International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

Toward a lightweight model for pulmonary tuberculosis classification

Raharinirina Eric Florent[1], Rahajaniaina Andriamasinoro[2], Ratiarison Adolphe Andriamanga[3]

[1]*Student, Department of Mathematics, Computer Scienceand Applications, University of Toamasina, Madagascar.*
[2]*Associate Professor, Department of Mathematics, Computer Science and Applications, University of Toamasina, Madagascar*
[3]*Professor Emeritus, Department of Physics and Applications, University of Antananarivo, Madagascar.*
Corresponding Author: Raharinirina Eric Florent

Date of Submission: 20-10-2024 Date of Acceptance: 30-10-2024

ABSTRACT: As in every developing country, Madagascar is still fighting against the tuberculosis (TB), the second leading cause of death from a communicable infectious disease. Its incidence remains high, exacerbated by limited medical infrastructure and unequal distribution of health resources. Several researches were conducted to automatically diagnosis disease to greatly lowering the total expenses. Deep learning approaches demonstrated its high effectiveness in the domain using chest x-ray images (CXR) that are the most cost-effective and widely utilized imaging technology to classify TB. Some preprocessing was necessary for getting better features from the images. Using a more lightweight model inspired by SSD Lite X, this work achieves the highest accuracy (training, validation and testing), recall, F1-Score and specificity of 100% each, making this model the most effective one in the state-of-the-art. It could be deployed in a low-resource environments, enabling automated diagnosis in every region of Madagascar and other developing countries.

KEYWORDS: Tuberculosis classification, Deep learning, Chest X-Ray, Medical imaging

I. INTRODUCTION

Tuberculosis remains a global health threat, affecting millions of people each year. According to the World Health Organization [1], in 2022, 10.6 million people have developed tuberculosis in the world, including 5.8 million men, 3.5 million women and 1.3 million children. In total, 1.3 million people died from this disease. Tuberculosis is therefore considered the second leading cause of death from an infectious disease, behind COVID-19. Additionally, it is a communicable infectious disease caused by a bacteria called Koch's bacillus (strains of Mycobacterium tuberculosis complex). The global effort on diagnosis and treatment, starting in 2000, helped save 75 million lives. In Madagascar, despite significant progress in the fight against this disease, the incidence of tuberculosis remains high. In fact, it[2] records an incidence of 233 per 100,000 people. This context imposes particular challenges in terms of diagnosis and treatment, exacerbated by limited medical infrastructure and unequal distribution of health resources.

Rapid and accurate diagnosis of TB is crucial to control the spread of the disease and improve cure rates. However, conventional diagnostic methods, such as smear microscopy and bacterial culture, are often slow and require specialized equipment that is not always available in rural areas of Madagascar. The emergence of multidrug-resistant tuberculosis (MDR-TB) [3] adds an additional layer of complexity, requiring more sophisticated testing and more complex treatment regimens. Typically, specialist doctors use chest x-ray images to examine whether a patient has tuberculosis. This process is both inevitably subjective and time-consuming [4][5][6]. Moreover, chest x-ray images of tuberculosis are often misdiagnosed as other diseases with comparable radiological patterns [7][8]. It would therefore be negligible that patients could obtain incorrect treatments which would deteriorate their health conditions instead of improving it [9]. In a country like Madagascar, the number of doctors, particularly radiologists, is far from sufficient.

In this context, the application of artificial intelligence techniques and machine learning offers a promising opportunity to improve the diagnosis

of tuberculosis. Indeed, they prove effective and efficient in the field of health [10], especially approaches based on deep learning [11][12] in the detection, segmentation and classification of different diseases [13 – 22]. To do this, these models often use chest x-ray images, which has the advantage of being both less expensive and easy to implement to diagnose tuberculosis, especially for children [23]. The emergence of the convolutional neural network (CNN) has greatly contributed to the analysis of images to derive characteristics allowing them to be recognized. It is widely used in image classification using publicly accessible databases. Many CNN-based classification models such as VGG, ResNet, DenseNet, MobileNet and SSD have been implemented for general image classification but they are also suitable for special images, for specific domains, such as chest x-ray images [24]. Indeed, the application of these models for the classification of tuberculosis, using transfer learning, shows quite high precision. But these models remain heavy and require operation on a fairly efficient environment [25][26].

Classification models based on machine learning and deep learning can quickly and accurately analyses medical images and clinical data, providing a potentially more accessible and less expensive alternative to traditional methods. In particular, the development of lightweight classification models, adapted to the constraints of resource-limited environments such as Madagascar, could revolutionize the fight against tuberculosis by allowing rapid diagnosis even in the most remote areas.

This work therefore aims to establish a tuberculosis classification application that is capable of working well on mobile devices such as smartphones. The model must be lightweight to respond effectively to constraints. Lightweight models often tend to gain speed but lose precision. The present model must reconcile both of these two criteria to circumvent this limit.

The rest of this paper is divided in the following sections: Section 2 presents relevant previous studies. Section 3 develops the methodology presenting the dataset, pre-processing steps and the model, while Section 4 gives a summary of the experimental results of the proposed approach. Finally, Section 5 concluded the study.

II. RELATED WORKS

The present work focuses on the binary classification of pulmonary tuberculosis. Several works have already been carried out in this direction, which continue to evolve gradually using various approaches.

Ahsan M [27] use VGG16 network to classify tuberculosis using 800 images obtained from a combination of Montgomery County dataset containing 58 tuberculosis and 80 healthy and the Shenzhen Hospital dataset having 336 tuberculosis and 326 normal images.75% of the dataset were used for training and the rest for testing. They achieved an accuracy of 81.25%.

Michael Norval, Zenghui Wang and Yanxia Sun [28] used images from chest X-rays to study the accuracy of two methods for detecting pulmonary tuberculosis. They implement both a color depth and preprocessing method, starting by taking a fixed random generator, testing the color depth and testing the preprocessing, and a hybrid approach by which they extract the lung region of interest at starting from a fixed random generator and applying deep convolutional neural networks. The hybrid method demonstrated a highest accuracy of 92.54%.

In their work, Jackson andal. [29] use two subsystems for tuberculosis detection. The first consists of data acquisition and the second consists of recognition. They found 95% accuracy using InceptionV3 Deep Net and SVM for classification. A public tuberculosis database was used [30].

Khan andal. [31] on the other hand use a database consisting of 12,638 patient tuberculosis records between 2016 and 2017 obtained from a tuberculosis laboratory in Khyber Pakhtunkhwa, Pakistan. Their model achieved a validation accuracy of 93% and an overall accuracy of 94% [30].

K. Munadi and al. [32] apply three types of image enhancement algorithms to bring out the best features of chest x-ray images for tuberculosis classification: Unsharp Mask algorithm, High-Frequency Emphasis Filtering Emphasis Filtering and Contrast Limited Adaptive Histogram Equalization. The enhanced images are then entered into two pre-trained models ResNet and EfficientNet for transfer learning, to have an accuracy of 89.92% accuracy and 94.8% AUC.

To go further, Xie andal. [33] use Faster R-CNN model to focus on detection of the TB-infected area Using a total of 6144 images, containing 3356 TB and 2788 normal images from a combination of 3 datasets (the Montgomery County dataset; the Shenzhen Hospital dataset; and the First Affiliated Hospitals Dataset (contains a total of 5,344 images with 2,962 TB and 2,382 control images) from Xian Jiao Tong University, Shanxi, China), and opted for a 80:20 train/test

split. They achieved a maximum accuracy of 92.60% and an AUC of 0.98.

To avoid the redundance of features generated by the CNN, Sahlolandal. [34] use a hybrid network, MobileNet-AEO (Artificial Ecosystem-based Optimization). The Shenzhen Hospital dataset and the Mendeley Dataset (UK), which contains atotal of 6,421 (3,883 TB and 2,538normal) CXR images were used and split into 80:20 ratios to achieve an accuracy of 94.10%.

Stefanus K. andal. [35] on the other hand uses the improved Canny edge detection algorithm on chest X-ray images. They combined the Montgomery and Shenzhen datasets together to have a total of 394 TB-infected lungs and 406 healthy lungs. To equalize the number of images in the 2 categories, 12 normal images were discarded. The dataset split into 90:10 ratios training/testing.The model achieves anaccuracy, sensitivity and specificity of 93.59%, 92.31% and 94.87% respectively.

Rahman andal. [36] present other approaches consisting of the use of 9 pre-trained models for transfer learning. The authors used ResNet18, ResNet50, ResNet101, ChexNet, InceptionV3, Vgg19, DenseNet201, SqueezeNet, and MobileNet on 700 chest x-ray images of tuberculosis and 3500 normal ones taken from multiple public databases combination. They were doing both segmentation with two U-net models, classification and lung segmentation. The classification demonstrates accuracy of 97.07%, precision of 97.34%, sensitivity and F1-Score and specificity of 97.07%, 97.14% and 97.36% respectively. But coupled with lung segmentation, the classification achieves an accuracy of 99.9% with DenseNet201.

Abideen andal. [37] perform a binary classification of CXR images into TB and normal using a B-CNN (Bayesian-based CNN) method. A combination of the Montgomery and Shenzhen datasets were used and split into 80:20 ratio to achieve an accuracy of 96.42%.

Ayaz andal. [38] use the same datasets as [37]. They combined hand-crafted features using Gabor filter and DL features with transfer learning-based pre-trained networks that are Inceptionv3,MobileNet, Xception, ResNet50,and Inception Resnetv2. A ten-fold cross-validation approach for training were performed to achieved an accuracy of 93.47%, with an AUC of 0.97.

Vasundhara A., Gaurav D. et al proposed a method for detecting and highlighting the region infected by tuberculosis in [39]. They used a combination of different datasets: the TBX11K (Tuberculosis X-ray) dataset, Shenzhen dataset, Montgomery dataset, Belarus, NIAID, SNA dataset, and the X-ray images dataset from the National Institute of Tuberculosis and Respiratory Diseases, New Delhi. So, a total of 8688 CXR images with 4936 TB and 3752 normal images were utilized for the experience. They use progressive resizing for automatic inference of TB with chest X-ray images. For classification, normalization-free neuralnetworks (NFNets) optimized on ImageNet are applied with an accuracy of 96.91%, an AUC of 99.38%, a sensitivity of 91.81% and a specificity of 98.42%. for multi-class classification; and 96% accuracy and 98% AUC for binary classification. As for highlighting the regions, the Score-cam algorithm was used.

Fati S., Senan E. et ElHakim N [40] use two AI techniques: CNN an ANN. Both of the approaches were applied to two datasets, the Shenzhen dataset and the TB Chest Radiography Database (created by researchers from Qatar University, Doha and Dhaka University Bangladesh in collaborationwith a group of physicians at the Hamad Medical Corporation consisting of 4200 chest X-rays in PNG format, with a resolution of 512 × 512 pixels for each X-ray: 3500 normal and 700 tuberculosis). The first approach hybridizes two CNN models, which are Res-Net-50 and GoogLeNet techniques applying the principal component analysis (PCA) algorithm before the SVM for the classification. The ANN approach fuses the features extracted by ResNet-50 and GoogleNet models then combines them with the features extracted by the gray level co-occurrence matrix (GLCM), discrete wavelet transforms (DWT) and local binary pattern (LBP) algorithms. Using both of the two datasets, the ANN approach achieves the highest scores: an accuracy of 99.2%, a sensitivity of 99.23%, a specificity of 99.41%, and an AUC of 99.78% with the first dataset and an accuracy of 99.8%, a sensitivity of 99.54%, a specificity of 99.68%, and an AUC of 99.82% with the second one.

SuciAulia and SugondoHadiyoso [41] propose the detection of tuberculosis using the architecture of Vgg-16. A database of chest x-ray images was used, including 700 normal images and 140 tuberculosis infected images which could be found at Kaggle. Having tested several hyperparameters for training the model, a batch size of 50 made it possible to have the maximum training accuracy of 99.64% including the general precision, recall and F1-score of 99 % each for binary classification.

Similarly, VggNet, ResNet50 and GoogleNet are used by Rohan Rana, Abhishek

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

Gaba [42] for tuberculosis detection. They use transfer learning from these models of which VggNet presents a validation accuracy, recall and F1-Score of 99.34%, 94.5% and 95.8% respectively. Although the accuracy is high, the other metrics presented seem to perform even worse, and VGG still contains a fairly high number of parameters to run smoothly on mobile devices [43].

Uma Maheswari andal. [44] propose a shallow convolutional network to have a lighter model and allow appropriate interpretation for the diagnosis of tuberculosis. Understanding how networks work would help radiologists better understand and accept models based on deep learning. The model consists of 4 layers of max pooling with several hyperparameters optimized by the Bayesian optimization technique. The precision, F1-Score, sensitivity and specificity are 95% and also an ROC of 97.6%.

As in [36], V. Sharma, Nillmani, S.K. Gupta andal. perform not only a classification but also a segmentation for accurate and precise detection of tuberculosis in chest X-ray images, with visualization of infection using gradient-weighted class activation mapping (Grad-CAM) heatmaps [45]. Using 704 chest X-ray images, they used UNet model for segmentation. So, UNet were trained with these images first, then implanted on 1400 tuberculosis to segment the lung region. The classification of the lung region into tuberculosis or normal is done with the segmented lung region using Xception model coupled with Grad-CAM heatmap visualization model to visualize the abnormalities. The classification achieved accuracy, precision, recall, F1-score, and AUC values of 99.29%, 99.30%, 99.29%, 99.29%, and 0.999, respectively. Lesions were shown in general in the upper part of the lungs.

James D., Hridayanand G. andal.[46] deals with multi-label classification approach to identify tuberculosis in CXR images. They use the publicly available Health Insurance Portability and Accountability (HIPPA) dataset split into 80:20 training/testing. They use U-Net model to segment the Region of Interest of the images, then apply EfficientNetB4 model for the classification. The AUC ranges from 0.95 to 1.00 with Calcification showing lowest AUC of 0.95.

III. METHODOLOGY

To perform a TB binary classification, a chest X-Ray images dataset, publicly available were used. Some preprocessing and data balancing were necessary for better results. When the data is balanced, it passes through the lightweight proposed model to classify it as normal or tuberculosis. The overall methodology is shown in figure 1.

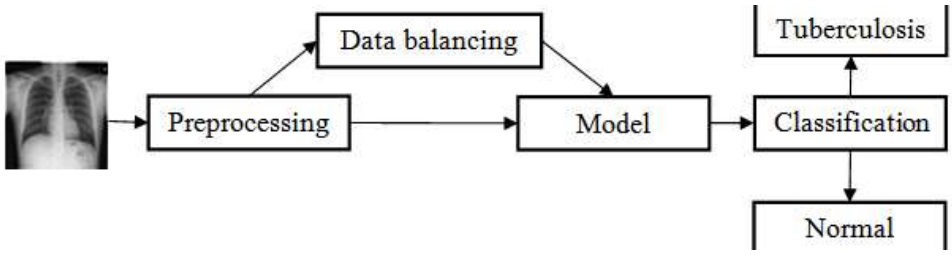

Figure 1-Overview of the proposed system

3.1 Dataset

For TB classification, a publicly accessible dataset is used in this present work: TB Chest Radiography Database (available on Kaggle). This dataset is a combination of three datasets: NLM dataset, Belarus dataset and RSNA dataset. It contains 700 TB images and 3500 normal images. The resolution of images is 512×512.

3.1.1 NLM dataset

The NML dataset [47] or National Library of Medicine dataset is the combination of two publicly available datasets: Montgomery and Shenzhen datasets; made by NLM in the U.S. The Montgomery dataset contains 138 posterior-anterior chest X-ray images divided into 58 TB images and 80 normal images, but Shenzhen holds 667 (336 from TB patients and 324 from normal subjects). The resolution of images in both datasets were not the same. For the Montgomery, the resolution was either 4,020×4,892 or 4,892×4,020 pixels but it was variable for the Shenzhen dataset (around 3000×3000 pixels). In total, there are 406 normal and 394 TB infected X-ray images.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

3.1.2 Belarus dataset

The National Institute of Allergy and Infectious Diseases, Ministry of Health, Republic of Belarus created the Belarus Set [48] from 169 patients. It contains 306 TB CXR images of 2248×2248 pixels resolution.

3.1.3 RSNA dataset

The RNSA or Radiological Society of North America pneumonia detection challenge dataset hold on 30,000 chest X-ray images [49]. But as its name, it's made for detecting pneumonia but not tuberculosis. 1000 images are normal and others are abnormal.

So, for the TB Chest Radiography Database, 3,094 normal images were taken from RNSA and rest of the 406 normal images were from NLM database. The total of TB images is 700, taken from NML and Belarus. The dataset was split into 90:10 train/test ratios and 10% is used for validation.

The Figure 2 gives some samples of CXR images from the dataset categorized into normal and tuberculosis.

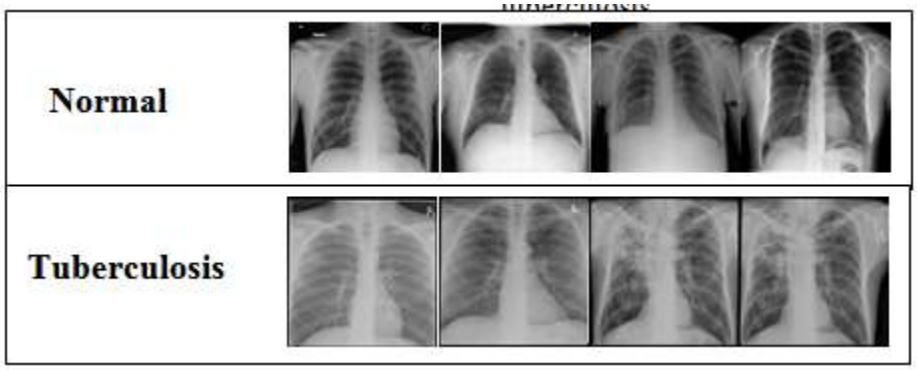

Figure 1-Samples of CXR images from the Dataset. (A) Normal and (B) Tuberculosis

3.2 Preprocessing and data balancing

Asworking with CNN using transfer learning, the input images must have a fixed size. A lightweight auxiliary stage inspired bySSDLiteX model with MobileNetV3Small as backbone (cf Model Section) is used, so the images were resized into 224×224 pixels. And then normalized using MobilenetV3 built-inpreprocessing function to ensure that the images are pre-processed in the same way as the images used to train MobileNetV3. This is crucial for the model to perform well on the data.

Then a sharpening method were used to the dataset to get higher features, combinedwith a histogram equalization methodfor getting further improvement.

The imbalanced dataset should be balanced to get higher performance of the model. The Resampling (upsampling) technique were used to balance it. At the same time, while balancing the dataset, three transformations were applied: rotation, flip left right and flip top bottom.The rotation used is a clockwise rotation with an angle of 30 degree.

Figure 2-Samples of augmented CXR images with resampling and transformation techniques

To optimize the performance, the CPU is kept busy and the training is improved throughput by using a prefetch function to overlap data preprocessing and model training by prefetching the next batch of data while the current batch is being processed. This operation ensures the training process is efficient and takes full advantage of available system resources.

3.3 Model

In this present work, a lightweight model inspired bySDDLiteX [50] model is used. SSDLiteX

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

is an enhancement of SSDLite [51] for small object detection. This work uses the potential of this model to enhance the classification ability provided by the backbone. In the SSDLite model, all the regular convolutions in SSD prediction layers are replaced by a depthwise followed by 1×1 projection. It reduces both parameter count and computational cost.The first layer of the auxiliary stage is attached to the last feature extractor layer that has an output stride of 16 and the second layer is attached tothe last feature extractor layer that has an output stride of32. These layers are referred as C4 and C5 by [52]. The auxiliary stage (stage after the backbone MobileNetV2) consists of a 1 × 1 convolutional layer reducing the number of channels, a 3 × 3 depthwise convolutional layer with the reduced number of channels, and a 1 × 1 convolutional layer restoring the number of channels [50]. SSDLiteX introduces an auxiliary stage consisting of a 3 × 3 depthwise convolutional layer and a 1 × 1 convolutional layer. To reduce the number of parameters, the present model uses the same structure for the auxiliary stage but with a smaller number of filters in each layer. Instead of using 256 filters at all stages, the present model uses a number of filters of 32, 32 and 16 at the first, the third and the last stage respectively.And the difference also is about the backbone. The proposed model uses the lighter version targeted for low resource use cases of MobileNetV3 called MobileNetV3Small [52] instead of using MobileNetV2.The figure 4 below shows the auxiliary stage of the 3 models:SSDLite, SSDLiteX and the proposed model.

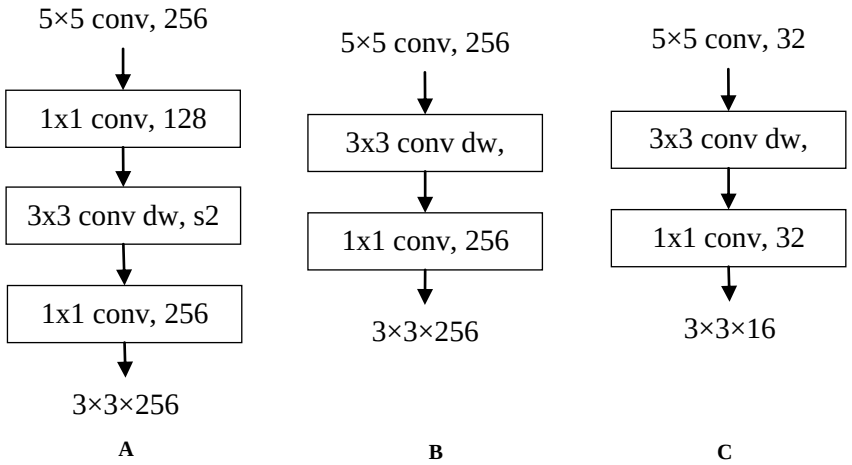

Figure 3-Auxiliary stage of (A) SSDLite, (B) SSDLiteX and (C) proposed model

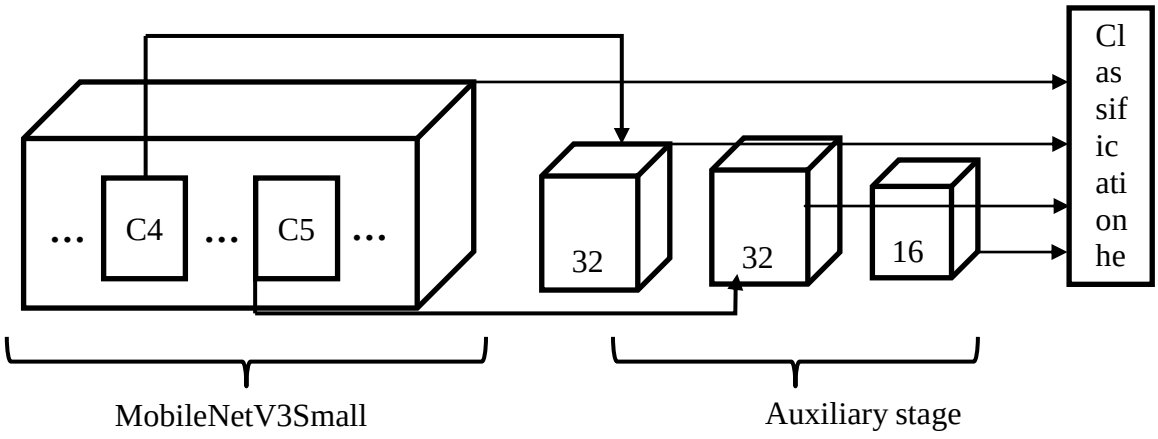

Figure 4-Architecture of the model

The proposed model uses a lightweight backbone to ensure itself to be so. Then a small number of filters are used, passing from 256 to 32 and 16 in the auxiliary stage. These operations ensure the reduction of the number of parameters and the computation cost.Besides, the proposed model has only 0.9 million (3.44 MB) of parameters. The model size is 6.03Mo.It

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

outperforms categorically MobileNetV3Small having 2.4 million of parametersand 9.83MB of size.The multiply-accumulate operations (MACs) of the proposed model is only 72 million.

IV. EXPERIMENTS AND RESULTS

4.1 Experiments

The presented model was trained in the imbalanced and balanced dataset. It was split into train and test. 90% of the images set was used for training and 10% for testing. Then 10% of the training data was used for validation. In another words, from the total of 3500 normal CXR images, 3150 (90%) were used for training and 350 (10%0 were used for testing. And from 3150 used for training, 10%, that is 315 were used for validating the model. Then, the split dataset is used to train and evaluate the model both with balanced and imbalanced dataset.

The Table 1 and Table 2 below give the repartition of the dataset into train and test and validation with both with and without any balancing techniques

Types	**Total**	**Training**	**Test**	**Validation**
Normal	3500	2835	350	315
Tuberculosis	700	567	70	63

Table 1-Details of training, testing and validation set for imbalanced classification set

Types	**Total**	**Training**	**Test**	**Validation**
Normal	3500	2835	350	315
Tuberculosis	700	2835	70	315

Table 2-Details of training, testing and validation set with balanced classification set

Two experiments were conducted to the model: using imbalanced and balanced dataset.

For first experiment with the imbalanced dataset, many batch-size were tested but the batch-size of 1 gave the better performance. The same batch-size is adopted for the balanced data. The implementation was doneusing TensorFlow and Keras with Python 3.10.13 on Intel(R) Core (TM) i5-5300U CPU @ 2.30GHz (4 CPUs), ~2.3GHz and 8 Gb RAM. During the training, the order of the data is randomized before each epoch to avoid bias and improve generalization. Then, to ensure the random operation produce the same result every time running the code, the seed value is set to 1000. It sets the initial state of the random number generator to ensure the reproducibility and consistency in results.

4.1 Evaluation metrics

To evaluate the model at the end of the training process, we adopted the following metrics:

The accuracy:it indicates how often the model correctly identifies both positive and negative instance. It gives the proportion of correct predictions made by the model.

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

The accuracy gives an overall effectiveness of the model by calculating the ratio of correctly predicted cases (both true positives and true negatives) to the total number cases.

The precision: it measures the proportion of true positive predictions out of all positive predictions made by the model.

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

The precision indicates how many of the predicted positive cases are actually positive. High precision

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

means when the model predicts a positive instance, it is actually correct.

Recall (Sensitivity or True Positive Rate): this measures the proportion of true positive predictions out of all actual positive instances.

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

The recall indicates how many of the actual positive cases the model correctly identifies.

F1-Score: it is the harmonic mean of precision and recall. It provides a single metric that balance both concerns.

$$\text{F1} - \text{Score} = 2\,\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

This metric is really useful to better measure the model's performance when there is an uneven class distribution.

The specificity (True Negative Rate): it measures the proportion of true negative predictions out of all actual negative instances.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

The specificity indicates how many of the actual negative cases the model correctly identifies. High specificity means the model can identify most of the negative instances.

4.1.1 Results

TB classification with imbalanced dataset

The optimizer used for training the model is the Adam optimizer from Keras. As a binary classification, binary cross entropy was used as training loss. The epoch was 50.

The model shows a great performance even with imbalanced data. It has an accuracy, precision, recall, F1-Score and specificity of 100% each. The validation accuracy is 99.74%. With the testing dataset, the model achieves a perfect performance of 100%.

The model demonstrates excellent performance with perfect precision, recall, and F1-scores. It shows perfect accuracy in predicting both NORMAL and TUBERCULOSIS cases. The confusion matrix below confirms the effectiveness of the model, without any errors in predictions.

350	**0**
0	**70**

The matrix shows that out of 350 NORMAL cases, and out of 70 TUBERCULOSIS cases, all are perfectly classified as NORMAL and TUBERCULOSIS respectively.

These results are confirmed by the ROC curve and an AUC of 1 that demonstrate a perfect classification at all threshold levels.

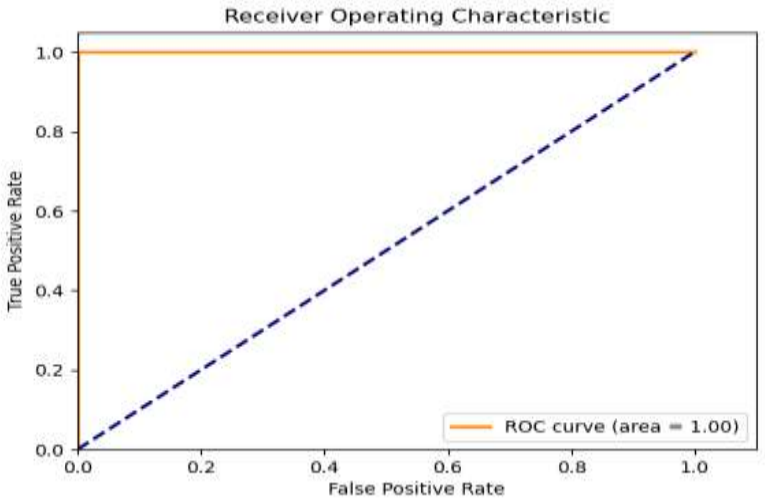

Figure 5-ROC Curve with Imbalanced dataset

The training loss, the validation loss and the test loss are 0%, 1.85% and 0% respectively. These values show that the model is learning training data very well and generalizing well to the

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

validation data. The graph of accuracy and loss of the model are shown in the figure 7.

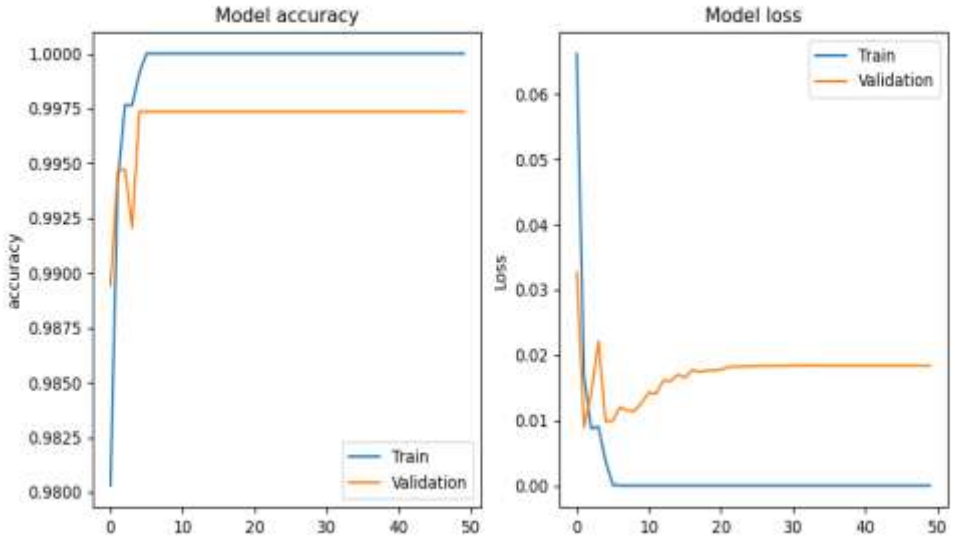

Figure 6-Imbalanced dataset - Model Accuracy and model Loss

The lightweight model presented in this work demonstrates a great performance that outperforms the state of the art even in an imbalanced dataset. All metrics are at their highest value of 100%, except the validation accuracy, which is 99.74%. However, this value still surpasses the state-of-art recorded to the best of our knowledge.

4.1.2 TB Classification with balanced dataset

The proposed model performs greatly with the imbalanced dataset. It could be improved with balanced dataset.

The same configuration was used about hyperparameters as in the imbalanced data. The epoch of 50 and a batch size of 1 were used. An improvement is expected in the validation accuracy to make it perfect.

Now, all metrics are at their highest levels, which is 100%. Thetraining accuracy, the validation accuracy and the test accuracy areall 100%. with a loss of 0.0%, 0.01% and 0.0% respectively. On the training, validation and testing sets, the model achieves perfect accuracy;showing that the model is correctly classified all instances in these sets. Then, the extremely low loss across all datasets suggests that the model fits the data perfectly.

Precision, recall, F1-score and specificity are also 100% each.That means the model, when predicts TUBERCULOSIS orNORMAL, it is correct 100% of time, with neither false positives nor false negatives. The model is effective at distinguishing between NORMAL and TUBERCULOSIS cases. The F1-Score confirms these results. It shows that the model achieves a good balance between precision and recall. It suggests that the model performs perfectly in both precision and recall.The ROC curve below (Figure 8) and the AUC of 1 confirms this effectiveness. The model discriminates between classeswith 100% accuracy without any false positives or false negatives.

The losses of the model are very low. Both training and validation data drops sharply and stabilizes close to 0 after few epochs. There's a very minimal difference between both losses. It indicates that the model is not overfitting, as both sets of data are being modeled with similar accuracy.

Figure 9 shows the graph of the accuracy and loss of the model.

With these results, not only the model performs exceptionally great in the imbalanced data, but it shows a perfect classification ability in the balanced one. The validation accuracy shows a nearly perfect generalization to unseen data. The balancing technique overcome this shortcoming by giving a perfect validation accuracy of 100%. With the balanced data, all metrics, that are the accuracy (training, validation and testing), the precision, the recall, the F1-Score and the specificity, are at 100% levels. This shows the effectiveness and robustness of the model in TB classification.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

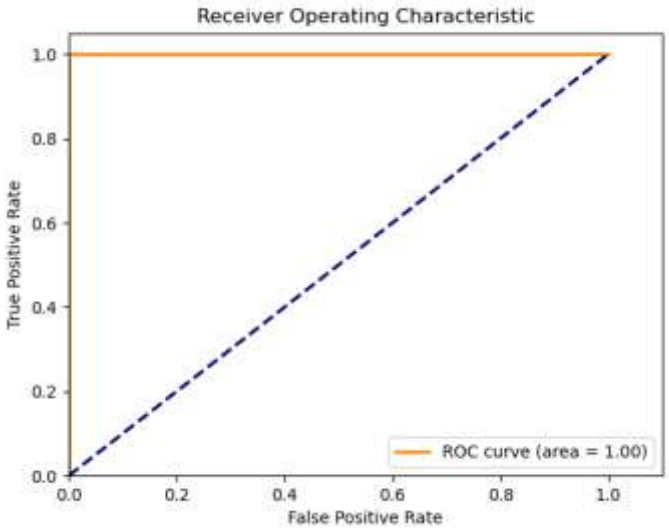

Figure 7-Balanced data:ROC curve and AUC

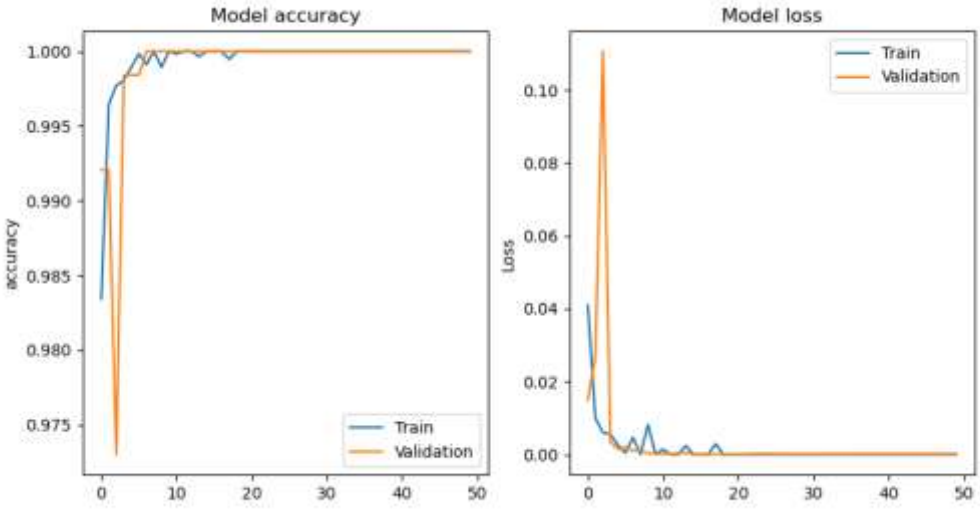

Figure 8-Balanced data:model accuracy and loss

The confusion matrix below is same as with imbalanced datathat shows perfect predictions between both classes:

350	**0**
0	**70**

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 10 Oct. 2024, pp: 308-321 www.ijaem.net ISSN: 2395-5252

So, all the 350 NORMAL are correctly classified as NORMAL and all 70 TUBERCULOSIS are also classified correctly as TUBERCULOSIS. There is no false positives nor false negatives recorded.

The Table 3 gives a comparison of the metrics values between the imbalanced and balanced datasets used in this paper to give an overview of both results. The balancing techniques contributed on completing the performance of the model making all metrics at the highest level.

Metrics	Imbalanced dataset (%)	Balanced dataset (%)
Accuracy	100	100
Validation accuracy	99.74	100
Precision	100	100
Specificity	100	100
Recall	100	100
F1-Score	100	100

Table 3-Metrics comparison between balanced and imbalanced datasets

CONCLUSION

The study provides a deep learning-based system for TB classification. It presents a lightweight model that balance both precision and speed. The perfect accuracy and validation accuracy of 100% demonstrate the effectiveness of the model. The precision, recall, F1-Scores and specificityof 100% for the balanced dataset respectively indicates that the model is highly reliable in identifying both classes. The confusion matrix demonstrates the robustness of the model, confirmed by an excellent diagnostic ability with perfect AUC by the ROC curve.

The model is highly effective for tuberculosis classification, showing both high accuracy and excellent generalization capabilities. Such high metrics are desirable in applications where it is important to correctly identify positive cases. It is highly reliable for practical use in sensitive and critical application like in the medical diagnosis here. The deployment of the application would resolve the lack of radiologists and would improve the diagnosis of tuberculosis in its early stage even in the most remote region of Madagascar.

The future works would apply the model with another datasets both for classification and object detection purpose.

REFERENCES

[1]. Official World Health Organization website, online, available:[WHO][Accessed on 07/04/2024 at 9:13 p.m.]

[2]. World Bank data, online, available: [World Bank Data][Accessed on 07/04/2024 at 9:13 p.m.]

[3]. James Devasia, Hridayanand Goswami and al., "Deep learning classification of active tuberculosis lung zones wise manifestations using chest X‑rays: a multi label approach", Scientific Reports, (2023) 13:887,https://doi.org/10.1038/s41598-023-28079-0

[4]. World Health Organization. Systematic Screening for Active Tuberculosis: Principles and Recommendations. Available from: [WHO][PubMed]

[5]. World Health Organization. Global Tuberculosis,Report2019. Geneva, 2019. [Google Scholar]

[6]. Altaf F, Islam SMS, Akhtar N, andal."Going deep in medical image analysis: concepts, methods, challenges, and future directions". IEEE Access 2019; 7:9954072.doi:10.1109/access.2019.2929365.

[7]. Sharma SK, Mohan A. "Tuberculosis: from an incurable scourge to a curable disease-journey over a millennium", Indian J Med Res 2013; 137: 455-493. [PMCfreearticle] [PubMed] [Google Scholar]

[8]. Silverman C. "An appraisal of the contribution of mass radiography in the discovery of pulmonary tuberculosis". Am Rev Tuberc 1949; 60: 466-482. [PubMed] [Google Scholar]

[9]. Ueda D, Yamamoto A, Shimazaki A, Walston SL, Matsumoto T, Izumi N, andal. "Artificial intelligence-supported lung cancer detection by multi-institutional readers with multi-vendor chest radiographs: a retrospective clinical validation study". BMC Cancer. 2021; 21:1120

[10]. Harris M, andal.: A systematic review of the diagnostic accuracy ofartificial intelligence-based computer programs to analyze chestx-

rays for pulmonary tuberculosis. PloS one 14:e0221339, 2019

[11]. Esteva A, Chou K, Yeung S, andal."Deep learning-enabled medical computer vision". NPJ Digit Med, 2021;4(1):5. doi:10.1038/s41746-020-00376-2.

[12]. Pandey A, Singh SK, Udmale SS, andal."Epileptic seizure classification using battle royale search and rescue optimization-based deep LSTM". IEEE J Biomed Health Inform, 2022;26(11):5494–505.doi:10.1109/jbhi.2022.3203454

[13]. Tripathi S, Verma A, Sharma N. Automatic segmentation of brain tumour in MR images using an enhanced deep learning approach. Comput Methods Biomech Biomed Engin 2020;9(2):121–30. doi:10.1080/21681163.2020.1818628.

[14]. Tandel GS, Tiwari A, Kakde OG. "Performance optimization of deep learning models using majority voting algorithm for brain tumor classification". Comput Biol Med, 2021;135:104564.doi:10.1016/j.compbiomed.2021.10456

[15]. Maji D, Sigedar P, Singh M. "Attention Res-UNet with guided decoder for semantic segmentation of brain tumors". Biomed Signal Process Control, 2022;71:103077. doi:10.1016/j.bspc.2021.103077

[16]. Nayak SR, Nayak DR, Sinha U, and al."Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study. Biomed Signal".Process Control, 2021;64:102365. doi:10.1016/j.bspc.2020.102365

[17]. Nillmani, Jain PK, Sharma N, andal. "Four types of multiclass frameworks for pneumonia classification and its validation in X-ray scans using seven types of deep learning artificial intelligence models." Diagnostics, 2022 ;12(3):652. doi:10.3390/diagnostics12030652.

[18]. Nillmani, Sharma N, Saba L, andal."Segmentation-based classification deep learning model embedded with explainable AI for COVID-19 detection in chest X-ray scans". Diagnostics, 2022;12(9):2132. doi:10.3390/diagnostics12092132.

[19]. Bhattacharyya A, Bhaik D, Kumar S, andal."A deep learning-based approach for automatic detection of COVID-19 cases using chest X-ray images". Biomed Signal Process Control, 2022;71:103182. doi:10.1016/j.bspc.2021.103182.

[20]. Debelee TG, Schwenker F, Ibenthal A, andal."Survey of deep learning in breast cancer image analysis". Evol Syst, 2019;11(1):143–63. doi:10.1007/s12530-019-09297-2.

[21]. Bai J, Posner R, Wang T, andal."Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: a review". Med Image Anal, 2021;71:102049. doi:10.1016/j.media.2021.102049.Roy S, Whitehead TD, Li S, andal."Co-clinical FDG-PET radiomic signature in predicting response to neoadjuvant chemotherapy in triple-negative breast cancer". Eur J Nucl Med Mol Imaging 2021;49(2):550–62. doi:10.1007/s00259-021-05489-8.

[22]. S. Jain, S. Andronikou, P. Goussard and al., "Advanced imaging tools forchildhood tuberculosis: potential applications and research needs,"Lancet Infect Dis., vol. 20, no. 11, pp. e289–e297, 2020

[23]. Blumenfeld A, Greenspan H, Konen E. "Pneumothorax detection in chest radiographs using convolutional neural networks". In: Mori K, Petrick N, editors. Medical Imaging 2018: Computer-Aided Diagnosis. Houston : SPIE ; 2018. p. 3.

[24]. Santosh K, Allu S, Rajaraman S, andal."Advances in deep learning for tuberculosis screening using chest X-rays: the last 5 years review". J Med Syst, 2022;46(11). doi:10.1007/s10916-022-01870-8.

[25]. Singh M, Pujar GV, Kumar SA, andal."Evolution of machine learning in tuberculosis diagnosis: a review of deep learning-based medical applications". Electronics, 2022;11(17):2634. doi:10.3390/electronics11172634

[26]. Ahsan M, Gomes R, Denton A. "Application of a convolutional neural network using transfer learning for tuberculosis detection". In: Proceedings of 2019 IEEE International Conference on Electro Information Technology (EIT);2019. https://doi.org/10.1109/eit.2019.8833768

[27]. Michael Norval, Zenghui Wang, Yanxia Sun, "Pulmonary Tuberculosis Detection Using Deep Learning Convolutional Neural Networks", ICVIP 2019: 2019 the 3rd International Conference on Video and Image Processing, DOI: 10.1145/3376067.3376068

[28]. Samuel, R. D. J., & Kanna, B. R. (2019). "Tuberculosis (TB) detection system using deep neural networks". Neural Computing and Applications, 31(5), 1533-1545.
[29]. B. Sandhiya1, R.Punniyamoorthy, Saravanan.B3, Vijay Prabhu.R, Subhash.V, "Deep Learning in Tuberculosis Diagnosis: A Survey", European Journal of Molecular & Clinical Medicine, Volume 7, Issue 4, 2020, ISSN 2515-8260
[30]. Khan, M. T., Kaushik, A. C., Ji, L., Malik, S. I., Ali, S., & Wei, D. Q. (2019). "Artificial neural networks for prediction of tuberculosis disease. Frontiers in microbiology".10.395
[31]. Khairu Munadi, Kahlil Muchtar, Novi Maulina and Biswajeet Pradhan, "Image Enhancement for Tuberculosis Detection Using Deep Learning", IEEE Acess, 2020, DOI:10.1109/ACCESS.2020.3041867
[32]. Xie Y, Wu Z, Han X, andal."Computer-aided system for the detection of multicategory pulmonary tuberculosis in radiographs". J HealthcEng.2020;2020:9205082.https://doi.org/10.1155/2020/9205082
[33]. Sahlol AT, Abd Elaziz M, Tariq Jamal A, and al."A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimization of deep neural network features". Symmetry, 2020;12(7):1146. https://doi.org/10.3390/sym12071146
[34]. Stefanus Kieu Tao Hwa, Abdullah Bade, Mohd Hanafi Ahmad Hijazi, Mohammad Saffree Jeffree, "Tuberculosis detection using deep learning and contrast enhanced canny edge detected X-Ray images", IAES International Journal of Artificial Intelligence (IJ-AI), Vol. 9, No. 4, December 2020, pp. 713~720, ISSN: 2252-8938, DOI: 10.11591/ijai.v9.i4.pp713-720
[35]. Tawsifur Rahman, Amith Khandakar, Muhammad Abdul Kadir andal., "Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization", IEEE Access, 2020, https://doi.org/10.1109/access.2020.3031384
[36]. Ul Abideen Z, Ghafoor M, Munir K, andal."Uncertainty assisted robust tuberculosis identification with Bayesian convolutional neural networks". IEEE Access.2020; 8:22812–25, https://doi.org/10.1109/ACCESS.2020.2970023
[37]. Ayaz M, Shaukat F, Raja G. "Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors". Phys Eng Sci Med 2021;44(1):183–94.https://doi.org/10.1007/s13246-020-00966-0
[38]. Vasundhara Acharya, Gaurav Dhiman, Krishna Prakasha et al, "AI-Assisted Tuberculosis Detection and Classification from Chest X-Rays Using a Deep Learning Normalization-Free Network Model", Computational Intelligence and Neuroscience, Hindawi, Volume 2022, Article ID 2399428, 19 pages, https://doi.org/10.1155/2022/2399428
[39]. Fati, S.M.; Senan, E.M.; ElHakim, N. "Deep and Hybrid Learning Technique for Early Detection of Tuberculosis Based on X-ray Images Using Feature Fusion". Appl. Sci. 2022, 12, 7092. https://doi.org/10.3390/app12147092
[40]. SuciAulia, SugondoHadiyoso, "Tuberculosis Detection in X-Ray Image Using Deep Learning Approach with VGG-16 Architecture",JurnalIlmiah Teknik ElektroKomputer dan Informatika (JITEKI), Vol. 8, No. 2, June 2022, pp. 290-297, ISSN: 2338-3070, https://dx.doi.org/10.26555/jiteki.v8i2.23994
[41]. Rohan Rana, Abhishek Gaba, "Tuberculosis Detection Using Deep Learning", International Journal of Advances in Engineering and Management (IJAEM), Volume 4, Issue 8 Aug. 2022, pp: 317-323 www.ijaem.net ISSN: 2395-5252, DOI: 10.35629/5252-0408317323
[42]. Leong Mei, Prasad Dilip, Lee Yong Tsui, Lin Feng, "Semi-CNN Architecture for Effective Spatio-Temporal Learning in Action Recognition" 2020, Applied Sciences, Vol 10, DOI:10.3390/app10020557
[43]. B. Uma Maheswari, Dahlia Sam, Nitin Mittal, Abhishek Sharma, Sandeep Kaur, S. S. Askar and Mohamed Abouhawwash, "Explainable deep-neural-network supported scheme for tuberculosis detection from chest radiographs", Maheswari andal. BMC Medical Imaging (2024) 24:32, https://doi.org/10.1186/s12880-024-01202-x
[44]. Vinayak Sharma, Nillmani, Sachin Kumar Gupta, Kaushal Kumar Shukla, "Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images", Intelligent Medicine 4 (2024)

104–113, https://doi.org/10.1016/j.imed.2023.06.001

[45]. James Devasia, Hridayanand Goswami, Subitha Lakshminarayanan, Manju Rajaram, Subathra Adithan, "Deep learning classification of active tuberculosis lung zones wise manifestations using chest X‑rays: a multi label approach", Scientific Reports, (2023) 13:887, https://doi.org/10.1038/s41598-023-28079-0

[46]. S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wáng, P.-X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," Quantitative imaging in medicine and surgery, vol. 4 (6), p. 475(2014)

[47]. B. P. Health. (2020). Belarus Tuberculosis Portal [Online]. Available: https://grantome.com/grant/NIH/AAI12021001-1-0-5[Accessed on 13 June 2024]

[48]. RSNA Pneumonia Detection Challenge [online],avalaible: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data, [Accessed on 20 June 2024]

[49]. Kang, H.-J. SSDLiteX: Enhancing SSDLite for Small Object Detection. Appl. Sci. **2023**, 13,12001. https://doi.org/10.3390/app132112001

[50]. Mark Sandler, Andrew Howard andal., MobileNetV2: Inverted Residuals and Linear Bottlenecks, arXiv:1801.04381v4 [cs.CV] 21 Mar 2019

[51]. Andrew Howard, Mark Sandler, Grace Chu andal., Searching for MobileNetV3, arXiv:1905.02244v5 [cs.CV] 20 Nov 2019