International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

EMPOWERING TUBERCULOSIS DIAGNOSIS THROUGH CHEST X-RAY IMAGING AND ADVANCED AI TECHNIQUES

Jagapathi Unnathi[1], Vojjala Kalyan Prasad[2], Bollam Stephen[3], Mr. G Sathish[4]
[1,2,3] UG Scholar, Dept. of IT, St. Martin's Engineering College, Secundrabad, Telangana, India, 500100
[4] Assistant Professor, Dept. of IT, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
unnathiii09@gmail.com

Abstract:

Tuberculosis (TB) remains a significant public health issue, particularly in India, which accounts for a large proportion of global TB cases. According to the World Health Organization (WHO), India reported over 2.6 million new TB cases annually, with an estimated 440,000 deaths each year. Chest X-ray imaging combined with advanced AI techniques can accelerate TB diagnosis and improve accuracy, making it a promising tool in public health. AI applications help in automating detection, ensuring early intervention, and supporting healthcare workers, especially in resource-limited areas. To leverage deep learning and AI-based approaches for improving the accuracy, speed, and scalability of tuberculosis diagnosis through chest X-ray imaging, thereby supporting India's goal to eliminate TB by 2025. Before the integration of AI, TB diagnosis relied on manual interpretation of chest X-rays by radiologists, sputum smear microscopy, and the Mantoux tuberculin skin test. These methods, though effective, are time-consuming, less precise, and prone to human error, particularly in regions with limited radiological expertise. Traditional TB diagnostic methods face challenges due to the reliance on skilled radiologists, time-intensive procedures, and limited accessibility in rural areas. The complexity and resource demands of traditional diagnostic methods create a pressing need for automated, scalable, and cost-effective diagnostic solutions. Leveraging AI can bridge this gap by facilitating early, reliable diagnosis, reducing workloads for healthcare providers, and supporting India's ambitious TB elimination target. Using Convolutional Neural Networks (CNN) and advanced AI algorithms, the proposed system aims to automate TB detection from chest X-ray images. By training models on vast datasets of labeled X-rays, AI can recognize patterns indicative of TB with high accuracy, aiding radiologists and reducing diagnostic times.

Keywords : Tuberculosis, AI-enhanced diagnosis, Deep learning, Chest X-ray imaging, Convolutional Neural Networks (CNNs), Public health, Automated detection, Healthcare.

1. INTRODUCTION

Tuberculosis (TB) is one of the oldest known infectious diseases, affecting millions worldwide, with India bearing a disproportionately high burden. As of recent WHO statistics, India reports over 2.6 million TB cases each year, resulting in approximately 440,000 deaths annually. TB diagnosis has historically relied on manual techniques, such as sputum smear microscopy, X-ray imaging, and tuberculin skin testing, which are effective but often labor-intensive, time-consuming, and prone to human error. Chest X-ray imaging combined with AI-based detection systems has emerged as a transformative tool, aiming to automate diagnosis, enhance detection accuracy, and reduce dependency on skilled radiologists, especially in resource-limited areas. AI applications have shown potential in processing large datasets to recognize TB-related patterns, enabling early intervention and more efficient TB control.

Despite extensive TB control measures, India remains one of the most affected countries by TB. Traditional diagnostic methods are challenging to scale, expensive, and slow, making rapid intervention difficult in high-burden regions. There is a critical need for a solution that enables automated, accurate, and scalable TB diagnosis accessible to rural and urban populations alike. AI-based TB diagnostic systems offer the potential to address these needs, reducing the workload for medical professionals and aligning with India's target of eliminating TB by 2025. AI can also ensure more consistent diagnosis, particularly valuable in areas with limited radiology expertise.

Before the advent of machine learning, TB diagnosis faced multiple limitations, including high reliance on experienced radiologists to interpret chest X-rays manually, often leading to inconsistencies and potential errors.

In addition, conventional tests such as sputum smear microscopy and tuberculin skin tests were time-intensive, and their diagnostic accuracy could vary. Rural and remote regions struggled with access to qualified medical personnel, contributing to delays in diagnosis and treatment. These challenges limited timely and accurate TB detection, causing prolonged transmission of the disease and increased mortality rates

2. LITERATURE SURVEY

1) Machine Learning Approaches in Diagnosing Tuberculosis Through Biomarkers, 2023
Authors:Vimala Balakrishnan, Yousra Kherabi, Ghayathri Ramanathan, Scott Arjay Paul, Chiong Kian Tiong

This articlehas proposed machine learning approaches in diagnosing tuberculosis through biomarkers. This review examines 19 studies on biomarker-based tuberculosis (TB) diagnosis using machine learning, following PRISMA guidelines. Supervised learning, particularly with Support Vector Machine and Random Forest algorithms, yielded high accuracy (97.0%), sensitivity (99.2%), and specificity (98.0%). Protein-based biomarkers were predominantly explored, with publicly available datasets commonly used.

2) A Tuberculosis Detection Method Using Attention and Sparse R-CNN, 2022

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

Authors : Xuebin Xu, Jiada Zhang, Xiaorui Cheng, Longbin Lu, Yuqing Zhao, Zongyu Xu, Zhuangzhuang Gu

The authors has proposed a tuberculosis detection method using attention and sparse RCNN.In this project the algorithm comprises CXTCNet for TB presence judgment, employing a channel attention mechanism within DenseNet, and CXTDNet, a Sparse R-CNN-based design for TB area detection without non-maximal suppression post-processing. Preprocessing involves CLAHE for noise reduction. On TBX11K dataset, CXTCNet achieves 99.10% accuracy, outperforming existing TB classification methods, while overall detection performance yields AP of 45.35% and AP50 of 74.20%. Additionally, a new dataset of 304 chest X-rays demonstrates diagnostic accuracy comparable to radiologists, aiming to advance TB detection efforts.

3) Exhaled Human Breath Analysis in Active Pulmonary TB Diagnosis by Comprehensive Gas Chromatography -Mass Spectrometry and Chemometric Techniques, 2018
Authors : Marco Beccaria , Carly Bobak, Boitumelo Maitshotlo, Theodore R Mellors, Giorgia Purcaro, Flavio A Franchina, Christiaan A Rees, Mavra Nasir, Wendy S Stevens, Lesley E Scott, Andrew Black, Jane E Hill

This article proposedcomprehensive gas chromatography -mass spectrometry and chemometric techniques for exhaled human breath analysis in active pulmonary TB diagnosis. This study explores using human breath analysis for diagnosing active TB, offering a non-invasive approach with machine learning algorithms achieving sensitivities of 0.82 and 1.00 and specificities of 0.92 and 0.60 in training and test data respectively, showing promise especially for patients with HIV comorbidities.

4) Diagnosis of Tuberculosis Using Colorimetric Gold Nanoparticles on a Paper-Based Analytical Device, 2017
Authors : Tsung-Ting Tsai , Chia-Yu Huang , Chung-An Chen , Shu-Wei Shen , Mei-Chia Wang , Chao-Min Cheng , Chien-Fu Chen

The research diagnosis of tuberculosis using colorimetric gold nanoparticles on a paper -based analytical device. In this project a colorimetric sensing strategy utilizing gold nanoparticles and paper-based platforms was developed for TB diagnosis, leveraging surface plasmon resonance to detect changes in nanoparticle aggregation due to DNA hybridization. This label-free technique, adaptable to resource-limited settings and compatible with smartphone measurement, achieved a detection limit of 1.95 × 10-2 ng/mL for TB DNA with rapid parallel results and low reagent consumption.

5) A Potential Method for TB Detection using Chest Radiography, 2017
Authors : Rahul Hooda, Sanjeev Sofat, Simranpreet Kaur, Ajay Mittal, Fabrice Meriaudeau

This research a potential method for TB detection using chest radiography. The paper proposes a deep learning-based method for TB detection from chest X-ray (CXR) images, achieving a 94.73% overall accuracy with the Adam optimizer, showcasing potential for early diagnosis and disease containment. Evaluation is conducted on Montgomery and Shenzhen datasets, signaling promise for computer-aided diagnosis systems in improving TB detection and treatment outcomes.

6) Automatic Detection of Mycobacterium TB using Artificial Intelligence
Authors : Yan Xiong, Xiaojun Ba, Ao Hou, Kaiwen Zhang, Longsen Chen, Ting Li

This article proposed a method of automatic detection of mycobacterium TB using artificial intelligence. This study developed a convolutional neural network (CNN) called TB-AI to detect acid-fast stained TB bacilli, achieving high sensitivity (97.94%) and specificity (83.65%) compared to diagnoses by human pathologists. TB-AI shows promise as a supportive tool in TB diagnosis, potentially reducing pathologists' workload and improving detection accuracy.

7) TB Diagnostics and Localization in Chest X-Rays Via Deep Learning Models
Authors : Ruihua Guo, Kalpdrum Passi, Chakresh Kumar Jain

This articledeveloped a model on TB diagnostics and localization in chest x-rays via deep learning models.The research proposes an integrated approach using convolutional neural networks (CNNs) to enhance TB diagnosis in chest X-rays, including model modification, fine-tuning, and ensemble methods, with class activation mapping for localization. This approach demonstrates superior performance in detecting lung abnormalities and diagnosing specific TB-related manifestations.

8) TB detection in chest radiography using CNN architecture and explainable AI, 2022
Authors : Saad I Nafisah , Ghulam Muhammad

This article has proposed TB detection in chest radiography using CNN architecture and explainable AI. This project proposes an automatic tuberculosis (TB) detection system utilizing advanceddeep learning models, incorporating sophisticated segmentation networks to extract the region of interest from chest X-rays (CXRs) and employing explainable artificial intelligence for visualization. Experimentation reveals that using segmented lung CXR images significantly enhances classification performance, with EfficientNetB3 achieving the highest accuracy of 99.1% among the tested convolutional neural network (CNN) models.

9) Image Enhancement for TB Detection Using Deep Learning
Authors : Khairul Munadi, Kahlil Muchtar, Novi Maulina, Biswajeet Pradhan

This article has developed an imahe enhancement method for TB detection using deeplearning. This study investigates the impact of image enhancement techniques, such as Unsharp Masking, High-Frequency Emphasis Filtering, and Contrast Limited Adaptive Histogram Equalization, on deep learning-based TB diagnosis from chest x-ray images. Employing these enhancements alongside pre-trained models like ResNet and EfficientNet, the study achieved notable classification accuracy (89.92%) and AUC scores (94.8%) on the Shenzhen dataset, offering promising avenues for improving TB diagnostic performance.

3. PROPOSED METHODOLOGY

To overcome the limitations of traditional image captioning methods, a deep learning-based automated caption generation system is proposed. This system leverages Convolutional Neural Networks (CNNs) for image feature extraction and Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, for natural language generation. It integrates Natural Language Processing (NLP) techniques to generate context-aware, fluent, and meaningful image captions.

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

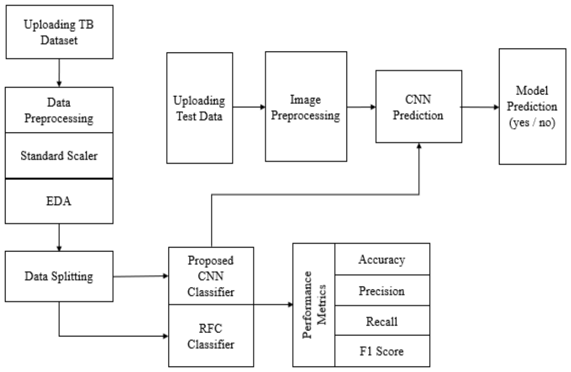


Fig 1: System Architecture

The proposed system uses machine learning and deep learning, particularly Convolutional Neural Networks (CNNs), to identify and classify TB from chest X-ray images. In this approach, CNN models are trained on large datasets of labeled chest X-ray images, allowing them to learn patterns associated with TB. With AI models, the system can provide real-time, accurate diagnostics, reducing the diagnostic load on radiologists. Research studies, such as those using deep learning architectures like ResNet or DenseNet for medical imaging, have shown high accuracy in X-ray-based TB classification, supporting AI's role in early and accessible TB diagnostics.

Dataset

The dataset consists of medical images categorized into two classes: Tumor (TB) and Normal. The dataset includes a total of 4200 images, with 700 images labeled as Tumor and 3500 labeled as Normal. This dataset serves as the foundation for building a binary classification model where the goal is to distinguish between tumor and normal conditions in medical scans.

Dataset Pre processing

Before training the model, the dataset undergoes preprocessing to ensure the quality of the input data. This step includes removing any **null values** that could potentially skew the model's learning process. Additionally, **label encoding** is performed to convert the categorical labels (Tumor, Normal) into numerical values, making them suitable for model training. Image resizing and normalization might also be applied to standardize the images for better model performance.

Proposed Algorithm - Convolutional Neural Network (CNN)

The proposed algorithm is a **Convolutional Neural Network (CNN)**, which is specifically designed for image classification tasks. CNNs are capable of automatically extracting features from images and are highly effective in medical image analysis. The model will consist of several convolutional layers, pooling layers, and fully connected layers to learn the spatial hierarchies and features of the medical images. This architecture is expected to outperform traditional machine learning methods, such as Random Forest, in terms of accuracy.

Performance Comparison

Once both the Random Forest classifier and the CNN model have been trained, their performance will be compared. The evaluation will focus on metrics like **accuracy**, **precision**, **recall**, and **F1-score**. It is anticipated that the **CNN model will yield higher accuracy** in detecting tumors compared to the Random Forest classifier, as CNNs are tailored for handling image data and capturing complex patterns within medical images. This comparison will demonstrate the effectiveness of CNNs over traditional methods in the context of tumor detection in medical images.

4. EXPERIMENTAL ANALYSIS

Figure 2 shows the GUI of website

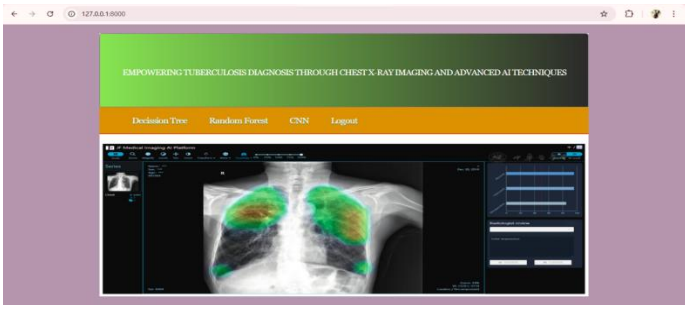


Figure 2: GUI

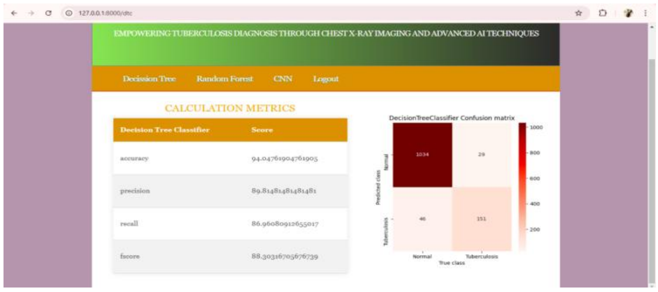
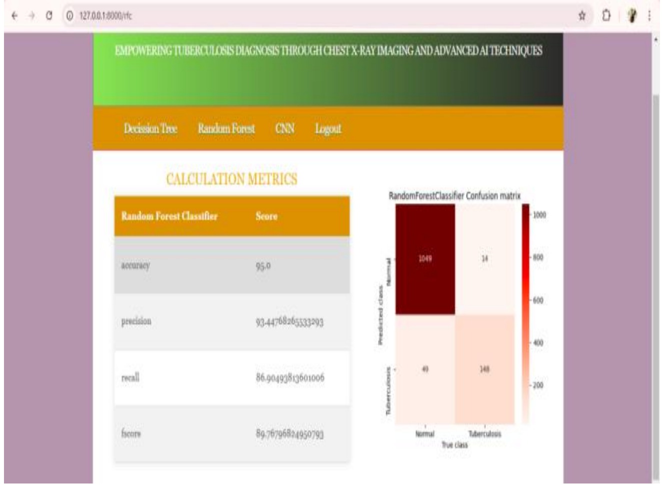


Figure 3: Performance Evaluation of DT

Figure 3 shows that The Decision Tree Classifier achieved an accuracy of **94.05%**, indicating that it correctly classified the majority of instances. Its **precision of 89.81%** suggests that when it predicts a positive class, it is correct most of the time, reducing false positives. The **recall of 86.96%** shows that it successfully identifies a high proportion of actual positive cases, though some may be missed. The **F1-score of 88.30%** balances precision and recall, reflecting the overall effectiveness of the model in classification tasks. These metrics suggest that the Decision Tree Classifier performs well, but there may be room for improvement in handling false positives and false negatives.



International journal of imaging science and engineering
ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

Figure 4: Performance Evaluation of RFC

Figure 4 shows that The **Random Forest Classifier** achieved an **accuracy of 95.0%**, indicating a strong overall performance in correctly classifying cases. Its **precision of 93.45%** suggests that when it predicts a positive case (Tuberculosis), it is correct most of the time, reducing false positives. The **recall of 86.90%** shows that it successfully identifies a high proportion of actual positive cases, though some may still be missed. The **F1-score of 89.77%** balances precision and recall, confirming its effectiveness.

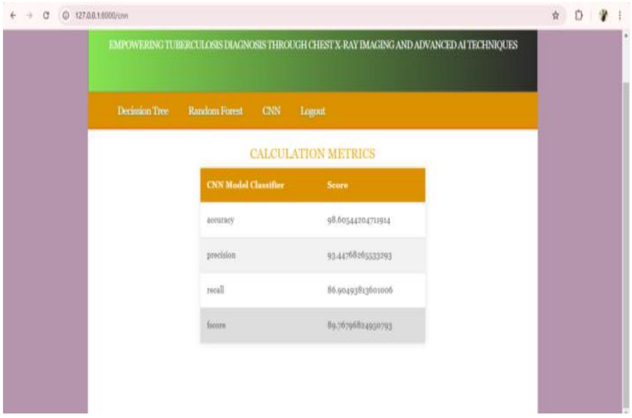


Figure 5: Performance Evaluation of CNN

Figure 5 shows that the CNN Model Classifier achieved an accuracy of 98.61%, significantly outperforming both the Decision Tree and Random Forest classifiers. This high accuracy suggests that the model is highly effective at distinguishing between Normal and Tuberculosis cases. The precision of 93.45% indicates that when the model predicts a positive case (Tuberculosis), it is correct most of the time, minimizing false positives. The recall of 86.90% shows that it successfully identifies a large proportion of actual Tuberculosis cases, though some may still be missed. The F1-score of 89.77% balances precision and recall, confirming its strong classification performance.

Figure 6: Predicted Output

Figure 6 shows that the X-ray image is normal.

5. CONCLUSION

In conclusion, the implemented image classification pipeline effectively demonstrates the use of machine learning and deep learning techniques for classifying images into predefined categories. The process involves several key steps: dataset loading, image preprocessing, feature extraction, model training, and evaluation.

By utilizing different classifiers like Decision Tree, Random Forest, and Convolutional Neural Networks (CNN), the model provides flexibility and allows for a comparison of performance metrics such as accuracy, precision, recall, and F1 score.

The CNN model, trained with a custom architecture, significantly enhances the model's ability to learn spatial features from images, providing superior results for image-based tasks. Additionally, saving and loading models using joblib and pickle facilitates easy model deployment and reuse.

The use of confusion matrices and classification reports ensures transparent evaluation of the model's performance.The future scope of this image classification project is vast and can be expanded in multiple directions. One area of improvement is the integration of more advanced deep learning models, such as ResNet, Inception, or EfficientNet, to improve classification accuracy, especially for complex and diverse image datasets.

Additionally, the model can benefit from data augmentation techniques to increase the size and variability of the training dataset, reducing overfitting and improving generalization. Another avenue for improvement is the optimization of the model for deployment in resource-constrained environments.

This can be achieved through model compression techniques like pruning or quantization, which would reduce the model's size and inference time. Furthermore, integrating the model into real-time applications, such as medical imaging platforms or surveillance systems, could lead to tangible societal impacts.

REFERENCES

[1] Vimala Balakrishnan, Yousra Kherabi, Ghayathri Ramanathan, Scott Arjay Paul, Chiong Kian Tiong, "Machine Learning Approaches in Diagnosing Tuberculosis Through Biomarkers", Progress in Biophysics and Molecular Biology, Volume 179, May 2023, Pages 16-25.

[2] Xuebin Xu, Jiada Zhang, Xiaorui Cheng, Longbin Lu, Yuqing Zhao, Zongyu Xu, Zhuangzhuang Gu, "A Tuberculosis Detection Method Using Attention and Sparse R-CNN", Vol. 16, No. 7, July 31, 2022 doi>10.3837/tiis.2022.07.001.

[3] Marco Beccaria , Carly Bobak, Boitumelo Maitshotlo, Theodore R Mellors, Giorgia Purcaro, Flavio A Franchina, Christiaan A Rees, Mavra Nasir, Wendy S Stevens, Lesley E Scott, Andrew Black, Jane E Hill, "Exhaled human breath analysis in active pulmonary TB diagnosis by comprehensive gas chromatography -mass spectrometry and chemometric techniques", PMID: 30394364 PMCID: PMC6537098 DOI: 10.1088/1752-7163/aae80e.

[4] Tsung-Ting Tsai , Chia-Yu Huang , Chung-An Chen , Shu-Wei Shen , Mei-Chia Wang , Chao-Min Cheng , Chien-Fu Chen, "Diagnosis of Tuberculosis Using Colorimetric Gold Nanoparticles on a Paper-Based Analytical Device", PMID: 28901134 DOI: 10.1021/acssensors.7b00450.

[5] Rahul Hooda; Sanjeev Sofat; Simranpreet Kaur; Ajay Mittal; Fabrice Meriaudeau, " A Potential Method for TB Detection using Chest Radiography", 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), DOI: 10.1109/ICSIPA.2017.8120663.

[6] Fei Deng, Jibing Huang, Xiaoling Yuan, Chao Cheng, Lanjing Zhang "Performance and Efficiency of Machine Learning Algorithms for Analyzing rectangular biomedical data", Laboratory Investigation volume 101, pages430–441 (2021)

[7] Yan Xiong, Xiaojun Ba, Ao Hou, Kaiwen Zhang, Longsen Chen, Ting Li , “Automatic Detection of Mycobacterium TB using Artificial Intelligence”, PMCID: PMC5906344 PMID: 29707349 doi: 10.21037/jtd.2018.01.91.

[8] Ruihua Guo, Kalpdrum Passi, Chakresh Kumar Jain, “TB Diagnostics and Localization in Chest XRays Via Deep Learning Models”, PMCID: PMC7861240 PMID: 33733221 doi: 10.3389/frai.2020.583427

[9] Khairul Munadi, Kahlil Muchtar, Novi Maulina, Biswajeet Pradhan, “Image Enhancement for TB Detection Using Deep Learning”, Page(s): 217897 – 217907 Electronic ISSN: 2169-3536 DOI: 10.1109/ACCESS.2020.3041867

[10] Huy, V.T.Q.; Lin, C.-M. An improved densenet deep neural network model for tuberculosis detection using chest x-ray images. IEEE Access 2023, 11, 42839–42849. [Google Scholar] [CrossRef]

[11] Le, T.; Thien, B.; Tat, N.; Ngo, V.M. Automated evaluation of tuberculosis using deep neural networks. Eur. Union Digit. Libr. 2022, 9, e4. [Google Scholar] [CrossRef]

[12] Visuña, L.; Yang, D.; Garcia-Blas, J.; Carretero, J. Computer-aided diagnostic for classifying chest X-ray images using deep ensemble learning. BMC Med. Imaging 2022, 22, 178.

[13] Goswami, K.K.; Kumar, R.; Kumar, R.; Reddy, A.J.; Goswami, S.K. Deep learning classification of tuberculosis chest x-rays. Cureus 2023, 15, e41583.

[14] Sharma, V.; Nillmani; Gupta, S.K.; Shukla, K.K. Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images. Intell. Med. 2023, 4, 104–113.

[15] Maheswari, B.U.; Sam, D.; Mittal, N.; Sharma, A.; Kaur, S.; Askar, S.S.; Abouhawwash, M. Explainable deep-neural-network supported scheme for tuberculosis detection from chest radiographs. BMC Med. Imaging 2024, 24, 32.

[16] Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology. Aug 2017;284(2):574-582.

[17] Ruihua Guo, Kalpdrum Passi, Chakresh Kumar Jain, “TB Diagnostics and Localization in Chest XRays Via Deep Learning Models”, PMCID: PMC7861240 PMID: 33733221 doi: 10.3389/frai.2020.583427

[18] Khairul Munadi, Kahlil Muchtar, Novi Maulina, Biswajeet Pradhan, “Image Enhancement for TB Detection Using Deep Learning”, Page(s): 217897 – 217907 Electronic ISSN: 2169-3536 DOI: 10.1109/ACCESS.2020.3041867

[19] Le, T.; Thien, B.; Tat, N.; Ngo, V.M. Automated evaluation of tuberculosis using deep neural networks. Eur. Union Digit. Libr. 2022, 9, e4. [Google Scholar] [CrossRef]

[20] Visuña, L.; Yang, D.; Garcia-Blas, J.; Carretero, J. Computer-aided diagnostic for classifying chest X-ray images using deep ensemble learning. BMC Med. Imaging 2022, 22, 178.

[21] Goswami, K.K.; Kumar, R.; Kumar, R.; Reddy, A.J.; Goswami, S.K. Deep learning classification of tuberculosis chest x-rays. Cureus 2023, 15, e41583.