Metaheuristic Optimization Review

♦ ISSN: 2692-4048 Vol: 01, No. 01, 2024 ♦ PP. 10–13

Article DOI: https://doi.org/10.54216/MOR.010103



REVIEW ARTICLE

Improving Tuberculosis Diagnosis and Forecasting Through Machine Learning Techniques: A Systematic Review

Abdelaziz Rabehi[1,*] **Pushan Kumar Dutta**[2]

[1] Telecommunications and Smart Systems Laboratory, University of Djelfa, PO Box 3117, Djelfa 17000, Algeria
[2] Amity University Kolkata, India & Amity School of Engineering and Technology, India

Emails: Abdelaziz.rabehi@univ-djelfa.dz · pkdutta@kol.amity.edu

Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024 ⋆ Corresponding author

ABSTRACT

Tuberculosis (TB) is ranked as one of the leading causes of death from infectious diseases in the present world, causing important health and economic consequences in developing countries. Traditional diagnostic approaches, although still used, are commonly associated with subjectivity, slowness, and operational burden, and they are often limited to visual observation and laboratory handling. The growing capacity of advanced machine learning (ML) is a promising area that has shown potential in improving TB diagnosis, identifying drug resistance, and supporting disease management. This review presents multiple aspects of using ML in diagnosing and managing TB, including deep learning, hybrid approaches, and metabolomics-based approaches. These methods have produced substantial diagnostic improvements in sensitivity, specificity, and accuracy. Furthermore, ML has been used to analyze the molecular profile of TB and identify potential drug targets for future targeted therapies. Although ML-based solutions offer substantial benefits, questions concerning data quality, model interpretation, generalization, and ethical issues continue to hinder broader application. This review concludes that ML can transform TB diagnosis and management and calls for further investment to overcome these barriers to global health.

Keywords: AI ▪ Machine Learning ▪ Tuberculosis ▪ Multiple Drug Resistance ▪ Advanced AI ▪ Deep Learning ▪ Metabolomics and TB Control

1. INTRODUCTION

Tuberculosis remains one of the most critical global health problems, and its effect is evaluated in terms of annual fatalities. As a persistent and highly contagious pulmonary disease caused by *Mycobacterium tuberculosis*, TB remains endemic in developing nations where health facilities are stretched to their limits. The World Health Organization estimated that, in 2018, ten million people became ill with TB and 1.5 million died, which underscores the necessity of combating this endemic disease. The fact that TB remains among the top killers indicates that social, economic, health, and genetic factors continue to expose populations to the disease, while new and more virulent strains challenge existing treatments [1, 2].

TB also has a major animal-health dimension, especially in cattle. *Mycobacterium bovis* infection affects animal health and productivity and poses risks to human health and food safety. Zoonotic transmission is a concern in farming regions where animals are a main source of livelihood and food. In Brazil, for example, bovine TB remains endemic and causes economic losses and public-health concerns; therefore, new control and management strategies are required [3].

For years, TB has been diagnosed using traditional methods such as direct observation and laboratory tests. However, these methods are often cumbersome, expensive, and prone to inaccuracy. Sputum-smear examination, while useful, de-

© 2024 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

pends heavily on the technician and may lead to variation in detection. These approaches may also fail to identify chronic or subclinical infection, resulting in delayed diagnosis and treatment. These limitations emphasize the need for more effective TB detection methods in humans and cattle [4].

New technologies and data-processing science make it possible to develop diagnostic methods based on ML algorithms. Artificial intelligence and machine learning can process extensive datasets and identify behavioral patterns that are difficult to detect manually. Incorporating ML approaches allows researchers to create models to diagnose TB faster and more successfully. Accordingly, using ML to identify TB is a viable way to improve disease detection and control [5].

This review analyzes the potential of machine learning for TB diagnosis and forecasting, including bovine TB prediction from epidemiological data. By using factors related to farms and cattle health conditions, ML studies can establish reliable hypotheses explaining TB risk in cattle. These tools go beyond diagnostic improvement by enabling farmers and public-health authorities to use actionable analytics models to fight disease spread [6, 7].

2. LITERATURE REVIEW

Recent studies demonstrate that ML and AI techniques can improve TB diagnosis, forecasting, drug-resistance analysis, and epidemiological control. Ensemble convolutional neural networks (CNNs), for example, have been used for chest X-ray analysis and have achieved high accuracy on the Montgomery and Shenzhen datasets through Bayesian-optimized voting and preprocessing variation [8]. Other work has addressed drug-resistant TB diagnosis by combining diagnostic methods with ML to improve diagnostic precision and reduce long diagnostic processes [9].

Support Vector Machine (SVM) models have also been used for preliminary TB diagnosis based on hospital datasets, showing the feasibility of automated recognition systems for early and accurate diagnosis [10]. In low-resource settings, AI-assisted chest-radiography systems such as TB-UNet with ResNeXt encoders improved sensitivity compared with radiologist-only diagnoses, showing potential for regions with limited specialists [11].

CNN configurations have been examined for binary TB classification using objective-function values and feature combinations. One reported model achieved a validation accuracy of 0.9671 and an AUC of 0.9733, indicating that CNN-based systems can capture signs of TB effectively [12]. Machine learning has also been applied to identify active molecules against *Mycobacterium tuberculosis*. Models such as SVM, Random Forest, Gradient Boosting, Deep Neural Networks, and XGBoost have been used for molecular-descriptor analysis, with consensus stacking producing strong external-test performance [13].

AI and ML methods have also been used to estimate multidrug-resistant TB (MDR-TB). High-throughput sequencing data containing approximately 10,000 mutations have been analyzed using Naive Bayes, K-nearest neighbors, SVM, and artificial neural networks, producing average accuracy values above 85% for genes associated with resistance to rifampicin, isoniazid, pyrazinamide, and fluoroquinolones [14].

Despite these promising results, large-scale adoption of computer-assisted TB detection software remains challenging. Issues include limited generalization to different populations, dataset distribution shift, and methodological weaknesses [15]. Evaluations of pretrained CNNs such as AlexNet, VGG16, and VGG19 have shown that image resolution, dataset splitting, and hyperparameter configuration strongly affect pulmonary TB detection performance [16]. CNN-based CAD models have also achieved useful validation accuracy in screening TB from X-ray images, particularly where radiologists are unavailable [17].

Other recent approaches include the Intelligent Emperor Penguin Optimized LightGBM (IEPO-LGBM) method for rapid TB diagnosis, which demonstrated superior sensitivity, specificity, accuracy, and precision compared with existing methods [18]. Hybrid machine-learning techniques combining an artificial immune recognition system (AIRS) with fuzzy logic controllers achieved high classification accuracy on patient epicrisis reports [19]. Automated image-analysis algorithms using mobile microscopes and SVM classification have also produced slide-level performance comparable to specialists in low-resource settings [21].

ML approaches have further been applied to anti-TB drug discovery and compound permeability. XGBoost, Random Forest, SVM, and Naive Bayes models can classify compounds as permeable or impermeable using enzymatic and cellular assay data, identifying key parameters that influence drug permeability [22]. Comparative CNN studies have reported that DenseNet121 can outperform other CNN architectures in TB detection tasks [23]. Metabolomics-based diagnosis has used serum metabolic fingerprints and ML-derived biomarkers to produce diagnostic AUC values of 97.8% for TB detection and 85.7% for classification [24].

For bovine TB, machine learning has been applied to epidemiological survey data from 1,743 farms in Sao Paulo, Brazil. With 77 predictors including farm type, lactation type, and herd size, gradient-boosted trees achieved 88.07% accuracy, with daily milk production, number of cows, farm type, breed, and slaughter of adult animals identified among relevant predictors [7].

Table 1 summarizes multiple analyses highlighting the use of ML and AI solutions in diagnosing and treating TB. These studies range from image-diagnostic CNNs for chest X-rays to ensemble ML models and AI systems that support radiologists in low- and middle-income countries. Other studies examine genomics data to detect drug resistance and classify molecules, while additional work aims to increase diagnostic efficiency through hybrid models and new algorithms. Overall, the literature shows that the speed, sensitivity, and cost-benefit of TB diagnosis can improve when ML and AI technologies are applied to overcome the drawbacks of conventional approaches.

3. DISCUSSION

The studies described in this review show significant potential for machine learning and AI to transform TB diagnosis. These technologies exploit large amounts of health, imaging, genomic, metabolomic, and epidemiological data to improve

Table 1. Summary of Literature Review

Ref.	Authors/Study	Focus Area	ML Approach	Key Findings
[8]	Ensemble CNN for TB detection	Chest X-ray analysis	Ensemble CNN, InceptionV3, Xception	Achieved 97.5–97.7% accuracy on Montgomery and Shenzhen datasets using Bayesian-optimized voting.
[9]	Diagnostic approaches for drug-resistant TB	Drug-resistant TB diagnosis	Combined diagnostic methods with ML	Enhanced diagnostic precision while addressing long diagnostic processes and accuracy issues.
[10]	SVM for preliminary TB diagnosis	Early and accurate TB diagnosis	Support Vector Machine	Demonstrated higher accuracy rates using local hospital data.
[11]	AI-based TB detection in low-resource settings	Underdeveloped regions	TB-UNet, ResNeXt encoder	Improved sensitivity detection by 11.8% compared with radiologist-only diagnoses.
[12]	Improving early TB diagnosis	Binary TB classification	CNN configurations with LVCEL, MVA, and TT features	Achieved validation accuracy of 0.9671 and AUC of 0.9733.
[13]	ML models for Mtb inhibitors	Drug-resistant TB and inhibitors	SVM, RF, Gradient Boosting, DNN, XGBoost	XGBoost with consensus stacking achieved an AUC of 0.942.
[14]	Estimating MDR-TB using AI and ML	MDR-TB analysis	Naive Bayes, K-NN, SVM, ANN	Over 85% accuracy in identifying genes associated with resistance.
[15]	ML for TB surveillance	CAD software for TB detection	Computer-assisted detection	Large-scale adoption remains limited by methodology and generalization issues.
[16]	CNNs in pulmonary TB detection	CXR analysis	AlexNet, VGG16, VGG19	Image-splitting and hyperparameter choices strongly affected detection performance.
[17]	CNN-based CAD model	Diagnostic support using CAD	Deep CNNs	Achieved validation accuracy of 87.1%.
[18]	IEPO-LGBM	Rapid TB diagnosis	Intelligent Emperor Penguin Optimized LightGBM	Demonstrated superior sensitivity, specificity, and precision compared with existing methods.
[19]	Hybrid ML diagnosis	TB case classification	AIRS with fuzzy logic	Achieved 99.14% classification accuracy.
[21]	Automated CellScope algorithm	Low-resource diagnosis	SVM-based image analysis	Obtained 89.2% $\pm$ 2.1% average precision, comparable to human specialists.
[22]	Drug permeability in TB	Anti-TB compound classification	XGBoost, RF, SVM, Naive Bayes	XGBoost identified key parameters influencing drug permeability.
[23]	CNN model comparison	Deep-learning TB detection	CNN architectures	DenseNet121 achieved the highest reported accuracy of 83.5%.
[24]	Metabolomics-based diagnosis	Metabolic fingerprinting	ML analysis of serum biomarkers	Diagnostic models achieved AUC values of 97.8% for detection and 85.7% for classification.
[7]	Predictive bovine TB analysis	Epidemiological animal TB	Gradient-boosted trees	Achieved 88.07% accuracy in predicting bovine TB cases in Sao Paulo farms.

the chances of correct diagnosis and early detection. In particular, deep learning is effective for chest X-ray interpretation, while ensemble models and hybrid ML systems can improve robustness and classification performance.

However, the reviewed literature also reveals several recurring limitations. Dataset quality, class imbalance, and population-specific distribution shifts can reduce generalizability. Model interpretability remains a challenge in clinical environments, where medical decisions require explanation and accountability. Ethical considerations, including privacy and equitable access to AI tools, must also be addressed to ensure that affected populations benefit from these technologies, especially in underprivileged areas.

4. CONCLUSION

TB remains a major health concern worldwide, especially in developing nations where complications of diagnosis and treatment are readily apparent. Traditional approaches can be slow, costly, and less accurate in some settings; therefore, better diagnostic approaches are required. This review shows that machine-learning techniques have the potential to overcome these challenges by increasing the accuracy, speed, and cost efficiency of TB diagnosis.

The reviewed studies demonstrate the broad scope of ML in TB control, including identifying drug-resistant strains, analyzing chest radiographs, discovering active molecules, using metabolomics biomarkers, and studying factors contributing to zoonotic TB. Modern approaches, including deep learning, hybrid models, and metabolomics-related ML methods, have shown good accuracy, sensitivity, and specificity. These improvements suggest that ML can contribute to early diagnosis and accurate management of TB, thereby reducing its spread.

Future challenges include data quality, model interpretability, generalization, privacy, and accessibility. Continued investigation and development of robust, explainable, and scalable AI solutions will be critical. This work is not only about machine learning; it is also about strengthening the fight against TB and improving global health outcomes.

REFERENCES

[1] K. Dheda et al., "The Lancet Respiratory Medicine Commission: 2019 update: epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant and incurable tuberculosis," *Lancet Respiratory Medicine*, vol. 7, no. 9, pp. 820–826, 2019, doi: 10.1016/S2213-2600(19)30263-2.

[2] World Health Organization, "End TB Strategy: Progress in implementing the global strategy and targets for tuberculosis prevention, care and control after 2015," accessed Oct. 25, 2024.

[3] T. Wingfield et al., "Challenges and opportunities to end tuberculosis in the COVID-19 era," *Lancet Respiratory Medicine*, vol. 9, no. 6, p. 556, 2021, doi: 10.1016/S2213-2600(21)00161-2.

[4] C. C. Cemeozkurt, "Improving tuberculosis diagnosis using explainable artificial intelligence in medical imaging," *Journal of Mathematical Sciences and Modelling*, vol. 7, no. 1, pp. 33–44, 2024, doi: 10.33187/JMSM.1417160.

[5] Y. Luo et al., "Development of diagnostic algorithm using machine learning for distinguishing between active tuberculosis and latent tuberculosis infection," *BMC Infectious Diseases*, vol. 22, no. 1, pp. 1–14, 2022, doi: 10.1186/S12879-022-07954-7.

[6] A. Abdulrahman, Z. M., A. M. Zaki, F. H. H. Rizk, M. M. Eid, and E.-S. M. El-Kenawy, "Exploring optimization algorithms: A review of methods and applications," *Journal of Artificial Intelligence and Metaheuristics*, vol. 7, no. 2, pp. 08–17, 2024, doi: 10.54216/JAIM.070201.

[7] L. E. C. Pereira et al., "Machine learning to predict tuberculosis in cattle from the state of Sao Paulo, Brazil," *European Journal of Public Health*, vol. 30, Supplement 5, 2020, doi: 10.1093/EURPUB/CKAA166.849.

[8] E. Tasci, C. Uluturk, and A. Ugur, "A voting-based ensemble deep learning method focusing on image augmentation and preprocessing variations for tuberculosis detection," *Neural Computing and Applications*, vol. 33, no. 22, pp. 15541–15555, 2021, doi: 10.1007/S00521-021-06177-2.

[9] T. Khan, I. H. Khan, and R. Ravindran, "Comprehensive approach to TB diagnosis powered by machine learning in endemic countries," *American Journal of Clinical Pathology*, vol. 162, Supplement 1, pp. S128–S128, 2024, doi: 10.1093/AJCP/AQAE129.282.

[10] A. Yahiaoui, O. Er, and N. Yumusak, "A new method of automatic recognition for tuberculosis disease diagnosis using support vector machines," accessed Oct. 23, 2024.

[11] M. Nijiati et al., "Deep learning assistance for tuberculosis diagnosis with chest radiography in low-resource settings," *Journal of X-Ray Science and Technology*, vol. 29, no. 5, pp. 785–796, 2021, doi: 10.3233/XST-210894.

[12] V. Singh, M. K. Gourisaria, and H. Gm, "Mycobacterium tuberculosis detection using CNN ranking approach," pp. 583–596, 2022, doi: 10.1007/978-981-16-4369-9_56.

[13] Q. Ye et al., "Identification of active molecules against Mycobacterium tuberculosis through machine learning," *Briefings in Bioinformatics*, vol. 22, no. 5, pp. 1–15, 2021, doi: 10.1093/BIB/BBAB068.

[14] S. Jamal, M. Khubaib, R. Gangwar, S. Grover, A. Grover, and S. E. Hasnain, "Artificial intelligence and machine learning based prediction of resistant and susceptible mutations in Mycobacterium tuberculosis," *Scientific Reports*, vol. 10, no. 1, pp. 1–16, 2020, doi: 10.1038/s41598-020-62368-2.

[15] S. Sathitratanacheewin, P. Sunanta, and K. Pongpirul, "Deep learning for automated classification of tuberculosis-related chest X-ray: Dataset distribution shift limits diagnostic performance generalizability," *Heliyon*, vol. 6, no. 8, e04614, 2020, doi: 10.1016/j.heliyon.2020.e04614.

[16] M. J. Norval, Z. Wang, and Y. Sun, "Evaluation of image processing technologies for pulmonary tuberculosis detection based on deep learning convolutional neural networks," 2021, doi: 10.12720/jait.12.3.253-259.

[17] M. Oloko-Oba and S. Viriri, "Diagnosing tuberculosis using deep convolutional neural network," *Lecture Notes in Computer Science*, vol. 12119, pp. 151–161, 2020, doi: 10.1007/978-3-030-51935-3_16.

[18] I. Alam, S. Sharma, S. Padmanabhan, A. Kannagi, S. Gupta, and J. Sahoo, "Assessment of tuberculosis using intelligent emperor penguin optimized light method," *Nanotechnology Perceptions*, vol. 20, no. S3, pp. 985–997, 2024, doi: 10.62441/NANO-NTP.V20IS3.74.

[19] S. Shamshirband et al., "Tuberculosis disease diagnosis using artificial immune recognition system," *International Journal of Medical Sciences*, vol. 11, no. 5, p. 508, 2014, doi: 10.7150/IJMS.8249.

[20] E. Elveren and N. Yumusak, "Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm," *Journal of Medical Systems*, vol. 35, no. 3, pp. 329–332, 2011, doi: 10.1007/S10916-009-9369-3.

[21] J. Chang et al., "Automated tuberculosis diagnosis using fluorescence images from a mobile microscope," *Lecture Notes in Computer Science*, vol. 7512, pp. 345–352, 2012, doi: 10.1007/978-3-642-33454-2_43.

[22] S. Nagamani and G. N. Sastry, "Mycobacterium tuberculosis cell wall permeability model generation using chemoinformatics and machine learning approaches," *ACS Omega*, vol. 6, no. 27, pp. 17472–17482, 2021, doi: 10.1021/ACSOMEGA.1C01865.

[23] J. Liu and Y. Huang, "Comparison of different CNN models in tuberculosis detecting," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 8, pp. 3519–3533, 2020, doi: 10.3837/TIIS.2020.08.021.

[24] Y. Liu et al., "Automated diagnosis and phenotyping of tuberculosis using serum metabolic fingerprints," *Advanced Science*, vol. 11, no. 39, 2406233, 2024, doi: 10.1002/ADVS.202406233.