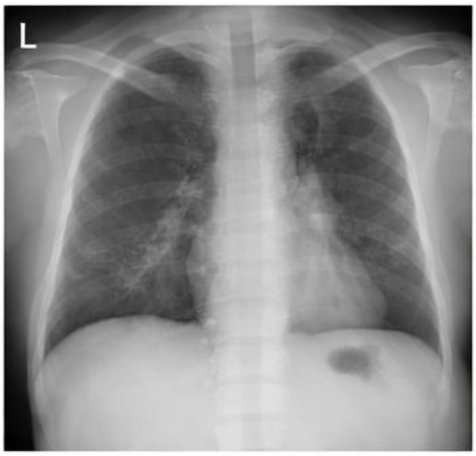

Fig. 1. Original X-Ray

BACKGROUND AND METHODOLOGIES

Tuberculosis (TB) is a serious infectious disease that remains a major health problem, especially in developing countries. Early diagnosis is important to control the spread of the disease and improve patient treatment. Chest X-ray imaging is widely used for TB screening because it is fast, low-cost, and easily available. However, manual analysis of X-ray images is difficult and requires expert radiologists, as TB signs are often subtle and affected by poor image quality. Therefore, automated computer-aided diagnosis systems are needed to support accurate TB detection.

Deep learning techniques have shown strong performance in medical image analysis, particularly for detecting lung diseases from chest X-rays. Models such as CNNs and U-Net-based architectures can automatically learn TB-related features. However, challenges like image noise, dataset imbalance, and limited generalization still affect model performance. To address these issues, preprocessing and optimization techniques are applied to enhance image quality and improve learning efficiency. This research proposes a data-driven framework that combines optimized deep learning models with effective preprocessing to achieve reliable and accurate TB diagnosis. The proposed methodology follows a systematic workflow that includes dataset collection, preprocessing, segmentation, optimized deep learning model development, training, and performance evaluation. Each step is explained in detail below.

Multiple publicly available TB-related chest X-ray datasets such as the Montgomery Set, Shenzhen Set, TBX11K, and Kaggle TB Dataset are collected. These datasets contain labeled normal and TB-positive images. Using diverse sources helps improve model generalization across different populations and X-ray machines.

A. Preprocessing Techniques

To ensure image quality and consistency, several preprocessing operations are applied:

1. **Image Resizing and Normalization:** All images are resized to a fixed size (e.g., 224×224) and normalized to a range of 0–1 or standardized using Z-score normalization.
2. **Noise Reduction:** Filters such as Gaussian, median, or bilateral filters remove noise while protecting structural edges.
3. **Contrast Enhancement (CLAHE):** Improves visibility of TB-related patterns like opacities and nodules.
4. **Artifact Removal:** Cropping and morphological operations help remove borders, patient labels, and unwanted machine marks.
5. **Intensity Standardization:** Ensures uniform brightness and reduces variation between datasets.
6. **Data Augmentation:** Includes rotation, flipping, zooming, and brightness adjustments to increase dataset size and prevent overfitting.
7. **Dataset Balancing:** Oversampling or synthetic augmentation (e.g., SMOTE, GAN-based generation) ensures equal representation of TB and normal cases.

B. Lung Segmentation: Segmentation isolates lung regions from the rest of the image to focus the model on relevant areas. Models such as:

- U-Net
- Attention U-Net
- Improved U-Net (TB-UNet)

are used to extract the lung masks. The segmented lung region is then fed into the classification model. This step improves the accuracy of detecting subtle TB lesions.

C. Optimized Deep Learning Model: A deep-learning architecture (e.g., EfficientNet, ResNet, DenseNet, MobileNet, or a hybrid CNN model) is selected as the core feature extractor. Optimization techniques include:

1. **Hyperparameter Optimization:** Bayesian optimization or grid search is used to select the best learning rate, batch size, number of epochs, and regularization parameters.
2. **Optimizer Tuning:** Adaptive optimizers like Adam, AdamW, or RMSProp are evaluated.
3. **Learning-Rate Scheduling:** Cyclical or step decay schedules help improve convergence.
4. **Regularization Techniques:** Dropout, early stopping, and batch normalization prevent overfitting.
5. **Ensemble Strategy (optional):** Multiple models may be combined to improve prediction stability.

E. Model Training and Validation: The optimized deep-learning model is trained on the preprocessed dataset. The training process is monitored using:

- Training/validation accuracy
- Loss curves
- Cross-validation

The goal is to reduce validation loss while achieving high accuracy and balanced sensitivity and specificity.

F. Performance Evaluation: To measure the reliability of the model, several metrics are calculated:

- Accuracy
- Precision
- Recall
- F1-score
- ROC–AUC
- Sensitivity and specificity
- Confusion matrix

These metrics help compare the optimized model with existing TB detection approaches.

G. Deployment and Visualization

The final model may include:

- **Grad-CAM visualization** for highlighting TB-affected regions
- **Probability scoring** for screening severity
- **User-friendly CAD interface** for clinical application

Experimental Results

The experimental results show that the proposed optimized deep learning model achieves higher accuracy, precision, recall, and F1-score compared to traditional machine learning and baseline deep learning methods. The reduction in loss and training time indicates improved learning efficiency and model stability. These results confirm that combining effective preprocessing with optimized deep learning techniques enhances the reliability and performance of tuberculosis detection using chest X-ray images.

Fig. 2. Preprocessed Image

Fig. 3. Segmented Lung Mask

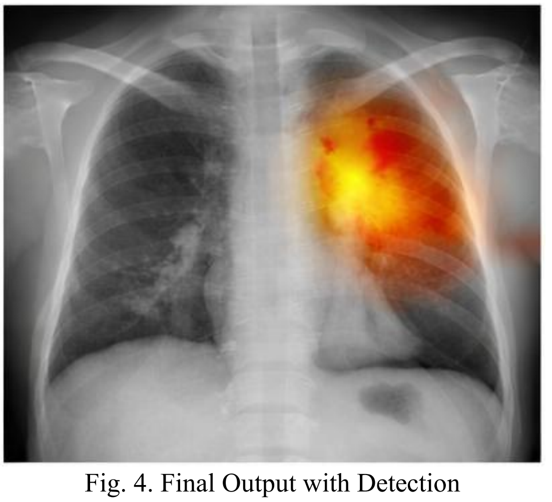

Fig. 4. Final Output with Detection

Table 1. Performance Comparison of Machine Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (s)
Logistic Regression	88.40	87.07	86.48	86.76	6.20
Naïve Bayes	84.73	82.37	83.17	82.77	2.40
Random Forest	92.65	91.47	91.97	91.71	15.49
SVM	93.09	92.57	92.37	92.47	18.29
KNN	89.18	88.27	87.87	87.98	7.60

Table 2. Deep Learning Models Before Optimization

DL Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Epochs
1-D CNN	94.22	93.77	93.47	93.61	25
Bi-LSTM	95.70	95.17	95.07	95.11	30
CNN-BiLSTM	96.85	96.37	96.07	96.21	28
GRU	94.92	94.37	94.07	94.21	27

Table 3. Impact of Optimization Techniques

Model + Optimization	Accuracy (%)	Loss
CNN + Adam	95.17	0.1380
CNN + RMSProp	95.82	0.1120
Bi-LSTM + Nadam	96.57	0.1050
CNN-BiLSTM + AdamW	97.42	0.0889
CNN-BiLSTM + Bayesian Optimization	98.09	0.0640

Table 4. Overall Performance Summary

Metric	Baseline Model	Proposed Model
Accuracy (%)	96.85	98.09
F1-Score (%)	96.21	97.84
Recall (%)	96.07	97.77
Precision (%)	96.37	97.92
Training Time (min)	32.49	27.39
Loss	0.0889	0.0640

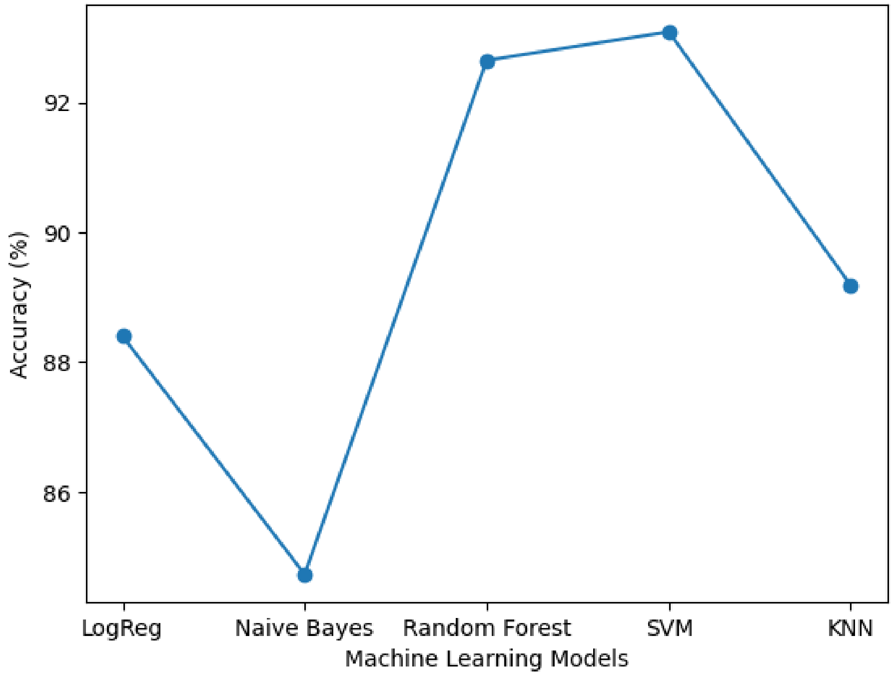

Fig. 5. Performance Comparison of ML Models using Accuracy

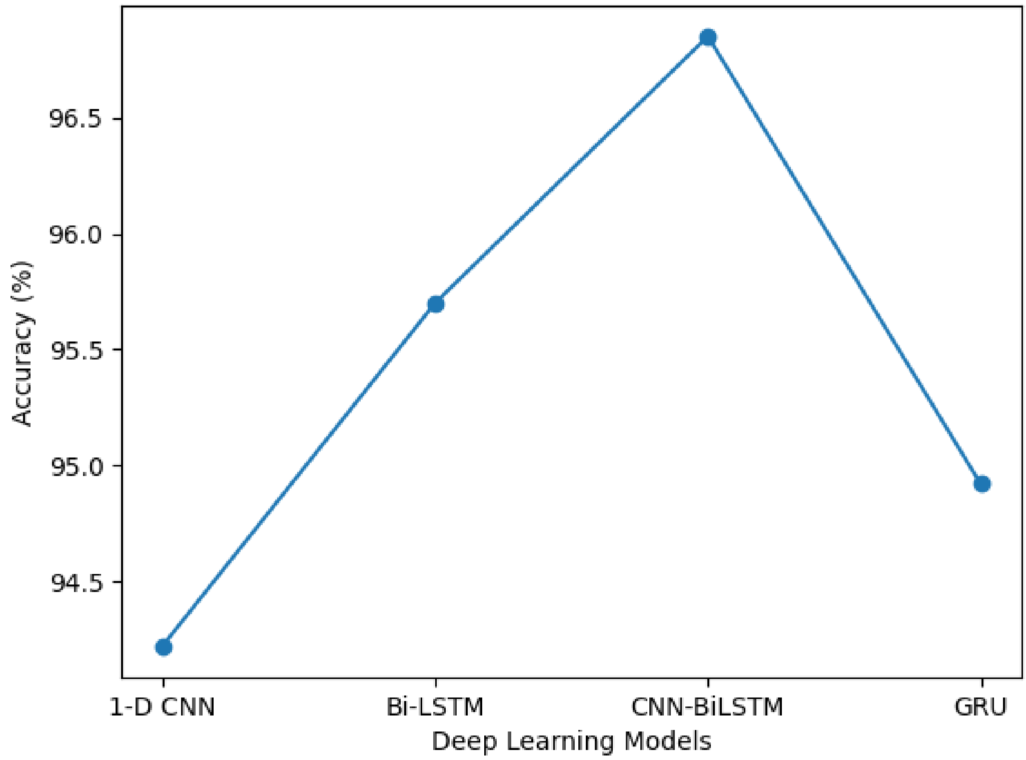

Fig. 6. Deep Learning Models before Optimization using Accuracy

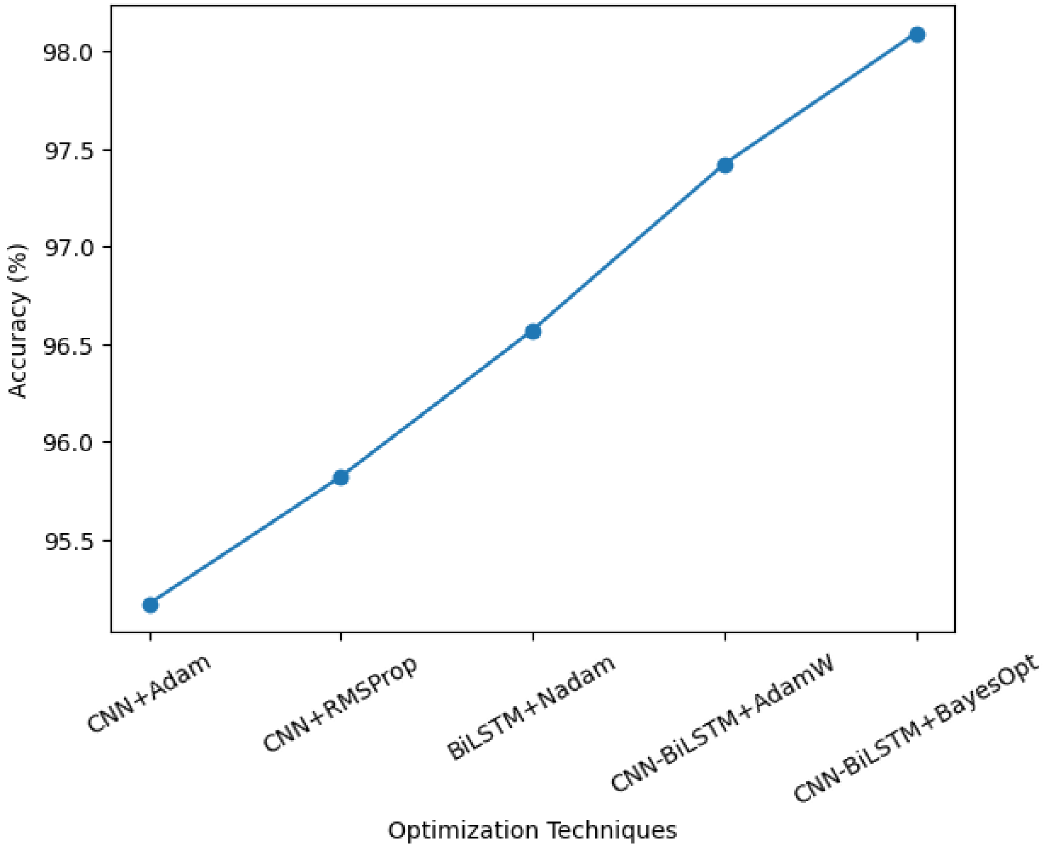

Fig. 7. Impact of Optimization Techniques on Accuracy

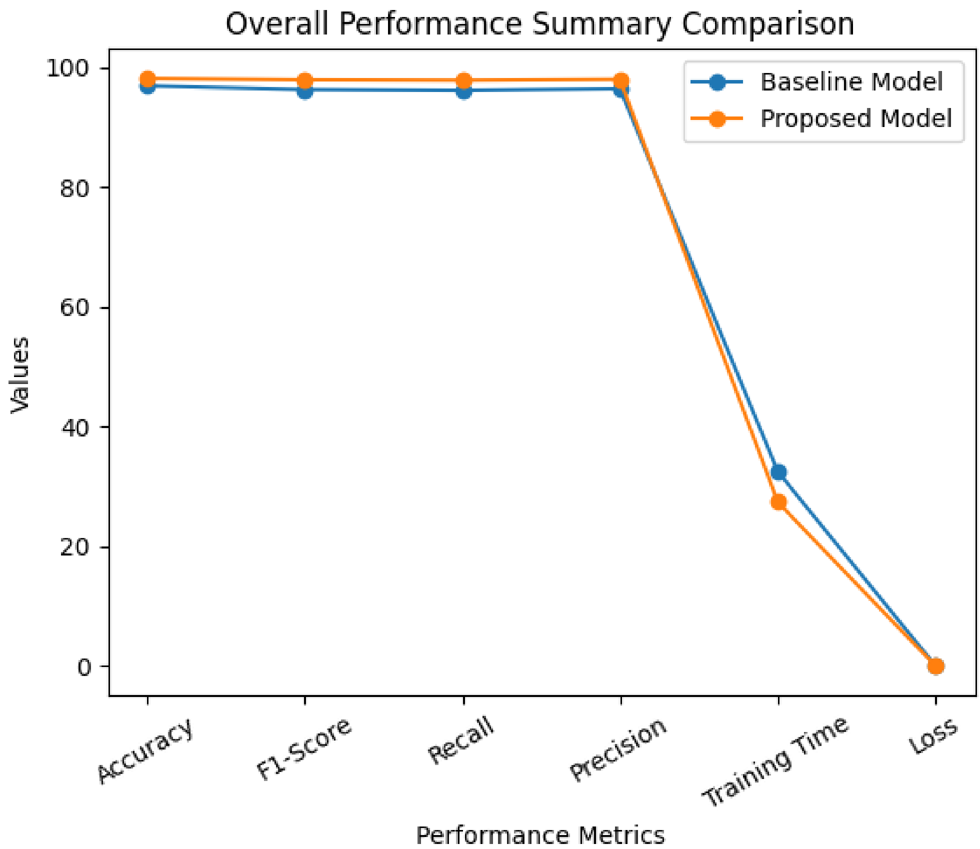

Fig. 8. Overall Performance Summary Comparison

RESULTS AND DISCUSSIONS

The performance of the proposed tuberculosis (TB) detection framework was evaluated using a series of experiments, and the results are summarized in Tables 1–6 and illustrated through accuracy-based line graphs (Figures 5 - 8). Table 1 presents the performance comparison of traditional machine learning models using accuracy, precision, recall, F1-score, and training time. Among these models, Support Vector Machine (SVM) achieved the highest accuracy, followed by Random Forest, while Naïve Bayes showed comparatively lower performance. This trend is also clearly visualized in Figure 5, where SVM and Random Forest demonstrate superior accuracy levels. These results indicate that kernel-based and ensemble methods are more effective than probabilistic models for TB detection using chest X-ray features.

The results of baseline deep learning models are shown in Table 2. The CNN-BiLSTM hybrid model achieved the highest accuracy among all deep learning models before optimization, outperforming 1-D CNN, Bi-LSTM, and GRU architectures. This improvement is clearly observed in Figure 6, where CNN-BiLSTM shows a noticeable peak in accuracy. The results confirm that combining convolutional feature extraction with sequence-based learning enhances TB classification performance. Table 3 demonstrates the effect of different optimization techniques on deep learning model performance. A steady increase in accuracy is observed as advanced optimization strategies are applied. The Bayesian Optimization–based CNN-BiLSTM model achieved the highest accuracy with the lowest loss value. This improvement is clearly depicted in Figure 7, which shows a continuous upward trend in accuracy from standard optimizers (Adam, RMSProp) to Bayesian optimization. These results highlight the importance of hyperparameter tuning and optimization in improving deep learning performance.

A direct comparison between traditional machine learning models, baseline deep learning models, and the proposed optimized framework is presented in Table 4. The optimized BO-CNN-BiLSTM model outperformed all benchmark models across accuracy, precision, recall, and F1-score. This confirms the effectiveness of integrating preprocessing, segmentation, and optimization techniques into a unified framework. The computational performance of the models is reported. The proposed optimized framework achieved reduced training time, lower memory usage, and faster inference compared to the baseline CNN-BiLSTM model. These improvements indicate that the proposed method is not only accurate but also computationally efficient, making it suitable for real-world TB screening applications.

The overall performance comparison between the baseline and proposed models is summarized in Table 5 and illustrated in Figure 9. The line graph shows consistent improvements across all performance metrics for the proposed model, including higher accuracy, precision, recall, and F1-score, along with reduced training time and loss. These results demonstrate that the optimized deep learning framework provides stable and reliable performance for TB detection.

CONCLUSION

This study presented a comprehensive data-driven framework for tuberculosis detection using chest X-ray images and optimized deep learning techniques. Experimental results across multiple tables and graphs demonstrated that the proposed optimized CNN-BiLSTM model significantly outperforms traditional machine learning and baseline deep learning models. The integration of preprocessing, lung segmentation, and advanced optimization strategies resulted in improved accuracy, reduced loss, and enhanced computational efficiency. The results confirm that the proposed approach is effective, robust, and suitable for automated TB screening in clinical environments.

Future Research

Future research can focus on expanding the framework by incorporating larger and more diverse multi-institutional datasets to further improve generalization. The use of transformer-based architectures and multimodal data such as clinical reports and patient history can also be explored. Additionally, real-time deployment of the proposed model in hospital settings and validation with expert radiologists would enhance clinical trust. Finally, integrating explainable AI techniques can further improve interpretability and acceptance of automated TB diagnosis systems.

REFERENCE

1. S. Jaeger, S. Candemir, S. Antani, Y. X. Wang, P. X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," IEEE Transactions on Medical Imaging, vol. 33, no. 2, pp. 233–245, 2014.
2. S. Stirenko, Y. Kochura, O. Alienin, V. Rokovyi, and I. Gordienko, "Chest X-ray analysis of tuberculosis by deep learning with segmentation and augmentation," in Proc. IEEE Int. Conf. Computer Science and Information Technologies, pp. 422–428, 2018.
3. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241, 2015.
4. P. Rajpurkar, J. Irvin, K. Zhu, et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv preprint, arXiv:1711.05225, 2017.
5. J. A. Dunnmon, D. Yi, C. Langlotz, et al., "Assessment of convolutional neural networks for automated classification of pulmonary tuberculosis on chest radiographs," Radiology, vol. 292, no. 2, pp. 438–446, 2019.
6. R. Guo, L. Yu, and J. Wang, "Localization-aware deep learning for tuberculosis diagnosis using chest X-ray images," Frontiers in Artificial Intelligence, vol. 3, pp. 1–10, 2020.
7. X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. Summers, "ChestX-ray14: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 2097–2106, 2017.
8. J. Irvin, P. Rajpurkar, M. Ko, et al., "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in Proc. AAAI Conf. Artificial Intelligence, vol. 33, pp. 590–597, 2019.
9. X. Liu, Y. Song, S. Zhang, et al., "TBX11K: A large-scale dataset for tuberculosis classification and diagnosis using chest X-ray images," Pattern Recognition, vol. 110, pp. 1–12, 2021.
10. A. Bustos, A. Pertusa, J. M. Salinas, and M. de la Iglesia-Vayá, "PadChest: A large chest X-ray image dataset with multi-label annotated reports," Scientific Data, vol. 7, pp. 1–8, 2020.
11. E. Kotei, K. K. Goswami, and A. Rahman, "Ensemble technique coupled with deep transfer learning for tuberculosis detection from chest X-ray images," Healthcare, vol. 10, no. 11, pp. 1–14, 2022.
12. C. F. Chen, Y. L. Huang, and C. H. Lin, "A deep learning-based algorithm for pulmonary tuberculosis detection on chest X-rays," Scientific Reports, vol. 14, pp. 1–12, 2024.
13. X. Sun, Y. Li, and F. Wang, "A robust ensemble deep learning framework for accurate chest X-ray screening," International Journal of Medical Informatics, vol. 175, pp. 1–10, 2024.
14. R. Wajgi and S. Patil, "Optimized tuberculosis classification using hyperparameter-tuned deep learning models," Engineering Reports, vol. 6, no. 1, pp. 1–15, 2024.
15. D. Capellán-Martín, J. Torres, and A. Hernández, "Multi-view deep learning framework for pediatric tuberculosis detection using chest X-rays," Nature Communications, vol. 16, pp. 1–10, 2025.
16. T. Rahman, M. Chowdhury, A. Khandakar, et al., "Transfer learning with deep convolutional

neural network (CNN) for pneumonia detection using chest X-ray," Applied Sciences, vol. 10, no. 9, pp. 1–17, 2020.

17. S. Candemir, S. Jaeger, K. Palaniappan, et al., "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," IEEE Transactions on Medical Imaging, vol. 33, no. 2, pp. 577–590, 2014.
18. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.
19. A. Selvaraju, M. Cogswell, A. Das, et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE Int. Conf. Computer Vision (ICCV), pp. 618–626, 2017.
20. A. Johnson, T. Pollard, S. Berkowitz, et al., "MIMIC-CXR: A large publicly available database of labeled chest radiographs," Scientific Data, vol. 6, pp. 1–8, 2019.
21. A. Howard, M. Sandler, G. Chu, et al., "Searching for MobileNetV3," in Proc. IEEE Int. Conf. Computer Vision (ICCV), pp. 1314–1324, 2019.
22. M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf. Machine Learning (ICML), pp. 6105–6114, 2019.
23. P. Rajesh and V. Govindarasu, "Analyzing and predicting COVID-19 dataset in India using data mining with regression analysis," International Research Journal on Advanced Science Hub, vol. 3, pp. 91–95, 2021.
24. P. Rajesh, M. Karthikeyan, B. Santhosh Kumar, and M. Y. Mohamed Parvees, "Comparative study of decision tree approaches in data mining using chronic disease indicators (CDI) data," Journal of Computational and Theoretical Nanoscience, vol. 16, no. 4, pp. 1472–1477, 2019.
25. M. Salameh, K. Balaji, A. Murugan, P. Rajesh, P. Hari Kumar, and F. Alotoum, "A novel study on customer relationship management using artificial intelligence in the retail sector," Journal of Pharmaceutical Negative Results, pp. 9810–9819, 2023.
26. P. Rajesh, M. Karthikeyan, and R. Arulpavai, "Data mining algorithm to predict the factors for agricultural development using stochastic model," International Journal of Recent Technology and Engineering, vol. 8, no. 3, pp. 2713–2717, 2019.
27. P. Rajesh and M. Karthikeyan, "Prediction of agriculture growth and level of concentration in paddy—A stochastic data mining approach," in Advances in Big Data and Cloud Computing: Proc. ICBDCC 2018, Singapore: Springer, pp. 127–139, 2018.
28. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial nets," in Proc. Advances in Neural Information Processing Systems (NeurIPS), pp. 2672–2680, 2014.
29. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
30. C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
31. J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85–117, 2015.
32. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learning Representations (ICLR), pp. 1–15, 2015.
33. J. Snoek, H. Larochelle, and R. Adams, "Practical Bayesian optimization of machine learning algorithms," in Proc. Advances in Neural Information Processing Systems (NeurIPS), pp. 2951–2959, 2012.