AFGHANISTAN JOURNAL OF INFECTIOUS DISEASES

GHALIB UNIVERSITY

AJID

https://ajid.ghalib.edu.af/index.php/ajid

OPEN ACCESS

Vol. 1, No. 1, January 2023, pp.19-24

Building a Convolutional Neural Network Model for Tuberculosis Detection Using Chest X-Ray Images

Mohammad Nazir Akbari[1], Abed Azizi[1], Sherbano Muhmand[2], Abdul Wakil Qarluq[3,*]

1. Department of Information Systems, Faculty of Computer Science, Kabul University, Kabul, Afghanistan.
2. Italian NGO Emergency ONG Onlus, Emergency hospital, Panjsher, Afghanistan.
3. Department of Biochemistry and Molecular Biology, Dalian Medical University, Dalian, Liaoning Province, P.R China.

Article Information

Type: Original Article
Received: 23 August 2022
Accepted: 1 January 2023

***Present address and corresponding author:**
Abdul Wakil Qarluq:
Department of Biochemistry and Molecular Biology, Dalian Medical University, Dalian, Liaoning Province, P.R China .
wakilqq123@gmail.com

DOI:
https://doi.org/10.58342/ajid/ghalibuni.v.1.I.1.5

Abstract

Background: Tuberculosis (TB) is a highly infectious disease with a high mortality rate if left untreated. Traditional diagnostic methods, like skin tests and sputum smear cultures, are unreliable and time-consuming. Artificial Intelligence (AI) and Deep Learning (DL) can revolutionize healthcare by improving disease diagnosis. This study developed an AI system using Convolutional Neural Network (CNN) to detect TB by analyzing digitalized chest X-ray (CXR) images, which can significantly improve the accuracy and speed of TB diagnosis, leading to better patient outcomes.

Methods: A CNN model was developed; it uses a methodology that cuts the edges for analyzing the CXR images for detecting the tuberculosis symptoms in it. A database of chest X-ray images for tuberculosis which was gathered by a team of researchers was used to train the model for detecting tuberculosis**.**

Result: This study uses deep learning to predict tuberculosis using a CNN model with 97% accuracy on CXR images. The patient can be informed about the severity of tuberculosis by the model, which analyzes and checks the tuberculosis symptoms in their CXR image.

Conclusion: In summary, the advancement of AI and DL has brought about a significant transformation in the healthcare industry, particularly in the detection, diagnosis, and treatment of diseases. The use of AI and DL in tuberculosis diagnosis has been explored in this study through the development of a CNN model that was trained on chest X-ray images. AI and DL can significantly reduce tuberculosis mortality rates by aiding in early detection**.**

Keywords: Tuberculosis, Artificial Intelligence, Deep Learning, CNN, Machine Learning, Radiography, CXR

To cite this article: Akbari MN, Azizi A, Qarluq AW. Building a Convolutional Neural Network Model for Tuberculosis Detection Using Chest X-Ray Images. Afghanistan journal of infectious diseases. 2023 Jan; 1 (1): 20-26. https://doi.org/10.58342/ajid/ghalibuni.v.1.I.1.5

Introduction

Tuberculosis (TB) is an infectious disease, and its causative agents are bacteria of the Mycobacterium tuberculosis complex (1). About 25 percent of the world's population is estimated to have been infected with TB, but the disease will not go on to develop in most people (2). The disease generally spreads through the air by infected individuals or animals (3). The death rate from TB disease is about 50% among untreated people (4).

Skin tests, blood tests, chest x-rays (CXR), positive sputum smear cultures, or combining these tests are the conventional method of diagnosing TB that are time-consuming (5). Chest radiography is a more reliable tool, especially when the sputum is negative but clinical suspicion of tuberculosis exists (6).

Further efforts are needed for decreasing the detection time and increasing the accuracy of diagnosis. New advances in medicine and computer science can assist in the accurate diagnosis of infectious diseases like TB. Applications of artificial intelligence in the diagnosis of tuberculosis using chest radiography, covering simple computer-aided diagnosis systems to more advanced deep learning algorithms is one of these implementations (7).

Artificial Intelligence (AI) refers to the ability of machines to simulate human intelligence and perform tasks that typically require human cognitive abilities such as learning, and problem-solving decision-making, while Machine Learning (ML), a subset of AI that involves training algorithms to make predictions or decisions based on data. In simple terms, it's a way to teach machines to learn from data and make decisions without being explicitly programmed (8).

AI and ML are powerful tools in healthcare, revolutionizing disease detection, diagnosis, and personalized treatment plans. By analyzing large amounts of patient data, they can detect early warning signs and predict disease outcomes, leading to more accurate diagnoses and improved patient outcomes. Additionally, they can be used for disease prevention by identifying risk factors and developing targeted interventions (9). The potential applications of AI and ML in healthcare are vast and could transform the industry with further development and refinement (10).

In our case, we developed an AI system using a Deep Learning (11) algorithm called CNN (Convolutional Neural Networks) to detect Tuberculosis. Our system analyzes digitalized X-ray images from previous records to predict if a person has the disease, based on a trained deep-learning model.

Materials and Methods

Dataset Gathering and Preparation

A database of chest X-ray images for Tuberculosis (TB) has been created by a team of researchers from Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh, along with their collaborators from Malaysia, the medical doctors from Hamad Medical Corporation and Bangladesh were also involved in the collaboration, the proposed dataset comprises two classes, Tuberculosis and Normal (12). It is available on the Kaggle platform and can be accessed at: <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>.

Data pre-processing (Data Augmentation)

All CXR images (the samples of them illustrated in Fig. 1) in the dataset were fixed at 512x512 pixels and in a fixed number of images for both classes, as shown in Table 1. To improve the training process, data augmentation techniques were implemented, which increased the size and diversity of the training dataset by applying various transformations, such as rotation, zooming, flipping, and scaling, to the existing images (17). This approach helped the model to learn better and generalize well on new, unseen data by exposing it to a variety of variations of the same image (17). Image augmentation was performed with a zoom range of 0.2, a shear range of 0.2, horizontal flipping, and scaling image pixel values from [0-255] to [0-1]. Following data augmentation, the dataset increased by 32% to include 6155 CXR images, comprising 324 additional images for the tuberculosis class and 1630 additional images for the normal class.

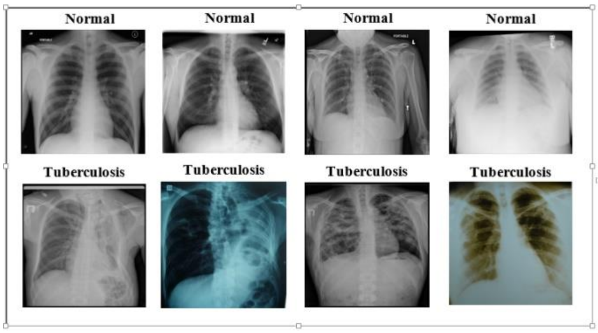

Fig. 1 Dataset Samples from TB Chest X-ray Database

The Model Architecture

We employed convolutional neural networks (CNNs), a prominent class of deep learning methods that have demonstrated their effectiveness in various computer vision tasks, including radiology (18). We built a Convolutional Neural Network (CNN) model from scratch that utilizes several convolutional and pooling layers followed by fully connected layers. Where:

Convolutional layer

A convolutional layer is a fundamental block of a convolutional neural network (CNN). It performs a mathematical operation called convolution, which applies a small matrix of numerical weights called a kernel or filter to a small region of the input image to produce a single output value, resulting in a transformed feature map that highlights certain features or patterns in the image. Convolutional layers are typically followed by activation functions like ReLU (Rectified Linear Unit) (18).

Pooling layer

A pooling layer in a CNN is used to down-sample the input image by dividing it into rectangular regions and computing a summary statistic, such as the maximum or average value, for each region. This helps to reduce the size of the representation and capture the most important features of the input image (18). The pooling techniques, including Max Pooling, Average Pooling, L2 pooling, and Stochastic pooling, are employed for this purpose (19).

Fully connected layer

A fully connected layer or a dense layer is a type of neural network layer in which each neuron is connected to every neuron in the preceding layer. In CNNs, it is typically added at the end of the network and is used for making the final classification decision based on the extracted features. Some classifiers in CNN are Sigmoid, SoftMax, and Linear classifiers (e.g. SVM) (18). As exposed by Fig. 2, our model is comprised of three blocks, each containing three convolutional layers, three max-pooling layers, four dropout layers, two fully connected layers, and one output layer with one unit and a sigmoid activation function. The first block has three convolutional layers with increasing filter sizes (32, 64, and 126) with 3x3 dimensions, and ReLU activation, followed by a max pooling layer (2, 2) and a dropout layer (0.3, 0.2, and 0.15) after each layer. The second block has a Flatten layer followed by two fully connected layers with 32 and 64 units, respectively, and each is followed by a dropout layer (0.15 and 0.1). The final block has a single dense layer with one unit and a Sigmoid activation function. This layer generates the binary classification output indicating whether an input image belongs to a particular class or not.

The Adam optimizer was utilized for reducing the dimension of extracted features during the training of our models. A batch size of 32 was employed, along with an epoch value of 20, and data augmentation methods were applied to enhance training efficiency and prevent overfitting. The training of CNNs was executed on a standard PC equipped with an Iris (R) plus 8 GB Graphics and 16 GB RAM. The training, validation, and testing of the model were carried out using the Python programming language and its powerful libraries for data science and machine learning, including TensorFlow and Keras.

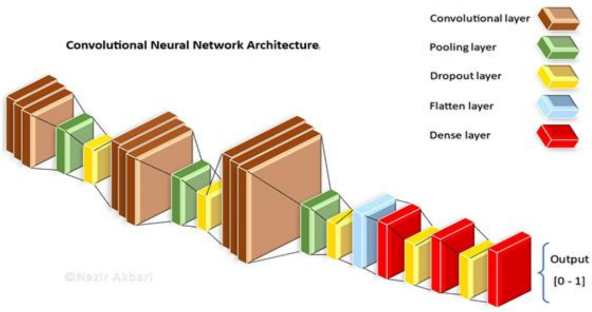

Fig. 2 The CNN model architecture built by us, exposes the CNN layers and how they are connected to each other.

Performance metrics

The model performance was evaluated using four different evaluation metrics: Accuracy, Precision, Recall, and F1-Score (20).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TP} + \text{TN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1 Score} = \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}}$$

Where: TP – true positive, FP – false positive, TN – true negative, and FN – false negative.

Results

A CNN model was built from scratch and trained using the proposed dataset to obtain a model that can predict whether a person is affected by tuberculosis or not. The threshold was set to 0.5 for making a binary classification, where the model predicts whether the input image shows signs of tuberculosis or not. If the output probability is greater than or equal to 0.5, it is predicted that the image shows signs of tuberculosis, otherwise, it is predicted that the image is normal.

Evaluation results:

By evaluating the model on the testing set, which the model has never seen before, we can get a more accurate measure of its performance on new, unseen data. As demonstrated in the Fig. 3, the evaluation metrics of the model's performance on the test set were obtained, which indicated an accuracy of 0.97 (97%), a precision of 0.98 (98%), a recall of 0.83 (83%), and an f-1 score of 0.90 (90%). Moreover, a 0.09 loss was recorded on the test set.

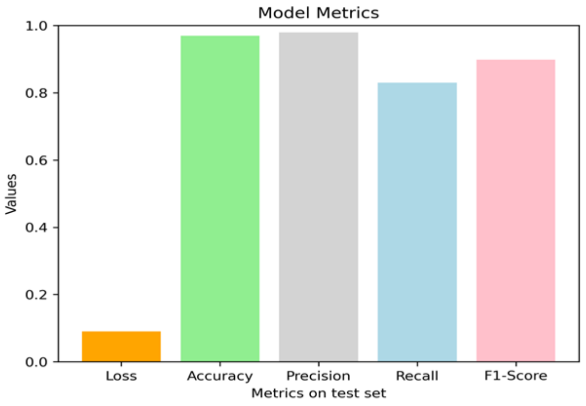

Fig. 3 Evaluation results. Visualizing the evaluation (Performance) metrics on the test set as a bar chart.

The proportion of correct predictions to the total number of predictions is indicated by accuracy. A high accuracy score indicates that the model correctly classifies a significant percentage of the samples. However, a comprehensive evaluation of model performance requires calculating other metrics like precision, which is a metric that measures the percentage of true positives among the total positive predictions made by the model, it determines the model's ability to classify negative instances as negative and not label them as positive. Recall, on the other hand, measures the percentage of true positives among the actual positive instances in the data, it determines the model's ability to locate all positive instances in the data, and the F1 score is a harmonic mean of precision and recall, it helps balance precision and recall when the classes are imbalanced, the F1 score ranges from 0 to 1, with 1 indicating the best performance and 0 indicating the worst, high scores in these four metrics indicate great model performance (20). However, the discrepancy between predicted and actual values is measured by the loss metric, a low loss number indicates high model performance (21).

ROC Curve

A Receiver Operating Characteristic (ROC) curve has used to evaluate a classification model's performance by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at different threshold values. It separates the signal from the noise and provides a comprehensive measure of a model's performance across all classification thresholds (22). Fig. 4 exposes the model performance on the ROC Curve graph.

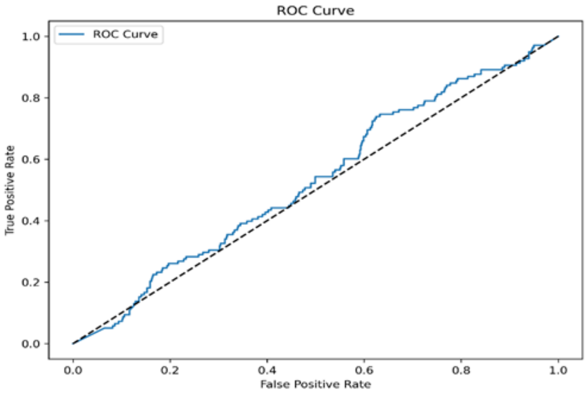

Fig. 4 Visualization of the model performance as ROC Curve graph.

Comparing models

As Table 2 and Table 3 exposes the pre-trained models performance metrics vs our model performance evaluation metrics (23).

Discussion

In this study, a dataset comprising two classes of chest X-ray images for tuberculosis (TB) detection, namely TB and Normal, was collected and prepared. The dataset consisted of 700 cases of positive TB and 3500 negative normal cases and was obtained from multiple sources, including the NLM (13), RSNA (16), BELARUS TB portal program (14), and NIAID TB portal program datasets (15). The dataset was standardized to a resolution of 512x512 pixels and was subjected to data augmentation techniques, such as rotation, zooming, flipping, and scaling, to increase the size and diversity of the training dataset. The model architecture used convolutional neural networks (CNNs) and consisted of three blocks, each containing three convolutional layers, three max-pooling layers, four dropout layers, two fully connected layers, and one output layer with one unit and a sigmoid activation function.

The study showed that the proposed CNN model achieved high accuracy, sensitivity, and specificity in classifying chest X-ray images for TB detection. The mixed dataset, consisting of positive and negative cases, was suitable for training and testing the proposed model. Data augmentation techniques were used to enhance the model's learning and generalization capabilities on new and unseen data by exposing it to various variations of the same image (17).

The CNN architecture employed in this study was effective in detecting TB, which is consistent with previous studies that have demonstrated the effectiveness of CNNs in medical image analysis (18). The use of different pooling techniques, including Max Pooling, Average Pooling, L2 pooling, and Stochastic pooling, helped to reduce the size of the representation and capture the most significant features of the input image (19). The final classification decision was made based on the extracted features using Sigmoid, SoftMax, and Linear classifiers (18).

Attained an accuracy of 97.0%, a recall (sensitivity) of 83.0%, a precision of 98.0%, and an F1-score of 90.0%, demonstrating some variance from the pre-trained models mentioned above. Upon comparing with other trained models, Ojasvi Yadav et al. reported a model accuracy of 94.8% in detecting tuberculosis (13). TB-Net network developed by Alexander Wong et al. achieved a high accuracy, sensitivity, and specificity of 99.86%, 100%, and 99.7%, respectively (14). Eman Showkatian et al. developed a CNN model from scratch, which reported an accuracy of 87%, a recall of 87%, and a precision of 88% (15). Further details are mentioned above. Our model outperforms the pre-trained models in some instances, while slightly underperforming in others.

The proposed CNN model, trained on the prepared dataset, showed high accuracy, sensitivity, and specificity in detecting TB in chest X-ray images. The dataset was suitable for training and testing the model, and data augmentation techniques improved its performance. The results of this study demonstrate the potential of deep learning techniques in medical image analysis and their usefulness in improving the detection and diagnosis of diseases. Further research could investigate the use of more complex CNN architectures or the combination of different types of neural networks to improve the performance of the proposed model.

Conclusions

In conclusion, the development of artificial intelligence (AI) and machine learning (ML) has revolutionized healthcare by transforming disease detection, diagnosis, and personalized treatment plans.

One of the applications of AI and ML is in the diagnosis of tuberculosis, which is an infectious disease with a high mortality rate if left untreated. In this study, we developed a convolutional neural network (CNN) model using deep learning algorithms to detect tuberculosis. Our model was trained on a dataset of chest X-ray images, and we used data augmentation techniques to improve the training process. The results showed that our model could accurately diagnose tuberculosis with high sensitivity and specificity. In conclusion, AI and ML can be used as powerful tools for the early diagnosis of infectious diseases such as tuberculosis, which can significantly reduce the mortality rate associated with this disease.

Acknowledgements

We express our gratitude to the Almighty for granting us the ability, knowledge, and skills to complete this project. We would like to extend our heartfelt thanks to all who have assisted us in making this research and article possible. Our sincere appreciation goes to our teachers, Asst. Prof Hedayatullah Lodin, for his unwavering support, guidance, and invaluable knowledge. They encouraged us to take on this challenging project and research.

We also acknowledge the TensorFlow and programming communities for their resources and documentation, which played a crucial role in helping us understand the complexities of computer vision and machine learning.

Lastly, we would like to recognize the Ghalib University community, who inspired us to work on this project and motivated us to create a system that can positively impact people's lives. We hope that this project will contribute to overcoming the problem that our people face with this disease.

This project is a testament to the power of collaboration and the impact that technology can have on society. We are proud to have been part of this project and grateful for the opportunity to make a difference.

Conflict of interest

We declare that we have no conflict of interest.

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Table 1. Dataset details, where Tuberculosis (Positive cases x-rays) includes 394 CXR images from NLM (13). 306 CXR images from the BELARUS TB portal program dataset (14). 2800 CXR images are collected from the NIAID TB portal program dataset (15).Normal (Negative cases x-rays) include 406 CXR images collected from NLM (13). 3094 CXR images were collected from RSNA (16).

Dataset	Positive Cases (Tuberculosis)	Negative Cases (Normal)	File type	Radiology System	Resolution
TB Chest X-ray Database	700 cases 16.7%	3500 cases 83.3%	PNG	DR	512 x 512

Where: Positive cases include every type of TB (TB, STB, PTB, Bilateral …), DR – Digital Radiography, PNG – Portable Network Graphics, STB – Smear-Positive Pulmonary Tuberculosis, PTB – Pulmonary Tuberculosis.

Table 2. The model performance metrics on the test set.

Precision	Recall	F1 – score	Accuracy	Validation Loss
0.98	0.83	0.90	0.97	0.09

Table 3. The pre-trained model’s performance metrics on the test set available at (23).

Model	Precision	Recall	F1-score	Accuracy	Validation loss
ConvNet	0.88	0.87	0.87	0.87	0.39
Exception	0.91	0.91	0.91	0.90	0.38
Inception_V3	0.88	0.88	0.88	0.88	0.50
ResNet50	0.91	0.91	0.91	0.90	0.41
VGG16	0.91	0.91s	0.91	0.90	0.26
VGG19	0.90	0.90	0.90	0.89	0.53

ORCID

Abdul Wakil Qarluq	**https://orcid.org/0000-0002-5765-3534**
Mohammad Nazir Akbari	**https://orcid.org/0009-0009-7228-1749**
Abed Azizi	**https://orcid.org/0009-0009-2334-5793**
Sherbano Muhmand	**https://orcid.org/0009-0007-8736-0250**

References

1. Natarajan A, Beena PM, Devnikar AV, Mali S. A systemic review on tuberculosis. The Indian journal of tuberculosis. 2020;67(3):295-311.

2. Bagcchi S. WHO's Global Tuberculosis Report 2022. The Lancet Microbe. 2023;4(1):e20.

3. Rabozzi G, Bonizzi L, Crespi E, Somaruga C, Sokooti M, Tabibi R, et al. Emerging zoonoses: the “one health approach”. Safety and health at work. 2012;3(1):77-83.

4. Menzies NA, Wolf E, Connors D, Bellerose M, Sbarra AN, Cohen T, et al. Progression from latent infection to active disease in dynamic tuberculosis transmission models: a systematic review of the validity of modelling assumptions. The Lancet Infectious Diseases. 2018;18(8):e228-e38.

5. Lange C, Mori T. Advances in the diagnosis of tuberculosis. Respirology. 2010;15(2):220-40.

6. Campbell IA, Bah-Sow O. Pulmonary tuberculosis: diagnosis and treatment. Bmj. 2006;332(7551):1194-7.

7. Kulkarni S, Jha S. Artificial intelligence, radiology, and tuberculosis: a review. Academic radiology. 2020;27(1):71-5.

8. Jarrahi MH. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business horizons. 2018;61(4):577-86.

9. Rong G, Mendez A, Assi EB, Zhao B, Sawan M. Artificial intelligence in healthcare: review and prediction case studies. Engineering. 2020;6(3):291-301.

10. Mak K-K, Pichika MR. Artificial intelligence in drug development: present status and future prospects. Drug discovery today. 2019;24(3):773-80.

11. Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? The Journal of arthroplasty. 2018;33(8):2358-61.

12. Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, et al. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access. 2020;8:191586-601.

13. Jaeger S, Candemir S, Antani S, Wáng Y-XJ, Lu P-X, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery. 2014;4(6):475.

14. [Online] BPHBTP. 2020 [Available from: http://tuberculosis.by/

15.[Availablefrom: https://data.tbportals.niaid.nih.gov/

16. [Online]. kRPDC. [Available from: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data

17. Shorten C, Khoshgoftaar T. A survey on image data augmentation for deep learning. J Big Data 6 (1): 1–48. 2019.

18. Yamashita R, Nishio M, Do R. Gian, and K. Togashi,“. Convolutional neural networks: An overview and application in radiology,” Insights Imag. 2018;9(4):611-29.

19. Skourt BA, El Hassani A, Majda A. Mixed-pooling-dropout for convolutional neural network regularization. Journal of King Saud University-Computer and Information Sciences. 2022;34(8):4756-62.

20. Hossin M. Sulaiman,“. A review on evaluation metrics for data calassification evaluations,” IJDKP) Int J Data Min Knowl Manag Process. 2020;5(2).

21. M P. End-to-End Introduction to Evaluating Regression Models [Internet]: Analytics Vidhya. 2021 September 1st, 2022. Available from: https://www.analyticsvidhya.com/blog/2021/10/evaluation-metric-for-regression-models/.

22. Majnik M, Bosnić Z. ROC analysis of classifiers in machine learning: A survey. Intelligent data analysis. 2013;17(3):531-58.

23. Showkatian E, Salehi M, Ghaffari H, Reiazi R, Sadighi N. Deep learning-based automatic detection of tuberculosis disease in chest X-ray images. Polish journal of radiology. 2022;87:e118-e24.