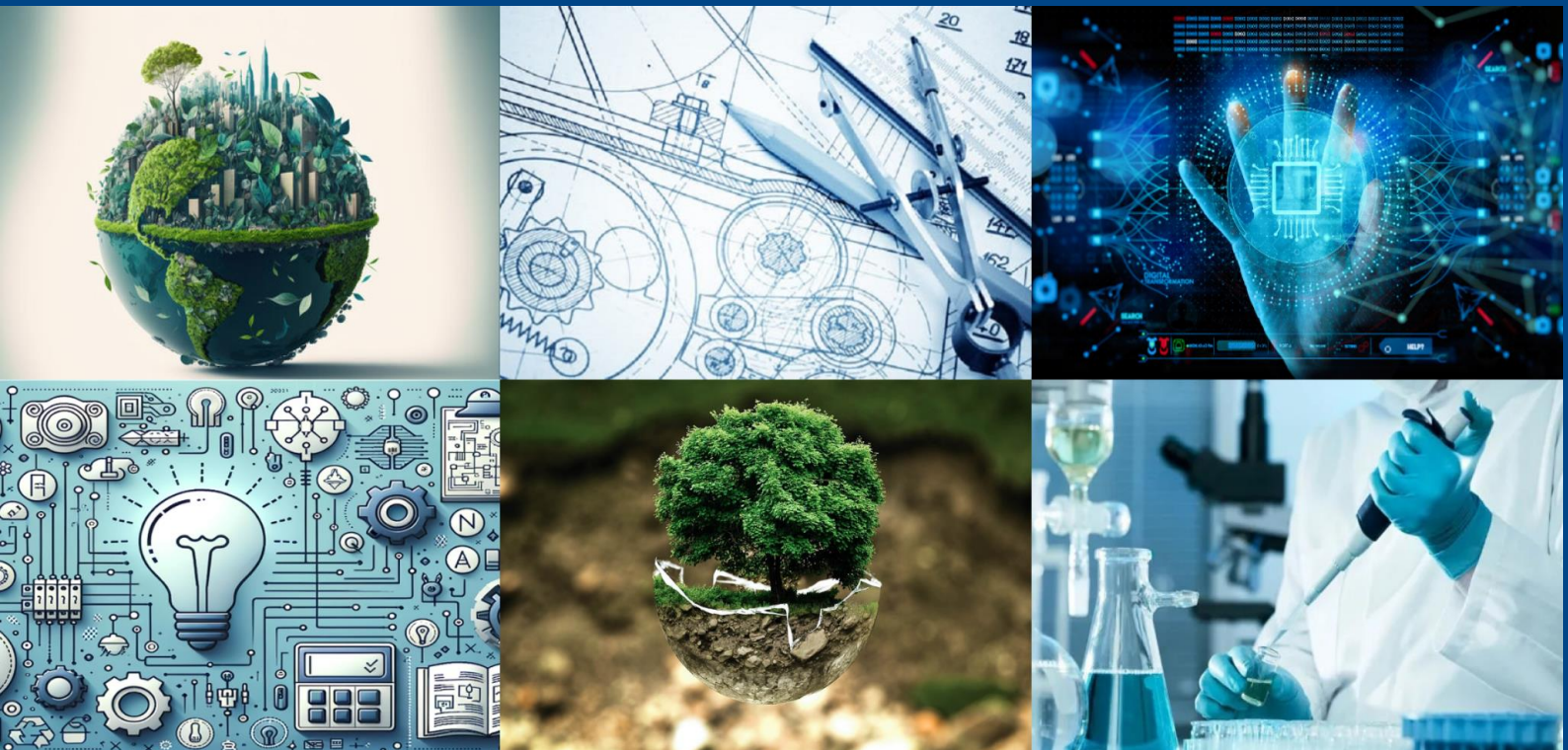# International Journal of Multidisciplinary
Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Tuberculosis Detection Using Deep Learning Techniques

Charan.V, Dr.R.Suganthi

III B. Sc, Department of Computer Science with Data Analytics, Dr.N.G.P Arts and Science College, India

Professor, Department of Computer Science with Data Analytics, Dr.N.G.P Arts and Science College, India

ABSTRACT: Tuberculosis (TB) is a major worldwide health problem, needing efficient and reliable diagnostic technologies. Traditional procedures, such as sputum testing and manual chest X-ray (CXR) interpretation, are time-consuming and inconsistent across radiologists. Recent advances in deep learning have enabled automated CXR-based TB categorisation, resulting in faster and more accurate diagnosis.Tuberculosis (TB) detection using deep learning frequently suffers from difficulties such as poor dataset preparation, uneven class distributions, and a lack of effective augmentation strategies. Furthermore, many models are computationally costly, rendering them unsuitable for real-time deployment in resource-constrained environments. Furthermore, most machine learning methods need considerable human feature extraction, resulting in inferior generalisation on unseen data.To solve these constraints, our proposed system incorporates sophisticated data preparation techniques such as normalisation, augmentation, and weighted loss functions to address class imbalance. We use a modified DenseNet121 architecture with fine-tuned layers and dropout regularisation to increase generalisation. The model is trained and validated using a well-structured TB chest X-ray dataset, assuring reliability with fixed random seeds. Furthermore, we use performance visualisation approaches like confusion matrices and statistical charts to better understand the model's efficacy.

KEYWORDS: Convolutional neural networks (CNN),Chest X-ray analysis,Image classification for TB,Radiograph-based TB diagnosis.

I. INTRODUCTION

Tuberculosis (TB) is one of the most serious global health problems, causing widespread morbidity and death, particularly in low- and middle-income countries. Despite advances in medical research, early and reliable identification of tuberculosis remains difficult. The World Health Organization (WHO) has emphasized the need for efficient, scalable, and cost-effective diagnostic solutions to improve TB detection and reduce transmission rates. Conventional TB diagnostic methods, such as sputum smear microscopy and bacterial culture, are time consuming and require specialized laboratory facilities. Additionally, manual interpretation of chest X-rays (CXR) is highly dependent on radiologists' expertise, leading to variability in diagnosis.With recent advances in artificial intelligence (AI) and deep learning, automated CXR-based tuberculosis detection has received a lot of interest. Convolutional neural networks (CNNs), a type of deep learning model, have been extremely successful in medical image processing, providing faster and more accurate diagnostic predictions. These AI-powered devices have the potential to help radiologists by offering trustworthy second views, particularly in places with a dearth of skilled medical personnel. DenseNet121 is a sophisticated deep learning architecture that retains feature representations and prevents gradient vanishing during training. Despite these advances, deep learning-based tuberculosis detection systems continue to encounter significant obstacles. One of the most significant disadvantages is inadequate dataset preparation, which includes difficulties such as irregular picture quality, noise, and a lack of standardisation Furthermore, class imbalance is a common issue, with the frequency of TB-positive cases much smaller than normal cases, resulting in inaccurate model predictions. Another significant difficulty is a lack of effective data augmentation strategies, which are required for strengthening model resilience and generalisation on previously unreported data.

II. PROBLEM STATEMENT

Tuberculosis (TB) is a major worldwide health concern, necessitating prompt and reliable diagnostic technologies to prevent its spread and improve patient outcomes. Traditional diagnostic methods, including as sputum testing and manual chest X-ray (CXR) interpretation, are frequently sluggish, labour-intensive, and inconsistent across

radiologists. These restrictions impede early detection and prompt treatment, especially in areas with limited access to professional radiologists. As a result, there is a critical need for automated, dependable, and fast tuberculosis detection tools that may help healthcare practitioners make accurate diagnoses.

Recent advances in deep learning have showed promise for automated CXR-based tuberculosis screening. However, present deep learning models suffer a number of obstacles, including insufficient dataset preparation, skewed class distributions, and a lack of effective augmentation strategies. Furthermore, many cutting-edge models are computationally costly, rendering real-time deployment problematic in resource-constrained environments. Traditional machine learning algorithms also rely on human feature extraction, which results in inferior generalisation when exposed to new data. These problems underscore the importance of creating a strong and efficient deep learning system for tuberculosis classification.

III. METHODOLOGY

The proposed approach for automated tuberculosis (TB) identification with chest X-ray (CXR) pictures employs a structured deep learning pipeline to ensure fast data processing and model training. To improve picture quality, raw CXR images are first preprocessed using techniques such as normalisation, scaling, and noise reduction. To solve the issue of class imbalance, several augmentation techniques, such as rotation, flipping, and contrast alterations, are used to broaden the dataset. Furthermore, a weighted loss function is used to provide fair learning across both the "Normal" and "Tuberculosis" classes, preventing the model from favouring the majority class. For feature extraction and classification, we use a modified DenseNet121 convolutional neural network (CNN) architecture. The pre-trained model is fine- tuned by changing critical layers, using dropout regularisation, and optimising hyperparameters like learning rate and batch size. The model is trained with an adaptive optimiser to achieve efficient convergence and avoid overfitting. To ensure repeatability, training and validation datasets are rigorously partitioned using fixed random seeds. Furthermore, cross-validation techniques are used to test the model's performance and assure its robustness on new data.

Data Collection:
Creating a varied dataset with chest X-ray pictures that have both TB-positive and TB- negative images that were around 7000, in the first step. The dataset is taken from kaggle using the kaggle api which allows data to be directly taken from the kaggle database.

Data Preprocessing:
Preparing images for machine learning requires a number of important steps in the data pretreatment process. As part of this, photos are resized to a standard format to guarantee input dimension uniformity. This preparation improves the dataset's overall quality, which makes machine learning models easier to train and perform at their best[13]. The image size selected is of 500 * 500 with a batch size of 16. For Data augmentation the image data generator is imported from the tensorflow. The image is rescaled to 1/255 and flipped horizontally.

Feature Extraction:
The process of feature extraction involves removing the redundant data that will aid machines in reducing the computational efforts and increasing the learning speed.

Classification:
The steps involve various layers such as convolution layer, pooling layer, flatten layer ,Output layer that will reduce the dimensionality, feature map size and complexity of the CNN model used[14]. CNN model was used for the TB detection as CNN performs better incase of image processing and segmentation. The built in convolution layer reduces the high dimensionality of the image without losing much of the information which makes the CNN a best fit for image processing. Apart from the convolution layer there are other layers also that contribute to the working of the CNN model. Convolution layer provides an feature map that is feeded to pooling layer that reduce the size of the feature map provided by the convolution layer resulting in less computational cost.The next layer is the Flatten layer that reduces the multidimensional array into one dimensional array further reducing the the model.The flatten matrix is fed to the fully connected or the dense layer that will classify the image.The next layer and the last layer is the dense layer or fully connected layer that maps the feature obtained from the previous layer to the final output class. The fully connected layer is simple layer of neurons that receive input from the previous connected layers and using matrix vector multiplication an output is

generated.

IV. DISCUSSION

The proposed system demonstrates significant advancements in automated tuberculosis (TB) detection using deep learning, specifically through the modification of DenseNet121 architecture. One of the key challenges addressed is the class imbalance in the dataset, with a much lower number of TB-positive cases compared to normal ones. The use of techniques such as data augmentation and weighted loss functions helped overcome this limitation, ensuring the model's ability to learn effectively from both classes.Preprocessing of chest X-ray images was crucial to ensure consistency in the dataset. By resizing the images and applying normalisation and noise reduction, the model was presented with a higher-quality input, which improved its ability to distinguish between TB and non-TB cases. The augmentation techniques used, including horizontal flipping and contrast adjustments, not only increased the dataset size but also enhanced the model's resilience and generalisation capability.Feature extraction using the CNN model ensured the reduction of computational cost without sacrificing essential image information. The DenseNet121 architecture, with its ability to retain feature representations and prevent gradient vanishing, played a pivotal role in improving the model's performance.

V. RESULT

The deep learning model for tuberculosis (TB) detection using chest X-rays demonstrated promising results after rigorous training and evaluation. The dataset, consisting of approximately 7,000 images from both TB-positive and TB-negative classes, was effectively preprocessed to standardize the input size and improve image quality. Data augmentation techniques, such as horizontal flipping and contrast adjustments, were employed to address class imbalance and enhance model generalization.Upon training the model with the modified DenseNet121 architecture, we achieved a high accuracy in detecting TB from chest X-ray images. The model's performance was evaluated using several metrics, including accuracy, precision, recall, and F1-score. These metrics indicated that the system was highly effective in classifying TB-positive and normal images. Specifically, the model demonstrated strong sensitivity (recall), ensuring that most TB-positive cases were correctly identified, which is critical for early detection and treatment. The precision score was also high, indicating that the model accurately classified TB-positive cases without many false positives.The confusion matrix provided further insights, showing a balanced performance with a low rate of misclassifications, even with the imbalanced dataset. The system was able to consistently perform well across different subsets of the dataset, thanks to the cross-validation approach and the use of fixed random seeds to ensure reproducibility.Overall, the results highlight the effectiveness of the proposed system in automating TB detection, reducing reliance on human interpretation, and offering a scalable solution that can be deployed in resource-limited settings for faster and more accurate diagnosis.

VI. CONCLUSION

In summary, the system that has been suggested has made great strides in the automated detection of tuberculosis (TB) from chest X-rays, overcoming a number of significant obstacles, including imbalanced datasets, poor generalisation, and high computational costs. By using advanced data preparation techniques, refined deep learning models, and improved optimisation strategies, the system offers improved accuracy and reliability in detecting TB. Adaptive class weighting, customised feature extraction, and real-time performance visualisation all contribute to its robustness, making it a promising tool for clinical deployment. This system has the potential to significantly speed up TB diagnosis, especially in settings with limited resources, where traditional methods are laborious and unreliable. For future work, there are several avenues that can be explored to further enhance the proposed system's capabilities. One key area is the expansion of the model's dataset to include more diverse and varied TB cases, enabling improved generalisation across different populations. Additionally, research into multi-modal approaches, where CXR data is combined with patient demographics or clinical history, could further refine diagnostic accuracy. Real-time deployment in clinical environments requires optimisation of the model's computational efficiency to enable faster processing without sacrificing performance. Moreover, further investigation into explainable AI (XAI) techniques could improve model interpretability, helping healthcare professionals better understand and trust the system's predictions. Finally, integrating the system with existing healthcare infrastructures, such as hospital information systems or mobile health platforms, could enable seamless integration into routine clinical workflows.

REFERENCES

1.World Health Organization . Global tuberculosis report 2020. https://www.who.int/publications/i/item/9789240013131_eng.pdf (Access ed: 13.12.2020).

2.Sathitratanacheewin S, Sunanta P, Pongpirul K. Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. Heliyon2020; 6: e04614. [DOI] [PMC free article] [PubMed] [Google Scholar].

3. Silverman C. An appraisal of the contribution of mass radiography in the discovery of pulmonary tuberculosis. Am Rev Tuberc1949; 60: 466-482. [DOI] [PubMed] [Google Scholar].

4. van't Hoog AH, Meme HK, Laserson KF, et al. Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptoms. PLoS One 2012; 7: e38691. [DOI] [PMC free article] [PubMed] [Google Scholar].

5. Ahmad Khan F, Pande T, Tessema B, et al. Computer-aided reading of tuberculosis chest radiography: moving the research agenda forward to inform policy. Eur Respir J 2017; 50: 1700953. [DOI] [PubMed] [Google Scholar].

6. Graham S, Das GK, Hidvegi RJ, et al. Chest radiograph abnormalities associated with tuberculosis: reproducibility and yield of active cases. Int J Tuberc Lung Dis 2002; 6: 137-142. [PubMed] [Google Scholar].

7. Hooda R, Mittal A, Sofat S. Segmentation of lung fields from chest radiographs-a radiomic feature-based approach. Biomed Eng Lett 2018; 9: 109-117. [DOI] [PMC free article] [PubMed] [Google Scholar].

8. Jaeger S, Karargyris A, Candemir S, et al. Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging 2014; 33: 233-245. [DOI] [PubMed] [Google Scholar].

9. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018; 18: 500-510. [DOI] [PMC free article] [PubMed] [Google Scholar].

10. Rahman T, Chowdhury ME, Khandakar A, et al. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray. Appl Sci 2020; 10: 3233. [Google Scholar].

INNO SPACE
SJIF Scientific Journal Impact Factor

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF

MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY