IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Diagnosis of Pulmonary Tuberculosis using Multi-Modal Deep Learning

Chaitanya Jwala Vegesna
Department of IT
S. R. K. R Engineering College
Bhimavaram, India

Krishna Chaitanya K
Department of IT
S. R. K. R Engineering College
Bhimavaram, India

Rajeswari M
Department of IT
S. R. K. R Engineering College
Bhimavaram, India

Jahnavi Lakshmi P
Department of IT
S. R. K. R Engineering College
Bhimavaram, India

Sofian Pitta
Department of IT
S. R. K. R Engineering College
Bhimavaram, India

Abstract-*Pulmonary tuberculosis is a bacterial infection caused by Mycobacterium tuberculosis bacteria (MTB). It primarily affects the lungs but also affect other parts of the body. The disease spreads from one person to another person through the air. In some areas where there aren't many specialized doctors and medical resources, finding pulmonary TB in its early stages is crucial. The idea is to teach the computer to recognize TB by using various multimodal deep learning techniques like VGG16, VGG19, and Xception to help quickly and accurately identify TB in places with limited medical resources. This paper how artificial intelligence and chest X-ray images can work together to identify whether a person has TB or not.*

***Index Terms*—** X-ray image, Deep learning, VGG16, VGG19, Xecption, Healthcare Accessibility, Medical Images.

I. INTRODUCTION

Mycobacterium tuberculosis bacteria (MTB) is the primary cause of tuberculosis, a bacterial infection illness that mainly affects the lungs but can potentially spread to other regions of the body [1]. Pulmonary tuberculosis is a kind of tuberculosis [2] that mostly affects the lungs. When an infected individual coughs, sneezes, or speaks, the disease spreads to the next person by airborne droplets.

Traditionally, pulmonary TB can be identified through symptoms, diagnostic tests and chest X-ray images. Symptoms include a cough last over 3 weeks, chest pain, fever, coughing up blood, weight loss, night sweats. Diagnostic testing and chest X-ray images are then used to confirm these symptoms. However, it can take a lot of time to analyse the chest X-ray images, particularly in places with little access to medical care.

Hence, there are some multi modal deep learning techniques like VGG16 [3], VGG19 [4], and Xception [5] that can be used to analyse the chest X-ray images and predict whether the person has TB or not based upon the different chest X-images and provide better care for patients.

II. LITERATURESURVEY

To predict Diagnosis of pulmonary tuberculosis so far so many multi modal deep learning algorithms were used and all the work done till now has improved the efficiency of the problem.

Malik et al. [6] introduced a Multi-modal DL method that classifies the chest diseases using several imaging and cough sounds. The proposed DL models also focused on classification of chest diseases very accurately. The main drawback of this approach is complexity and computational cost associated with multi-modal integration.

Agarwal et al. [7] presented the classification ML-MF of DNN for Chest X-Ray Images. The proposed approach is also called as MultiFusionNet that classifies the abnormal conditions in the input images. The main drawback of the proposed approach is more complex based on the fusion that may increase computational cost and model complexity.

Ahmed et al. [8] proposed the new model that detects the tuberculosis from Chest X-ray Images. The proposed approach addresses the need for simultaneous detection of multiple lung diseases. While the study obtains the low performance which lacks detailed discussion or interpretation of these results.

Priya et al. [9] proposed the Multi Modal Smart Diagnosis (MMSD) of Pulmonary Diseases. The MMSD is the integration of multi-modal data for pulmonary disease diagnosis. Their drawbacks complexity in data integration, pre-processing, and alignment.
Wang et al. [10] presented the Enabling chronic obstructive pulmonary disease (COPD) diagnosis using chest X-rays. COPD was diagnosed using chest X-rays from numerous sites and modalities, which increased the generalisability and robustness of the results. The paper's disadvantages include potential problems in data consistency.

Ahmed et al. [11] described the Multimodal stacking machine learning (MS-ML) technique for the patient of COVID-19 mortality risk based on chest X-ray pictures and clinical data. The proposed approach, BIO_CXRNET, is a powerful MS-ML technique for predicting COVID-19 mortality risk. The proposed approach MS-ML integrates the chest X-ray pictures with clinical data to improve prediction accuracy. The drawback of this paper is limited explanation or validation of the stacking ML technique may hinder reproducibility and understanding by other researchers. Potential challenges in data integration, preprocessing, and feature selection across multiple modalities may affect model performance and generalizability.

Oloko-Oba et al. [12] discussed several DL techniques for Tuberculosis Detection from Chest Radiograph. The study followed the PRISMA guidelines, resulting in a rigorous and transparent method to identifying relevant material. The Demerits of this paper is he systematic review focused exclusively on CAD systems developed between 2017 and 2021 for the classification of TB using DL models from chest X-ray images.

Karki et al. [13] investigated the challenge of detecting drug-resistant tuberculosis using chest X-rays. It overcomes the extension concerns in detecting drug-resistant tuberculosis using chest X-rays. The disadvantages include a lack of complete solutions or tactics for overcoming extension issues.

Dasanayaka et al. [14] described the DL models for Screening of TB Using Chest X-rays. The DL pipeline achieved an impressive classification accuracy of 97.1% indicating it’s potential as an effective tool for tuberculosis screening using chest X-rays.

Li et al. [15] indicated a system for diagnosing spinal TB from CT images using DL and multimodal feature fusion (MM-FF). The proposed DL-based methodology automates the diagnosis of spinal TB using CT images. Their drawbacks are complexity of MM-FF methods may require extensive computational resources and optimization efforts.

Lee et al. [16] proposed DL based on the activity of pulmonary tuberculosis on chest radiographs. The DL model was highly accurate and consistent in detecting active pulmonary tuberculosis on chest radiographs. While the deep learning model showed performance in the tested cohorts and clinical settings may be limited.

Norval et al. [17] proposed the Image Processing approach for the detection of Pulmonary Tuberculosis Detection. The paper systematically investigates four factors affecting the precision of detecting pulmonary tuberculosis based on patient’s chest X-ray (CXR) images using Convolutional Neural Networks (CNNs). The study finding may have limited generalizability due to the focus on a specific set of factors and datasets.

Angappan, K[18] their paper titled Prediction of multi-lung disease. Addresses the prediction of multi-lung disease, which is a significant and clinically relevant problem. The drawback of the paper lack of specificity regarding the methodology employed for prediction, potentially hindering reproducibility and understanding.

Khan et al. [19] proposed the DL based approach that detects and diagnosis the accurate TB model.
Rahman et al. [20] introduced the DL based segmentation and visualization. The utilization of multiple deep CNN, image pre-processing techniques, and segmentation models increases the complexity and resource requirements of the proposed method. Implementation and maintenance of such a system may be challenging.

Guo et al. [21] described new DL based model to diagnose and localize the tuberculosis in chest X-rays. The suggested approach uses CNN and DL models to provide an effective detection of abnormal regions.

Ma et al. [22] proposed using DL to detect the active pulmonary TB in multi-stage CT images. These models may require a considerable quantity of labeled data to train, which can be time-consuming and resource-intensive.

Munadi et al. [23] presented the improved images for the detection of tuberculosis using DL. DL approaches, notably CNNs, when paired with image enhancement algorithms, have showed promise in enhancing the accuracy and efficacy of tuberculosis identification from chest X-ray pictures. Despite reaching excellent classification accuracy and AUC ratings, the study may have limitations due to the large dataset used.

Hussain et al. [24] described an automated diagnosis system for lung cancer detection using multimodal feature extraction techniques. One potential disadvantage of the proposed approach is the lack of explicit explanation or confirmation of the selected feature extraction procedure.

III. MULTI-MODALDEEP LEARNING MODELS

A Multi-Modal DL is a type of ML where AI is supposed to train itself on relating between various modalities of information multimodal kind of Deep Learning, whose core aim involves using AI and Deep Learning principles in processing both different kinds of information. They involve combing audios, images, videos, texts, etc.

Classification of models: Multimodal deep learning models are classified based on how different types of data or modalities are processed, fused, and learned to perform a task. These classifications are essential for understanding the architectural design choices, data processing techniques, and task-specific considerations in multimodal systems. The following modals were used in our paper to improve the efficiency of diagnosis.

A. ***VGG16***- It is a DCNN architecture which is belongs to DL algorithms. It is more popular and high performance in computer vision applications such as image categorisation and object detection. This was proposed by Karen Simonyan and Andrew Zisserman in 2014. This model has 16 layers, 13 convolutional and 3 fully connected layers (FCL), organised in a sequential architecture with small convolutional filters and max-pooling layers. This architecture allows the model to learn hierarchical visual information efficiently, making it more robust and accurate. It gained popularity after winning the ImageNet Large Scale Visual Recognition Challenge, which required categorising images into 1,000 separate groups and properly detecting items from 200 classes. This enhanced its reputation as a robust and adaptable deep-learning model. The below Fig. 1 describes the architecture of VGG16.

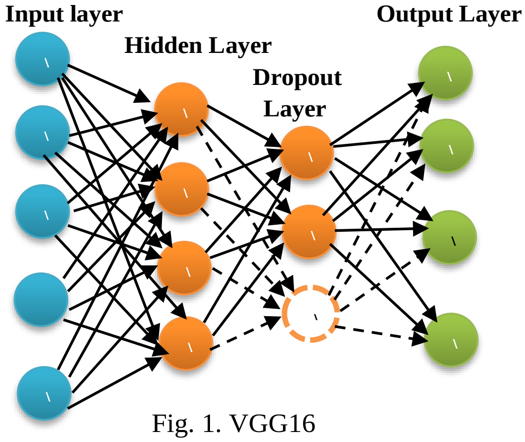

Fig. 1. VGG16

B. ***VGG19***- It is a DCNN of 19 weighted layers, includes 16 convolutional layers and three FCLs. It uses a simple and repeated approach, making it easier to grasp and apply. To maintain spatial resolution, the convolutional layers use 3x3 filters with a stride of 1 and a padding of 1. To induce nonlinearity, no ReLU activation function follows each convolutional layer. The max polling layers reduce the spatial dimensions of the feature maps by a factor of two and a filter size of 2x2. Finally, three fully connected layers are employed for classification, followed by a final softmax layer to output class probabilities. The below Fig. 2 shows the representation of VGG19 workflow.

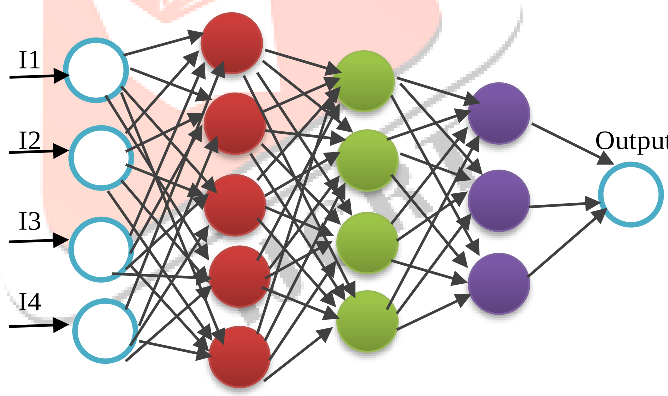

Fig.2. VGG19

C. ***Xception***- Xception stands for Extreme Inception. It has been designed by Francois Chollet in 2017. That particular network, an even higher yet version of Inception, is based on this idea that is depth-wise separable convolutions, which, for each of the inputs distinct input channels, determine a single fill. One by the other sequentially, depth-wise convolutions come first, following the one described thoroughly in a previous section by pointwise convolutions, which perform linear combinations of the results. This stripping out of the standard convolution means that computation gets decreased by a high percentage, and in the same way, such good model performance is sustained. Xception comprises 36 layers of convolution that are organised around entry, middle, and exit flows, making it particularly efficient and effective for tasks of computer vision, such as object recognition and detection. By its very simplicity, it captures a lot about the layout and channel-wise characteristics of the data, which makes the method very powerfully designed. The below Fig. 3 depicts the workflow of Xception.

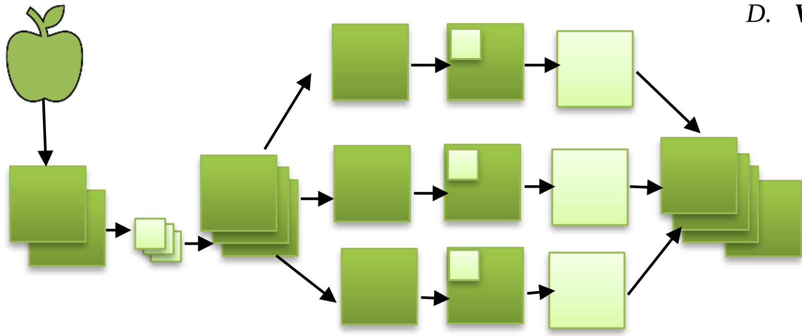

Fig. 3. Xception

IV. METHODOLOGY

A. *Chest X-Ray Analysis (Image-Based Detection):*

- **Image Preprocessing**: Basic image enhancement methods such as noise reduction, contrast adjustment, and resizing will be applied to chest X-rays (CXR) to improve image quality.
- **Feature Extraction**: Simple Convolutional Neural Network (CNN) models, such as VGG16 or MobileNet, will be used to identify key features in X-rays that indicate TB (e.g., lung lesions, cavities).
- **Classification**: The model will classify the X-ray into "TB Positive" or "TB Negative" based on the extracted features.

B. *Integration of Clinical Data*

- **Use of Basic Patient Information:** Clinical data such as age, gender, symptoms (e.g., fever, cough duration), and medical history will be collected in a simple table format.
- **Data Analysis**: A simple machine learning model like Logistic Regression, Decision Tree, or Support Vector Machine (SVM), would analyze the clinical data to predict TB likelihood.

C. *System Workflow*

- The system works by analyzing two key pieces of information, those are: a chest X-ray image and patient's clinical details, such as symptoms, age etc.
- The system examines the chest X-ray using a CNN model and processes the patient's clinical details with a basic machine learning algorithm to help with the analysis.
- The system combines the results from both analyses to make final diagnosis, identifying whether the patient has TB(Positive) or not (Negative).
- The system will give a diagnosis with confidence score shows how certain the result.

D. *Validation and Testing*

- **Dataset**: Public TB chest X-ray datasets, like NIH or Shenzhen datasets, will be utilized for training and testing
- **Performance Metrics**: The system is tested for its accuracy, sensitivity, and specificity in order to produce accurate outcomes.

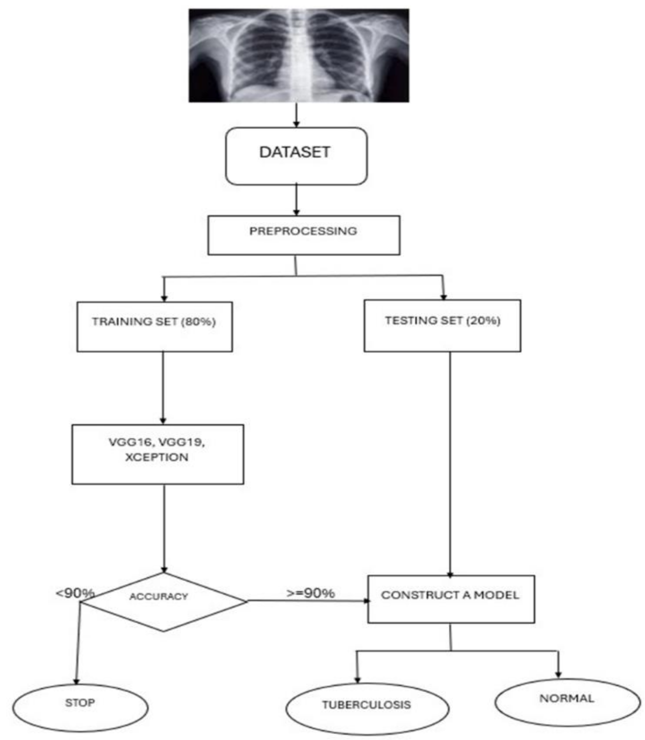

Fig. 4. System Architecture

This Fig. 4 represents the architecture of a deep learning model designed to detect tuberculosis from chest X-ray images. Here's a simple explanation of how it works:

1)Dataset Collection:

The process starts with a dataset of chest X-ray images, which includes both tuberculosis-infected and normal images.

2)Preprocessing:

The images undergo preprocessing, which may include resizing, normalization, and noise reduction to make them suitable for model training.

3)Splitting the Data:

The dataset is divided into two parts:

Training Set (80%): Used to train the model.

Testing Set (20%): Used to evaluate the model's performance.

4)Model Selection:

Different deep learning models, such as **VGG16, VGG19, and**

Xception, are used for training. These are pre-trained convolutional neural networks (CNNs) known for their effectiveness in image classification.

5)Accuracy Evaluation:

The trained model is tested, and its accuracy is measured.

If the accuracy is **less than 90%**, the process stops, indicating t the needs improvement. If the accuracy is **90% or higher**, the m model is considered reliable.

6)Model Prediction:

The final trained model classifies new chest X-ray images into

Tuberculosis (if the X-ray indicates infection).

Normal (if the X-ray shows no signs of tuberculosis).

This architecture ensures a structured approach to developing an effective tuberculosis detection system using deep learning.

The dataset includes 1,700 chest X-ray images, normal and tuberculosis positive, which were sourced from publicly accessible data storage centers such as Kaggle and NIH. The datasets are split into 80% training and 20% testing sets. Resizing, pixel normalization, and removal of noise are performed in the images for preprocessing purposes to ensure consistency and improve model performance. Using convolutional neural networks (CNNs), the system identifies patterns indicative of tuberculosis in X-ray images. Clinical data, such as patient symptoms and history, is also integrated through fully connected neural networks to enhance diagnostic accuracy. The model is trained on the combined data, validated using the test set, and fine-tuned for optimal performance. Once validated, the system predicts and classifies new user-submitted cases, providing healthcare professionals with accurate and timely TB diagnosis.

V. RESULTS AND DISCUSSION

The results obtained after using the VGG16, VGG19, and Xception models were analyzed on the dataset. The comparison was based on accuracy and validation accuracy achieved by each model over several epochs.

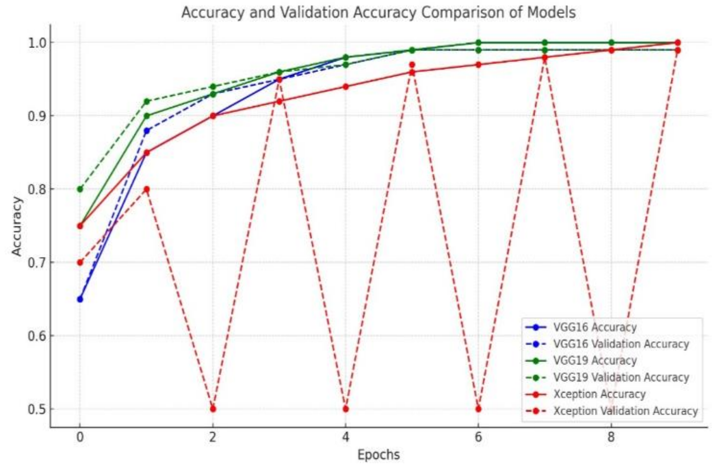

Fig. 5. Comparison graph based on accuracy and validation

Fig.5 depicted the performances of these three models. Along the X-axis, the epochs are denoted, while along the Y-axis, the accuracy and validation accuracy levels are represented.

1)VGG16 achieved a steady increase in accuracy, reaching close to 100% after a few epochs, with its validation accuracy stabilizing at 99%.

2)VGG19 performed similarly but showed slightly better accuracy at earlier epochs, with both accuracy and validation accuracy converging to 99% and above.

3)Xception showed similar accuracy to the other models, but its validation accuracy kept fluctuating a lot, which suggests it might have issues with overfitting.

Overall, VGG19 and VGG16 were more stable, while Xception's accuracy was less consistent. This comparison suggests that VGG-based models such as VGG-16 and VGG-19 might be better for this dataset.

VI. CONCLUSION

In conclusion, this paper demonstrates how deep learning can be employed to complement diagnosis of pulmonary tuberculosis, particularly in rural communities where advanced medical care is not readily available. Using models such as VGG-16, VGG-19, and Xception, the system reads clinical data and medical images to offer precise and quick results. The system assists medical staff to easily diagnose the disease and demonstrates how the integration of AI and various forms of data can make medical testing quicker, more precise, and available in remote locations.

VII. REFERENCES

[1] Tuberculosis. Retrieved from https://en.wikipedia.org/wiki/Tuberculosis

[2] Pulmonary Tuberculosis. Retrieved from https://www.healthline.com/health/pulmonary-tuberculosis

[3] VGG16. Retrieved from https://www.geeksforgeeks.org/vgg-16-cnn-model/

[4] VGG19. Retrieved from https://www.geeksforgeeks.org/vgg-net-architecture-explained/

[5] Xception. Retrieved from https://iq.opengenus.org/xception-model/

[6] Malik, H., & Anees, T. (2024). Plos one, 19(3), e0296352.

[7] Agarwal, S., Arya, K. V., & Meena, Y. K. Multi-modal deep learning methods for classification of chest diseases using different medical imaging and cough sounds.

(2024). MultiFusionNet: Multilayer Multimodal Fusion of Deep Neural Networks for Chest X-Ray Image Classification. arXiv preprint arXiv:2401.00728.

[8] Ahmed, M. S., Rahman, A., AlGhamdi, F., AlDakheel, S., Hakami, H., AlJumah, A., ... & Basheer Ahmed, M. I. (2023). Joint Diagnosis of Pneumonia, COVID-19, and Tuberculosis from Chest X-ray Images: A Deep Learning Approach. Diagnostics, 13(15), 2562.

[9] Priya, S. U., Tarun, S. G., Shamitha, S., Rao, A. S., & Prasad, V. B. (2023, May). Multi Modal Smart Diagnosis of Pulmonary Diseases. In 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT) (pp. 33-40). IEEE.

[10] Wang, R., Chen, L. C., Moukheiber, L., Seastedt, K. P., Moukheiber, M., Moukheiber, D., ... & Celi, L. A. (2023). Enabling chronic obstructive pulmonary disease diagnosis through chest X-rays: A multi-site and multi-modality study. International Journal of Medical Informatics, 178, 105211.

[11] Ahmed, M. S., Rahman, A., AlGhamdi, F., AlDakheel, S., Hakami, H., AlJumah, A., ... & Basheer Ahmed, M. I. (2023). Joint Diagnosis of Pneumonia, COVID-19, and Tuberculosis from Chest X-ray Images: A Deep Learning Approach. Diagnostics, 13(15), 2562.

[12] Oloko-Oba, M., & Viriri, S. (2022). A systematic review of deep learning techniques for tuberculosis detection from chest radiograph. Frontiers in Medicine, 9, 830515.

[13] Karki, M., Kantipudi, K., Yang, F., Yu, H., Wang, Y. X. J., Yaniv, Z., & Jaeger, S. (2022). Generalization challenges in drug-resistant tuberculosis detection from chest X-rays. Diagnostics, 12(1), 188.

[14] Dasanayaka, C., & Dissanayake, M. B. (2021). Deep learning methods for screening pulmonary tuberculosis using chest X-rays. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 9(1), 39-49.

[15] Li, X., Zhou, Y., Du, P., Lang, G., Xu, M., & Wu, W. (2021). A deep learning system that generates quantitative CT reports for diagnosing pulmonary tuberculosis. Applied Intelligence, 51, 4082-4093.

[16] Lee, S., Yim, J. J., Kwak, N., Lee, Y. J., Lee, J. K., Lee, J. Y., ... & Yoon, S. H. (2021). Deep learning to determine the activity of pulmonary tuberculosis on chest radiographs. Radiology, 301(2), 435-442.

[17] Norval, M., Wang, Z., & Sun, Y. (2021). Evaluation of image processing technologies for pulmonary tuberculosis detection based on deep learning convolutional neural networks. Journal of Advances in Information Technology Vol, 12(3).

[18] Angappan, K. (2021). Prediction of multi-lung disease. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(10), 365-372.

[19] Khan, F. A., Majidulla, A., Tavaziva, G., Nazish, A., Abidi, S. K., Benedetti, A., ... & Saeed, S. (2020). Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. The Lancet Digital Health, 2(11), e573-e581

[20] Rahman, T., Khandakar, A., Kadir, M. A., Islam, K. R., Islam, K. F., Mazhar, R., ... & Chowdhury, M. E. (2020). Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access, 8, 191586-191601.

[21] Guo, R., Passi, K., & Jain, C. K. (2020). Tuberculosis diagnostics and localization in chest X-rays via deep learning models. Frontiers in Artificial Intelligence, 3, 583427.

[22] Ma, L., Wang, Y., Guo, L., Zhang, Y., Wang, P., Pei, X., ... & Lure, F. Y. (2020). Developing and verifying automatic detection of active pulmonary tuberculosis from multi-slice spiral CT images based on deep learning. Journal of X-ray Science and Technology, 28(5), 939-951.

[23] Munadi, K., Muchtar, K., Maulina, N., & Pradhan, B. (2020). Image enhancement for tuberculosis detection using deep learning. IEEE Access, 8, 217897-217907.

[24] Hussain, L., Rathore, S., Abbasi, A. A., & Saeed, S. (2019, March). Automated lung cancer detection based on multimodal features extracting strategy using machine learning techniques. In Medical imaging 2019: physics of medical imaging (Vol. 10948, pp. 919-925). SPIE.