Journal of Computer Science

Original Research Paper

Classification of X-Ray Images Using Convolutional Neural Network and Automatic Hyper-Parameter Selection to Detect Tuberculosis (TB)

[1]Biswaranjan Debata, [1]Rojalina Priyadarshini, [2]Sudhir Kumar Mohapatra and [3]Tarikwa Tesfa Bedane

[1]Department of Computer Science and Engineering, C.V. Raman Global University, India
[2]Faculty of Emerging Technologies, Sri Sri University, Cuttack, India
[3]Department of Software Engineering, Addis Ababa Science and Technology University, Ethiopia

Article history
Received: 09-07-2024
Revised: 10-09-2024
Accepted: 06-11-2024

Corresponding Author:
Tarikwa Tesfa Bedane
Department of Software Engineering, Addis Ababa Science and Technology University, Ethiopia
Email: tarikwa.tesfa@aastu.edu.et

Abstract: Tuberculosis (TB) is a major public health issue in India, contributing significantly to the global burden of respiratory diseases. This study introduces a Convolutional Neural Network (CNN)--based model for the early and cost-effective detection of TB using chest X-ray images. The proposed model, featuring 13 layers and automated hyperparameter selection, classifies images as infected or not infected. It is evaluated on three open datasets: Chest X-ray Masks and Labels, Tuberculosis X-ray (TB ×11 K), and Shenzhen. The model achieves an accuracy of 99.42% on the chest X-ray masks and label dataset, 99.27% on the TB ×11 K dataset, and 97.73% on the Shenzhen dataset, outperforming six existing models in terms of F1 score and precision. Unlike existing models that are tested on a single dataset, our model demonstrates consistent and robust performance across multiple datasets, highlighting its generalizability.

Keywords: Tuberculosis, TB, TB Detection Using CNN, CNN

Introduction

Pulmonary disease is a group of lung diseases that prevents the airflow and makes it difficult to breathe. Asthma, Tuberculosis, COVID-19, and lung cancer are examples of pulmonary illnesses. The condition known as Chronic Obstructive Pulmonary Disease (COPD) is brought on by the lungs and airways destruction and inflammation. It is frequently linked to long-term exposure to hazardous chemicals like cigarette smoke. Tuberculosis, asthma, occupational lung disorders, and lung cancer are among the most common. Other troubles, besides tobacco smoking, comprise air pollution, occupational toxins and dust, and recurring lower breathing illnesses throughout infancy. In India Tuberculosis is a major health challenge. It is a pulmonary illness, instigated by the bacteria named Mycobacterium tuberculosis. TB germs not only bouts the lungs but also affects other regions of the body which include the kidney, brain, and spine. It is not necessary that every infected tuberculosis patient will show symptoms of illness. Because two COVID-19 waves in India resulted in a slight drop in TB registrations yet National Tuberculosis Elimination Program (NTEP) notified a figure that a rise of 19% over the previous year, with 19,33,381 in the year 2021 compared to 16,28,161 in the year 2020. The foresight of the National Strategic Plan for Tuberculosis Elimination (NSP 2017-25) will have pervaded to the district and state level once more to include further targets in the year 2021 (Division, 2022).

Co-morbidities such as malnutrition, diabetes, HIV, cigarette smoking, and alcohol all influence TB propensity and severity. To tackle this, the program implemented a slew of measures. In 2021, data was collected for 72% of all communicated patients, with 7% admitting to alcohol use. Similarly, 12% of the TB sufferings appeared to be tobacco consumers, even though 74% of all TB patients were known to use tobacco. Thirty percent of individuals examined were referred to cigarette cessation assistance. Childhood tuberculosis is a massive concern in India, accounting for around 31% of the global burden. Over the previous decade, children have consistently constituted 6-7% of all NTEP patients treated annually, indicating a 4-5% gap in overall notification vs the projected incidence (Division, 2022).

Detecting Tuberculosis (TB) from chest X-rays is vital in a clinical context for numerous reasons. TB is an infectious and possibly fatal illness if not treated. Early identification by chest X-rays enables rapid diagnosis and treatment, preventing the illness from spreading to others and lowering the chance of serious sequelae.

Science Publications

© 2025 Biswaranjan Debata, Rojalina Priyadarshini, Sudhir Kumar Mohapatra and Tarikwa Tesfa Bedane. This open-access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

Tuberculosis is a major worldwide illness, especially among low- and middle-income nations. Chest X-rays are a commonly available screening technique that may be utilized for widespread testing in TB-endemic individuals, aiding in preventing infection and frequency reduction. Chest X-rays remain one of the non-invasive and less expensive ways of identifying possible cases of TB. This is of particular importance in resource-poor areas where more sophisticated diagnostic approaches such as molecular testing might not be easily accessible. Chest X-rays reveal certain abnormalities of TB like lung infiltrate cavitations and pleural effusions. These results may assist clinicians in determining whether to perform further tests or provide an empirical treatment in patients with clinical TB. In the case of patients undergoing treatment for TB, high-quality chest X-rays are also needed for the evaluation of disease progress and the results of therapy. They assist in evaluating the healing of the lesions and the presence of complications or relapses.

Finding Tuberculosis (TB) using chest X-rays has numerous obstacles that might affect the precision and efficacy of detection. Physicians' knowledge and expertise are primarily relied upon when interpreting chest X-rays. Variability in skill levels can result in inconclusive diagnosis, with some cases of tuberculosis being undetected or misinterpreted. Tuberculosis can have a variety of diagnostic patterns that match with various other lung illnesses, such as pneumonia, lung cancer, or fungal infections. This lack of precision can make it hard to identify tuberculosis from other illnesses. In most low-resource situations, X-ray devices and scans are of inferior quality, which leads to poor pictures that make comprehension difficult. Limited optical technology and upkeep worsen such problems. While improvements in machine learning (AI) and Computer-Aided Detection (CAD) systems offer potential for improving TB diagnosis using chest X-rays, accessibility to such tools is still limited in many areas. The high expenses, inadequate facilities, and inadequate education in using these technologies limit adoption.

The machine learning model is the most widely used model for the prediction purpose. Nowadays day machine learning methods are mostly used in the area of health science for the prediction of diseases. Convolution Neural Networks (CNN) is a part of the deep learning advanced in the present day. It is applied in the field of image processing. The CNN comprises feature extraction as well as a mapping layer. A model proposed by Li *et al*. (2018) describes how CNN is used for the analysis of TB from CT images. In the present proposed model, he describes how CNN works. In this model, he combines two features that is features of an auto encoder with the features of Conv. Here he also used an eigenvector which is used to instruct a SoftMax classifier. The result is a six-dimensional vector that identifies the possibility of diseases. Using a self-encoder in this proposed model it has been depicted that the classification model has enhanced. Junaedi *et al*. (2019) describe a model that is used to detect TB from CXR images exploiting optimized Gray Level Co-occurrence Matrix (GLCM) facets as given as the input. GLCM is advanced putting the use of the Principal Component Analysis (PCA) and then cataloged with the help of a Support Vector Machine (SVM). Multi-class SVM method is used for the classification of this model. C-SCV is used to separate different classes. Here cross validation method is used to validate the classification system. In the SVM classification problem, PCA is more accurate in optimizing the GLCM features.

Oltu *et al*. (2021) proposed a system, that can spontaneously segregate unaffected, and Chest X-Ray (CXR) images are produced for the scheme and exact inspection. In this model, different networks are applied and the execution is examined to discover the extractor that is better than the other. As the size of data is limited, therefore CNN architecture is chosen. In this model, SVM is provided with various architectures for classification. This model provides a result accuracy of 96.6% with 0.99 AUC.

The major challenge of our research is the availability of the data set. The local health center keeps the data without any labeling. Even they are very reluctant to share the data. Openly available data set is very small which is a major challenge in this research. As machine learning models are data-hungry so large amount of data is a necessity for the accuracy of the model. Another challenge is when the center agrees to provide unlabeled data, there is a lot of paperwork that needs to be done. This paperwork required more time. The research gap of the existing work is studied and listed in the related work section.

Tuberculosis detection from CXR images nowadays takes human interference which is a time-consuming process. The main objective of our research:

1- To automatically detect the disease correctly
2- To identify the disease without any expert intervention

Related Work

This section focuses on the existing research related to Tuberculosis (TB) prediction using machine learning techniques. Several relevant studies have been reviewed from repositories like Scopus and Web of Science.

In the study by Rustum *et al*. (2022), a model was proposed to identify patients with tuberculosis and prioritize critical genes for further biological examination among a large set of candidates. The model utilizes Otsu's binarization method followed by the application of a Convolutional Neural Network (CNN) algorithm trained on a dataset comprising 700 X-ray images of TB-infected patients and 400 images of healthy individuals. The model's performance was evaluated using a confusion matrix and dict. Keys () method, achieved an accuracy of 95%.

Similarly, Mehrrotraa *et al*. (2022) presented an ensemble approach combining efficient deep convolutional

networks and machine learning algorithms for TB detection. The model incorporates multiple deep convolutional networks, where salient features extracted by these networks serve as input to machine learning classifiers for final prediction. This model was validated using both 5 and 10-fold cross-validation techniques, yielding accuracies of 87.9 and 99.1%, respectively, with AUC scores of 0.94 and 1.0 for normal and TB-infected images.

Lu *et al*. (2022) introduced TBNet, a novel TB diagnosis model based on a context-aware graph neural network. The model extracts sample-level features while leveraging context evidence from neighboring features in the feature space. Through 5-fold cross-validation, TBNet achieved an overall accuracy of 98.93%, a sensitivity of 100%, a precision of 97.86%, and an F1-score of 98.91%.

In another study, Duong *et al.* (2021) explored TB detection from chest X-ray images using a dual-stage model combining Deep Convolutional Neural Networks (DCNN) with efficient net and attention-based vision transformers. The study addressed four key research questions:

1. Which network family provides the best prediction performance?
2. Which transfer learning technique is most effective?
3. How does the proposed classifier compare with the baseline?
4. How reliable are the models in forecasting outcomes?

Dey *et al*. (2022) developed a fuzzy ensemble CNN model, utilizing VGG19, ResNet50, and DenseNet121 as pre-trained models for feature extraction in the detection of TB from chest X-ray images. The confidence scores from these models were combined using a Sugeno integral-based ensemble method to generate the final prediction. For optimization, eight different algorithms were employed and the model achieved an accuracy of 99.75% using the Ant Lion Optimization (ALO) technique in a 5-fold cross-validation.

An *et al*. (2022) proposed a model incorporating an Efficient Channel Attention (ECA) mechanism alongside a novel network architecture for chest X-ray image analysis. Although this model integrated ECA blocks for enhanced feature representation, it yielded a relatively lower accuracy of 85.0% compared to other models.

Additionally, Support Vector Machines (SVM) have been employed by various researchers for classification tasks related to COVID-19 and other pulmonary diseases (Mohapatra *et al*., 2022; Debata *et al*., 2023).

Sharma *et al*. (2024) suggested the technique uses segmentation and classification algorithms for identifying TB in CXR pictures. The UNet segmentation model is used to segment CXR pictures and pick regions of interest, such as lung parts while eliminating noise. The Xception DL classification model is used to separate lung parts into TB and normal categories. They employed a five-fold cross-validation approach for segmentation and an 80:20 training/testing ratio and acquired 99.29% classification accuracy and an AUC of 0.999. Goswami *et al.* (2023) The project attempted to create a TB detection model using chest X-ray pictures from Kaggle.com and Google's collaboration platform. The model was developed using 1196 chest X-ray pictures from both TB-positive and normal patients. The program was taught to identify abnormalities in TB chest X-rays, allowing patients to receive timely treatment. Gupta *et al*. (2024) proposed that the FLQ-HRM test is a Mol DST approach for detecting FLQ-resistant tuberculosis in sputum samples, outperforming earlier HRM assays for clinical strains. Using sputum samples directly reduces medication resistance diagnosis time and aids in initial therapy commencement. Genetic sequencing is able to be utilized as Mol-DST. The WHO wants to examine its efficacy and endorse it as the benchmark for genotypic DST. The purpose of this article by Singh *et al*. (2024) is to use sputum pictures to assess the seriousness of tuberculosis. Previously, sputum pictures were pre-processed using dynamic bilateral filters to remove ring distortion from input images. Bacilli categorization is performed using pre-processed pictures and U-Net. The divided raw sputum picture is then sent to the element analysis step, where it is extracted for texture and GLCM characteristics. DenseNet uses the collected characteristics to determine the severity of TB. The SMRO is used for DenseNet training. The SMRO integrates WSO and MRO. The suggested SMRO-based DenseNet enhanced efficiency and accuracy by 94.7 percent. This study by Vats *et al*. (2024) proposes an intelligent system that uses incremental learning to detect and locate tuberculosis infections. The above framework employs a multi-layered strategy to identify infections and understand the effect of emerging TB strains. The ILCM model outperforms both local demographics (10% not included in training) and a demographic dataset (Golden standard prepared from various geographic regions). Additionally, the gradual technique lowers the need for persons to prepare training datasets on a regular basis. The summary of the related work is presented in Table (1).

The major limitations of the current work can be summarized as:

1. Current models often use limited or non-diverse datasets, hindering their ability to generalize across broader populations and various TB manifestations. Expanding dataset diversity is crucial for improving model robustness and real-world applicability
2. Many models struggle with poor-quality or noisy chest X-ray images, affecting their accuracy. Future research should focus on developing preprocessing techniques or more adaptive models to handle low-resolution or inconsistent image data effectively
3. There is a need for models that focus on early TB detection and the identification of drug-resistant strains, which are currently underexplored. Enhancing model capabilities in these areas is vital for improving treatment strategies and outcomes

Table 1: Summary of related works

Author	Title	Performance	Dataset	Open issue
Rustum *et al*. (2022)	Detection of tuberculosis using machine learning techniques and image processing	Accuracy 95%	Kaggle dataset	NA
Mehrrotraa *et al*. (2022)	Ensembling of efficient deep convolutional networks and machine learning algorithms for resource-effective detection of tuberculosis using thoracic (chest) radiography	Accuracy 98%	JSRT, KIT, etc.	NA
Miranda *et al*. (2022)	Pulmonary tuberculosis detection using digitally photographed chest X-RAY images	Accuracy 87.9 and 99.1%	TB ×11 K dataset, Shenzhen dataset,	Limited dataset, so do not capture all the symptoms of diseases
Lu *et al*. (2022)	TBNet-A context-aware graph network for tuberculosis diagnosis	Accuracy of 98.93%, a sensitivity of 100.00%, a precision of 97.86%	Fourth hospital of Huai'an	Fusion of distinct image can be progress the classification execution
Duong *et al*. (2021)	Detection of tuberculosis from the chest x-ray images: boosting the performance with vision transformer and transfer learning	Accuracy of 97.92%	MC dataset, Shenzhen dataset,	To incorporate more baseline for comparison with the conceived tools
Dey *et al*. (2022)	An optimized fuzzy ensemble of convolutional neural networks for detecting tuberculosis from chest X-ray images	Accuracy of 99.75%	NLM dataset, Belarus data set, RSNa chest X-ray data set	Poor-quality images cannot give the result
An *et al*. (2022)	E-TBNet: Light deep neural network for automatic detection of tuberculosis with x-ray DR imaging	Accuracy of 85.0%	NLM dataset, Shenzhen dataset	Accuracy should be maximized
Sharma *et al*. (2024)	Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images	Accuracy of 99.29%, AUC 99.9%	NIAID TB portal program dataset	This model requires larger or more diverse datasets
Goswami *et al*. (2023)	Deep learning classification of tuberculosis chest x-rays	Accuracy of 94%	Kagel dataset	Additional testing investigations and studies are required to evaluate the model's generalizability early diagnosis and improved control of tuberculosis
Gupta *et al*. (2024)	Direct detection of fluoroquinolone resistance in sputum samples from tuberculosis patients by high resolution melt curve analysis	Sensitivity 90.8%, Specificity 86.9%	archived DNA of 25 MDR from (NJIL and OMD)	Validating techniques on an extensive sample is crucial for evaluating their effectiveness
Singh *et al*. (2024)	Dense net with shark mud ring optimization for severity detection of tuberculosis using sputum image	accuracy of 94.7%, TPR of 93.3%, TNR of 90.6%, PPV of 89.4% and NPV of 88%	Tuberculosis image dataset	Future study might include testing the model's practicality with more datasets
Vats *et al*. (2024)	Incremental learning-based cascaded model for detection and localization of tuberculosis from chest x-ray images	Accuracy 83.32%, F1 81.24%	The golden standard dataset	Requires more data set for future study

This article is going to address the above limitation by using more datasets. The researcher also proposed a model with high accuracy.

Proposed Model

This model is based on CNN. In this proposed model we have taken 3500 datasets (Dey *et al*., 2022) from which 2800 numbers datasets are normal data sets and 700 are TB images. With setting different splitting for training and testing which are divided as 2650 for training purposes, 450 for validating purpose of the prototype and the left are for explicit singlet test if required. These training and testing split allocations are 70-30% respectively. The output layer used the sigmoid activation function combined with the loss function. Learning rates are tuned between 0.001 and 0.0001 with the Adam gradient descent method.

The model consists of 13 layers. The Conv2D layer is the first layer. It takes the chest image input. 26 features are used with a rectifier activation function. The size of the chest image input should be 180×180 pixels. The next layer of Conv2D is the MaxPolling layer. These layers are successive uses up to the 5th layer which is a Conv2D layer. These layers reduce the 26×26 feature matrix to a 3×3 feature matrix. The sixth layer is a flattened layer. It converts the 2D matrix to a vector. From the 7th-13th layer dropout and dense layers are used. Lastly, the model produces the output. The proposed model is given in Fig. (1). The details of each layer are discussed in the following section.

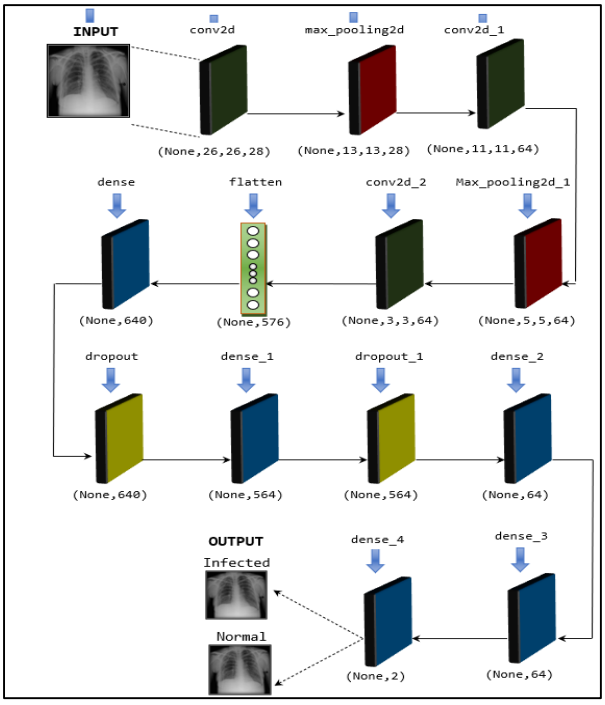

Fig. 1: The proposed model

The convolutional layer 1 applies convolutional operation with 28 filters of size 3×3, resulting in 28 feature maps. None: This dimension is typically used to represent the batch size. In this context, "None" means that the batch size can vary and the network can process multiple samples in parallel during training or inference 26: This dimension represents the height of the output tensor. It indicates that the convolutional operation has reduced the height of the input by 2 pixels. This reduction is common when using a 3×3 convolutional kernel without padding 26: This dimension represents the width of the output tensor. Similar to the height, it indicates that the convolutional operation has reduced the width of the input by 2 pixels 28: This dimension represents the number of channels in the output tensor. It signifies that this convolutional layer has 28 filters and each filter produces a feature map. These feature maps collectively form the output tensor. The convolutional layer applies 28 filters of size 3×3 to the input, resulting in 28 feature maps with dimensions of 26×26 each.

The next layer is a max pooling layer. It Performs max pooling with a pool size of 2×2, reducing the spatial dimensions by half. The "Output Shape: (None, 13, 13 and 28)" for a MaxPooling2D layer signifies the shape of the output tensor produced by that particular pooling layer. None: This dimension is typically used to represent the batch size. In this context, "None" means that the batch size can vary and the network can process multiple samples in parallel during training or inference. 13: This dimension represents the height of the output tensor. It indicates that the max pooling operation has reduced the height of the input by a factor of 2. Max pooling is a down-sampling operation that retains the maximum value within each 2×2 region. 13: This dimension represents the width of the output tensor. Similar to the height, it indicates that the max pooling operation has reduced the width of the input by a factor of 2.28: This dimension represents the number of channels in the output tensor. It signifies that the MaxPooling2D layer is applied independently to each of the 28 channels from the previous layer, resulting in 28 output channels. In summary, the MaxPooling2D layer with an output shape of (None, 13, 13, and 28) performs spatial down-sampling on the input tensor by retaining the maximum value in each 2×2 region for each channel independently. This helps reduce the spatial dimensions of the feature maps, making the network more robust to variations in position and reducing computational complexity.

The convolutional layer 2 d-1 applies convolutional operation with 64 filters of size 3×3, resulting in 64 feature maps. None: This dimension is typically used to represent the batch size. In this context, "None" means that the batch size can vary and the network can process multiple samples in parallel during training or inference.11: This dimension represents the height of the

output tensor. It indicates that the convolutional operation has modified the height of the input feature maps.11: This dimension represents the width of the output tensor. Similar to the height, it indicates that the convolutional operation has modified the width of the input feature maps. 64: This dimension represents the number of channels in the output tensor. It signifies that this Conv2D layer has 64 filters and each filter produces a feature map. These 64 feature maps collectively form the output tensor. The Conv2d-1 layer with an output shape of (None, 11, 11, and 64) applies 64 filters of size, for example, 3×3, to the input feature maps. The convolutional operation involves sliding these filters across the input, producing 64 feature maps, each with spatial dimensions of 11×11. The number of channels is determined by the number of filters applied in the convolution operation.

The MaxPooling2D layer with an output shape of (None, 5, 5, and 64) performs spatial down-sampling on the input tensor by retaining the maximum value in each 2×2 region for each channel independently. This helps reduce the spatial dimensions of the feature maps, making the network more robust to variations in position and reducing computational complexity. The Conv2D layer with an output shape of (None, 3, 3, and 64) applies 64 filters (or kernels) of a certain size (e.g., 3×3) to the input feature maps. The convolutional operation involves sliding these filters across the input, producing 64 feature maps, each with spatial dimensions of 3×3. The number of channels is determined by the number of filters applied in the convolution operation.

Next, the flattening layer with an output shape of (None, 576) takes the output tensor from the preceding convolutional layers and reshapes it into a one-dimensional vector with 576 elements, which is then ready to be fed into the subsequent fully connected layers of the neural network. The dense layer is a fully connected layer where each neuron receives input from every neuron in the previous layer. In this case, the output layer has 640 neurons and the output shape (None, 640) indicates that the network produces a vector with 640 elements for each input sample in the batch. The values in this output vector are the result of weighted combinations and activation functions applied to the features learned by the preceding layers in the neural network. This output is typically used as input to subsequent layers in the network. The dropout layer with an output shape of (None, 640) does not change the shape of the input tensor but introduces randomness by zeroing out a fraction of the input units during training. The values in the output tensor are either the original values or zero, depending on the dropout mask applied during each training iteration. Subsequently, the dense layer is a fully connected layer with 564 neurons. It is followed by the dropout Layer, which applies dropout regularization. Again, the next dense layer is a fully connected layer with 64 neurons followed by another dense layer, which is another fully connected layer with 64 neurons. Finally, the last dense layer is a layer with 2 neurons, representing the final classes. Each layer in the model contributes to the overall processing and learning of features from the input data, ultimately leading to the final classification.

Data Set Preparation

In this section, we are going to discuss how the data set is being prepared. Medical image data set preparation is quite difficult as compared to other image data sets. The CXR data is the alone data which are used in this suggested model.

In this proposed model we have used three openly available data sets chest X-ray masks and labels (Jaeger *et al*., 2014a), Tuberculosis X-ray (TB ×11 K) dataset (Liu *et al*., 2020), and Shenzhen dataset (Jaeger *et al*., 2014b) respectively.

The First data set is chest x-ray masks and label dataset consisting of 4200 CXR images in total of resolution 512*512 pixels of Portable Network Graphics (PNG) file format. Out of which 3500 are normal data and 700 are TB data. The image set is segregated into training and testing data of the ratio 80:20 that is 2800 images are cast off for training and 700 images are cast off for testing purposes. From testing images, 560 images out of 700 images are negative which means not infected and the rest 140 images are infected.

The second data set is the Tuberculosis X-Ray (TB ×11 K) dataset consisting of 11200 CXR images in a total of resolution 512*512 pixels of Portable Network Graphics (PNG) file format. Out of which 5000 are normal data 1200 are TB data and 5000 are non-TB data. The image set from 6200 is segregated into training and testing data of the ratio 80:20 that is 4960 images are cast-off for training and 1240 images are cast off for testing purposes. From testing images, 240 images out of 1240 images are positive which means infected and the rest 1000 images are normal images.

The Third data set is the Shenzhen dataset which consists of 662 CXR images in a total of resolution 3k*3k pixels approximately of Portable Network Graphics (PNG) file format. Out of which 326 are normal data and 336 are TB data. The image set from 662 is segregated into training and testing data of the ratio 80:20 that is 530 images are cast-off for training and 132 images are cast-off for testing purposes. From testing images, 27 images out of 132 images are positive which means infected and the rest 105 images are normal images.

Data preparation (pre-processing) is a common difficulty that researchers encounter as the first step in deep learning techniques, preparing raw data in a way that the machine learning algorithm can accept as input. Deep learning, on the other hand, does not require explicit data

preparation since algorithms may use raw pictures, and patterns are obtained autonomously. The acquired image data comes in a range of sizes. However, the sizes of the obtained photographs have to match the size of the picture input layer used in CNN methods.

As an outcome, we have to standardize (resize) images with the appropriate pixels so as to improve desirable qualities or reduce variables (bias) that impact the prediction models. For example, before training our model, we adjusted the collected images to 512×512 pixels, eliminated unwanted noise from each image, manually cropped some images, performed sequential renaming, and data balancing for classes, and finished other important initial processing tasks.

To solve the issue of varying picture resolutions among the Shenzhen dataset (3×3 k pixels) and the other datasets (512×512 pixels) we have done the downscaling of the image to a common resolution. To be identical to the other dataset's resolution, resize the Shenzhen photos from 3×3 k pixels to 512×512 pixels. This guarantees that all pictures given in the model have identical dimensions, resulting in an improved training process. This is shown in the Fig. (2).

When combining numerous datasets for tuberculosis diagnosis using chest X-rays, it is necessary to harmonize data from various sources. This guarantees that the system can adapt effectively and does not acquire dataset-specific errors. Here's a thorough technique for harmonizing datasets such as chest x-ray masks and labels, TB ×11 K, and Shenzhen.

Augmentation

In this study, we used image data generator methods from the Keras package to generate enough images for the suggested approach using actual photographs. We enhanced the data with image data generator parameters such as rotation range, width shift range, height shift range, zoom range, and others before feeding it into our models.

There are almost 213 unique data and the data is enhanced into around 2113 images, with the help of 5 feature factors, which are horizontally flipped, vertically flipped, gamma Contrast, Sigmoid Contrast, and liner Contrast.

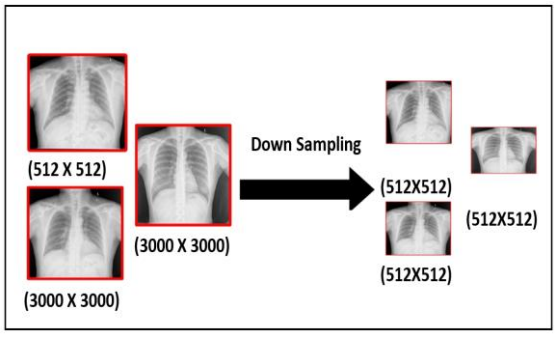

Fig. 2: A down sampling of original images to a uniform size

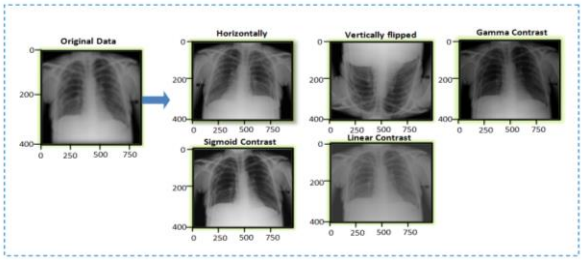

Fig. 3: Demonstration of unique image before augmentation and augmented images using unique image

These features are employed for every individual image in the unique image, so they have yielded more than 2000 images from those unique images and saved them in the same format. The next thing is to upsurge the data set size (amount) with the help of the 10 feature factors, as defined in the above part. Here are the samples of enhanced images of the unique image. The augmented dataset is given in the Fig. (3).

Data Integration

Data integration is a method of merging data from several sources to create a single, consistent perspective. In the framework of machine learning, especially for addressing problems such as TB diagnosis using chest X-rays, integrating data entails combining numerous datasets to form an extensive set of data that may be employed to train and test models. This process guarantees that data from many different places is harmonized, which means that it is structured and presented in a manner that renders it consistent and usable for evaluation or forecasting.

For our model, we combine the datasets such as chest X-ray masks and labels, TB ×11 K, and Shenzhen for Tuberculosis (TB) diagnosis using chest X-rays, the objective is to build a uniform dataset (each image size 180*180 pixels) that enables a machine learning model to learn successfully from disparate data sources.

Results and Discussion

Results and Discussion of Chest X-Ray Masks and Label Dataset, Tuberculosis X-Ray (TB ×11 K) Dataset and Shenzhen Dataset

In this current study, the interpretation of our proposed model is grounded on the various extraction layers used in our model. In our proposed model we have taken three publicly available data sets. In our model, tuberculosis has been predicted from the given datasets and also the performance of the model has been evaluated through a confusion matrix. The confusion matrix of all the three data sets is given in Figs. (4-6).

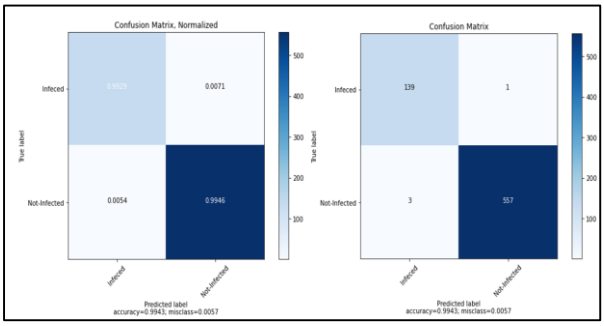

Fig 4: Confusion matrix of chest x-ray masks and label data set

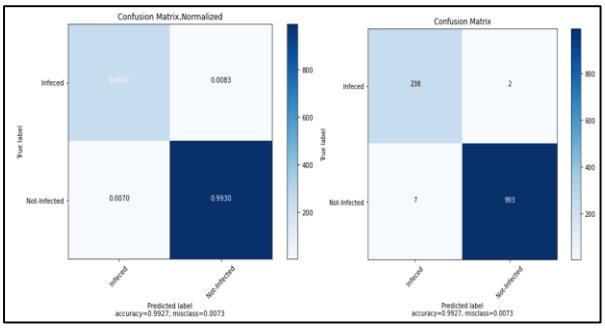

Fig. 5: Confusion matrix of tuberculosis x-ray (TB ×11 K) data set

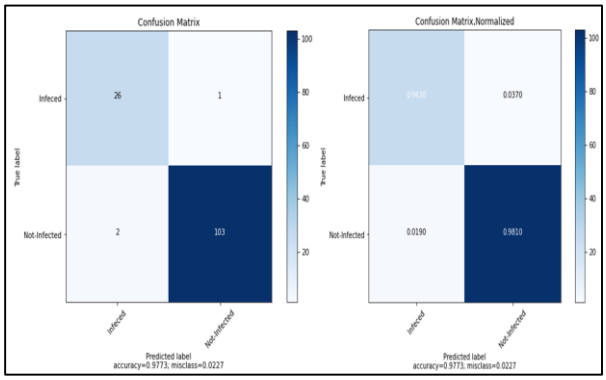

Fig. 6: Confusion matrix of Shenzhen data set

a) Data set: Chest X-ray masks and label
b) Dataset: Tuberculosis x-ray (TB ×11 K) dataset
c) Dataset: Shenzhen

The Table (2) present comparison of our model is compared with some other models. Our model has the highest accuracy of 99.42% in the chest x-ray masks and label data set as the data set is large enough as compared to another dataset.

Table 2: Comparison of the proposed model with the existing model

Author	Dataset	Accuracy %	AUC %	Precision %	F1 score %	Recall %
Rustum *et al*. (2022)	Kaggle dataset	95	96	95.13	91.23	0.896
Mehrrotraa *et al*. (2022)	Tuberculosis x-ray (TB ×11 K) dataset, Shenzhen dataset	99.10	94	88	91	0.873
Miranda *et al*. (2022)	TB ×11 K dataset	92.39	99.23	49	47	0.969
Lu *et al*. (2022)	Fourth hospital of Huai'an	98.93	95	97.86	98.91	0.909
Duong *et al*. (2021)	MC dataset	97.92	92.50	93.24	90.23	0.829
Fati, Suliman Mohamed et al.	NLM dataset	98.5	99.78	99.50	95.2	0.956
	Belarus data set	99.80	99.82	100	97.99	0.923
Sharma *et al*. (2024)	NIAID TB portal program dataset	99.29	99.9	99.30	99.29	0.982
Goswami *et al*. (2023)	Kagel dataset	94	93.4	94.1	94.1	0.934
Singh *et al*. (2024)	Tuberculosis image dataset	94.7	91.95	89.4	91.37	0.933
Vats *et al*. (2024)	The golden standard dataset	83.32	NA	83.57	82.24	0.829
Proposed model	chest x-ray masks and label	99.42	99.78	99.29	98.58	0.979
	Tuberculosis x-ray (TB ×11 K) dataset	99.27	99.56	99.17	98.14	0.971
	Shenzhen dataset	97.73	99.81	97.73	94.55	0.929

The proposed model achieves an accuracy of 99.42% on the chest X-ray masks and label dataset, 99.27% on the Tuberculosis X-ray (TB ×11 K) dataset, and 97.73% on the Shenzhen dataset. These accuracy values are consistently higher than most of the other models listed, except for a few cases (e.g., the model by Mehrrotraa *et al.* with 99.10% on the TB ×11 K dataset and the Belarus dataset with 99.80%). This consistent high performance across different datasets suggests that the proposed model is robust and generalizes well. The proposed model shows very high precision across different datasets: 99.29% on the chest x-ray masks and label dataset, 99.17% on the TB ×11 K dataset, and 97.73% on the Shenzhen dataset. These precision values are among the highest compared to the other models. A high precision indicates that the proposed model effectively minimizes false positives, making it highly reliable when it predicts a positive case.

Additionally, the proposed model exhibits an excellent F1 score (98.58% on the chest x-ray masks and label dataset and 98.14% on the TB ×11 K dataset). The F1 score is a harmonic mean of precision and recall, balancing both false positives and false negatives. The high F1 score of the proposed model reflects its overall effectiveness in accurately identifying cases. The proposed model achieves high recall values, particularly on the chest x-ray masks and label dataset (0.979) and the TB ×11 K dataset (0.971). A high recall indicates the model's effectiveness in identifying all true positive cases, minimizing the number of false negatives. This performance is particularly crucial in medical imaging, where missing a positive case (e.g., a tuberculosis diagnosis) could have severe consequences.

While some other models excel in one or two metrics, the proposed model demonstrates a balanced performance across all key evaluation metrics, including accuracy, precision, F1 score, and recall. This balance suggests that the proposed model is not only precise but also sensitive, making it highly reliable and effective for practical applications.

The statistical analysis of the proposed model is based on the mean value of the existing model and a comparison of our model with the mean value. The mean performance scores of the existing models are, Accuracy-95.72%, AUC-96.16%, Precision-89.92%, F1 score-88.96%, and Recall-0.65. In comparison, the proposed model shows significantly better results across all metrics. For accuracy, the proposed model achieves 99.42, 99.27 and 97.73%, surpassing the mean accuracy of the existing models. The AUC values of 99.78, 99.56, and 99.81% for the proposed model are notably higher than the mean AUC of 96.16%. Similarly, the Precision scores of 99.29, 99.17, and 97.73% are much greater than the average precision of 89.92%. The F1 scores of the proposed model, at 98.58, 98.14, and 94.55%, outperform the mean F1 score of 88.96%. The proposed model's recall values (0.979, 0.971, and 0.929) are higher than the mean recall of the existing models. Overall, these results indicate that the proposed model is superior in performance across all key evaluation metrics.

Conclusion

This article presents a novel CNN-based model for detecting Tuberculosis (TB) by classifying lung X-ray images. The proposed model has been rigorously evaluated on three diverse open datasets of varying sizes: The chest X-ray masks and label dataset, the Tuberculosis X-ray (TB ×11 K) dataset, and the Shenzhen dataset. Compared to ten existing models, our model consistently demonstrates superior performance, achieving an accuracy as high as 99.42% on the chest x-ray masks and label dataset. While several existing models reach an accuracy of around 99%, they have been tested on only a single dataset, limiting their generalizability.

In contrast, our proposed model not only achieves high accuracy but also maintains robust performance across multiple datasets, as evidenced by its high F1 scores and precision rates. This indicates that our model is well-suited to handle different data scales and types effectively. The consistency of our model's results across diverse datasets highlights its potential for broader application in real-world settings.

In future work, we plan to validate the model further using datasets acquired from local hospitals to enhance its applicability and reliability in clinical environments.

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

This research did not receive any funding.

Author’s Contributions

Biswaranjan Debata: Conceptualization, methodology, conducted the experiment(s), writing-original draft preparation

Sudhir Kumar Mohapatra: Conceptualization, methodology, conducted the experiment(s), writing original draft preparation, writing review and edited.

Rojalina Priyadarshini: Data curation, formal analysis, writing review and edited.

Tarikwa Tesfa Bedane: Data curation, formal analysis.

All authors reviewed the manuscript.

Ethics

This research did not involve human participants or animals, and all datasets were used with proper permissions. The study adheres to ethical research practices, with no conflicts of interest declared.

References

An, L., Peng, K., Yang, X., Huang, P., Luo, Y., Feng, P., & Wei, B. (2022). E-TBNet: Light Deep Neural Network for Automatic Detection of Tuberculosis with X-Ray DR Imaging. *Sensors*, *22*(3), 821. https://doi.org/10.3390/s22030821

Debata, B., Mohapatra, S. K., & Priyadarshini, R. (2023). A Systematic Literature Review on Pulmonary Disease Detection Using Machine Learning. *Proceedings of the International Conference on Cognitive and Intelligent Computing*, 515–522. https://doi.org/10.1007/978-981-19-2358-6_48

Dey, S., Roychoudhury, R., Malakar, S., & Sarkar, R. (2022). An Optimized Fuzzy Ensemble of Convolutional Neural Networks for Detecting Tuberculosis from Chest X-Ray Images. *Applied Soft Computing*, *114*, 108094. https://doi.org/10.1016/j.asoc.2021.108094

Division, C. T. B. (2022). *National TB Prevalence Survey India 2019-2021*.

Duong, L. T., Le, N. H., Tran, T. B., Ngo, V. M., & Nguyen, P. T. (2021). Detection of Tuberculosis from Chest X-Ray Images: Boosting the Performance with Vision Transformer and Transfer Learning. *Expert Systems with Applications*, *184*, 115519. https://doi.org/10.1016/j.eswa.2021.115519

Fati, Suliman Mohamed, Ebrahim Mohammed Senan, and Narmine ElHakim. "Deep and hybrid learning technique for early detection of tuberculosis based on X-ray images using feature fusion." Applied Sciences 12.14 (2022): 7092.

Goswami, K. K., Kumar, R., Kumar, R., Reddy, A. J., & Goswami, S. K. (2023). Deep Learning Classification of Tuberculosis Chest X-Rays. *Cureus*, *15*(7), e41583. https://doi.org/10.7759/cureus.41583

Gupta, R. K., Anthwal, D., Bhalla, M., Tyagi, J. S., Choudhary, S., & Haldar, S. (2024). Direct Detection of Fluoroquinolone Resistance in Sputum Samples from Tuberculosis Patients by High Resolution Melt Curve Analysis. *Current Microbiology*, *81*(1), 27. https://doi.org/10.1007/s00284-023-03519-2

Jaeger, S., Candemir, S., Antani, S., Wang, Y.-X. J., Lu, P.-X., & Thoma, G. (2014a). Two Public Chest X-Ray Datasets for Computer-Aided Screening of Pulmonary Diseases. *Quantitative Imaging in Medicine and Surgery*, *4*(6), 475–477. https://doi.org/10.3978/j.issn.2223-4292.2014.11.20

Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., Zhiyun, X., Palaniappan, K., Singh, R. K., Antani, S., Thoma, G., Yi-Xiang, W., Pu-Xuan, L., & McDonald, C. J. (2014b). Automatic Tuberculosis Screening Using Chest Radiographs. *IEEE Transactions on Medical Imaging*, *33*(2), 233–245. https://doi.org/10.1109/tmi.2013.2284099

Junaedi, I., Yudaningtyas, E., & Rahmadwati, R. (2019). Tuberculosis Detection in Chest X-Ray Images Using Optimized Gray Level Co-Occurrence Matrix Features. *2019 International Conference on Information and Communications Technology (ICOIACT)*, 95–99. https://doi.org/10.1109/icoiact46704.2019.8938584

Li, L., Huang, H., & Jin, X. (2018). AE-CNN Classification of Pulmonary Tuberculosis Based on CT Images. *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, 39–42. https://doi.org/10.1109/itme.2018.00020

Liu, Y., Wu, Y.-H., Ban, Y., Wang, H., & Cheng, M.-M. (2020). Rethinking Computer-Aided Tuberculosis Diagnosis. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2643–2652. https://doi.org/10.1109/cvpr42600.2020.00272

Lu, S.-Y., Wang, S.-H., Zhang, X., & Zhang, Y.-D. (2022). TBNet: A Context-Aware Graph Network for Tuberculosis Diagnosis. *Computer Methods and Programs in Biomedicine*, *214*, 106587. https://doi.org/10.1016/j.cmpb.2021.106587

Mehrrotraa, R., Ansari, M. A., Agrawal, R., Tripathi, P., Bin Heyat, M. B., Al-Sarem, M., Muaad, A. Y. M., Nagmeldin, W. A. E., Abdelmaboud, A., & Saeed, F. (2022). Ensembling of Efficient Deep Convolutional Networks and Machine Learning Algorithms for Resource Effective Detection of Tuberculosis Using Thoracic (Chest) Radiography. *IEEE Access*, *10*, 85442–85458. https://doi.org/10.1109/access.2022.3194152

Miranda, F. M., Fajardo, A. C., & Medina, R. P. (2022). Pulmonary Tuberculosis Detection using Digitally Photographed Chest X-RAY Images. *2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T)*, 1–5. https://doi.org/10.1109/icpc2t53885.2022.9776987

Mohapatra, S. K., Assefa, B. G., & Belayneh, G. (2022). A SVM Based Model for COVID Detection Using CXR Image. *Advances of Science and Technology*, *411*, 368–381. https://doi.org/10.1007/978-3-030-93709-6_24

Oltu, B., Guney, S., Dengiz, B., & Agildere, M. (2021). Automated Tuberculosis Detection Using Pre-Trained CNN and SVM. *2021 44th International Conference on Telecommunications and Signal Processing (TSP)*, 92–95. https://doi.org/10.1109/tsp52935.2021.9522644

Rustum, R., Babu, G. C., Praveen, P., Lakshmi, V., Indupalli, T., & Sowjanya, A. (2022). Detection of Tuberculosis using Machine Learning Techniques and Image Preprocessing. *2022 International Conference on Inventive Computation Technologies (ICICT)*, 226–229. https://doi.org/10.1109/icict54344.2022.9850909

Sharma, V., Nillmani, Gupta, S. K., & Shukla, K. K. (2024). Deep Learning Models for Tuberculosis Detection and Infected Region Visualization in Chest X-Ray Images. *Intelligent Medicine*, *4*(2), 104–113. https://doi.org/10.1016/j.imed.2023.06.001

Singh, J., Ramya, R., & Vijay, M. (2024). Dense Net with Shark Mud Ring Optimization for Severity Detection of Tuberculosis Using Sputum Image. *Biomedical Signal Processing and Control*, *91*, 105929. https://doi.org/10.1016/j.bspc.2023.105929

Vats, S., Sharma, V., Singh, K., Katti, A., Mohd Ariffin, M., Nazir Ahmad, M., Ahmadian, A., & Salahshour, S. (2024). Incremental Learning-Based Cascaded Model for Detection and Localization of Tuberculosis from Chest X-Ray Images. *Expert Systems with Applications*, *238*, 122129. https://doi.org/10.1016/j.eswa.2023.122129