Detection, Localization of Cardiomegaly and TB Disease of CXR Images using Deep Learning

Ganesh Pradeep P V[1] , R.Dinesh[2], Anwesh P R[2], and Narayana Darapaneni[3]

[1] PES University, Bangalore, Karnataka, 560050, India
[2] Great Learning, Hyderabad, Telangana, 500089, India
[3] Northwestern University, Evanston, IL 60208, United States

Abstract

INTRODUCTION: Tuberculosis (TB) continues to pose a significant worldwide public health concern, as it stands as the primary contributor to mortality stemming from infectious illnesses. Cardiomegaly, characterized by an enlarged heart, poses medical concern as well.

OBJECTIVES: Timely identification of Cardiomegaly is vital for effective management. Chest X-ray diagnosis is an easily available method with less radiation exposure to detect several lung infections and heart enlargement. Utilizing computer-aided diagnostic systems can aid in the early detection of lung conditions and the enlargement of the heart.

METHODS: We worked on different state-of-the-art CNN architectures such as VGG, DenseNet and EfficientNet with customization over dataset generated from combination of multiple publicly available datasets, which consists of 12939 annotated images across three different categories, one being normal and other two being TB and cardiomegaly diseases..

RESULTS: EfficientNetB5 with optimization has shown excellent results amongst others in classifying Tuberculosis and Cardiomegaly with a remarkable accuracy of 97%.

CONCLUSION: The proposed model is ready for clinical diagnosis and triaging of X-ray images. Our solution also offers efficient ways to show the presence of the above diseases using Grad-CAM technique.

*Corresponding author. Email: anwesh@greatlearning.in

1. Introduction

Tuberculosis (TB) is a pressing global health issue affecting numerous countries worldwide. It ranks among the foremost infectious disease-related fatalities globally, alongside conditions such as HIV/AIDS. TB is triggered by the bacterium Mycobacterium tuberculosis, primarily afflicting the respiratory system (pulmonary TB), but it can also manifest in other body parts (extrapulmonary TB). The World Health Organization (WHO) aggregates worldwide TB data. In recent times, there has been an annual report of approximately 10 million fresh TB cases, encompassing both new and relapse cases. TB persists as a primary contributor to mortality stemming from infectious illnesses, with an estimated 1.5 million individuals succumbing to it in 2020. Cardiomegaly, also recognized as an enlarged heart, is a medical condition defined by the enlargement of the heart. Rather than being a standalone ailment, it serves as an indicator or consequence of an underlying cardiac ailment or other health concerns. Cardiomegaly can impact the heart's chambers, walls, or both, and its enlargement can be prompted by various factors such as chronic medical conditions, heart-related illnesses, hypertension (high blood pressure), heart valve irregularities, coronary artery disease, congenital heart anomalies, or specific infections. Utilizing chest X-rays, which offer a readily available diagnostic method with minimal radiation exposure, can help identify several lung infections and the presence of an enlarged heart.

Additionally, CAD systems can facilitate the early detection of these conditions.

Deep convolutional neural network (CNN) architectures have displayed significant potential within the realm of medical image analysis, particularly in the domain of lung disease detection and diagnosis. The application of deep learning models for diagnosing lung diseases holds the potential to enhance diagnostic accuracy, expedite the detection process, and ultimately contribute to improved patient outcomes. Through harnessing the computational capabilities of neural networks and sophisticated algorithms, these deep learning models can efficiently process vast volumes of medical data, revealing intricate patterns that may elude human observation. In this context, a deep learning model tailored for lung disease diagnosis has the potential to redefine our approach to the detection and management of lung-related conditions.

Introducing our state-of-the-art AI/ML model, specifically designed to address the growing concern of lung diseases. Utilizing sophisticated machine learning algorithms and deep neural networks, our model can predict the presence of lung and heart diseases with an exceptional level of accuracy. Whether it's Tuberculosis or Cardiomegaly, our model has been trained on reasonable amounts of data and is capable of providing reliable and actionable insights to healthcare professionals and patients alike. By leveraging the power of AI, we hope to revolutionize the way tuberculosis and cardiomegaly diseases are diagnosed and treated, ultimately leading to better outcomes for patients worldwide.

2. Literature Review

Chest X-rays are frequently employed as an initial imaging examination to assess lung disorders due to their ability to provide a comprehensive view of the lungs and surrounding anatomical structures. They serve in the detection of various lung ailments, including pneumonia, lung cancer, tuberculosis, chronic obstructive pulmonary disease (COPD), and other conditions affecting the chest and lungs. A chest X-ray is capable of revealing alterations in lung tissue, such as the presence of fluid or air, the development of nodules or masses, and the enlargement of lymph nodes. Additionally, it can highlight variations in the size or configuration of the heart, which may indicate specific cardiac conditions. Artificial intelligence (AI) can play a pivotal role in the analysis of chest X-rays by assisting radiologists in the identification and diagnosis of lung diseases. AI algorithms can be trained using extensive datasets of chest X-rays to detect irregularities such as pneumonia, lung opacities, or other lung conditions that may suggest the presence of a pulmonary ailment.

In their research, Nafisah et al. [1] introduced an automated tuberculosis (TB) detection system utilizing a deep learning model based on EfficientNetB3. This system incorporates advanced segmentation networks to identify regions of interest in chest X-rays (CXRs). Their findings showed considerable promise, with EfficientNetB3 achieving an impressive accuracy rate of 98.7% and an ROC score of 99.9%. Nonetheless, there is an opportunity for enhancement in terms of the F1-Score, and it is important to note that the model's training dataset size was relatively limited. D. Capellán-Martín et al. [2] introduced a multi-view deep learning solution aimed at automating the identification and extraction of lung and mediastinal regions of interest in pediatric chest x-rays, particularly in cases with potential TB findings. This approach employed Yolov5 for lung region detection, followed by the nnUnet algorithm for lung region segmentation. The subsequent classification of images as normal or indicative of tuberculosis was performed using a ResNet deep learning model. Although this method shows promise, it should be noted that it was trained and tested on a limited number of images.

Research by Alexander Wong et al. [3] came up with a specialized deep convolutional neural network (CNN) called TB-Net, which incorporates self-attention mechanisms explicitly tailored for tuberculosis (TB) screening. TB-Net underwent training and testing using chest X-ray (CXR) data collected from a diverse patient population across multiple countries. The authors adopted a machine-driven design exploration approach, employing generative synthesis techniques to optimize the neural network architecture. Rigorous performance evaluation of TB-Net revealed outstanding accuracy, sensitivity, and specificity, achieving values of 99.86%, 100.0%, and 99.71%, respectively. In a separate study, conducted by S. Stirenko et al. [4], the focus was on the analysis of CXR images for tuberculosis detection through computer-aided diagnosis using deep learning, emphasizing 2D image analysis techniques. The study showcased the effectiveness of lung segmentation and various data augmentation methods, even with a limited and imbalanced dataset. Importantly, the deep CNN demonstrated the ability to mitigate overfitting when applied to the pre-processed dataset obtained after lung segmentation, as opposed to the original dataset without segmentation.

Wasunan Chokchaithanakul et al. [5] conducted a study that delved into the examination of dataset inconsistencies within the realm of chest radiography. In their research, they introduced the lung balance contrast enhancement (lung BCET) method, which automatically identifies lung regions and normalizes images to enhance resilience when dealing with out-of-domain data. Their experiments revealed that when lung BCET was combined with conventional augmentation techniques, it led to performance enhancements in both in-domain and out-of-domain scenarios. Meanwhile, V. Acharya et al. [6] directed their efforts toward the development of Normalization-free networks (NFNets) initially trained on the ImageNet dataset and subsequently fine-tuned for Tuberculosis classification. They employed the Score-Cam algorithm to emphasize specific regions within images for precise disease detection. This approach exhibited promising outcomes in binary and multi-class classification, achieving high levels of AUC, precision, recall, accuracy sensitivity and specificity.

T. Rahman et al. [7] introduced a transfer learning strategy employing a variety of CNN models to automate tuberculosis detection from chest radiographs. Their investigation highlighted a notable enhancement in classification accuracy, particularly through the process of image segmentation, with a specific focus on lung segmentation, a fact supported by Score-CAM visualizations. Abbas Jafar et al. [8] introduced a novel segmentation network called CardioNet, specifically designed for fine segmentation of anatomical chest structures. This network demonstrated excellent accuracy for heart and lung segmented regions, contributing to the calculation of the cardiothoracic ratio (CTR).

S. S. Sarpotdar et al. [9] conducted an investigation into the practicality of employing transfer learning techniques in the realm of Cardiomegaly detection using X-ray images. They crafted a customized and retrained U-Net model, which yielded an impressive diagnostic accuracy rate of 94%, surpassing the performance of pre-existing models in identifying Cardiomegaly disease. On a related note, I. Chamveha et al. [10] introduced an algorithm designed to automate the computation of the cardiothoracic ratio (CTR) from chest X-ray films. This method harnessed the capabilities of deep learning to extract lung and heart masks and subsequently calculate the CTR, leading to a significant reduction in the time and effort required by radiologists.

These studies collectively contribute to the evolving landscape of medical image analysis, offering innovative solutions to improve the accuracy and efficiency of disease diagnosis and prognosis through deep learning and image processing techniques. However, there is still scope of improvement in the performance metrics of such models using balanced datasets, latest deep learning models, tuning hyper parameters and better data augmentation techniques. In the further sections we discuss about our dataset, approach to build a deep learning model using which we can segment the chest x-ray images, detect the Tuberculosis and Cardiomegaly disease, determine the Cardio thoracic ratio from mask images and visualize the disease localized areas on the Chest X-ray images.

3. Materials And Methods
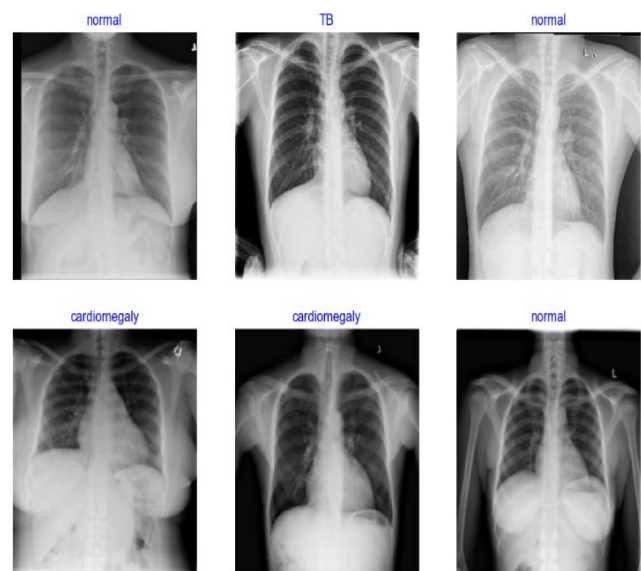
3.1 Dataset

In this study, we have used two different datasets, one dataset for segmentation model and the other for classification model training and validation. Chest X-ray segmentation dataset is with raw images and corresponding mask images. We have created a dataset for classification model training by pulling images from multiple publicly available datasets such as National Institute of Health (NIH) USA [17], TB Radiography database [7], SCH and Montgomery datasets [13][14] and TBX11K dataset [19] of variable sizes 128*128 to 4020*4892.

The classification dataset consists of below number of annotated images as follows:

- 5000 Normal images

- 4626 Tuberculosis images

- 3313 Cardiomegaly images

In the NIH database, the number of cardiomegaly annotated chest x-ray pictures were 1094. In order to reduce this imbalance in cardiomegaly dataset, 2219 cardiomegaly annotated images from a Kaggle dataset were added to our dataset. These images were treated with CLAHE algorithm to increase the amount of brightness and contrast. We divided our dataset as per ratio 70:15:15 for training, testing and validation to facilitate the training and validation. Figure 1 shows the chest x-ray images belong to different classes from our dataset.

Figure 1 Chest x-ray Images from classification dataset

For segmentation model training, we created heart and lung segmentation datasets. Each of the dataset comprises of raw chest x-ray images and organ specific mask images. For e.g., heart segmentation dataset comprises of chest x-ray images and heart mask images. The lung segmentation dataset is created by pulling images from Shenzhen and Montgomery datasets [13][14]. The dataset is segregated as 800 Normal images and 704 Lung mask images. Figure 2 shows the Lung Image and corresponding mask image for both dataset sources.

The heart segmentation dataset is created by pulling images from JSRT dataset [15] and PadChest dataset [16]. These datasets only contained raw images. Maks images were not present. Therefore we have utilized the annotated landmarks from the research HybridGNet[17] and generated the mask images for the JSRT and Padchset images programmatically. The heart segmentation dataset is split equally as 383 normal and 383 mask images. Figure 3 shows the Chest x-ray image and corresponding mask image from the dataset sources. Both the heart and lung segmentation datasets were split as 80% for training and 20% for test.
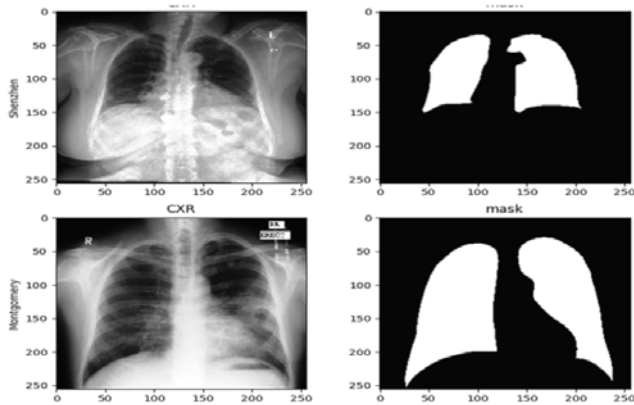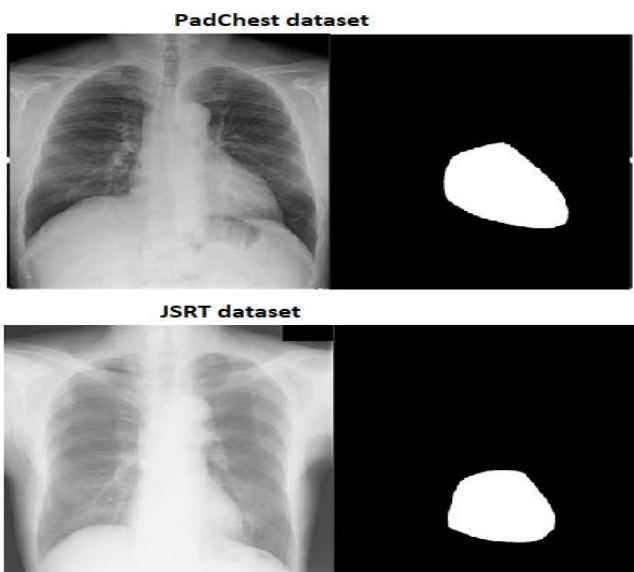
Figure 2 Lung Segmentation dataset

Figure 3 Heart segmentation dataset

3.2. Development Environment

For the classification and segmentation deep learning model training we have used Google Colab Pro and below GPU configuration.
- NVIDIA Tesla T4, for segmentation
- NVIDIA Tesla V100, for classification model

Pytorch 2.0.1 version has been used for segmentation model training and Tensorflow 2.9.1 has been used for classification model training. Python 3.10.2 has been the programming language for both the models. Additional frameworks used are numpy, opencv and matplotlib

3.3 Data Pre-processing

Data augmentation and data generation are crucial techniques in machine learning and data science as they can enhance the size and variety of a dataset, ultimately improving the performance of machine learning models. Data augmentation involves modifying existing data samples by applying transformations such as flipping, rotating, cropping, zooming, or adding noise to images or audio signals, creating new but similar samples. The model can then learn to recognize patterns in various contexts, reducing overfitting and improving its generalization ability. Data generation, on the other hand, involves creating entirely new data samples that resemble the original data but are not identical copies. This approach is particularly useful in situations where there is limited or no real-world data available, such as in the development of synthetic datasets or simulations. Generating new data samples can help train models that are robust to a wider range of scenarios, improving their accuracy and reliability. Ultimately, data augmentation and data generation are essential tools for enhancing the quality and diversity of datasets, leading to better performance and more accurate predictions from machine learning models.

In our proposed classification model, we have reduced the shape of the input images to 224*224*3 to meet the input shape requirement of underlying base model without compromising in quality and resolution of the image. We have applied the following augmentation techniques such as horizontal flip, rescale to add the variety in the training of the model. In addition, we rescaled the test and validation data as well. Figure 1 shows the chest x-ray images for each of the pathology.

In our proposed segmentation model, we have reduced the shape of the input images to 224*224*3 to meet the input shape requirement of underlying base model without compromising in quality and resolution of the image. We have applied the following augmentation techniques such as resizing the image with bilinear interpolation, adjusting brightness and contrast to 0.35, rotating the images to 35 degrees. Scaling and shifting the images by 0.35 and horizontally flipping the images.

3.4 Proposed Methodology

TB can affect the human respiratory system very badly, and cardiomegaly can affect the heart function. Both these diseases can be diagnosed using chest x-ray images. Recently, deep learning algorithms have become increasingly important in identifying and classifying these diseases, thus saving time for healthcare providers. To address this issue, our study has proposed a multi class deep learning classification model that aims to identify Tuberculosis and cardiomegaly. The research work focuses on designing a deep learning framework to classify these diseases. In addition, a segmentation model is designed to generate the lung and heart masks for better diagnosis of cardiomegaly and tuberculosis.

To develop the most efficient and high-performance classification and segmentation models, we began with some of the top-performing models that have dominated the ImageNet dataset with high accuracy, while also requiring minimal computation resources. These models come with pre-trained weights that can be readily accessed using transfer learning techniques. During our literature survey, we discovered various models, but with our optimized and innovative approach, we were able to surpass some of the most renowned architectures. Our method as shown in figure

4 resulted in a significant increase in accuracy and f1-score of classification model while reducing the validation loss, despite having only a few million trainable parameters and training for just 30 epochs for classification model.

We implemented U-Net based segmentation model with ResNet34 as the encoder using Pytorch framework on Google Colab with T4 GPU. We applied the Adam optimizer with a learning rate set at 0.005 and employed a learning rate scheduler known as CosineAnnealingLR. The output layer was equipped with a sigmoid activation function, and we utilized the dice loss as our chosen loss function. During the training phase, the segmentation model was trained on batches of 64 images for a total of 100 epochs. To assess the model's performance during training, we relied on the IOU metric.

For evaluating the predictions against their corresponding ground truth masks in the test images, we considered a range of metrics, including the dice coefficient, IOU, accuracy, precision, and recall. The performance results for the heart and lung segmentation models can be found in [Table 1]. Using heart and lung masks we have also calculated the Cardiothoracic ratio (CTR) to check if the patient chest x-ray shows any signs of heart enlargement. Figure 10 shows the CTR ratio of a patient with heart enlargement.

With regard to classification model, we implemented customized models based on some of the popular ImageNet pre-trained models such as VGG19, DenseNet201 and EfficientNetB5. In the process of customization, we unfroze few layers of the pre-trained model and added our custom new layers at the top. We trained all our models with the optimized batch size of 16. We compiled the previously mentioned models with highly effective optimizers, specifically Adam and SGD, utilizing a learning rate of 0.0001. Our chosen loss function was categorical cross-entropy, which quantifies the dissimilarity between the predicted probability distribution and the actual probability distribution of the classes. This loss function imposes a significant penalty on the model when it assigns a low probability to the correct class while assigning high probabilities to other classes.

We used several optimization, regularization and memory optimization technique such as ReduceLROnPlateau, Early stopping callback and Model Checkpoint respectively for faster convergence and reduce overfitting. We implemented all these models by adding custom layers at the top followed by above mentioned optimization and regularization techniques. Our investigation revealed that the optimized model built upon EfficientNetB5 surpassed the performance of the other models, particularly in terms of test accuracy and F1-score. This achievement led to impressive outcomes, including an overall accuracy rate of 97% and an F1-score of 99% for TB, 98% for Viral Pneumonia, 95% for Cardiomegaly, and 96% for Normal images. [Table 2] shows the EfficientNetB5 based optimized model performance results and [Table 3] shows the multiple classification model performance results.

Furthermore, we incorporated Gradient Class Activation Maps (Grad CAM) [12] into our approach. This technique facilitates the identification of pivotal image regions responsible for a Convolutional Neural Network (CNN) predicting a particular class. Grad CAM represents an advanced iteration of the Class Activation Map (CAM) method, as it improves the localization of crucial regions by integrating gradient information related to the target class within the feature maps. Figures 7,8 show the Grad CAM heatmaps of TB and cardiomegaly classified chest x-ray images.
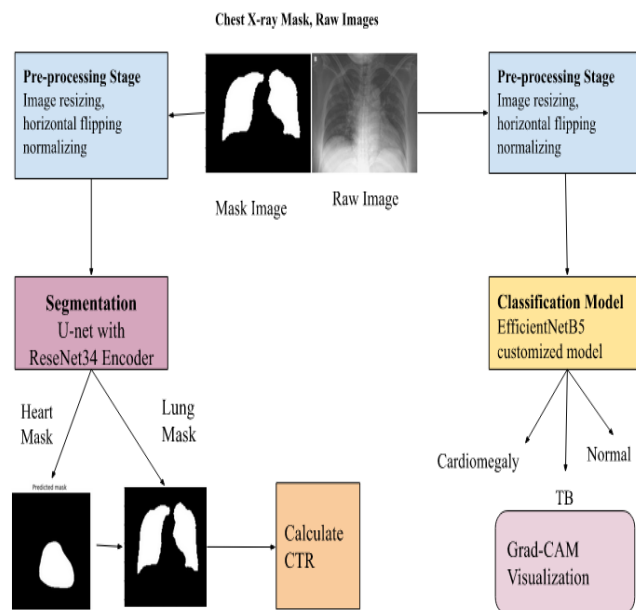

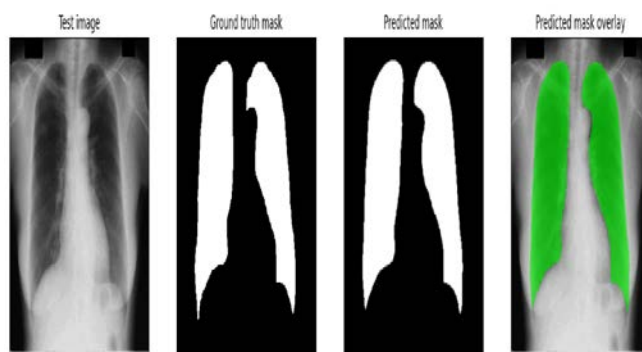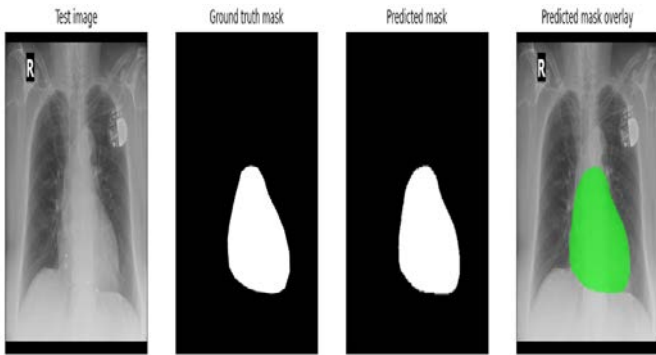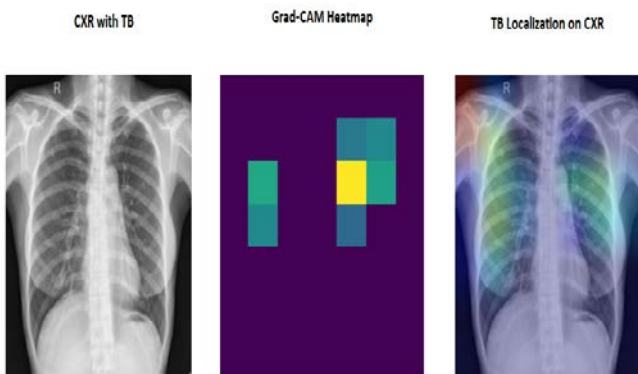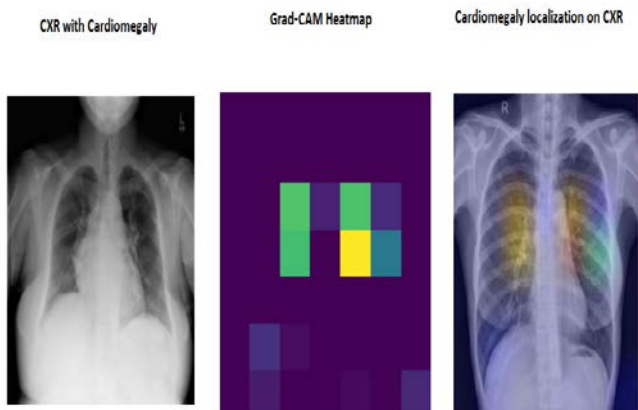**Figure 4** Proposed System for Cardiomegaly and TB detection

Figure 5 Lung Segmentation model prediction

Figure 6 Heart segmentation model prediction

Figure 7 Grad-CAM visualization of TB patient

Figure 8 Grad-CAM visualization of Cardiomegaly patient

4. Results

To construct a deep learning model designed for the segmentation of the heart and lung regions within chest X-rays, we leveraged Python 3 and the PyTorch framework. Our implementation took place on Google Colab. In the preprocessing phase, we harnessed PyTorch's Dataset class for tasks such as data augmentation and conversion to tensors.

The resulting preprocessed data served as the input for our proposed deep learning model, which was specifically

engineered for segmenting the lung and heart areas within chest X-rays.Our model underwent training and validation using suitable training algorithms and optimization techniques, with the Adam optimizer utilizing a learning rate of 0.0005. The segmentation model was trained for a total of 100 epochs, with each training batch consisting of 64 images. [Table 1] shows our model performance metrics against test dataset.

To create a deep learning model for classifying tuberculosis and cardiomegaly diseases, we utilized Python 3 and the Keras framework, and implemented our proposed model on Google Colab. During the pre-processing stage, we used the ImageDataGenerator class in Keras for data augmentation and conversion to a data array. The pre-processing step's output was utilized to create the input for the suggested deep learning model for multi-classification of tuberculosis and cardiomegaly illnesses. We trained and validated the model using appropriate fit algorithms and optimizers such as Adam and R-Adam. Number of epochs our base model trained was 15 and with batch size of 16 images. For our final model based on EfficientNetB5, we employed the R-Adam optimizer with a learning rate of 0.0001 (LR) and conducted training over the course of 6 epochs. The evaluation of the model's performance was carried out by assessing metrics such as Precision, Recall, F1-Score, accuracy, and validation loss. [Table 2].

Table 1 U-Net Segmentation model performance metrics

Model (U-Net with ResNet34 Encoder)	Dice mean	IOU mean	Accuracy Mean	Precision Mean	Recall Mean
Heart Segmentation	0.9607	0.9132	0.988	0.924	0.986
Lung Segmentation	0.9414	0.8820	0.9612	0.9193	0.956

Table 2 EfficientNetB5 Optimized model performance results

Class	Model	F1	Precision	Recall
TB	EfficientNetB5 customized model with Accuracy 0.97	0.99	0.99	0.98
Cardiomegaly		0.95	0.94	0.95
Normal		0.96	0.95	0.95

Cardiothoracic Ratio : 0.6291980808773132

Figure 9 CTR of patient with increased heart

Table 3 Classification Model comparison

Model	Optimizer	Best Epoch	Test Accuracy	F1-TB	F1-Card
VGG19	Adam	15	0.78	0.83	0.77
DenseNet 201	SGD	13	0.85	0.9	0.81
EfficientNetB5	Adam	10	0.96	0.99	0.93
EfficientNetB5	SGD	4	0.963	0.99	0.94
EfficientNetB5	R-Adam	1	0.965	0.99	0.95

5. Discussions And Conclusions

The objective of this research was to develop a deep convolutional neural network (CNN) model that could accurately distinguish between Tuberculosis and Cardiomegaly diseases using chest X-ray images. To achieve this, we implemented various state-of-art Computer Vision architectures, pre-trained models. Out of the whole lot shown in [Table3], EfficientNetB5 with RAdam optimization demonstrated outstanding results with a remarkable overall accuracy of 97% and F1-scores of 99% for Tuberculosis, 95% for Cardiomegaly, 96% for normal class. Grad CAM technique was implemented to visualize the infected area of the lung. In addition to the objective was to develop a segmentation model to generate the lung and heart masks and later use these masks to confirm if the patient has enlarged heart condition. U-Net with ResNet34 encoder has been used to generate heart and lung masks. The proposed model was able to accurately generate the heart and lung masks with an accuracy of 0.98 for heart and 0.96 for lung portion from chest x-ray. Enlarged heart condition was successfully detected using the masks.

Although most of the targeted goals have been met, there is still scope of improvement of the classification model accuracy and better visualization disease localized areas on the heatmaps.

Future directions of the work include

- Apply data pre-processing techniques such as center crop, resize, enhance the quality of the images by increasing brightness, contrast
- Improve the classification model accuracy using latest deep learning architectures and tuning hyper parameters
- Improve the Visualization of disease localized areas with latest Class activation map methods.

References

[1] S. I. Nafisah and G. Muhammad, "Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence," Neural Comput. Appl., pp. 1–21, 2022.

[2] D. Capellán-Martín et al., "Deep learning-based lung segmentation and automatic regional template in chest X-ray images for pediatric tuberculosis," arXiv [cs.CV], 2023.

[3] A. Wong, J. R. H. Lee, H. Rahmat-Khah, A. Sabri, A. Alaref, and H. Liu, "TB-Net: A tailored, self-attention deep convolutional neural network design for detection of tuberculosis cases from chest X-ray images," Front. Artif. Intell., vol. 5, p. 827299, 2022.

[4] S. Stirenko et al., "Chest X-ray analysis of tuberculosis by deep learning with segmentation and augmentation," in 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO), 2018.

[5] W. Chokchaithanakul, P. Punyabukkana, and E. Chuangsuwanich, "Adaptive image preprocessing and augmentation for tuberculosis screening on out-of-domain chest X-ray dataset," IEEE Access, vol. 10, pp. 132144–132152, 2022.

[6] V. Acharya et al., "AI-assisted tuberculosis detection and classification from chest X-rays using a deep learning Normalization-free network model," Comput. Intell. Neurosci., vol. 2022, p. 2399428, 2022.

[7] T. Rahman et al., "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization," IEEE Access, vol. 8, pp. 191586–191601, 2020.

[8] A. Jafar, M. T. Hameed, N. Akram, U. Waqas, H. S. Kim, and R. A. Naqvi, "CardioNet: Automatic semantic segmentation to calculate the cardiothoracic ratio for cardiomegaly and other chest diseases," J. Pers. Med., vol. 12, no. 6, p. 988, 2022.

[9] S. S. Sarpotdar, "Cardiomegaly detection using Deep convolutional neural network with U-Net," arXiv [eess.IV], 2022.

[10] I. Chamveha, T. Promwiset, T. Tongdee, P. Saiviroonporn, and W. Chaisangmongkon, "Automated cardiothoracic ratio calculation and cardiomegaly detection using deep learning approach," arXiv [eess.IV], 2020.

[11] I. Allaouzi and M. Ben Ahmed, "A novel approach for multi-label chest X-ray classification of common thorax diseases," IEEE Access, vol. 7, pp. 64279–64288, 2019.

[12] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in 2017 IEEE International Conference on Computer Vision (ICCV), 2017.

[13] S. Jaeger et al., "Automatic tuberculosis screening using chest radiographs," IEEE Trans. Med. Imaging, vol. 33, no. 2, pp. 233–245, 2014.

[14] S. Candemir et al., "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," IEEE Trans. Med. Imaging, vol. 33, no. 2, pp. 577–590, 2014.

[15] J. Shiraishi et al., "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," AJR Am. J. Roentgenol., vol. 174, no. 1, pp. 71–74, 2000.

[16] A. Bustos, A. Pertusa, J.-M. Salinas, and M. de la Iglesia-Vayá, "PadChest: A large chest x-ray image dataset with multi-label annotated reports," arXiv [eess.IV], 2019.

[17] N. Gaggion, L. Mansilla, C. Mosquera, D. H. Milone, and E. Ferrante, "Improving anatomical plausibility in medical image segmentation via hybrid graph neural networks: applications to chest x-ray analysis," arXiv [eess.IV], 2022

[18] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[19] Y. Liu, Y.-H. Wu, Y. Ban, H. Wang, and M.-M. Cheng, "Rethinking computer-aided tuberculosis diagnosis," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.