



International Journal of Marketing Management

ISSN 2454 - 5007



www.ijmm.net

Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

Automated Tuberculosis Drug Resistance Prediction Using Computerized Tomography Images: A Deep Learning Framework

¹ Ms.Mubeen Begum, ² Md.Abdul Malik

¹ Assistant Professor, Department of CSE-AIML, Lords Institute of Engineering & Technology

² Student Department of CSE-AIML, Lords Institute of Engineering & Technology.

Abstract—

The goal of this study is to use genetic data to predict TB medication resistance and sensitivity using a deep learning Convolutional Neural Network (CNN). The current methods for identifying drug-resistant TB strains are time-consuming and involve cultivating germs while they are exposed to medications. It may take a long time before you get the results of this. Furthermore, these methods aren't always precise and might generate false positives or negatives. To remedy these deficiencies, the proposed study would provide a more rapid and accurate method for identifying TB strains with medication resistance. When it came to making predictions, the suggested method was 97.27 percent accurate. Improved patient outcomes after treatment could result from using this model in clinical settings to detect drug resistance in TB more rapidly. The following terms are used as index terms: TB, drug sensitivity, drug resistance, deep learning, CNN, CT.

I. INTRODUCTION

Tuberculosis is a bacterial illness that typically affects the lungs but has the ability to spread to other areas of the body. It is caused by the bacterium *Mycobacterium tuberculosis* [1]. Infected people transmit TB germs by tiny respiratory droplets that they exhale when they cough, sneeze, or speak. Similar infections may be caused by various related mycobacteria, however TB is mostly caused by mycobacterium tuberculosis, the most well-known and prominent of them. According to the globe Health Organization (WHO), 10.6 million people would be sick with TB in the globe in 2021, with 1.6 million people losing their lives on the disease. There was also a 3.6% spike in the TB incidence rate from 2020 to 2021. Tuberculosis most often affects the lungs, although it may also cause harm to the kidneys, spine, brain, and lungs. The need for early diagnosis, prognosis, and identification of resistant TB cases is the main issue that requires fixing. When the mycobacterium tuberculosis bacteria enter a person's respiratory system, it may lead to a lung infection. Following this, macrophages, which are immune cells, take in the bacteria, allowing them to survive and multiply. Microscopically thick clumps of bacteria and immune cells called granulomas may form as a consequence of this [2]. Although it has been around for a long time, TB is still a major problem today, especially in places where people don't have easy access to healthcare and public health services. The battle against TB includes vaccination, early diagnosis, effective treatment, contact tracing, and public health activities to minimize transmission. Important concepts in TB treatment include medication sensitivity and resistance [3]. These terms characterize the responses of the mycobacterium tuberculosis bacteria to the anti-TB drugs. If we want to manage TB well and stop the spread of drug-resistant strains, we must understand these concepts. TB produced by strains of *Mycobacterium TB* that are sensitive to the commonly used antibiotics is known as drug-sensitive tuberculosis. The bacteria are efficiently killed or inhibited by commonly administered TB medications such as isoniazid, rifampin, ethambutol, and pyrazinamide. The bacteria that cause TB, *Mycobacterium tuberculosis*, may develop resistance to any or all of the drugs used to treat the disease. While TB is often curable, it may cause major problems if not addressed properly. A combination of drugs is usually prescribed to treat TB; however, treatment has become increasingly challenging due to the increasing prevalence of drug-resistant forms of the illness. Similar to how drug-susceptible TB is spread, drug-resistant tuberculosis utilizes the same vectors. Transmission of drug-resistant TB may occur via the respiratory secretions of an infected person. Drug resistance in TB may develop for many reasons, including inefficient treatment, inadequate drug regimens, poor patient adherence to medicine, and the transmission of drug-resistant strains from one individual to another. The drug-sensitivity or drug-resistance status of a TB strain is an important consideration when choosing an appropriate treatment approach. Bacterial culture in an antibiotic medium is the gold standard for detecting drug resistance.

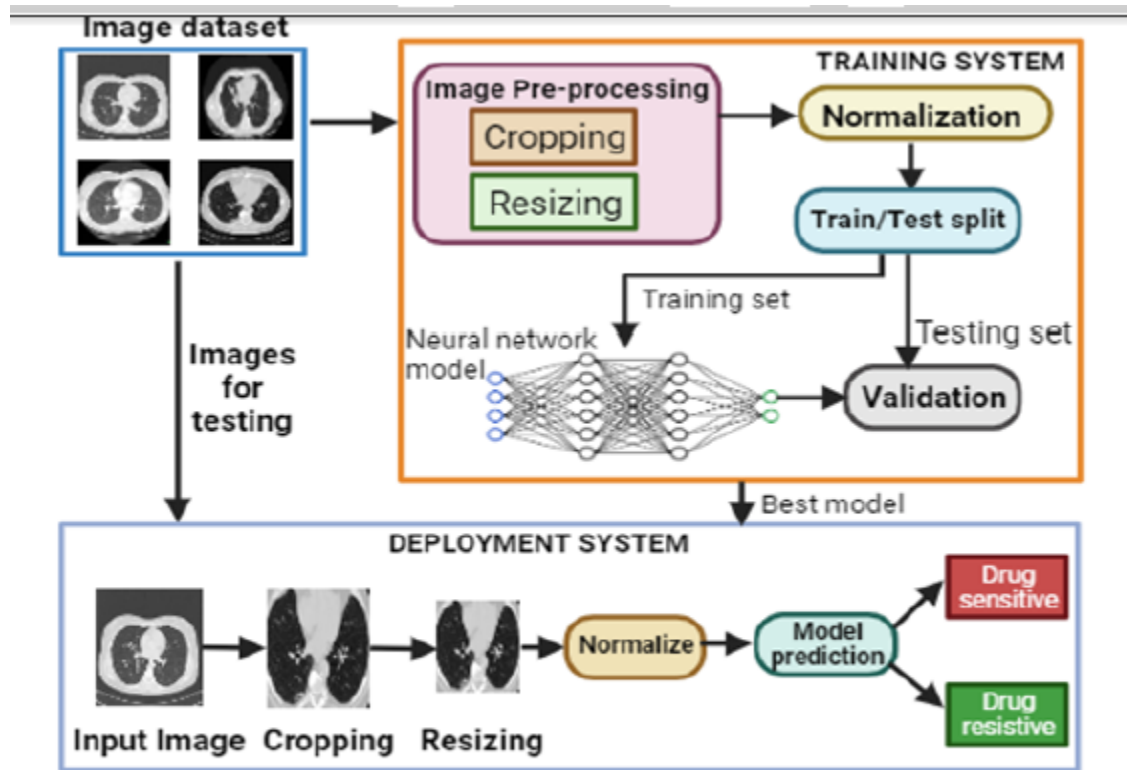


Fig. 1. Block representation of the proposed model for drug sensitivity and drug resistance prediction

while monitoring their progress [4]. However, you may not notice any benefits for many days—if not weeks—because of how long it takes. Recent developments in AI have shown promise in the accurate detection of TB strains resistant to therapy. Automatic detection of TB medication resistance is possible with the use of deep learning and machine learning algorithms [5]. Predicting the treatment susceptibility or resistance of a particular TB strain by evaluating genetic and medical imaging data (e.g., chest X-ray, CT scan, etc.) is one approach. Many published papers on the subject of TB have made predictions using machine learning and deep learning algorithms. To distinguish between the two types of TB, Karki M et al. [6] created a convolutional neural network (CNN) model. Using NIAID portal chest X-ray images, a CNN InceptionV3 model was pre-trained. The created model achieved impressive results in terms of accuracy and classification strength. Kouchaki S et al. [7] used real-time sample data to identify medication resistance using several machine learning classifiers and linear dimension reduction approaches. Eleven distinct drugs were evaluated using a group of 13,402 samples gathered from sixteen nations across six continents. When compared to more conventional molecular diagnostic procedures, the authors' machine learning classifiers produced more accurate predictions. Information on the possible existence of drug-resistant TB may be derived from computational analysis of a chest X-ray. Models for TB type categorization using computed tomography (CT) lung images have been developed by Gao X et al. [8]. In order to apply the features of CT TB pictures with restricted areas collecting abnormalities, a Deep Convolutional Neural Network (CNN) and a Support Vector Machine (SVM) classifier were used [9]. Ye Q et al. [10] used four machine learning algorithms—Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Deep Neural Networks (DNN)—to differentiate between mycobacterium tuberculosis inhibitors and noninhibitors using different types of chemical representations. With the DNN model, they were able to attain the best classification accuracy. An automated classification method for identifying TB treatment responses using CT images is to be built utilizing a deep learning CNN architecture; this is the aim of the proposed study. The format of the research paper is as follows.

- The first section provides an overview of the study, including the disease, its frequency, the treatment method, the

importance of measuring medication response, and the developing role of AI in diagnosing TB. The methodology of the study is detailed in Section 2. The dataset, pre-processing techniques, architecture, and training procedures of the deep learning model are detailed in this section. Section 3 presents the findings, user interface, and explanations of the suggested method's operation on the test photos. In Section 4, we summarize the whole project.

II. PROPOSED METHODOLOGY

The use of AI to detect, predict, and treat medication sensitivity and resistance in TB is very promising. Through analysis of mycobacterium TB strain genetic makeup, artificial intelligence may aid in treatment resistance forecasting. An additional

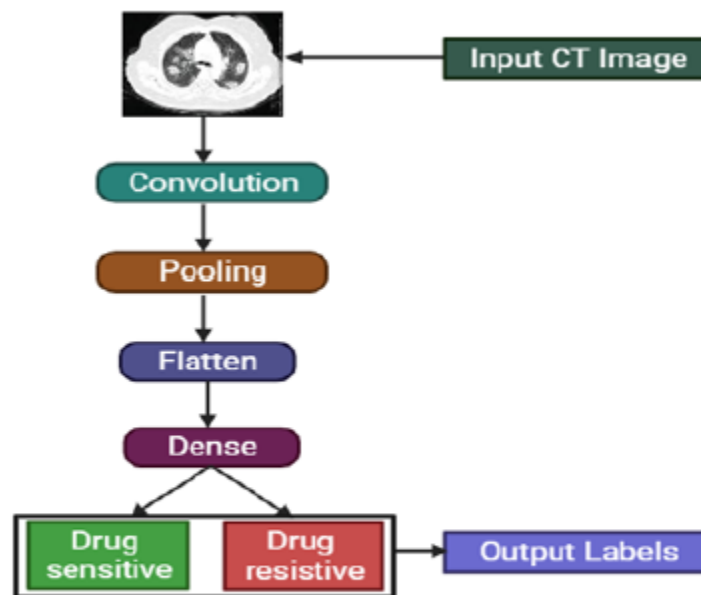


Fig. 2. Architecture diagram of the CNN model used for feature extraction and classification in the proposed work

By identifying genetic alterations associated with treatment resistance, machine learning algorithms enable quick and accurate diagnosis. Algorithms powered by artificial intelligence can sift through CT and chest X-ray images to identify patterns associated with TB and the extent of lung damage. This information may be useful for determining the severity of the disease and making therapy recommendations. In this study, we employed convolutional neural networks (CNNs) to automatically detect TB drug resistance in CT scans. Images are preprocessed before the dataset is prepared, the CNN model is developed, trained, validated, and finally, the result is deployed into an interface. As shown in Figure 1, the whole project is depicted in block form. When designing convolutional neural network (CNN) architectures for medical image processing, it is essential to consider the specific requirements and characteristics of medical pictures. Deep learning systems such as convolutional neural networks (CNNs) use many layers of convolutional operations to find patterns in data. Convolutional layers are the building blocks of a convolutional neural network (CNN). A number of filters, also called kernels, comprise these layers, and they traverse the input picture to detect various aspects [11]. certain filters may identify certain patterns or characteristics, such as edges, corners, or textures. By using convolutional layers, the network is able to learn hierarchical representations of the input data. The downsampling process begins with feature maps provided by convolutional

layers and continues with max-pooling layers. The computational efficiency and probability of over-fitting in the network are both enhanced by pooling, which lowers the spatial dimensions. In order to transform the 2D feature maps into a 1D vector, a flatten layer is included after the convolutional and pooling layers, as seen in the architectural diagram of the CNN model used in this study (Fig. 2). Getting the data ready for fully linked layers is what this does. A standard neural network has these layers, in which

TABLE I
NUMBER OF SAMPLES PRESENT PER EACH CLASS IN THE TRAINING AND TESTING SETS OF THE DATASET

Class	Volume		Image slices	
	Training	Testing	Training	Testing
Drug resistive	134	99	4020	2970
Drug sensitive	125	137	3750	4110
Total	259	236	7770	7080

Nodes in one layer may communicate with any node in any other tier. It is common practice to use the extracted characteristics to make predictions or classifications when using fully linked layers. This work employed the CLEF-2018 TB dataset [12], which contains 3D NIFTI CT volumes from drug-sensitive and drug-resistant individuals. We created a 2D image database by removing the center 30 slices from each CT volume. Table 1 shows the amount of data for each class's training and testing sets. The 2D slices were cropped to the desired area, and then scaled to the standard size of (224,224) for CNN training. In addition, a user interface is created to automatically analyze the CT images and generate predictions.

III. RESULTS AND DISCUSSIONS

The likelihood of treatment resistance in TB is often assessed using genetic or molecular data analysis. For this investigation, researchers selected a dataset of patients who had already completed medication resistance testing. There is information in this collection about the drug sensitivity and resistance of every sample. Using the genetic data, a CNN-friendly format is generated. This requires the use of encoding methods to represent the data numerically. Utilizing convolutional, pooling, and fully connected classification layers, we have implemented CNN in this study [13]. The possibility of medication resistance is produced by feeding encoded genetic data into the CNN. Convolutional neural networks (CNNs) are developed using training datasets, an optimization approach, and an appropriate loss function. By adjusting the CNN's weights to narrow the gap between the actual and predicted drug resistance states of each train in the training set, back-propagation trains the CNN on the training set. To train the CNN architecture, we use an Adam optimizer with a learning rate of 0.001 and a binary cross entropy loss function. Using a batch size of 16, the model is trained for 30 epochs. Using the validation set, we can monitor the CNN's progress during training and tweak its hyperparameters as needed to get better results. In order to compare the performance, we have also built and tested a classical machine approach using the identical input signals. When it comes to categorization and feature extraction, the conventional machine learning approach employs separate algorithms. We have ex- for the traditional ML model

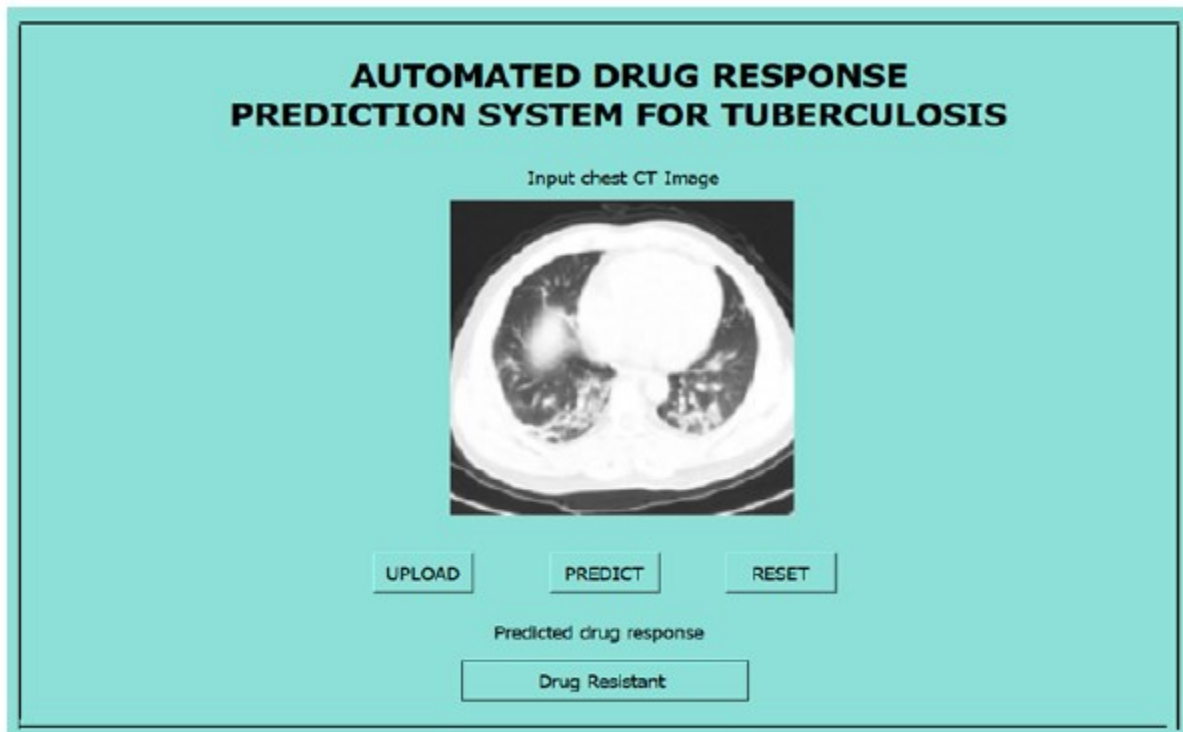


Fig. 3. Screenshot of the designed user interface showing the predictions made for a test sample

TABLE II
COMPARISON OF PERFORMANCE VALUES OF PROPOSED DEEP LEARNING
APPROACH AND CONVENTIONAL MACHINE LEARNING APPROACH

Parameters	Conventional Approach	CNN based model
Accuracy (%)	89.01	97.27
Sensitivity	0.928	0.983
Specificity	0.852	0.961
Precision	0.867	0.967
Error Rate	0.109	0.027
F1 Score	0.896	0.975
MCC	0.783	0.945

retrieved the CT image's textural properties by use of the gray-level co-occurrence matrix (GLCM) method [14]. The next step is to apply a support vector machine classifier to these characteristics. Our team has identified the typical metrics for assessment, such as precision, sensitivity, specificity, F1, score, error rate, and Matthews Correlation Coefficient (MCC) values. To get these metrics, one uses the confusion matrix-generated data on false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN) [15]. Table 2 displays the ROIs for both the conventional model and the proposed convolutional neural network (CNN) model. The suggested model scored 97.27 percent accuracy, 96.7 percent precision, 98.1 percent sensitivity, and 96.1 percent specificity on the test dataset. The conventional method that relies on machine learning achieved an accuracy of 89.01% and a rate of error of 0.109. This demonstrates how much more efficient the suggested model is when compared to the standard

techniques and approaches. Radiologists and physicians have a hurdle when trying to do analysis in a coding environment. So, making sure the interface is easy for doctors and radiologists to use is crucial. Using PyQt5, a Python user interface design tool, the project's user interface was created. One part of the interface is a file upload button that, when clicked, brings up a dialogue box asking for the image's directory path. When you submit a picture, all of the pre-processing is done, and the trained model receives the results in the background. Upon completion of the prediction, the user interface displays a text box indicating the outcome. A snapshot of the user interface generated using the PyQT tool and the anticipated output for a test sample are shown in Fig. 3. The suggested deep learning model only required 3.284 seconds to compute. This is quite low in comparison to more conventional machine learning approaches. Keep in mind that developing accurate models to predict medicine sensitivity or resistance is a difficult task requiring extensive research, expertise in the field, and collaboration with medical professionals. Validating their predictions via rigorous testing and clinical trials is crucial since these models are affected by several variables that impact their usefulness and reliability.

IV. CONCLUSIONS

The proposed study uses a deep learning model trained on computed tomography (CT) images to categorize TB medication response. A CNN model that uses deep learning makes the prediction. The suggested CNN-based model achieved a 97.27% classification accuracy, proving the efficacy of deep learning approaches in this investigation. CNN may expedite the assessment of medication sensitivity and resistance, which in turn speeds up the diagnosis and treatment of TB. More targeted and effective treatment programs may be the outcome of this. The suggested model performed well in the tests and did not show any signs of overfitting. The development and deployment of the user interface also aids radiologists in making real-time TB diagnoses. When it comes to assessing the treatment response of TB samples, the suggested technique may substantially assist radiologists in making better educated medical judgments. You should know that convolutional neural networks (CNNs) are just a part of the equation when it comes to drug resistance prediction. Having a high-quality dataset and picking the correct model architecture are also crucial for producing accurate predictions. Working together with medical professionals and subject specialists is crucial for making sure the model's predictions are in line with real-world clinical practice and decisions.

REFERENCES

- [1] Acharya, B., Acharya, A., Gautam, S., Ghimire, S.P., Mishra, G., Parajuli, N. and Sapkota, B., 2020. Advances in diagnosis of Tuberculosis: an update into molecular diagnosis of Mycobacterium tuberculosis. *Molecular biology reports*, 47, pp.4065-4075.
- [2] Ranjitha, J., Rajan, A. and Shankar, V., 2020. Features of the biochemistry of Mycobacterium smegmatis, as a possible model for Mycobacterium tuberculosis. *Journal of infection and public health*, 13(9), pp.1255-1264.
- [3] Ernest, J.P., Strydom, N., Wang, Q., Zhang, N., Nuermberger, E., Dartois, V. and Savic, R.M., 2021. Development of new tuberculosis drugs: translation to regimen composition for drug-sensitive and multidrugresistant tuberculosis. *Annual review of pharmacology and toxicology*, 61, pp.495-516.
- [4] Singh, V. and Chibale, K., 2021. Strategies to combat multi-drug resistance in tuberculosis. *Accounts of chemical research*, 54(10), pp.2361- 2376.
- [5] Jamal, S., Khubaib, M., Gangwar, R., Grover, S., Grover, A. and Hasnain, S.E., 2020. Artificial Intelligence and Machine learning based prediction of resistant and susceptible mutations in Mycobacterium tuberculosis. *Scientific reports*, 10(1), p.5487.
- [6] Karki, M., Kantipudi, K., Yu, H., Yang, F., Kassim, Y.M., Yaniv, Z. and Jaeger, S., 2021, November. Identifying drug-resistant tuberculosis in chest radiographs: Evaluation of CNN architectures and training strategies. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 2964-2967). IEEE.
- [7] Kouchaki, S., Yang, Y., Walker, T.M., Sarah Walker, A., Wilson, D.J., Peto, T.E., Crook, D.W., CRYPTIC Consortium and Clifton, D.A., 2019. Application of machine learning techniques to tuberculosis drug resistance analysis. *Bioinformatics*, 35(13), pp.2276-2282.

- [8] Gao, X., Comley, R. and Khan, M.H.M., 2020, October. An enhanced deep learning architecture for classification of tuberculosis types from CT lung images. In 2020 IEEE International Conference on Image Processing (ICIP) (pp. 2486-2490). IEEE.
- [9] Bhaskar, N. and Suchetha, M., 2019, July. An approach for analysis and prediction of CKD using deep learning architecture. In 2019 International Conference on Communication and Electronics Systems (ICCES) (pp. 1660-1664). IEEE.
- [10] Ye, Q., Chai, X., Jiang, D., Yang, L., Shen, C., Zhang, X., Li, D., Cao, D. and Hou, T., 2021. Identification of active molecules against Mycobacterium tuberculosis through machine learning. Briefings in Bioinformatics, 22(5), p.bbab068.
- [11] Navaneeth, B. and Suchetha, M., 2019. PSO optimized 1-D CNNSVM architecture for real-time detection and classification applications. Computers in biology and medicine, 108, pp.85-92.
- [12] Kozlovski, S., Liauchuk, V., Dicente Cid, Y., Tarasau, A., Kovalev, V. and Müller, H., 2020. Overview of ImageCLEF tuberculosis 2020 automatic CT-based report generation. In Proceedings of the CLEF 2020-Conference and labs of the evaluation forum. 22-25 September 2020.
- [13] Shinde, S., Tupe-Waghmare, P., Chougule, T., Saini, J. and Ingalthalikar, M., 2021. Predictive and discriminative localization of pathology using high resolution class activation maps with CNNs. PeerJ Computer Science, 7, p.e622.
- [14] Iqbal, N., Mumtaz, R., Shafi, U. and Zaidi, S.M.H., 2021. Gray level cooccurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. PeerJ Computer Science, 7, p.e536.
- [15] Bhaskar, N., Suchetha, M. and Philip, N.Y., 2020. Time series classification-based correlational neural network with bidirectional LSTM for automated detection of kidney disease. IEEE Sensors Journal, 21(4), pp.4811-4818.