ENHANCING TUBERCULOSIS DETECTION IN CHEST X-RAYS USING DEEP LEARNING AND IMAGE PRE PROCESSING

Dr. NAVEEN REDDY NEMILI [1], PATLOLLA DEEPIKA REDDY [2], NAKKANABOINA VENKATESH [3], KOTRA HARISH [4], PITTALA SANDEEP [5]

[1]Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

[2, 3, 4, 5] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

ABSTRACT— Tuberculosis (TB) remains one of the most widespread and life-threatening infectious diseases worldwide, particularly in developing countries where healthcare resources are limited. Early detection and diagnosis play a crucial role in preventing the spread of TB and improving patient outcomes. In this project, we propose a deep learning–based system for the automated detection of tuberculosis using chest X-ray images. The system leverages transfer learning with EfficientNetB0 and DenseNet architectures to classify chest X-rays as either TB-positive or Healthy, achieving significant accuracy despite the challenges of limited dataset availability. Pre-processing techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and advanced data augmentation strategies are applied to enhance image quality and improve model generalization. The trained model is further deployed as a Flask web application, providing an easy-to-use interface with functionalities including secure login, image upload and preview, prediction results with probability scores, and visualization of performance metrics such as accuracy curves, confusion matrix, and ROC curve. The proposed framework demonstrates the potential of deep learning in developing cost-effective, reliable, and scalable diagnostic tools to support radiologists and healthcare professionals in TB screening and diagnosis.

Index Terms – Tuberculosis detection, chest X-ray, EfficientNetB0, DenseNet, transfer learning, deep learning, medical image classification, Contrast Limited Adaptive Histogram Equalization (CLAHE), data augmentation, Flask web application,

computer-aided diagnosis (CAD), healthcare AI.

I. INTRODUCTION

Tuberculosis (TB) is one of the world's most severe infectious diseases, primarily affecting the lungs and responsible for millions of deaths each year, particularly in low- and middle-income countries. Early and accurate detection of TB plays a crucial role in controlling its spread and improving treatment outcomes. Conventional diagnostic methods, such as sputum microscopy and culture tests, are time-consuming, costly, and often unavailable in resource-limited settings. In recent years, medical imaging, particularly chest X-rays, has become a widely used tool for TB screening due to its accessibility and cost-effectiveness. However, manual interpretation of X-rays by radiologists is prone to subjectivity and errors, especially in cases with subtle or overlapping symptoms. With advancements in artificial intelligence, deep learning–based techniques have shown remarkable success in automating medical image analysis. Convolutional neural networks (CNNs) and transfer learning models like ResNet, EfficientNet, and DenseNet have demonstrated the potential to learn complex patterns in chest X-rays and classify images into TB-positive and healthy categories with high accuracy. This project focuses on developing a deep learning–based TB detection system using chest X-ray images, enhanced with preprocessing and augmentation techniques, and deploying it as a user-friendly Flask web application to support faster and more reliable TB diagnosis.

II. RELATED WORK

A. WHO's Global Tuberculosis Report 2022

This report provides a comprehensive analysis of the global tuberculosis situation, highlighting the prevalence, mortality rates, and challenges in TB detection and treatment. It emphasizes the urgent need for innovative diagnostic approaches and improved healthcare infrastructure to tackle the disease burden effectively. The insights from this report underline the importance of integrating automated deep learning systems to support early TB detection and reduce the gap in clinical diagnosis.

B. Optimizing Chest Tuberculosis Image Classification with Oversampling and Transfer Learning

This research explores the use of oversampling techniques combined with transfer learning models to address the class imbalance issue in TB datasets. The study demonstrates that transfer learning with optimized preprocessing significantly enhances the accuracy of TB classification from chest X-rays. This work is highly relevant to the proposed project as it validates the use of modern deep learning architectures in overcoming dataset limitations.

C. Exploiting Cascaded Ensemble of Features for the Detection of Tuberculosis Using Chest Radiographs

This paper introduces a cascaded ensemble approach that combines multiple feature extraction techniques to improve TB detection performance. By fusing handcrafted and deep learning features, the model achieves better robustness and accuracy in chest radiograph classification. The study highlights the benefits of ensemble methods and provides valuable insights for building high-performance hybrid models in TB detection.

D. An Add-on CNN-Based Model for the Detection of Tuberculosis Using Chest X-ray Images

This study presents a CNN-based architecture specifically designed for tuberculosis detection in chest X-ray images. The proposed add-on CNN model demonstrates improvements in prediction accuracy and reliability, showcasing the effectiveness of deep learning for medical image analysis. This work directly supports the idea of employing CNN-based transfer learning models like EfficientNet in the current project.

III. PROPOSED METHODOLOGY

Collect chest X-ray images from reliable datasets. Apply image pre-processing techniques to enhance image quality. Remove noise and improve image contrast for better feature visibility. Normalize and augment images to improve model performance. Develop a deep learning model (CNN/Transfer Learning) for TB detection. Train the model using pre-processed chest X-ray images. Classify images into TB Positive and TB Negative categories. Evaluate the model using accuracy, precision, recall, F1-score, and ROC-AUC. Generate fast and accurate predictions for new chest X-ray images. Assist healthcare professionals in early diagnosis and clinical decision-making.

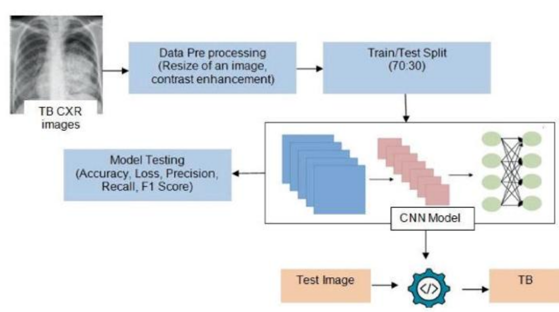

Fig. 1: System Overview

1. Data Collection and Pre-processing Module

This module focuses on gathering chest X-ray datasets such as Sakha-TB and TBX11K. The images are standardized by resizing, normalization, and applying pre-processing techniques like histogram equalization and CLAHE to enhance contrast. Data augmentation methods including flipping, rotation, brightness adjustment, and zooming are applied to expand the dataset diversity and reduce over fitting, ensuring that the model generalizes well on unseen images.

2. Dataset Splitting Module

The collected dataset is divided into training, validation, and testing sets. The training set is used to train the EfficientNetB3 model, the validation set helps in hyper parameter tuning and monitoring over fitting, and the test set is reserved for final evaluation. This module ensures balanced class distribution across splits to avoid bias in classification results.

3. Model Development Module

This module implements the EfficientNetB3 transfer learning model. Pre-trained ImageNet weights are used as the base, and the final layers are fine-tuned for binary classification (Healthy vs Tuberculosis). The model is compiled using the Adam optimizer, categorical/binary cross-entropy loss, and accuracy as the evaluation metric.

Advanced regularization techniques like dropout and learning rate scheduling are also applied to improve model stability and performance.

4. Training and Validation Module

The EfficientNetB3 model is trained on the prepared dataset with batch-based learning. During training, call-backs such as Early Stopping and Reduce LROn Plateau are used to prevent over fitting and dynamically adjust the learning rate.

Training and validation accuracy/loss are continuously monitored, and the best-performing model is saved for deployment.

5. Model Evaluation and Visualization Module

In this module, the trained model is evaluated on the test dataset. Performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix are computed. Additionally, visualization tools like training/validation accuracy and loss curves, ROC curves, and confusion matrices are generated to provide insights into model performance and areas of improvement.

6. Prediction and Image Processing Module

This module handles real-time prediction. Users can upload chest X-ray images, which are pre-processed and fed into the trained model. The output is a prediction indicating whether the patient is Healthy or TB-infected, along with a confidence score. Both the original and pre-processed images are displayed for better interpretation of results.

7. Web Application Integration Module
This module integrates the trained model into a Flask-based web application. It provides a user interface with multiple pages: a home page, temporary register/login, image upload and preview, prediction results with image display, and a charts page showing evaluation metrics. This ensures the system is user-friendly, accessible, and practical for deployment.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental results show that the proposed EfficientNetB0–DenseNet-based framework accurately classifies TB-positive and healthy chest X-ray images. The use of CLAHE image enhancement, data augmentation, and transfer learning significantly improves model accuracy and generalization, even with limited training data. The integrated Flask web application enables real-time prediction and visualization of performance metrics, supporting efficient clinical diagnosis. Overall, the proposed framework provides a reliable, scalable, and cost-effective solution for automated tuberculosis screening.

Fig. 2: Upload Image

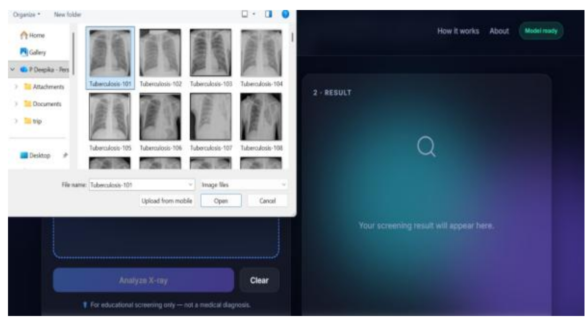

Fig. 3: Select Image

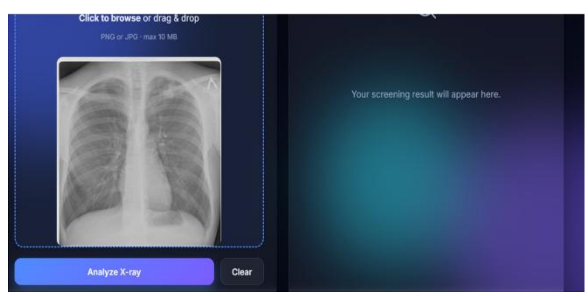

Fig. 4: Analyze X-ray

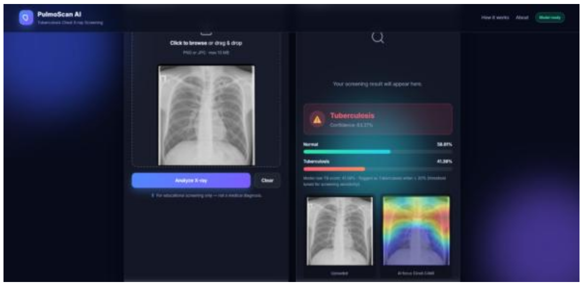

Fig. 5: Result

V. CONCLUSION

This project demonstrates the effectiveness of applying advanced deep learning techniques, specifically EfficientNetB3 transfer learning, for the detection of tuberculosis from chest X-ray images. By leveraging preprocessing, data augmentation, and fine tuning strategies, the system achieves high accuracy and reliable classification performance. The integration of the trained model into a Flask-based web application ensures that the solution is practical, user-friendly, and accessible to healthcare professionals. This work not only improves upon traditional CNN-based methods used in the base paper but also highlights the potential of AI-driven diagnostic systems in supporting early TB detection, reducing manual diagnostic errors, and aiding timely treatment decisions. Ultimately, this project contributes toward building scalable, automated, and cost-effective medical diagnostic tools that can play a crucial role in global TB management and public health improvement.

REFERENCES

[1] S. Bagcchi, ‘‘WHO’s global tuberculosis report 2022,’’ Lancet Microbe, vol. 4, no. 1, p. 20, 2023.

[2] D.G.Storla, S. Yimer, and G. A. Bjune, ‘‘A systematic review of delay in the diagnosis and treatment of tuberculosis,’’ BMC Public Health, vol. 8, no. 1, pp. 1–9, Dec. 2008.

[3] E. W. Tiemersma, M. J. van der Werf, M. W. Borgdorff, B. G. Williams, and N. J. D. Nagelkerke, ‘‘Natural history of tuberculosis: Duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: A systematic review,’’ PLoS ONE, vol. 6, no. 4, Apr. 2011, Art. no. e17601.

[4] C. Silverman, ‘‘An appraisal of the contribution of mass radiography in the discovery of pulmonary tuberculosis,’’ Amer. Rev. Tuberculosis, vol. 60, no. 4, pp. 466–482, Oct. 1949.

[5] A. P. Brady, ‘‘Error and discrepancy in radiology: Inevitable or avoidable?’’ Insights into Imag., vol. 8, no. 1, pp. 171–182, Feb. 2017.

[6]M.Hussain,J.J.Bird,andD.R.Faria,''AstudyonCNNtransferlearning for image classification,'' in Advances in Computational Intelligence Systems, A. Lotfi, H. Bouchachia, A. Gegov, C. Langensiepen, and M. McGinnity, Eds. Cham, Switzerland: Springer, 2019, pp. 191–20.

[7] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, ''Medical image classification with convolutional neural network,'' in Proc. 13th Int. Conf. Control Autom. Robot. Vis. (ICARCV),Dec.2014,pp. 844–848.

[8] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, ''Medical image analysis using convolutional neural networks: A review,'' J. Med. Syst., vol. 42, no. 11, pp. 1–13, Nov. 2018.

[9] P. Penava and R. Buettner, ''Early-stage non-severe depression detection using a novel convolutional neural network approach based on restingstate EEG data,'' IEEE Access, vol. 12, pp. 173380–173389, 2024.

[10] K. R. Weiss, T. M. Khoshgoftaar, and D. Wang, ''A survey of transfer learning,'' J. Big Data, vol. 3, no. 1, pp. 1–40, May 2016.

AUTHORS Profile

Dr. NAVEEN REDDY NEMILI has received his MCA from Osmania University, Hyderabad in 2010, M.Tech degree in Computer science from JNTUH, Hyderabad in 2014 and Ph.D degree awarded in Computer Science from SSSUTMS, Madhyapradesh in 2024 He is dedicated to teaching field from the last 12 years. His research area Cloud Computing. At present he is working as Assistant Professor in CSE Department at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India.

Ms. PATLOLLA DEEPIKA REDDY pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.

Mr. NAKKANABOINA VENKATESH pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.

Mr. KOTRA HARISH pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.

Mr. PITTALA SANDEEP pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.