TBXNet Pro: A Hybrid CNN-LSTM Architecture for Tuberculosis Detection from Chest X-Ray Images

Akshwin T
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology
Vellore, India
akshwint.2003@gmail.com

Kathiravan S
School of Computer Science and Engineering
Vellore Institute of Technology
Vellore, India
kathiravan.srinivasan@vit.ac.in

***Abstract*— Thie study proposes a TBXNET Pro, a hybrid deep learning model for tuberculosis detection using a combination of Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BI-LSTM) networks. The architecture utilizes CNNs for extracting relevant features from Chest X-ray images and Bi-LSTM layers to capture the temporal dependencies within these features. To enhance model robustness, data augmentation techniques such as rotation, zooming, shifting, and flipping were employed, significantly improving generalization and reducing the risk of overfitting. The model was trained using the Adam optimizer with a learning rate of 1e-4 and binary cross-entropy as the loss function. Evaluation using metrics like accuracy, precision, recall, and F1 score, along with a confusion matrix, indicated a high classification performance, achieving an accuracy of 99%. The findings highlight the model's effectiveness in detecting tuberculosis, offering a promising tool for clinical diagnosis and enhancing screening capabilities in medical imaging.**

Keywords —Tuberculosis detection, Deep learning, Hybrid CNN-RNN model, Medical Imaging, Classification.

I. INTRODUCTION

Tuberculosis remains one of the most significant global health challenges, accounting for an estimated 1.5 million deaths in 2020 alone [1]. The traditional diagnostic methods for TB, including sputum smear microscopy and culture tests, often fall short in terms of sensitivity and specificity, especially in cases involving extrapulmonary TB or co-infection with HIV [2]. Consequently, the development of efficient and reliable diagnostic tools is essential for timely treatment and control of the disease.

Recent advances in AI and Deep Learning have opened new avenues for improving Tuberculosis diagnostics. CNNs, in particular, have demonstrated exceptional performance in various medical image classification tasks [3]. These models can automatically learn to extract intricate features from images, reducing the dependency on human interpretation and the associated risks of misdiagnosis [4]. This shift towards automated diagnostics has the potential not only to enhance accuracy but also to facilitate early detection, ultimately leading to better patient outcomes.

Despite the promising capabilities of deep learning models, their performance is heavily reliant on the availability and quality of training data. A comprehensive dataset that captures the variability in TB presentations and the characteristics of normal chest X-rays is crucial for developing robust models. This research utilizes a dataset comprising 2,494 chest X-ray images from TB patients and 514 normal images from the source [15]. The images have been meticulously processed and resized to ensure uniformity, thereby providing a solid foundation for further analysis.

In addition to leveraging a high-quality dataset, this study employs data augmentation techniques to enhance the model's generalization capabilities. Techniques such as rotation, zooming, and horizontal flipping are applied to the training dataset while maintaining the integrity of the validation and test datasets. This approach aims to mitigate overfitting and improve the model's performance in real-world scenarios.

The primary objectives of this research are given below.

- Develop TBXNET Pro, a robust hybrid CNN-RNN model that accurately classifies Chest X-ray images, distinguishing between TB-infected and healthy cases.
- Harness the power of deep learning and optimize the model architecture

Abbreviations

Abbreviation	Full Form
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
BI-LSTM	Bidirectional Long Tern Short Memory
ADAM	Adaptive Moment Estimation
TB	Tuberculosis
AI	Artificial intelligence
HIV	Human Immunodeficiency Virus
CXR	Chest X-Ray
CAD	Computer Aided Diagnosis
AUC	Area under Curve
ROC	Receiver Operating Characteristic
VGG 16	Visual Geometry Group 16
CEED	Contrast-Enhanced Canny Edge detection
PCA	Principal Component Analysis
FC-SVNN	Fractional Crow Search-Based Neural Network
AFC	Adaptive Fractional Crow
CAPSNET	Capsule Neural network
ResNet	Residual Neural Network
I-CNN	Image Convolutional Neural Network
D-CNN	Demographic Variables Convolutional Neural network
ReLu	Rectified Linear Unit
RMSProp	Root Mean Square Propagation
DenseNet	Densely Connected Convolutional Networks

II. LITERATURE REVIEW

Ahmed Iqbal et al. (2022) introduced a novel CAD system called TBXNet for TB detection in CXR images [5]. The research aimed to address the challenges of manually interpreting CXR images, which often led to variability in diagnoses among radiologists. TBXNet utilized five dual convolution blocks with varying filter sizes (32, 64, 128, 256, and 512) integrated into a fusion layer that employed pre-trained layers to enhance feature extraction and knowledge transfer. The model was evaluated on multiple datasets, achieving high accuracy rates of 0.98 on Dataset A and 0.99 on Dataset B, with excellent performance on Dataset C, which included normal, tuberculous, pneumonia, and COVID-19 cases. The study reported high precision of 0.956, recall of 0.95, F1-score of 0.95, and overall accuracy of 0.95, demonstrating the superiority of TBXNet over other state-of-the-art methods. This work highlighted the potential of deep learning frameworks in providing cost-effective, scalable, and reliable solutions for TB detection, ultimately reducing diagnostic errors and dependency on radiologists in clinical settings [5].

Suci Aulia et al. (2022) proposed a system for TB detection based on CXR images using the VGG-16 CNN architecture [6]. The study addressed the challenge of variability in X-ray image quality, which could lead to diagnostic errors such as true negatives or false negatives in TB screening. To tackle this, they implemented a CNN model using the VGG-16 architecture, trained on a dataset consisting of 700 normal and 140 TB-affected chest X-ray images. The results of their study demonstrated a high classification accuracy of 0.997 in distinguishing between normal and TB-infected lungs. The study highlighted the importance of optimizing model parameters, such as batch size, to achieve peak performance, with the highest accuracy being obtained using a batch size of 50. This research contributed to the development of a robust machine learning model for assisting radiologists in the detection of TB from CXR images, aiming to reduce the error rate in initial TB screening and enhance early diagnosis and treatment [6].

Stefanus Kieu Tao Hwa et al. (2020) presented a novel approach to improving TB detection through the use of ensemble deep learning techniques combined with enhanced imaging [7]. The research focused on utilizing CEED to produce more precise lung CXR images, emphasizing edge details that could be crucial in identifying TB-affected areas. They generated two sets of features—one from the enhanced CXR images and another from the edge-detected images. This approach enhanced the diversity of error patterns among the base classifiers, leading to improved diagnostic accuracy. It addressed the limitations of previous methods that relied on training classifiers with similar features. The ensemble method proposed in this study achieved a high level of performance, with an accuracy of 0.936, sensitivity of 0.923, and specificity of 0.948. This demonstrated the effectiveness of using varied feature sets to enhance the capabilities of deep learning models in detecting tuberculosis, highlighting the importance of diverse feature extraction in ensemble learning [7].

Chithra et al. (2020) focused on tackling the challenges of accurate TB detection by leveraging deep learning techniques. The study began by transforming the images from the RGB (Red, Green, Blue) color space to LUV, enhancing image segmentation [8]. Adaptive thresholding was employed to segment the images more effectively, followed by the extraction of relevant features such as coverage, density, color histogram, area, length, and texture. PCA was used to optimize the features for better classification. The study introduced a Fractional Crow Search-Based Neural Network (FC-SVNN) to classify tuberculosis-infected regions and proposed an Adaptive Fractional Crow (AFC), a deep CNN model. This model adapted the fractional crow search method using self-adaptive concepts for improved accuracy in predicting TB severity. Image-level features like bacilli count and scattering coefficients played a crucial role in assessing infection severity. The proposed AFC-Deep CNN algorithm demonstrated high performance with a maximum accuracy of 0.935, indicating its effectiveness in enhancing TB detection and severity assessment compared to traditional methods. This study highlighted the significance of combining feature extraction techniques with advanced deep learning models to improve diagnostic accuracy in medical imaging [8].

Karnkawinpong et al. (2019) explored the effectiveness of CNNs in detecting TB in CXR images [9]. They focused on three deep neural network architectures—AlexNet, VGG-16, and CapsNet—to evaluate their performance in TB classification. The models were trained using datasets obtained from the National Library of Medicine and private Thai datasets. To enhance model performance and address the challenge of overfitting, data augmentation techniques such as shuffle sampling and affine transformations were employed. These techniques ensured that the models could generalize better when confronted with variant instances in test data that were not present in the training set. The study measured the effectiveness of the classifiers using accuracy, sensitivity, and specificity, finding that all models showed improved performance when trained on augmented datasets. This research highlighted the potential of deep learning approaches, particularly CNNs, in the early detection and classification of TB from CXR images, emphasizing the importance of data augmentation in achieving higher accuracy and robustness in medical image analysis [9].

Heo et al. (2019) investigated the effectiveness of CNNs in detecting TB using CXRs and demographic information [10]. They compared the performance of two models. First one using only image data (I-CNN) and another integrating demographic variables (D-CNN). The study employed multiple deep learning architectures, including VGG19, InceptionV3, ResNet50, DenseNet121, and InceptionResNetV2, trained on 1000 chest X-ray images. The findings revealed that D-CNN models, which incorporated demographic factors like age, weight, height, and gender, consistently outperformed I-CNN models in terms of the AUC and ROC curves. Specifically, the AUC for VGG19 increased by 0.0144 in the training set and 0.0138 in the test set when demographic data were included,

indicating statistically significant improvements. Furthermore, the D-CNN models demonstrated greater sensitivity and robustness compared to I-CNN models, especially at higher specificity levels. The study concluded that incorporating demographic information alongside image data enhanced the accuracy and sensitivity of TB detection, highlighting the potential of machine learning to improve diagnostic processes in clinical settings [10].

Ghorakavi et al. (2019) introduced TBNet, a deep learning system designed to enhance TB diagnosis using CXR images [11]. This approach addressed the limitations of existing methods that often relied on unnecessary features by employing the ResNet architecture coupled with effective data augmentation strategies, including Haar and Local Binary Patterns. These techniques improved the focus on TB-affected regions, significantly boosting detection accuracy. Previous research had demonstrated the efficacy of deep CNNs in TB detection, reporting high classification accuracies and emphasizing the importance of robust data augmentation for model generalization. By leveraging publicly available datasets, TBNet contributed to the evolving landscape of computer-aided diagnosis, promising enhanced screening capabilities in the ongoing fight against TB [11].

Hooda et al. (2019) presented a deep-learning-based system aimed at improving TB detection from CXR images [12]. Recognizing that traditional CAD systems faced significant challenges in accurately identifying TB due to the disease's varying impact on X-ray images, they proposed an ensemble method combining three well-established architectures. They are AlexNet, GoogleNet, and ResNet. By training these models from scratch and creating a tailored ensemble, the study showcased a method that not only enhanced classification performance but also achieved an impressive accuracy of 0.88 and an AUC of 0.93. This performance surpassed many existing TB detection methods, highlighting the potential of ensemble learning in addressing the complexities of TB diagnosis in medical imaging. The research underscored the importance of integrating multiple deep learning architectures to enhance diagnostic capabilities and assist radiologists in the effective identification of TB [12].

Liu et al. (2017) tackled the critical issue of TB diagnosis in low- and middle-income countries, where challenges such as healthcare inequalities and limited resources hindered effective management of the disease [13]. They proposed a novel approach utilizing a CNN designed to address the issues of unbalanced and less-categorized CXR images. The authors emphasized the significance of shuffle sampling combined with cross-validation in improving the model's training process, leading to enhanced accuracy in classifying various TB manifestations. Their method achieved a remarkable classification accuracy of 0.85 on a large TB image dataset, outperforming existing state-of-the-art techniques. The findings of this research highlighted the potential of advanced computer techniques to facilitate faster and more accurate TB diagnoses in resource-poor healthcare settings, ultimately contributing to the global effort to combat the TB epidemic [13].

Hwang et al. (2016) developed a CAD system aimed at improving TB screening [14]. The authors noted that TB remained a significant health threat, particularly in developing countries where diagnostic resources were limited. Traditional CAD systems often relied on handcrafted features to differentiate lesion types in CXR images, but these methods were not always effective. In contrast, the team employed a deep CNN for automatic TB screening, taking advantage of the network's ability to perform end-to-end training from feature extraction to classification, thus removing the need for manual feature engineering. Their results showed impressive TB screening performance, achieving AUC scores of 0.96, 0.93, and 0.88 across three field datasets. The authors also highlighted the role of transfer learning in boosting the model's accuracy and adaptability. This study demonstrated the potential of CNNs in developing more accurate and scalable TB screening solutions, which could significantly enhance early diagnosis efforts in regions with limited healthcare resources [14].

Table 1 Literature Review

Ref	Year of publication	Methodology	Accuracy
4	2022	TBXNet	99.17
5	2022	VGG-16	99.76
6	2020	Ensemble CNN	93.59
7	2020	AFC Deep CNN	93.5
8	2019	AlexNet VGG-16 CapsNet	92.9 94.56 90.33
9	2019	DCNN (Image CNN +Demographic variable)	Sensitivity=0.815 Specificity=0.962 AUC = 0.92
10	2019	Resnet-18	65.77
11	2017	CNN	85.68
12	2019	Ensemble of (AlexNet, GoogleNet &ResNet)	88.24
13	2016	AlexNet with transfer learning AlexNet without transfer learning	0.90 0.71

III. TBXNET PRO

A. System Architecture

The Figure 1 illustrates the system architecture TBXNET Pro architecture used to detect the tuberculosis from the CXR Images. The process begins with data collection, where relevant datasets are gathered, followed by data preprocessing to clean and prepare the data for analysis. Once the data is ready, data generators are created to efficiently manage large datasets in batches. The workflow

proceeds with the setup of the TBXNET Pro Model which is then compiled with appropriate configurations such as optimizer and loss function. The model is trained and compiled, learning patterns from the data. During training, an accuracy and loss curve is monitored to visualize the model's performance across epochs. The evaluation phase includes generating a classification report and a confusion matrix to assess metrics like accuracy, precision, and recall. This structured workflow supports iterative refinement, looping back to the initial input for further improvement or to handle new data.

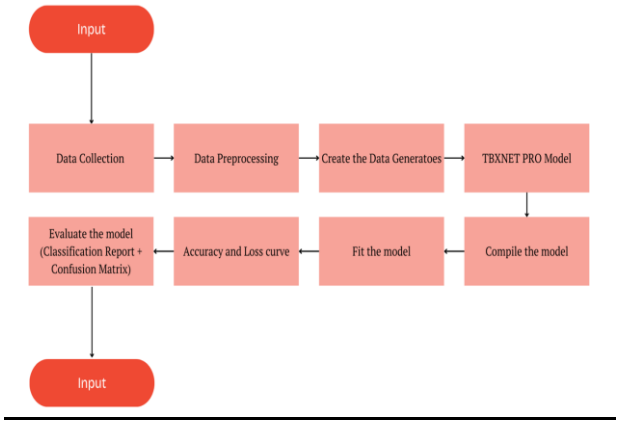

Figure 1 System Architecture

B. *Dataset and Preprocessing*

The dataset utilized in this study consists of 2,494 chest X-ray images from patients diagnosed with TB and 514 normal chest X-ray images, it's an open-source dataset from the open source [15] with comprehensive representation of TB-affected CXR, offering a diverse set of cases for analysis. All images have been uniformly resized to ensure consistency and facilitate efficient processing during model training.

To ensure that the model learns robust and generalizable features from the CXR images, a series of preprocessing steps are applied to the training dataset. This is crucial for enhancing the model's ability to handle variations in the data and avoid overfitting. The key preprocessing techniques include data augmentation and normalization.

Table 2 Dataset content

Class	Number of images
Normal Chest X-Ray	514
TB Chest X-Ray	2494

To account for variations in the angles at which X-ray images are captured, a random rotation of up to 20 degrees is applied. A zoom range of 15 percent simulates changes in camera distance or image magnification. Additionally, both the width and height of the images are shifted by up to 20 percent to emulate off-center captures, which makes the model less sensitive to the precise placement of the X-ray subject. Shear transformations are employed to introduce slight distortions, enhancing the model's invariance to geometric changes in real-world data. Random horizontal flips further augment the diversity of training samples, thereby reducing the risk of overfitting to specific image orientations. These transformations are applied randomly to each image in the training dataset during model training, improving the model's generalization ability by exposing it to a broad range of potential distortions and orientations.

After augmentation, all images are rescaled by dividing pixel values by 255, normalizing them to a range of 0 and 1. This normalization ensures that all input data is on a consistent scale, which helps improve the convergence speed during model training by reducing the variance in pixel intensity values [16].

To maintain the integrity of the validation and test sets, no augmentation is applied to these datasets [16]. Instead, the images are only rescaled by a factor of 1/255, ensuring that the evaluation of the model is performed on unaltered data, thus providing a realistic assessment of model performance.

C. *Proposed Model Architecture*

The proposed TBXNET Pro architecture for the tuberculosis detection model is designed as a streamlined and efficient pipeline for processing CXR images. At the outset, raw images undergo a series of preprocessing steps, including resizing to a uniform dimension and normalization, ensuring that all images are suitable for input into the neural network. This is followed by data augmentation techniques to enhance the dataset's diversity, enabling the model to generalize better by preventing overfitting to specific patterns.

The core of the architecture consists of a CNN, which is responsible for feature extraction from the CXR images. The CNN comprises several convolutional layers that progressively learn hierarchical features, followed by pooling layers that reduce the spatial dimensions of the data, leading to computational efficiency. The model is trained using the Adam Optimizer and Binary Cross-Entropy as the loss function, facilitating optimal convergence during the training phase. Overall, TBXNET Pro architecture is aimed at providing accurate and reliable predictions for tuberculosis detection from CXR images, thus aiding in timely diagnosis and treatment. Figure 3 shows the model Architecture of TBXNET Pro.

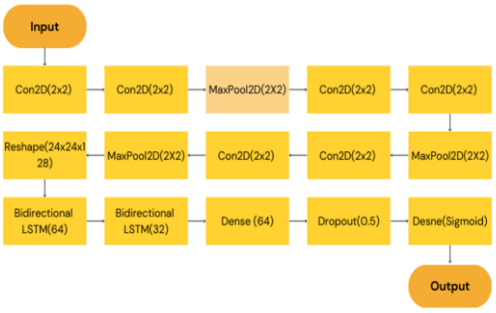

Figure 3 TBXNET Pro Model Architecture

In this subsection, a detailed analysis and explanation of the methodological development and implementation of our proposed architecture TBXNET Pro, a hybrid CNN-BiLSTM model designed for tuberculosis detection from CXR images.

The TBXNET Pro architecture integrates two powerful deep learning models. They are Convolutional Neural Networks and Bidirectional Long Short-Term Memory networks. CNNs are responsible for extracting key spatial features from the input images, capturing important patterns and visual hierarchies. After three convolutional blocks, the model applies a flattening layer to transform the feature maps into a suitable format for the LSTM layers. Subsequently, two bidirectional LSTM layers are employed to capture both forward and backward dependencies in the sequential data, improving the model's ability to learn temporal relationships from the extracted features. This is followed by a dense layer, a dropout layer to reduce overfitting, and a final dense output layer, which provides the prediction.
This hybrid TBXNET Pro architecture enhances the model's capability in processing image data, especially when sequential dependencies within the features need to be considered for accurate disease detection.

CNNs are used to extract spatial features from the input X-ray images. The architecture comprises three blocks of convolutional and pooling layers. In first block, the input image is processed by two 2-dimensional convolution layers, each with 32 filters and a 3x3 kernel. This block is followed by a max-pooling layer that reduces the spatial dimensions by half. The output is a feature map of size (112, 112, 32), capturing local patterns such as edges and textures. The second block contains 64 filters in each convolutional layer. Max-pooling reduces the size to (56, 56, 64). These filters detect more complex patterns, such as parts of the lungs or other anatomical features. The third block has 128 filters in each convolutional layer. After max-pooling, the feature map size is (24, 24, 128), encoding even higher-level abstractions such as anomalies in the lung structure. After the final CNN blocks, the 3-dimensional feature maps are flattened and reshaped into a 2D sequence of size (576, 128), where each of the 576 steps represents a spatial region of the image, and the 128 features represent learned characteristics of that region. This transformation allows the Bi-LSTM layers to process the spatial information sequentially.

LSTMs are specialized for handling sequential data, and in this architecture, they process the sequence of flattened features generated by the CNN layers. Two BiLSTM layers are used. The first BiLSTM layer has 64 units and returns a sequence, preserving the temporal relationships between the spatial features of the image. The second BiLSTM layer has 32 units and provides a final condensed sequence representation of the X-ray. By processing the data in both forward and backward directions i.e., bidirectional, the BiLSTMs capture the full context of the sequence, improving the model's ability to detect subtle patterns associated with pneumonia.

The output from the BiLSTM layers is fed into a fully connected neural network for the final classification. This consists of a fully connected layer with 64 neurons and ReLU activation. This layer processes the sequence output from the BiLSTM layers and prepares it for classification. The ReLU is used to introduce non-linearity and help the network to learn more complex patterns. The ReLU function takes an input value and outputs the value itself if it's greater than or equal to zero, and zero if it's negative. The ReLu function is given by the Equation 1.

$$R(x) = \max(0, x) \qquad (1)$$

The dropout layer is a widely-used regularization technique in deep learning to prevent overfitting, especially in models that are prone to memorizing the training data. It works by randomly dropping out a fraction of the neurons during each training iteration, effectively setting their output to zero. This process forces the network to learn more robust features by preventing it from relying on specific neurons. Mathematically, the dropout process can be represented by the Equation 2.

$$y = \frac{1}{1-p}.W.x \qquad (2)$$

where,

- p is the dropout rate
- W is the weight matrix,
- x is the input,
- y is the output of the layer after applying dropout.

The factor 1/(1-p) ensures that the outputs are appropriately scaled during training, so the magnitude of the outputs remains consistent even with some neurons being dropped. During testing, no neurons are dropped, and the entire network is used for predictions. A dropout rate of 50% is applied to reduce overfitting by randomly ignoring half the neurons during each training iteration. This is because the regularization parameter, $p(1\text{-}p)$ in Equation above is maximum at $p = 0.5$.
The final layer has a single neuron with a sigmoid activation function, which outputs a probability value between 0 and 1. This represents the binary classification of either pneumonia i.e., class 1 or normal i.e., class 0. The sigmoid function used is given by the Equation 3.

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (3)$$

where,

- x represents the input to the function.
- $\sigma(x)$ represents the sigmoid function
- e represents the base of the natural logarithm.

D. *Hyperparameter*

The proposed TBXNET Pro model is compiled using the Adam optimizer, with binary cross-entropy as the loss function and accuracy as the primary evaluation metric. The learning rate is set to 1e-4 to ensure a balance between convergence speed and training stability. Table 3 shows the hyperparameters used during the experiment.

Table 3 Hyperparameters

Hyperparameter	**Value**
Optimizer	Adam
Loss Function	Binary Cross Entropy
Learning Rate	1-e4

a) Adam Optimizer

The Adam optimizer is employed due to its efficiency and adaptability in handling sparse gradients and noisy data. Adam combines the advantages of two traditional optimization methods, Momentum and RMSProp. The recursive update equation for a parameter θ_t in Adam optimizer is given by the Equation 4

$$\theta_t = \theta_{t-1} - \frac{(\alpha * m_t)}{\sqrt{v_t} + \epsilon} \qquad (4)$$

where,

- θ_t represents the current parameter value at time step t.
- α represents the learning rate.
- m_t represents the first moment estimate i.e., the moving average of the gradients, which helps smooth out oscillations.
- v_t represents the second moment estimate which accounts for the gradient variance.
- ϵ represents a small constant added to prevent division by zero.

Adam adjusts the learning rate for each parameter based on the moments of the gradients, allowing for faster convergence with reduced risk of overshooting the minimum.

b) Binary Cross-Entropy Loss Function

Since it is a binary classification problem binary cross-entropy (BCE) is used as the loss function. The binary cross entrpy loss function is represented by the Equation 5.

$$L_{bce} = -\left(y * \log\left(y_{pred}\right) + (1-y).\log(1-y_{pred})\right) \quad (5)$$

where,

- L_{bce} represents the binary cross entropy loss.
- y represents the true label which can be 0 or 1.
- y_{pred} represents the predicted probability.
- log denotes the natural logarithm.

This loss function calculates how far the predicted probabilities are from the true binary labels and helps guide the model's learning process to minimize the error. BCE is particularly well-suited for binary classification problems, ensuring the model is penalized proportionally to the confidence of incorrect predictions.

c) Learning Rate

The learning rate, a critical hyperparameter, controls the step size during optimization. A smaller learning rate, like the chosen 1e-4, ensures that updates are made gradually, promoting stable convergence while avoiding overshooting minima. This value was empirically chosen after experimentation to provide a good trade-off between speed of convergence and model stability during training.

IV. EVALUATION METRICS

A. *Accuracy curve*

The accuracy curve tracks the model's performance over time, showing an upward trend in both training and validation accuracy. The gradual improvement indicates that the model is learning effectively and generalizing well to unseen data without overfitting. Figure 4 and Figure 5 shows the accuracy curve of the model for 25, 50, 75 and 100 epochs.

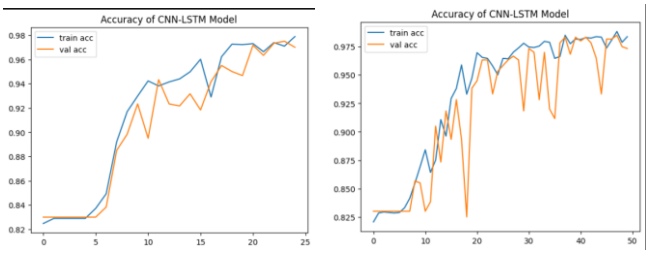

Figure 4 Accuracy curve for 25 and 50 epochs

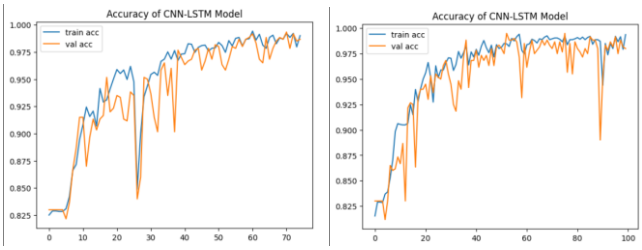

Figure 5 Accuracy curve for 75 and 100 epochs

B. *Loss curve*

The loss curve, based on the binary cross-entropy function, shows a steady decrease in both training and validation loss across epochs. This suggests the model is minimizing errors effectively while maintaining a balance between learning and preventing overfitting. Figure 6 and Figure 7 shows the loss curve of the model for 25, 50, 75 and 100 epochs.

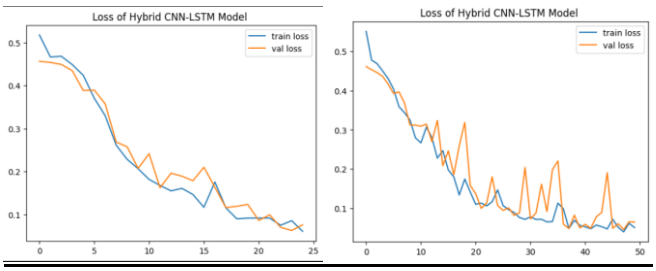

Figure 6 Loss curve of the model for 25 and 50 epochs

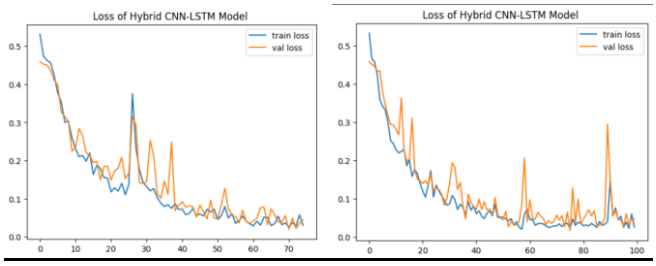

Figure 7 Loss curve of the model for 75 and 100 epochs

C. *Accuracy score*

Accuracy is defined as the ratio of correctly predicted instances i.e., true positives and true negatives to the total number of instances in the dataset. Mathematically, accuracy is expressed in the Equation 6.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (6)$$

D. *Precision*

Precision is defined as the ratio of correctly predicted positive instances i.e., true positives to the total number of instances predicted as positive i.e., the sum of true positives and false positives. Mathematically, precision is expressed in the Equation 7.

$$Precision = \frac{TP}{TP + FP} \qquad (7)$$

E. *Recall*

Recall is defined as the ratio of correctly predicted positive instances i.e., true positives to the total number of actual positive instances i.e., the sum of true positives and false negatives. Mathematically, recall is expressed in the Equation 8.

$$Recall = \frac{TP}{TP + FN} \qquad (8)$$

F. *F1-score*

F1 Score is a harmonic mean of precision and recall, providing a single metric that balances both metrics in situations where there is an uneven class distribution. Mathematically, recall is expressed in the Equation 9.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (9)$$

G. *Confusion matrix*

A confusion matrix is a table used to evaluate the performance of a classification algorithm by comparing the actual target values with those predicted by the model, displaying TP, TN, FP, FN. The figure 8 shows the confusion matrix and the terms in it.

	Positive	Negative
Positive	TP	FP
Negative	FN	TN

where,

- *TP refers to True Positive*
- *TN refers to True Negative*
- *FP refers to False Positive*
- *FN refers to False Negative*

. Figure 8 and Figure 9 shows the confusion matrix of the model for 25, 50, 75 and 100 epochs.

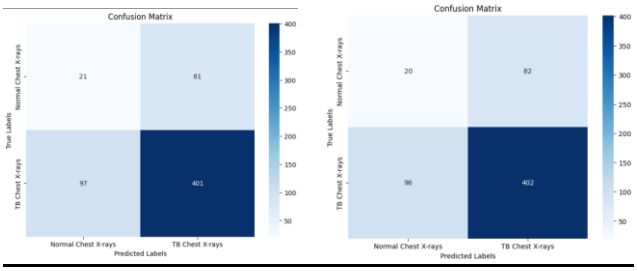

Figure 9 Confusion matrix for 25 and 50 epochs

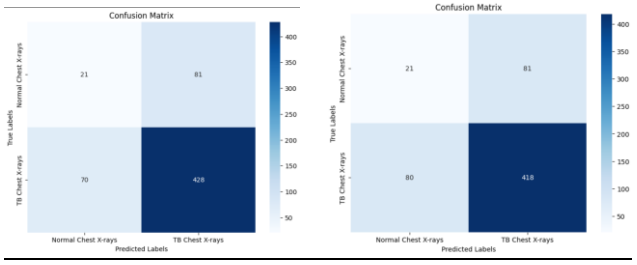

Figure 10 Confusion matrix for 75 and 100 epochs

The table 4 shows the performance of the model with the values of the various evaluation metrics used.

Table 4 Performance of the model

S.No	Epochs	Accuracy	Precision	Recall	F1 Score
1	25	0.97	0.72	0.70	0.71
2	50	0.96	0.72	0.70	0.71
3	75	0.99	0.74	0.75	0.74
4	100	0.99	0.73	0.73	0.73

V. Conclusion and Future Work

In conclusion, this study demonstrates the effectiveness of deep learning techniques in the detection of tuberculosis from chest X-ray images. The TBXNET Pro model achieved a high accuracy of 99% after rigorous training and evaluation, showcasing its potential as a reliable tool for aiding radiologists in the early diagnosis of TB. The preprocessing techniques applied, including data augmentation and normalization, significantly contributed to the model's performance by enhancing its ability to generalize across diverse image conditions. This underscores the importance of robust data handling in machine learning applications in healthcare.

For future work, the dataset can be expanded to include a broader range of chest X-ray images, encompassing various stages of tuberculosis and other respiratory conditions. The number of normal CXR images can be added for better precision, recall and f1-score. This will help improve the model's robustness and accuracy further. Additionally, exploring advanced architectures, such as transfer learning with pre-trained models, could enhance performance. Implementing real-time analysis tools and integrating this model into clinical workflows will also be a priority, providing healthcare professionals with efficient diagnostic support in their practice.

VI. References

1. World Health Organization. (2021). Global Tuberculosis Report 2021.
2. Boehme, C. C., Nabeta, P., Hillemann, D., Nicol, M. P., Shenai, S., Krapp, F., ... & Perkins, M. D. (2010). Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, *363*(11), 1005-1015.
3. Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014, December). Medical image classification with convolutional neural network. In *2014 13th international conference on control automation robotics & vision (ICARCV)* (pp. 844-848). IEEE.
4. Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, *284*(2), 574-582.
5. Hooda, R., Mittal, A., & Sofat, S. (2019). Automated TB classification using ensemble of deep architectures. *Multimedia Tools and Applications*, *78*, 31515-31532.
6. Liu, C., Cao, Y., Alcantara, M., Liu, B., Brunette, M., Peinado, J., & Curioso, W. (2017, September). TX-CNN: Detecting tuberculosis in chest X-ray images using convolutional neural network. In *2017 IEEE international conference on image processing (ICIP)* (pp. 2314-2318). IEEE.
7. Ghorakavi, R. S. (2019). TBNet: pulmonary tuberculosis diagnosing system using deep neural networks. *arXiv preprint arXiv:1902.08897*.
8. Hwang, S., Kim, H. E., Jeong, J., & Kim, H. J. (2016, March). A novel approach for tuberculosis screening based on deep convolutional neural networks. In *Medical imaging 2016: computer-aided diagnosis* (Vol. 9785, pp. 750-757). SPIE.
9. Heo, S. J., Kim, Y., Yun, S., Lim, S. S., Kim, J., Nam, C. M., ... & Yoon, J. H. (2019). Deep learning algorithms with demographic information help to detect tuberculosis in chest radiographs in annual workers' health examination data. *International journal of environmental research and public health*, *16*(2), 250.
10. Karnkawinpong, T., & Limpiyakorn, Y. (2019, April). Classification of pulmonary tuberculosis lesion with convolutional neural networks. In *Journal of Physics: Conference Series* (Vol. 1195, No. 1, p. 012007). IOP Publishing.
11. Chithra, R. S., & Jagatheeswari, P. (2020). Severity detection and infection level identification of tuberculosis using deep learning. *International Journal of Imaging Systems and Technology*, *30*(4), 994-1011.
12. Hwa, S. K. T., Bade, A., Hijazi, M. H. A., & Jeffree, M. S. (2020). Tuberculosis detection using deep learning and contrastenhanced canny edge detected X-Ray images. *IAES International Journal of Artificial Intelligence*, *9*(4), 713.
13. Aulia, S., & Hadiyoso, S. (2022). Tuberculosis detection in X-Ray image using deep learning approach with VGG-16 architecture. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, *8*(2), 290-297.
14. Iqbal, A., Usman, M., & Ahmed, Z. (2022). An efficient deep learning-based framework for tuberculosis detection using chest X-ray images. *Tuberculosis*, *136*, 102234.
15. https://data.mendeley.com/datasets/8j2g3csprk/2
16. Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, *3*(1), 91-99.