
BER-CNN: Bio-Inspired Evolutionary Refinement of CNNs for Rapid and Accurate Tuberculosis Detection

Marwa M. Eid^{1,2}

¹*Faculty of Artificial Intelligence Delta University for Science
& Technology Mansoura 35111, Egypt*

²*Jadara University Research Center Jadara University
Jordan mmm@ieee.org*

Abstract.

In this study, we present BER-CNN, a novel deep learning framework optimized through bio-inspired evolutionary refinement for automated TB detection. The model was trained and evaluated on a curated dataset of 928 sputum smear images containing 3,734 annotated bacilli. Comprehensive benchmarking against four optimization-driven CNN models—GWO-CNN, PSO-CNN, WOA-CNN, and GA-CNN—demonstrates that BER-CNN achieves superior performance across all metrics, with an accuracy of 97.22%, sensitivity of 96.86%, specificity of 97.53%, and F-score of 97.12. Beyond predictive accuracy, BER-CNN reduces computational cost by converging faster than competing approaches, making it practical for large-scale deployment. These results highlight the potential of BER-CNN to transform TB diagnostics by providing an accurate, efficient, and scalable solution that bridges the gap between advanced machine learning and clinical practice..

Keywords. Tuberculosis Detection, Deep Learning, Bio-Inspired Optimization, Convolutional Neural Networks (CNNs), Medical Image Analysis

1. INTRODUCTION

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (Mtb), remains one of the world's deadliest infectious diseases. Despite the availability of effective treatment, TB continues to affect millions each year and is responsible for over a million deaths annually. Artificial intelligence (AI), particularly deep learning (DL), has revolutionized multiple areas of healthcare by enabling the automated analysis of complex biomedical data [1]. DL has demonstrated strong performance in radiology, pathology, genomics, and infectious disease detection, offering improvements in both speed and accuracy compared to human experts. In the field of TB diagnostics, DL models have been applied to several modalities, including sputum smear microscopy, chest radiography, Raman spectroscopy, and metabolomic profiling. These approaches have shown the potential to improve diagnostic yield, reduce human workload, and provide scalable solutions for low-resource environments. Importantly, DL frameworks have the capacity to integrate heterogeneous data sources, enabling multi-modal TB detection strategies that go beyond the limitations of single diagnostic techniques [2] [3]. While the promise of deep learning in TB diagnostics is well recognized, several research gaps remain. First, many DL approaches are limited to small or homogeneous datasets, raising questions about their robustness and generalizability in real-world clinical settings. Second, despite reporting high performance, issues of model interpretability, reproducibility, and integration into clinical workflows remain

underexplored. Third, most studies focus on a single optimization strategy or architecture, leaving opportunities for further improvement in accuracy and computational efficiency. Finally, there is limited comparative benchmarking of deep learning models optimized with bio-inspired algorithms for TB detection, which could reveal the most effective frameworks for real-world deployment [4]. To address these gaps, we develop and evaluate a set of deep learning models enhanced by bio-inspired optimization techniques for TB detection. Specifically, we benchmark the performance of the following models: Grey Wolf Optimizer CNN (GWO-CNN), Particle Swarm Optimizer CNN (PSO-CNN), Whale Optimization Algorithm CNN (WOA-CNN), and Genetic Algorithm CNN (GA-CNN). Against these baselines, we introduce our proposed model, BER-CNN, which integrates an advanced optimization mechanism to enhance feature selection and classification accuracy [5].

2. LITERATURE REVIEW

To address this, one study proposed an automated examination framework using an adapted UNet model to localize *Mycobacterium tuberculosis* (Mtb) and quantify their lipid content and dimensions, achieving improved accuracy and efficiency compared to manual evaluation [6]. Similarly, another group developed a deep learning pipeline for bacterial quantification in sputum samples using fluorescence microscopy. Their approach combined generative adversarial networks, patch extraction, classification, and regression, reducing error rates and bypassing the variability of manual counting [7]. Deep learning has also been employed as a supportive tool for pathologists, automating the identification of acid-fast bacilli in microscopy images, and significantly reducing workload without compromising diagnostic accuracy [8]. Alternative modalities have also emerged for rapid and reliable detection. Surface-enhanced Raman spectroscopy (SERS) coupled with deep learning was shown to effectively distinguish between sputum samples with or without Mtb, as well as classify strains based on site of infection and antibiotic resistance profiles, demonstrating high accuracy with handheld devices [9]. Metabolic phenotyping through urine analysis further demonstrated potential in differentiating TB from nontuberculous mycobacteria (NTM), identifying distinct metabolite signatures—particularly involving amino acid metabolism—that could be exploited for diagnostic biomarkers [10]. Likewise, bacterial volatilome analysis using HS-SPME-GC-MS revealed discriminant volatile organic compounds (VOCs) that allow differentiation of *Mycobacterium* species, highlighting the potential for clinically translatable breath- or sputum-based biomarkers [11].

3. DATASET DESCRIPTION AND PREPROCESSING

Tuberculosis (TB) is an infectious disease caused primarily by the bacterium *Mycobacterium tuberculosis* (Mtb). While TB most commonly affects the lungs, it can also spread to other organs, leading to extrapulmonary manifestations. A large proportion of TB infections remain latent, in which individuals carry the bacteria without showing symptoms. However, active TB presents with characteristic symptoms such as a persistent cough (sometimes blood-streaked), weight loss, night sweats, and fever. Transmission occurs when infected individuals expel aerosolized droplets through coughing or sneezing.

In total, the dataset contains 3,734 annotated bacilli across the 928 images. The annotation details are stored in associated XML files, where each file corresponds to an image and encodes the bounding box coordinates for all bacilli present. These structured annotations

enable supervised training of deep learning models, allowing them to learn both classification (presence/absence of bacilli) and localization (precise position of bacilli) tasks.

A. Data Preprocessing

Prior to training, the dataset underwent a series of preprocessing steps to ensure consistency, reduce noise, and enhance model performance through steps like Image normalization, Color channel adjustment, Annotation parsing, Data augmentation, Dataset splitting. Through these preprocessing steps, the dataset was prepared for the development of robust deep learning models capable of accurately detecting and localizing TB bacilli in sputum smear images. This structured approach ensured both data quality and consistency, laying the foundation for the subsequent training and evaluation stages.

B. Deep Learning Models

1) Convolutional Neural Network (CNN): Convolutional Neural Networks (CNNs) form the backbone of many computer vision tasks, including medical image analysis, due to their ability to automatically extract hierarchical features from input images. In the context of tuberculosis (TB) detection, CNNs have been employed to process sputum smear images, capturing spatial patterns that correspond to the presence of *Mycobacterium tuberculosis* bacilli. Initial experiments with a baseline CNN demonstrated promising results, but the performance was limited by manual tuning of hyperparameters and network design choices.

4. DEEP LEARNING PERFORMANCE RESULTS

From the results, it is evident that BER-CNN achieved the highest accuracy of 97.22%, surpassing all other models by a significant margin. In particular, it recorded an improvement of more than 5% over the weakest baseline (GA-CNN) and approximately 2.5% over the strongest competing baseline (GWO-CNN). This superior performance was consistent across other metrics as well, with BER-CNN reaching a sensitivity of 96.86% and specificity of 97.53%, effectively minimizing both false negatives and false positives. The positive predictive value (97.37%) and negative predictive value (97.08%) further reinforce the robustness of the model in distinguishing between TB-positive and TB-negative cases.

To further explore the relationships between the different evaluation metrics, a pairplot of the performance results was generated, as illustrated in Figure 1. This visualization shows pairwise comparisons of accuracy, sensitivity, specificity, precision (PPV), negative predictive value (NPV), and F-score, with regression lines added to highlight correlations. The diagonal plots display the distribution of each individual metric, while the off-diagonal scatter plots with fitted regression lines provide insight into how improvements in one metric relate to others. From the pairplot, it is evident that the metrics are strongly and positively correlated, particularly between accuracy, sensitivity, and F-score, which reinforces the conclusion that models achieving higher accuracy also tend to maintain balanced predictive performance across other key criteria.

This study presented a comparative evaluation of deep learning models enhanced with bio-inspired optimization strategies for the detection of tuberculosis (TB) from sputum smear images. The proposed BER-CNN framework consistently outperformed competing models (GWO-CNN, PSO-CNN, WOA-CNN, and GA-CNN) across all evaluation metrics. BER-

CNN achieved the highest accuracy (97.22%), sensitivity (96.86%), specificity (97.53%), and F-score (97.12%), representing a significant improvement over baseline approaches. The superior balance between sensitivity and specificity demonstrates BER-CNN's effectiveness in reducing both false negatives and false positives, which is essential in medical diagnostics. Furthermore, BER-CNN demonstrated lower computational costs, achieving faster convergence compared to other optimization-based CNNs. This establishes BER-CNN as both an accurate and efficient framework suitable for early TB detection.

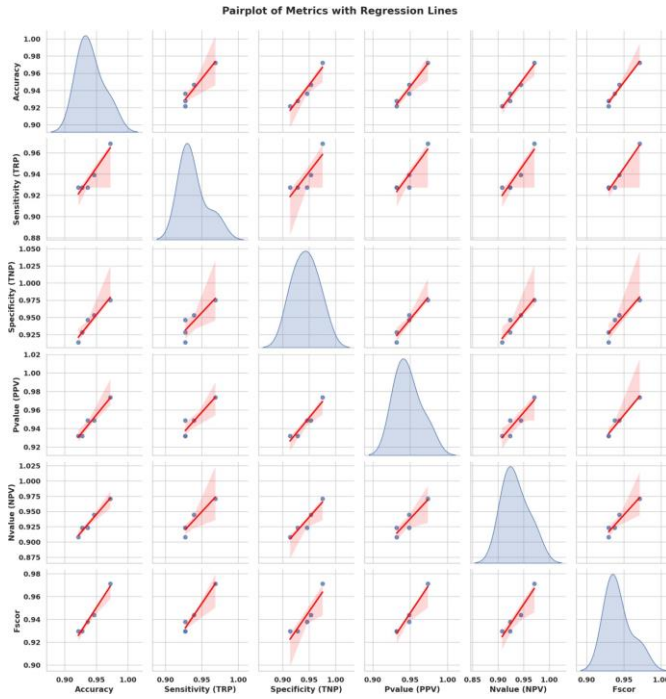


Fig. 3. Pairplot of model performance metrics with regression lines, showing relationships among accuracy, sensitivity, specificity, PPV, NPV, and F-score.

While the results demonstrate the effectiveness of BER-CNN, several avenues for future research remain. First, larger and more diverse datasets should be employed to further validate the model's generalizability across different patient populations, imaging devices, and clinical settings. Second, explainability and interpretability techniques should be incorporated to make the decision-making process of BER-CNN more transparent to clinicians, fostering trust and clinical adoption. Third, integration of multimodal data—such as combining sputum imaging with radiographic, genomic, or clinical features—could provide more comprehensive diagnostic insights. Fourth, exploring lightweight architectures and model compression methods could make the framework even more suitable for mobile or handheld diagnostic devices. Finally, longitudinal studies assessing the impact of deploying such models in real-world healthcare systems would provide critical evidence for large-scale implementation.

5. CONCLUSION

BER-CNN offers a promising pathway toward accurate, efficient, and scalable TB diagnostics. With continued refinement, validation, and integration into health-care workflows, optimized deep learning models such as BER-CNN have the potential to play a transformative role in global TB control efforts.

REFERENCES

- [1] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [2] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, “Neural network-based classification of beamforming matrices,” *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El- Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [5] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [6] R. K. Ratnesh, R. K. Chauhan, R. Kumar, M. Afroz, and J. Singh, “Recent advancements of plasmonic sensor technologies in healthcare and electronics,” *Journal of Molecular Structure*, 2025, Art. no. 142945, doi: 10.1016/j.molstruc.2025.142945.
- [7] Marios Zachariou, Ognjen Arandjelovic', Wilber Sabiiti, Bariki Mtafya, and Derek Sloan. Tuberculosis bacteria detection and counting in fluorescence microscopy images using a multi-stage deep learning pipeline. *Information*, 13(2), 2022.
- [8] S. Arora, S. Sharma, R. Anand, R. Kumar et al., “Integration of metamaterials for quintuple band-notched ultra-wideband antennas,” *Analog Integrated Circuits and Signal Processing*, vol. 123, 2025, doi: 10.1007/s10470-025-02355-7.
- [9] Liang Wang, Xue-Di Zhang, Jia-Wei Tang, Zhang-Wen Ma, Muhammad Usman, Qing-Hua Liu, Chang-Yu Wu, Fen Li, Zuo-Bin Zhu, and Bing Gu. Machine learning analysis of sers fingerprinting for the rapid determination of mycobacterium tuberculosis infection and drug resistance. *Computational and Structural Biotechnology Journal*, 20:5364–5377, 2022.
- [10] Nguyen Ky Anh, Nguyen Ky Phat, Nguyen Quang Thu, Nguyen Tran Nam Tien, Cho Eunsu, Ho-Sook Kim, Duc Ninh Nguyen, Dong Hyun Kim, Nguyen Phuoc Long, and Jee Youn Oh. Discovery of urinary biosignatures for tuberculosis and nontuberculous mycobacteria classification using metabolomics and machine learning. *Scientific Reports*, 14(1):15312, 2024.