Metaheuristic Optimization Review (MOR) *Vol. 01, No. 01, PP. 35-44, 2024*

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

Email: Abdelaziz.rabehi@univ-djelfa.dz; pkdutta@kol.amity.edu

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 the different developing countries. The practices of traditional diagnostic approaches, although still expected, are associated with relativity, slowness, and organs, besides being confined to visual observations and touch. The new and increased capacity in advanced machine learning is a promising area that has shown potential in improving the diagnosis of TB, as well as identifying drug resistance and disease management. This review presents various aspects of using ML in diagnosing and managing TB disease based on its various categories of models, including deep learning, hybrid approach, and the metabolomics approach. Some of these methods have been very effective, with high diagnostic performance improvements in sensitivity, specificity and accuracy; Furthermore, ML has been used to analyze the molecular picture of TB and to find drug targets of the disease toward future targeted therapies. As seen with the integration of ML, substantial benefits are provided by the solutions proposed. However, questions concerning the quality of data, interpretations of ML models and ethical problems hinder further application. This review concludes with the idea that ML can transform the diagnosis and management of TB and calls for more investment in developing this field 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 is still one of the most critical global health problems to date, and its effect is evaluated in terms of annual fatalities. Persistent and inherently contagious pulmonary disease, TB is a leading killer resulting from Mycobacterium tuberculosis; this illness remains endemic in the third world and developing nations where health facilities are stretched to their limits. However, the WHO estimates that in 2018, ten million people became ill with TB, and 1.5 million died, which speaks about the necessity to combat this endemic disease. The fact that TB still ranks among the top killers means that there are still social, economic, health and even genetic factors that compel people to suffer and die from the disease and new, more virulent strains are challenging existing treatments [1], [2].

It is essential for human beings and animals, especially cattle, to have a high prevalence of TB, as shown in this study. Mycobacterium bovis infection also affects animal health and productivity and poses a risk to

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

human health and food safety. Zoonotic communication of Tuberculosis, especially from animals to humans, especially in farming places where animals are the main livelihood and source of food, causes so much worry. In Brazil, for example, bovine TB remains endemic, leading to economic loss and potential public health problems; therefore, new ways of controlling and managing bovine TB are required [3].

For years, Tuberculosis has been diagnosed using traditional methods such as direct observation and laboratory tests. However, these methods are often cumbersome, expensive, and prone to high levels of inaccuracy. For instance, sputum smear examination for AE, while efficient, is dependent on the technician, leading to variations in detection. Moreover, these approaches often fail to identify chronic or subclinical instances of infection, resulting in delayed diagnoses and treatment. The limitations of these traditional diagnostic methods underscore the urgent need for more effective TB detection methods in humans and cattle [4].

New technologies and the science of data processing make it possible to develop new approaches and diagnostics methods based on ML algorithms. Artificial intelligence, a machine learning division, could consist of extensive data sets that can be interpreted for usually unforeseen behavior patterns. Incorporating machine learning approaches allows researchers to create models to diagnose TB faster and more successfully. In this regard, affixing machine learning to identify TB is a viable way of improving disease identification and control [5].

This work aims to analyze the potential of machine learning to predict bovine Tuberculosis based on a new data set from an epidemiological investigation in São Paulo State, Brazil. With the help of several factors based on farm and cattle health conditions, the study aims to establish reliable hypotheses that could explain the risk of developing TB in cattle. These therapy assets go beyond the mere improvement of the diagnostic capabilities of the algorithm but also allow farmers and public health authorities to have an actionable analytics model in the fight against the spread of the disease [6].

This study employs a rigorous quantitative research method, utilizing gradient-boosted trees on data obtained from 1,743 farms. The comprehensive data set enables the investigation into different causes of bovine TB, including farm type, lactation practice, and herd size. Through ordinary cross-validation for tuning the model parameters, the research aims to identify the best predictors of the prevalence of TB among the cattle population, enhancing our understanding of the disease's progression.

The relevance of the proposed study is based not only on the possible contribution to enhanced diagnostic opportunities but also on the consequences for public health. Diagnostic tests should be performed early enough, and treatment of infected cattle should be promptly done to reduce transmission and endangerment to human health. Furthermore, this research seeks to understand factors contributing to TB in cattle to contribute to policies and practices in human health and animal husbandry to fight against this infectious zoonotic disease [7].

Therefore, there is a pressing need to equip patients diagnosed with Tuberculosis and working physicians with better disease diagnosis models to enhance public health strategies. These research strategies are crucial as the world grapples with the continued impact of TB, improving disease identification and fate and contributing to the global battle against Tuberculosis. The implications of the results derived from this research are profound, underscoring the need for increased resources to explore the potential of existing technological tools in the fight against diseases. This re-emphasizes that more resources should be expended to discuss the feasibility of the existing technological tools in the fight against diseases.

2. Literature Review

Tuberculosis (TB) is a current major health threat that contributes to high levels of morbidity and mortality. Even with the development in health technologies, early diagnosis of TB, therefore accurate diagnosis, remains a significant issue, especially where resources are limited. This literature review aims to review the most recent studies on advancing TB diagnosis by focusing on machine learning and artificial intelligence. Mainly, we focus on literature that employs different machine-learning approaches and imaging technologies to enhance the sensitivity, speed, and affordability of diagnosing TB.

Tuberculosis is determined in the abstract as an infectious disease that can be fatal and impacts the lungs mainly. Awareness of the symptoms should be made, and an early diagnosis should be recommended to help

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

prevent mortality and its spreading. As described in [8], CXR continues to be the most common medical imaging modality for diagnosing thoracic diseases. Cadet solutions increase diagnostic accuracy and improve the speed of the decision-making process for physicians. The current study presents a new ensemble CNN model for TB detection involving voting and preprocessing differences. It works with fine-tuned CNN architectures borrowed from InceptionV3 and Xception and with added CLAHE and 10 image transformations as augmentation data. Based on the selection of 40 variations, the study identifies the best models for voting using Bayesian-optimization-based weighted voting and soft voting across two TB CXR datasets. Performance is proven as efficiency with scores of 97.500% and 97.699% on the Montgomery and Shenzhen datasets, showing that our proposed model performs better than existing techniques in the literature.

According to WHO data, TB is one of the world's leading killers. Current approaches to diagnostics have specific problems, especially when it comes to drug-resistant infections or latent TB. As noted in [9], these include entailing long diagnostic processes and accuracy problems affecting efficient disease management. To achieve these critical challenges, the study presents a new approach that combines more diagnostic methods with the machine learning algorithm to enhance the precision of TB diagnosis.

In the abstract, the authors point out that TB is an infectious disease caused by Mycobacterium tuberculosis that poses a considerable threat to global health and, more so, to low-income countries where cases that are either untreated or treated poorly can result in high mortality. The following has been noted in the paper [10] for controlling the disease: early diagnosis and accurate diagnosis are essential to treat the disease and minimize fatalities. SVM was applied in this research for the first time to make the preliminary diagnosis of TB using a machine-learning approach. The system created through the local hospital patient data analysis showed promising diagnostic prowess, earning it higher accuracy rates than similar specialist systems used in diagnosing TB.

Tuberculosis, as described in [11], is still a significant threat all over the world, and its death rates are high. Recent studies devoted to reducing the burden of TB diagnosis based on AI are numerous; however, most of the investigations concentrate on developed, urbanized populations, leaving the possibility of AI application in low-resource contexts explored. This study attempts to fill that gap by deploying an automated AI-based detection system on a large population in an underdeveloped region and investigating how it may help local radiologists diagnose TB using CXR images. For experimental work, the researchers TRAINED the model using 2627 TB-positive control samples along with 7,375 TB-negative control samples and TESTED the model using 276 TB-positive and 619 TB-negative samples. This study's AI system development processes were image labeling, preprocessing, training, and testing. TB-UNet, a segmentation model with ResNeXt as the encoder, was used to locate the diseased areas. Because of the confidence scores produced for each test case, the AI indicated how likely each was to be TB-positive. The comparative experiments showed that in the absence of the AI system, the detection rate dropped to 62% from 85% achieved with the help of the system. Furthermore, the integration of AI support enhanced the ability of local radiologists, specifically in diagnostic sensitivity detection, by 11.8%. Such results imply that TB could be improved substantially through AI in the context of fields and countries that lack resources and are most heavily affected by the disease.

In this paper [12], it is pointed out that the leading cause of human fatalities is Mycobacterium tuberculosis (TB), which can be in its active and latent forms. It has also been argued that any misdiagnosis with TB can be very disastrous, no matter the degree. This research proposes to improve early diagnosis of TB, hence facilitating early treatment and controlling the increasing incidence. To this end, the researchers used several CNN configurations to perform binary classification, as shown below. Experimental Setup Binary classification was based on objective function values obtained from models incorporating the LVCEL, MVA, and TT features. Experimental Setup these objective function values may be a helpful tool to aid medical professionals in diagnosing TB patients with better precision. The experimental evaluation has established that the architecture, which we aimed to select, showed the maximum objective function value of 6.503; it has a validation accuracy of 0.9671 (as explained in Fig. 11) and an AUC of 0.9733 in the ROC analysis presented in Fig. 12. The following measures show that the proposed model captures the signs that indicate the existence of TB effectively.

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

From the discovery made in [13], TB, which is a bacterial infection caused by Mtb, is still a top ten killer disease all over the world today and has even been complicated by drug-resistant forms like the XDR-TB. To tap this emerging crisis, the present study explored new lead compounds for classifying Mtb inhibitors using ML based on a multiplicity of molecular descriptors. Four ML models were developed: Support Vector Machine, Random Forest, Gradient Boosting, and Deep Neural Network. Of all, the XGBoost model demonstrated better prediction accuracy. Additionally, the study has improved the predictive performance through the use of two approaches for consensus, with the stacking method producing the best prediction from RF, XGBoost, and DNN models with an AUC ROC of 0.84 for the 10-fold CV training data set and 0.942 for the external test data set. Further, the Shapley additive explanations method was used to explain the interaction of critical molecular descriptors with bioactivity, consequently increasing the model's explainability. To make the proof of concept more applicable, the researchers have created an online resource, which the researchers named Chem TB and is available to use to identify potential MTB inhibitors as a free computational tool.

Published in paper [14], Tuberculosis (TB), which is due to Mycobacterium tuberculosis (M.TB), continues to be the leading cause of death in the world from bacterial diseases, suggesting the increasing require for novel approaches to diagnosis and therapy of the disease. The study offers a computational platform for the estimation of MDR-TB incorporating AI and ML methods for the analysis of genes. These genes are associated with resistance to the most critical first-line anti-TB drugs, rifampicin, isoniazid, pyrazinamide and fluoroquinolones. The analysis of approximately 10,000 mutations was conducted using high-throughput sequencing data, and mutations were described using various sequence and structural characteristics that may indicate the impact of mutations on the protein targets of such genes. Naïve Bayes, K-NN, SVM, and ANN were used, resulting in over 85% accuracy on average for all genes under consideration. Evaluations were based on external datasets, and the proposed models were further validated. Moreover, molecular docking and dynamics simulation studies were performed on the wild type and mutant protein-drug complexes to establish how mutations affects protein folding and conformations and to confirm the calculated resistance profiles.

Published in the paper [15], Machine learning has proved helpful in several aspects of infectious diseases, including tuberculosis surveillance. Promisingly, the World Health Organization declined to issue guidelines for the large-scale adoption of computer-assisted TB detection software. This is because most of the studies still need to be expanded due to an adequate number of studies and methodological problems to overcome; in addition, there are questions about the ability to generalize the findings to different populations and contexts.

Ref. [16] explains that TB (Tuberculosis) is an infectious ailment caused by the bacterium Mycobacterium Tuberculosis that primarily attacks the lungs. TB is a highly contagious disease that spreads when a patient coughs or snails and releases minute droplets into the air. The present research aims to assess the performance of Convolutional Neural Networks (CNNs) in detecting pulmonary Tuberculosis by using CXR images of the patients. This study is concerned with four main aspects: the resolution of the image dataset, the pre-trained networks such as AlexNet, VGG16, and VGG19, and how changing the hyperparameters affects the outcome. Moreover, an experiment was performed using the first 406 regular and 394 abnormal images and later with a second data set of 239 regular and 554 abnormal images. It was found that the image-splitting approach led to the highest accuracy and confirmed the feasibility of CNN models in enhancing the TB detection processes.

As noted in the analysis performed in [17], Tuberculosis (TB) is still one of the prominent infectious diseases and a leading cause of mortality, which is caused by Mycobacterium tuberculosis bacillus. This could be because TB is highly contagious, and there is early diagnosis deficiency, together with a lack of radiologists, mainly in the developing world, which leads to increased cases and deaths from TB. Therefore, TB can be treated if it is diagnosed early. Although many screening techniques are available, chest game examination is instrumental in diagnosing lung diseases, but it depends solely on the radiologist's experience. The research aims to address the problem of the excessive burden of TB in Montgomery County (MC) – using a CAD model incorporating deep convolutional neural networks (CNN) to examine X-ray images of known TB patients. This result achieved a validation accuracy of 87.1%. Other indices, such as a confusion matrix and overall accuracy, proved their ability to improve TB screening effectiveness, mostly in areas where radiologists are unavailable.

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

As stated in [18], The efficient and effective control of Tuberculosis (TB) in all affected regions of the world is an unattainable goal, even with the available medical diagnostic technologies. Conventional diagnostic procedures are often fraught with problems of low sensitivity and specificity, resulting in long waits for definitive diagnosis and poor treatment results. In order to mitigate these drawbacks, there is an urgent need for the development of novel diagnostic methods that will aid in the rapid and precise diagnosis of TB to improve the management of patients and enhance the implementation of TB control measures. This paper proposes a novel intelligent Emperor Penguin Optimized LightGBM (IEPO-LGBM) technique for TB diagnosis with an emphasis on improving the speed and accuracy of detection. The proposed method was implemented in Python tools and tested using a Kaggle-sourced dataset against existing methods. The results showed that the IEPO-LGBM method was superior to the current systems in all aspects of the evaluation with enhanced sensitivity, specificity, accuracy, and precision. The research underlines the potential role of the IEPO-LGBM approach in reducing the time taken for TB diagnosis and treatment, thus aiding the fight against the dreadful disease.

The research detailed in [19] indicates that conventional methods for diagnosing Tuberculosis (TB) have a significant possibility of misdiagnosis, hence the need for more dependable solutions. This research aimed to improve the accuracy of TB diagnosis using hybrid machine learning techniques. Old patient epicrisis reports provided by the Pasteur Laboratory in northern Iran were used for the study, and 175 samples, each with twenty features, were included. In conjunction with an artificial immune recognition system (AIRS), fuzzy logic controllers were utilized in the study for classification purposes. Regular and TB classes were classified using the AIRS algorithm after normalizing the features with the help of the fuzzy rule-based labeling system. The results established that the best classification accuracy was recorded at a learning rate (α) of 0.8. It was observed that the diagnostic performance of AIRS combined with fuzzy logic was the best, whereby a classification accuracy of 99.14% was recorded with the corresponding sensitivity and specificity of 87.00% and 86.12%, respectively, which was better than the performance results of the classical empirical methods.

In the article marked by reference [20], an innovative method developed to diagnose Tuberculosis (TB) is aimed at low-resource settings. The standard method includes smear microscopy and visual identification of TB bacilli. In this way, the paper presents an automated algorithm for TB analysis using images obtained from CellScope, a handheld digital microscope working in brightfield and fluorescence modes. This development seeks to foster access to the healthcare system in the backward terrains without laboratory facilities by rendering said diagnostic procedures possible robotically and on-site. The algorithm further presents potential TB bacilli employing morphological and Gaussian kernel template matching operations, feature extraction by Hu's moments, geometric and photometric features, and histograms of oriented gradients. Classification is then conducted using a support vector machine (SVM). Trained on a dataset of 594 CellScope photograph images of sputum smears from 290 patients in Ukraine, the algorithm obtained 89.2%±2.1% Average Precision at the object level. The slide-level classification performance is comparable to that of human specialists, increasing the prospects of TB diagnosis improvement in areas with limited resources.

As emphasized in [21], there is a remarkable increase in the emergence of Mycobacterium tuberculosis (M.TB) drug-resistant strains, which calls for the development of new antitubercular agents. However, the small molecules used also show poor efficacy for other reasons, even in the absence of mutations of target genes or very snug binding pockets – the coat of M. tuberculosis is known to be complex, and there are influx-efflux transport systems, which are notable permeability limiting factors. To address these challenges, the present study sought to construct machine learning models to classify compounds as either permeable or impermeable using data measured from enzymatic (IC50) and cellular (minimal inhibitory concentration) assays. This was based on the conjecture that those compounds, which possess high potency in both the enzymes and cellular assays, would be favorable in penetrating the cell wall of E. coli. The various models tested found that the XGBoost algorithm was superior to the random forest, support vector machine, naïve Bayes and other methods. The XGBoost model was tested using a database of 21 permeable and 19 impermeable compounds. It was found that parameters like molecular weight, the type of atoms in the compound, electron topological state, number of hydrogen bond donors/acceptors, and topochemical extension of atoms influence M. tuberculosis cell permeability and inhibition. Besides, the advancement of drug repurposing was also done, as competent drugs available with the same action against M. tb from

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

DrugBank were pass-screened through the permeability model in case of possible use nuisance and pharmacologically active. Eventually, those selected drugs were further docked to some of the active sites of 10 identified targets that are potentially antimycobacterial. The results of the current research extend an understanding of M. tuberculosis permeability and activity, which potentially is helpful in the development of new medicines targeting mycobacterium.

In the source [22], the authors present Tuberculosis (TB) as a long-term illness predominantly affecting the young population and a significant contributor to global mortality rates. In the same year, 2018, the World Health Organization (WHO) reported that there were about 10 million new cases of TB, out of which 1.5 million were deaths from the disease, with 251,000 such individuals being HIV positive. Moreover, since TB is the most lethal microbiological agent in history, the purpose of this study is to help facilitate the ability of physicians to diagnose patients with the once-dreaded disease by comparing the classification accuracy of six popular convolutional neural networks' architectures on the same dataset with intending to find the best one for the highest result possible. The research included, among other things, the optimization of three main parameters of the CNN in order to improve the results. The activation function was changed from a step function to a sigmoid function, the binary cross-entropy cost was used rather than the conventional quadratic performance cost function, and the optimization method used was stochastic gradient descent (SGD). The results of the experiments showed that the most suitable model for tuberculosis diagnosis was Densenet121, which produced the highest accuracy level of 83.5%. Here, the author believes that if Densenet121 is fine-tuned further with a more extensive dataset, the performance will be helpful in the clinical setting, enhancing the war against Tuberculosis.

According to the study referenced [23], Tuberculosis (TB) is essentially the second most killer infectious disease next to COVID-19, and the treatment remains highly dependent on correct diagnosis and profiling. While elucidating differential metabolites is the basis of understanding metabolic processes in any system, the challenge has been the reframing around a workable metabolic strategy on how TB can be diagnosed and profiled. The current study proposes a diagnostic method based on metabolomics to achieve a fast turnaround time for detecting TB using a metabolic fingerprint pattern in serum through an automated nanoparticle-enhanced laser desorption/ionization mass spectrometry system. This method is advantageous due to the speed of detection (in seconds), small volumes used (viral loads in nanoliters), and cheapness (costing about $3 per one). Regarding the TB diagnosis, 14 m/z features were identified as clinically relevant biomarkers, while in TB phenotyping, 4 m/z features served as biomarkers. Cutting-edge machine-learning techniques were employed to construct TB metabolic models using these markers. The created metabolic model proved to be effective diagnostically with AUC values of 97.8% (95% confidence interval (CI), 0.964-0.986) for TB detection and 85.7% (95% CI, 0.806-0.891) for TB classification. The current work provided serum metabolic biomarker panels and addressed the feasibility of diagnosing and phenotyping TB, betraying the goodness of fit of the end TB strategy.

According to the publication [24], Tuberculosis is a common zoonosis, particularly in Brazil, which has recorded over 1,500,000 cases since 2001, including a death toll of 8,226 in the State of Sao Paulo. The present study examines the application of machine learning algorithms for the predictive analysis of Bovine Tuberculosis within a dataset obtained from an epidemiological survey on 1743 farms in the State of Sao Paulo. The response variable was the apparent prevalence of positive cases, while 77 predictors included farm type, lactation type and herd size, among other factors. The methodology included splitting the data into training (75%) and testing (25%) sets, which were then followed by preprocessing, model learning and evaluation using gradient-boosted trees. The models had an accuracy of 88.07%, a training error of 3% and a 12% error rate during validation (n = 436). Among the predictors identified were daily milk production, number of cows, farm type, breed of bovine, and slaughter of adult animals, with 2.06% false positives and 9.86% false negative rates. This indicates that better surveillance and data collection might help develop better learning models, thus helping control the disease and increasing outreach programs on animal health.

Table 1 gives a detailed outline of multiple analyses highlighting the use of ML and AI solutions in diagnosing and treating TB. These studies crosscut mapping image diagnostic methods CNNs for chest X-rays to working with ensemble ML models and state-of-the-art techniques such as implementing AI systems to support radiologists in LMICs. Furthermore, some studies examine the possibility of using genomics data to detect drug resistance and classify molecules. In contrast, others aim to increase diagnostics efficiency by employing quantitative techniques like hybrid models and new algorithms. In conclusion, the study shows

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

that both speed, sensitivity and cost-benefit of TB diagnosis may improve by applying ML and AI technologies to overcome the drawbacks of conventional approaches and offer practical solutions on the large scale of examined parameters in clinical and epidemiological fields.

Table 1: Summary of Literature Review,

Ref. No.	**Authors/Study**	**Focus Area**	**Machine Learning Approach**	**Key Findings**
[8]	Ensemble CNN for TB Detection	Chest X-ray (CXR) analysis	Ensemble CNN, InceptionV3, Xception	We 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, addressing long diagnostic processes and accuracy issues.
[10]	SVM for Preliminary TB Diagnosis	Early and accurate TB diagnosis	Support Vector Machine (SVM)	Demonstrated higher accuracy rates using local hospital data.
[11]	AI-Based TB Detection in Low-Resource Settings	Application in underdeveloped regions	TB-UNet, ResNeXt as encoder	AI system improved sensitivity detection by 11.8% compared to radiologist-only diagnoses.
[12]	Improving Early TB Diagnosis	Binary classification for TB detection	CNN configurations, LVCEL, MVA, and TT features	Achieved high validation accuracy (0.9671) and AUC (0.9733).
[13]	ML Models for Mtb Inhibitors	Drug-resistant TB and Mtb inhibitors	SVM, Random Forest, Gradient Boosting, Deep Neural Network, XGBoost	XGBoost with consensus stacking achieved an AUC of 0.942.
[14]	Estimating MDR-TB using AI and ML	Multi-drug resistant TB (MDR-TB) analysis	Naïve Bayes, K-NN, SVM, ANN	Over 85% accuracy in identifying genes associated with resistance.
[15]	Machine Learning for TB Surveillance	CAD software for TB detection	Computer-assisted detection (CAD)	WHO declined large-scale adoption due to methodological and generalization issues.
[16]	CNNs in Detecting Pulmonary TB	CXR analysis	Pre-trained CNNs (AlexNet, VGG16, VGG19)	The image-splitting approach led to the highest accuracy for detecting pulmonary TB.

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024

[17]	CNN-based CAD Model for TB in Montgomery County	Diagnostic support using CAD	Deep Convolutional Neural Networks (CNNs)	Achieved validation accuracy of 87.1% using CAD model.
[18]	Intelligent Penguin Optimized LightGBM (IEPO-LGBM)	Rapid TB diagnosis	IEPO-LGBM	She demonstrated superior sensitivity, specificity, and precision compared to existing methods.
[19]	Hybrid ML Techniques for TB Diagnosis	Classification of TB cases	AIRS with fuzzy logic controllers	Achieved 99.14% classification accuracy.
[20]	Automated Algorithm Using CellScope	TB diagnosis in low-resource settings	SVM-based image analysis	Obtained 89.2%±2.1% average precision, comparable to human specialists.
[21]	ML Models for Drug Permeability in TB	Classification of anti-TB compounds	XGBoost, Random Forest, SVM, Naïve Bayes	XGBoost model identified key parameters influencing drug permeability.
[22]	Comparative Analysis of CNNs for TB Detection	TB detection using deep learning	CNN architectures	Densenet121 achieved the highest accuracy of 83.5%.
[23]	Metabolomics-based TB Diagnosis	TB diagnostic using metabolic fingerprint	ML-based analysis of serum biomarkers	The diagnostic model achieved an AUC of 97.8% for detection and 85.7% for classification.
[24]	Predictive Analysis for Bovine TB	Animal TB using epidemiological data	Gradient-boosted trees	88.07% accuracy in predicting bovine TB cases in Sao Paulo farms.

Finally, the studies described in this review show a significant possibility of using machine learning and AI to transform the diagnosis of TB. These technologies harness a significant amount of health data to increase the chances of correct diagnosis and early detection of TB in FRU Zones. Nevertheless, studies applying timeline follow-back methods in different settings and strengthening internal validity with standardized data collection techniques, quality of data, and integration of methods in clinical practice require further improvement. Future developments are expected to increase AI's contribution to reducing the burden of Tuberculosis globally.

3. Conclusion

TB is a significant health concern worldwide today, especially in developing nations where complications of diagnosis and treatment are readily apparent. In some cases, these algorithms can be slow in terms of wall clock time or iterations and sometimes less accurate than traditional approaches; thus, better diagnosis approaches are required. This review also shows how machine learning (ML) techniques have the potential to overcome these challenges by increasing the accuracy, speed, and cost efficiency of TB diagnosis. All the presented studies show the broad scope of using ML in Tuberculosis, including identifying drug-resistant strains or studying the factors contributing to zoonotic TB. Modern methods, including deep learning, hybrid models and metabolomics-related approaches, have shown good accuracy, sensitivity and specificity. All these improvements suggest that ML can contribute to the early diagnosis and accurate management of TB to reduce its spread. However, future challenges include data quality issues and problems related to model interpretability and generalization. Two other issues are crucial in the current society where data like health information is being used; data privacy and accessibility to AI solutions are also essential aspects that should be considered to ensure that all affected populations benefit from AI solutions in their societies, especially in the underprivileged areas. Therefore, this paper concerning is a revolutionary approach to diagnosing TB, enabling more efficient, accurate, and scalable solutions. The importance of this cannot be overstated. Further attention and investigation into this area and the development of solutions to existing difficulties will be critical. This work is not just about machine learning; it is about the fight against TB and general health enhancement.

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 Respir Med, vol. 7, no. 9, pp. 820–826, Sep. 2019, doi: 10.1016/S2213-2600(19)30263-2.

[2] “End TB Strategy Progress in implementing the global strategy and targets for tuberculosis prevention, care and control after 2015 (the End TB Strategy)”, accessed: Oct. 25, 2024. [Online]. Available: https://www.who.int/teams/global-

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

[4] C. C. Cem¨ozkurt, “Improving Tuberculosis Diagnosis using Explainable Artificial Intelligence in Medical Imaging,” Journal of Mathematical Sciences and Modelling, vol. 7, no. 1, pp. 33–44, May 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 Infect Dis, vol. 22, no. 1, pp. 1–14, Dec. 2022, doi: 10.1186/S12879-022-07954-7/FIGURES/5.

[6] A. Abdulrahman, Z. M., A. M. Zaki, F. H. H.Rizk, M. M. Eid, and E.-S. M. EL 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,” Eur J Public Health, vol. 30, no. Supplement_5, Sep. 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 Comput Appl, vol. 33, no. 22, pp. 15541–15555, Nov. 2021, doi: 10.1007/S00521-021-06177-2/TABLES/5.

[9] T. Khan, I. H. Khan, and R. Ravindran, “Comprehensive Approach to TB Diagnosis Powered by Machine Learning in Endemic Countries,” Am J Clin Pathol, vol. 162, no. Supplement_1, pp. S128–S128, Oct. 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. [Online]. Available: www.biomedres.info

[11] M. Nijiati et al., “Deep learning assistance for tuberculosis diagnosis with chest radiography in low-resource settings,” J Xray Sci Technol, vol. 29, no. 5, pp. 785–796, Jan. 2021, doi: 10.3233/XST-210894.

[12] V. Singh, M. Kumar Gourisaria, H. Gm, V. Singh, V. Singh, and · M K Gourisaria, “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,” Brief Bioinform, vol. 22, no. 5, pp. 1–15, Sep. 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 2020 10:1, vol. 10, no. 1, pp. 1–16, Mar. 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, p. e04614, Aug. 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 (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12119 LNCS, pp. 151–161, 2020, doi: 10.1007/978-3-030-51935-3_16/TABLES/3.

[18] I. Alam, S. Sharma, S. Padmanabhan, A. Kannagi, S. Gupta, and J. Sahoo, “Assessment of Tuberculosis Using Intelligent Emperor Penguin Optimized Light Method,” Nanotechnol Percept, vol. 20, no. S3, pp. 985-997–985–997, May 2024, doi: 10.62441/NANO-NTP.V20IS3.74.

[19] S. Shamshirband et al., “Tuberculosis Disease Diagnosis Using Artificial Immune Recognition System,” Int J Med Sci, vol. 11, no. 5, p. 508, Mar. 2014, doi: 10.7150/IJMS.8249.

[20] E. Elveren and N. Yumuşak, “Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm,” J Med Syst, vol. 35, no. 3, pp. 329–332, Jun. 2011, doi: 10.1007/S10916-009-9369-3/TABLES/1.

[21] J. Chang et al., “Automated Tuberculosis Diagnosis Using Fluorescence Images from a Mobile Microscope,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 7512 LNCS, 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, Jul. 2021, doi: 10.1021/ACSOMEGA.1C01865/ASSET/IMAGES/MEDIUM/AO1C01865_M009.GIF.

[23] J. Liu and Y. Huang, “Comparison of Different CNN Models in Tuberculosis Detecting,” KSII Transactions on Internet and Information Systems (TIIS), vol. 14, no. 8, pp. 3519–3533, Aug. 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, p. 2406233, Oct. 2024, doi: 10.1002/ADVS.202406233.

DOI: https://doi.org/10.54216/MOR.010103
Received: August 15, 2023 Revised: December 07, 2023 Accepted: January 08, 2024