

Detection of Tuberculosis Based on Hybridized Pre-Processing Deep Learning Method

Mohamed Ahmed Elashmawy¹, Irraiyan Elamvazuthi², Lila Iznita Izhar³, Sivajothi Paramasivam⁴, Steven Su⁵

Smart Assistive and Rehabilitative Technology (SMART) Research Group,

Department of Electrical and Electronic Engineering,

Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 32610, Malaysia^{1, 2, 3}

School of Engineering, UOWMKDU University College, Shah Alam, 40150, Malaysia⁴

School of Biomedical Engineering, University of Technology Sydney, Ultimo, 2007, Australia⁵

Abstract—The disease, tuberculosis (TB) is a serious health concern as it primarily affects the lungs and can lead to fatalities. However, early detection and treatment can cure the disease. One potential method for detecting TB is using Computer Aided Diagnosis (CAD) systems, which can analyze Chest X-Ray Images (CXR) for signs of TB. This paper proposes a new approach for improving the performance of CAD systems by using a hybrid pre-processing method for Convolutional Neural Network (CNN) models. The goal of the research is to enhance the accuracy and Area Under Curve (AUC) of detection for TB in CXR images by combining two different pre-processing methods and multi-classifying different manifestations of the disease. The hypothesis is that this approach will result in more accurate detection of TB in CXR images. To achieve this, this research used augmentation and segmentation techniques to pre-process the CXR images before feeding them into a pre-trained CNN model for classification. The VGG16 model managed to achieve an AUC of 0.935, an accuracy of 90% and a 0.8975 F1-score with the proposed pre-processing method.

Keywords—Tuberculosis; CNN; pre-processing; CXR images; augmentation; segmentation

I. INTRODUCTION

Tuberculosis, also known as TB, is an infectious disease brought on by the bacterium *Mycobacterium tuberculosis*. Although it mostly affects the lungs, it can also have an impact on the kidneys, the spine, and the brain.

When an infected individual coughs, sneezes, or talks and another person inhales the bacteria, TB is transmitted through the air. It is crucial to remember that TB cannot be transmitted via innocuous touch such as handshakes or sharing of utensils. Chronic cough, chest pain, blood in the cough, exhaustion, fever, night sweats, and weight loss are some of the signs of TB. Yet, some TB patients may not even exhibit any symptoms [1]. Although TB is a treatable and curable illness, a lengthy antibiotic treatment regimen is necessary. TB can be fatal if neglected.

According to the World Health Organization (WHO), tuberculosis (TB) is one of the top 10 causes of death worldwide. In 2020, there were an estimated 10 million new cases of TB globally with an estimated 1.5 million deaths from TB in 2020. The global TB treatment success rate was 85% in 2019 and disproportionately affects vulnerable populations,

such as people living in poverty, people who use drugs, and prisoners [2].

Chest X-rays and CT (Computed Tomography) scans are two different medical imaging techniques used to diagnose and monitor various conditions related to the chest, such as lung diseases. However, considering the disproportionality of TB affecting impoverished areas, chest X-rays have the advantage regarding its cost, availability, convenience and quick results [3].

Sputum microscopy, Chest X-rays (CXR), and culture in solid and liquid media can all be used to diagnose and find TB. CXR, one of the most popular and cost-effective imaging tests worldwide, can be utilized for TB early detection [3]. Although CXRs are helpful for making an early diagnosis, radiologists may encounter difficulties, such as the inability to tell TB from other symptoms in some situations [4], [5].

Computer aided diagnosis (CADx) and computer aided detection (CADE) systems also known as CAD systems, have been shown to improve the accuracy of medical diagnoses, reduce false positives and false negatives, and potentially reduce the time needed for interpretation. However, like any diagnostic tool, CADE/x systems are not infallible and must be used in conjunction with clinical judgment and expertise. Overall, CADE/x systems are a valuable tool in modern medicine, helping medical professionals to make more accurate diagnoses and improve patient outcomes [6]. A CAD system traditionally consisted of four main stages; pre-processing, segmentation, feature extraction and classification [7].

To optimize the system and produce higher accuracy, a wide variety of techniques and algorithms can be applied at each step, in various combinations [8]. Classifiers make the final determination regarding the patient's health state among the four stages of a CAD system. Information from earlier stages is compressed and filtered to acquire the information that is most pertinent to the patient's health and is then fed to classifiers [9]. Machine learning emerged in the field of computer science that enabled computers to classify data without being explicitly programmed. As the field of machine learning (ML) research developed, many classification algorithms, including Decision Trees, Support Vector Machines (SVM), Genetic Algorithms (GAs), and Fuzzy Algorithms (FA), flourished [10-12]. However, Convolutional

Neural Networks (CNNs), perhaps more significantly, have recently demonstrated their reliability [13].

CNN is a deep learning technique which makes it ideal for big data. The arrangement of the visual cortex in the brain served as the model for CNN construction. It is made up of multiple layers of linked neurons that work in a hierarchical fashion to interpret information ranging from basic features like edges and corners to more intricate forms and patterns. [14].

The objective of this research is to investigate the effects of combining two main methods of pre-processing methods such as shear, zoom and flipping for augmentation and the robustness of U-Net mask for segmentation on the training set of a pre-trained CNN classifier in hopes to a high accuracy and AUC with the proposed method of pre-processing on binary and multi-classification. Hence, being able to classify different types of manifestations of TB in CXR images.

The literature review is discussed in Section II. Section III describes the chest X-ray dataset that is used and the proposed methodology for pre-processing, the performance measurements are also explained in this section. Section IV discusses the experiments conducted and results obtained by applying the proposed methodology on the CXR dataset and the comparison with other papers. Finally, the proposed research work is concluded.

II. RELATED WORK

To identify any lung-related disease on CXRs, Computer Assisted Diagnosis (CAD) has been extensively used. Methods involving machine learning alternatives to CNN have been researched as CAD systems. Some methods of TB detection make use of the segmentation of the lungs. Here, scientists attempt to separate the heart or lung structures before assessing them for any anomalies. Other research implements augmentation methods onto the CXR images, changing the parameters of the images within a designated range before feeding it into the machine learning model to achieve more reliable results.

Related works show research on other models used aside from CNN with pre-processing for detection of TB in CXR images. Antony B. et al. [15] attempted eliminating background noise by applying two segmentation methods, the CANNY algorithm and a median filter on 662 images. The results reported only an 80% accuracy as the highest accuracy while using a K-NN classifier in his paper, the highest accuracy among the other two classification methods (SMO and SLR).

To detect Tuberculosis in patients, Muhathir [16] used the K-NN classification method and HOG feature extraction technique. The results indicate that 70.90% of positive cases were correctly identified, with 234 out of 330 samples, while 72.72% of negative cases were correctly identified, with 240 out of 330 samples. The study shows that using the K-NN and HOG feature approach, the X-ray Set TB can be classified with an accuracy of 71.81% when using cross-validation.

Three alternative deep-learning models—AlexNet, ResNet-18, and DenseNet121—were tested in Jared et al. [17] study to

see which was most effective in identifying tuberculosis (TB) in CXR. For their training set, the researchers used 180,000 images, but they made no mention of data pre-processing. According to the findings, DenseNet121 had the highest accuracy (91%), with an Area Under Curve (AUC) ranging from 0.94 to 0.96. Although Jared et al. [17] did not rely on augmented photos to increase the size of the dataset, the high accuracy and AUC were probably caused by the numerous training images employed.

For the purpose of TB identification in CXR pictures, Syeda et al. [18] developed an ensemble model using the pretrained deep-learning models VGG-16, VGG-19, ResNet50, and GoogleNet. The ensemble's accuracy was 86.7% with an AUC of 0.92 after training on 600 images and 200 more; however, if any pre-processing techniques were used, the accuracy might have been enhanced with a bigger data set and variance to prevent overfitting problems.

Gordienko et al. [19] reported an increase in accuracy and loss after segmentation of 247 TB CXR images with a U-Net CNN and Bone Shadow Exclusion and using them for training a self-made seven-layer CNN. No reports of augmentation have been applied to the image, and with a low number of images overfitting could occur. It has only been reported that the test accuracy has increased, and the test loss has decreased.

Erdal [20] proposes and compares in his study, three methods of segmentation: bounding box, lung mask with black background and lung mask with white background. AlexNet, VGG16 and VGG19 deep-learning architecture were utilized for feature extraction and a Random Forest algorithm for classification. Accuracy reached 88.3% and AUC reached 0.93 as reported. Erdal [20] also reported that the lack of contrast enhancements and augmentation have acted as a limitation in his research.

Only using the Montgomery County (MC) TB CXR images dataset, with 138 images total of both abnormal and normal CXR images, Mustapaha & Serestina [21], have managed to multiply the total images up 5000 images through augmentation. Through their proposed CNN model a 87.1% accuracy was achieved.

Oposing previous studies that have utilized segmentation as a pre-processing technique for lung segmentation before feeding the images to the training model, such as the studies done by Erdal [20] and Gordienko et al. [19], Ahsan et al. [22] displayed that it is possible to achieve a comparable accuracy using the VGG-16 model without pre-processing segmentation. An accuracy of 80% was reached by the VGG-16 model and an accuracy of 81.25% when partial augmentation was applied.

Marcio et al. [23] combined the Montgomery, Shenzhen and PadChest CXR datasets with a total of 290 images, generating the training and test datasets using a HDF5 dataset generator, then applied augmentation. Marcio et al. [23] tested out three different pretrained CNN models being AlexNet, GoogleNet and ResNet50, achieving results between 0.78 and 0.84 AUC.

Similarly, Eman et al. [24] has used Montgomery and the Shenzhen Datasets and has used augmentation to multiply the dataset images up to 2040 images. However, Eman et al. [24]

have explored a more specialized and advanced array of CNN models; ConvNet, Exception, ResNet50, VGG16 and VGG19. All CNN models have achieved an accuracy above 87% and a maximum of 90% and an AUC of 0.91.

To summarize the presented related work, it can be concluded that there is emphasis on augmentation for improving classification models, while sometimes in other related works there is neglect of preprocessing. However, the benefits of segmentation in combination with augmentation remain unexplored as it is deemed more time-consuming and difficult, particularly when dealing with big or complicated data sets [22]. Addressing this underutilization of segmentation alongside augmentation is key to unlocking their full potential for classification performance.

The main contribution of this research is to develop a hybrid pre-processing method to enhance data quality by reducing noise and outliers, while normalizing data for more effective CNN learning. This study demonstrates that combining segmentation and augmentation as pre-processing enhances model accuracy.

III. METHODOLOGY

A. Datasets

The data sets that will be used during the experimental stage are all obtained from online open-source image databases:

1) *Shenzhen dataset*: The dataset was collected in collaboration with Shenzhen No. 3 People's Hospital, Guangdong Medical College, Shenzhen, China. It contains 662 cases of chest X-rays, including 326 normal cases and 336 tuberculosis cases [25].

2) *Montgomery (MC) dataset*: The dataset was collected from the Department of Health and Human Services in partnership with Montgomery County, Maryland in the United States. The group consisted of 138 frontal chest radiographs from the Montgomery County Tuberculosis Screening

Program, of which 80 were normal and 58 were tuberculosis [25].

The datasets were divided into training/validation dataset and test dataset with a respective ratio of 80:20.

B. Flowchart

The flowchart in Fig. 1 outlines the process of training a VGG16 CNN model for image classification. The training dataset is pre-processed with ZCA, normalization, U-net segmentation, and augmentation techniques in that specific order before being used to train the model. The validation dataset is used to monitor model performance during training, and the test dataset is used for prediction and calculating model performance metrics.

C. Data Pre-Processing

Image resizing to 227x227, Mean Normalization and Standardization were applied to the set before segmentation and augmentation. Regarding segmentation, U-net CNN segmentation is being utilized for the pre-processing segmentation.

1) *Normalization and ZCA*: Normalization is used to scale the pixel values of an image to a range between 0 and 1. This is done to ensure that the pixel values are in a consistent range and to prevent the dominance of certain pixel values. Normalization is often done by dividing each pixel value by the maximum pixel value in the image [26]. ZCA is used to remove the correlation between the different color channels in an image. This is important because the color channels may be correlated, which can lead to redundant information and increased computational complexity. ZCA whitening transforms the image data so that each pixel value is uncorrelated with every other pixel value. This is done by performing eigenvalue decomposition of the covariance matrix of the pixel values and then transforming the data using the eigenvectors [27]. Both normalization and ZCA are useful for improving the performance of machine learning models on image datasets.

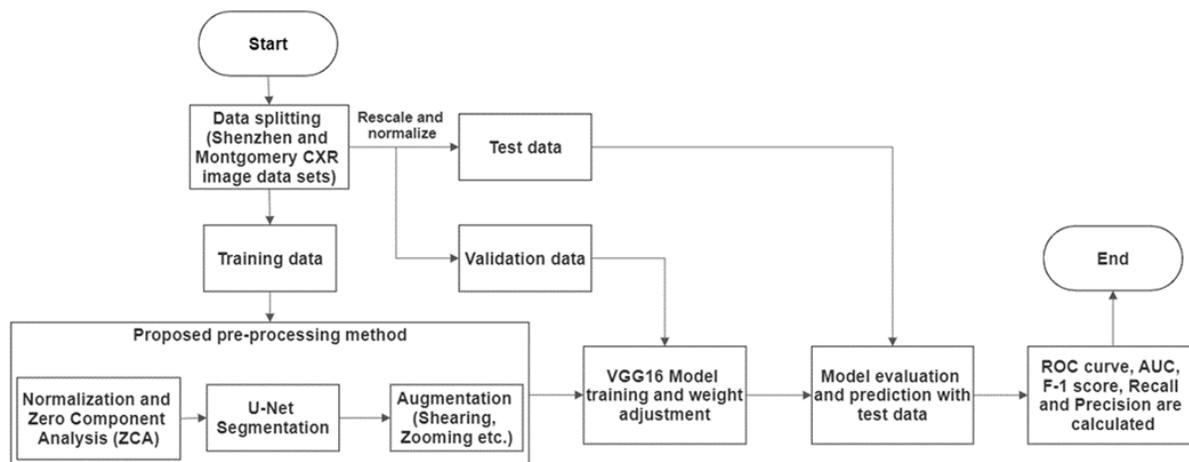


Fig. 1. Flowchart of simulation including the proposed method.

2) *U-net segmentation*: U-net CNN is a CNN designed for segmentation of biomedical images. What makes U-net stand out from other general convolutional neural networks is that general CNNs need to focus on image classification in biomedical cases, hence they require the user to identify the presence of a disease and localize the area of said disease. U-net eliminates this issue by applying classification to each pixel and can distinguish and localize borders by itself [28]. This method of segmentation could help eliminate the need for manual segmentation.

An expanding path and a contracting path make up the U-Net architecture. The contracting path is composed of convolutional layers followed by max-pooling layers, much like a conventional convolutional neural network. This path's goal is to minimize the spatial dimensionality of the input image while capturing its context.

The segmented image is localized using the expanding path, and the low-resolution feature maps created by the contracting path are unsampled. This path concatenates the corresponding feature maps from the contracting path with transposed convolutions to increase the spatial dimensionality of the features. This process is repeated several times to recover the original resolution of the input image.

The U-Net architecture also incorporates skip connections, which connect feature maps from the contracting path to feature maps from the expanding path. These connections aid in protecting the input image's small features, which can be lost during down sampling. That is where it gets its U shape from as seen in Fig. 2.

Many image segmentation tasks, including medical picture segmentation, cell segmentation, and object detection, have proven to be successful when using U-Net. It works particularly effectively in cases when the input images are tiny and the borders of the items that need to be divided are clearly defined [28].

3) *Augmentation*: Several transformations of augmentation were applied to the segmented images to increase the dataset's variability. The following augmentations were performed:

- Rescale transform every pixel value from range [0,255]
- Random rotations were applied up to 0.2 radians.

- Random shifts were applied to the width and height dimensions up to 0.1.
- Random shearing was applied for a maximum shear of 0.2.
- Zoom range up to 0.2.
- Horizontal and vertical flips.

D. VGG16 CNN Classifier

The VGG-16 architecture consists of 16 layers, including 13 convolutional layers, and three fully connected layers. It uses small 3x3 convolutional filters with a stride of 1 pixel, and max pooling layers with a 2x2 filter and stride of 2 pixels, which helps reduce the spatial dimensionality of the features.

The output of the last convolutional layer is flattened and fed into the fully connected layers, which perform classification on the input image. The final layer uses SoftMax activation to produce a probability distribution over the image classes.

Object detection, picture segmentation, and style transfer are just a few of the computer vision applications for which the VGG-16 architecture has been extensively employed as a pre-trained model [29]. Its popularity is a result of its efficiency and simplicity, which make it a reliable benchmark model for comparison with other designs.

1) *Fine-tuning*: The VGG-16 CNN model's weights are pretrained on a massive image dataset known as ImageNet. ImageNet is mostly comprised of natural and colorful pictures of animals and food, which deviates from what monochromatic Chest X-rays are. Fine-tuning the pretrained weights of the VGG-16 can be accomplished by unfreezing the weights of the 5th block of layers of the model. The unfreezing will increase the computation time needed for training the model, but it might result in more optimized weights [30].

2) *ADAM optimizer*: In order to calculate the difference between a neural network's predicted and actual output during training, the Adaptive Moment Estimation (ADAM) optimizer utilizes a loss function [31]. The optimizer then modifies the neural network's weights in accordance with the gradient of the loss function relative to the weights. Table I displays the hyperparameters used during experimentation.

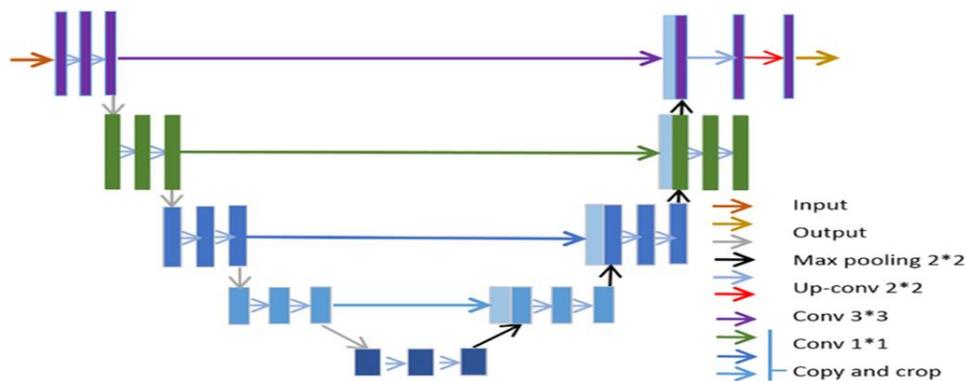


Fig. 2. The U-Net architecture displayed with the arrows denoting different operations [28].

TABLE I. HYPERPARAMETERS USED DURING EXPERIMENTATION

Hyperparameter	Values considered
Batch size	32,64
Number of epochs	30,40,50
Learning rate	0.001, 0.0001, 0.00001
Embedded dropout rate	0.3,0.5,0.7
Embedded activation function	Sigmoid
Optimizer	ADAM
Number of unfrozen layers	4 layers (5 th block)

E. Performance Evaluation Metrics

1) *Confusion matrix*: A confusion matrix is a table used to evaluate the performance of a classification algorithm by comparing the predicted and actual values of a dataset as seen in Table II. The matrix is organized into rows and columns, with each row representing the instances in a predicted class, and each column representing the instances in an actual class. The diagonal of the matrix represents the instances that were correctly classified, while the off-diagonal elements represent the instances that were misclassified.

The confusion matrix provides a useful way to visualize the performance of a classifier, and it can be used to calculate various metrics such as accuracy, precision, recall, and F1 score.

TABLE II. CONFUSION MATRIX

Actual values	Predicted values	
	Yes (Predicted class)	No (Predicted class)
Yes (Actual class)	True Positive (TP)	False Negative (FN)
No (Actual class)	False Positive (FP)	True Negative (TN)

- True Positive (TP): Number of patients correctly classifies as having TB.
- True Negative (TN): Number of patients correctly classified as not having TB.
- False Positive (FP): Number of patients incorrectly classified as having TB.
- False Negative (FN): Number of patients incorrectly classified as not having TB.

2) *Accuracy*: Accuracy of the model can be calculated from the confusion matrix by the following mathematical equation:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

3) *Precision*: Precision measures how many of the model's positive predictions were right. A high precision indicates that the model made few incorrect positive predictions and is a good indicator of how well the model detects positive cases. The precision mathematical equation is as follows:

$$P = \frac{TP}{TP + FP} \quad (2)$$

4) *Recall*: The model's recall evaluates how successfully it detects positive cases out of all actual positive cases. A high recall indicates that the model is good at recognising positive examples, whereas a low recall indicates that the model is missing many positive cases. The recall mathematical equation is as follows:

$$R = \frac{TP}{TP + FN} \quad (3)$$

5) *F1 score*: The F1-score is a measure of the balance between precision and recall. The F1-score is a useful metric for evaluating the overall performance of a binary classification model, especially when the classes are imbalanced. The F1-score equation is as follows:

$$F1 = \frac{2(Recall * Precision)}{(Recall + Precision)} \quad (4)$$

6) *ROC curve and AUC*: An ROC (Receiver Operating Characteristic) curve is a plot used to visualize the performance of a binary classification model. It is created by plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) at different classification thresholds, as seen in Fig. 3.

AUC (Area under the ROC Curve) is a valuable metric for evaluating the performance of binary classification models and is commonly used in a variety of machine learning applications. AUC represents the area under this curve, which ranges from 0 to 1. A perfect classifier would have an AUC of 1, while a random classifier would have an AUC of 0.5.

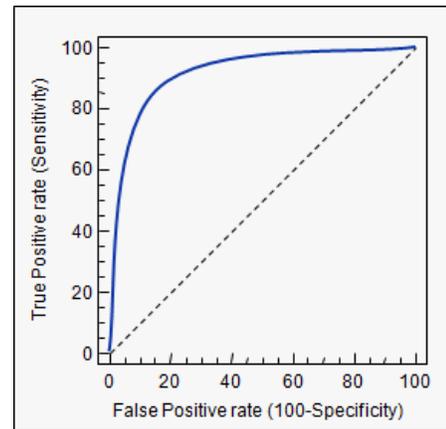


Fig. 3. ROC curve - A plot of true positive rate vs false positive rate.

IV. RESULTS AND DISCUSSION

A. Experiments

A total of three experiments have been conducted in this study to observe the effects of the hybridization of the pre-processing methods, such as no segmentation and augmentation (experiment 1), with segmentation only (experiment 2) and with augmentation and segmentation (experiment 3), where the results are discussed in terms of the confusion matrix and the other performance measures

discussed in Section III. The test data set consists of 160 CXR images with 80 TB positive cases and 80 normal cases.

B. Proposed Pre-Processing Method Performance

The hybridized pre-processing method used in experiment 3 has achieved a higher performance in comparison to experiment 1 and 2, as seen in the metrics achieved in the confusion matrix in table 3 and the performance measures in table 4, with an accuracy of 90%, a recall of 87.5%, a precision of 92.11% and a F1-score of 0.8975. Regarding the AUC shown in Fig. 4, out of the 3 experiments the proposed method has produced the best ROC curve and an AUC of 0.935, followed by 0.89 from experiment 2 and 0.87 from experiment 3, as seen from the confusion matrix table in Table III and the performance metrics table in Table IV.

C. Result Discussion

In this section, the discussion will progress from a baseline with no preprocessing to segmentation alone, and finally to the combined use of both techniques, the study systematically highlights their individual and joint impacts on model accuracy. This approach effectively highlights the research's

focus on demonstrating the effectiveness of segmentation and the combination of both segmentation and augmentation.

The first experiment has an accuracy of 82.5% and an AUC of 0.87 and does not use segmentation or augmentation. For the performance of the model, this approach is regarded as the standard. The input data is supplied straight to the CNN model without segmentation or augmentation, which may cause the data to be overfitted or underfitted.

The accuracy and AUC of the second experiment, which only uses segmentation, are 86.25% and 0.89, respectively. This method significantly outperforms the baseline method in terms of accuracy, proving that segmentation can increase the quality of the input data used to train the CNN model.

The third experiment involves both segmentation and augmentation and achieved 90% accuracy and an AUC of 0.935. This method shows the highest accuracy and AUC compared to the other two experiments. The combination of both segmentation and augmentation techniques could have helped the model to isolate important features and improve the robustness of the model.

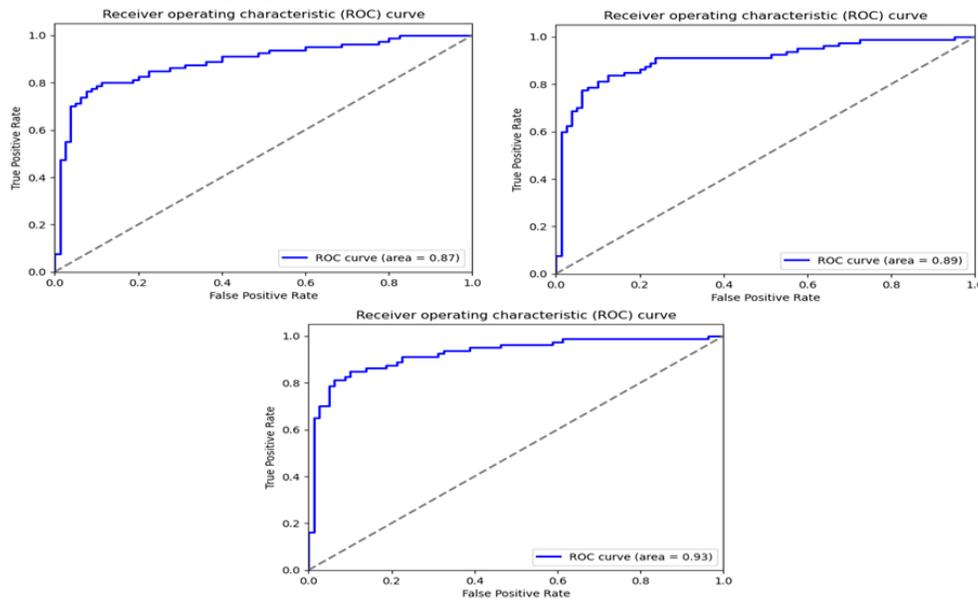


Fig. 4. ROC curves of experiment 1 (top left), experiment 2 (top right) and experiment 3 (bottom middle).

TABLE III. THE CONFUSION MATRIX OF EXPERIMENT 1-3 REPRESENTED IN NUMBER OF IMAGES

Experiment No.	Confusion matrix			
	True positive	True negative	False positive	False negative
1	56	76	4	24
2	64	74	6	16
3	70	74	6	10

TABLE IV. PERFORMANCE MEASURES CALCULATED FROM THE CONFUSION MATRIX

Experiment No.	Performance measures			
	Accuracy	Recall	Precision	F1-score
1	82.5%	70%	93.33%	0.8000
2	86.25%	80%	91.43%	0.8533
3	90%	87.5%	92.11%	0.8975

TABLE V. OVERALL CLASSIFICATION ACCURACY AND AUC COMPARISON WITH PREVIOUS RESEARCH WORK

Reference	CNN model	Dataset	Pre-processing	Overall Accuracy Rate and AUC
Proposed pre-processing method	VGG-16	Shenzhen and Montgomery	Augmentation and U-Net segmentation	90% and 0.935
Syeda et al. [18]	Ensemble (VGG-16, VGG-19, ResNet50 and GoogleNet)	Shenzhen and Montgomery	None	86.7% and 0.92
Ahsan M et al. [22]	VGG-16	Shenzhen and Montgomery	Augmentation	81.25% and 0.89
Erdal [20]	Ensemble (VGG-16, VGG-19, AlexNet, Random Forest)	Shenzhen	Manual segmentation	88.3% and 0.92

D. Comparative Analysis

The overall accuracy rate and AUC comparison of the various pre-processing methods from previous work is represented in Table V. For a valid comparison, the author has compared the proposed pre-processing method with other applications of pre-processing from the related works in Section II. It's crucial to maintain the experimental parameters as consistently as feasible.

The proposed pre-processing method, which included VGG-16, augmentation, and U-Net segmentation, achieved the maximum accuracy of 90% and AUC of 0.935, illustrating the efficiency of these techniques in improving the model's ability to generalize and detect essential traits in CXR images. Syeda et al.'s [18] technique showed lower accuracy and AUC even when employing multiple architectures, demonstrating that augmentation and segmentation are still beneficial and necessary in improving model performance.

As demonstrated by Ahsan M. et al.'s [22] method, augmentation alone was unable to improve model performance with a small data set size. While manual segmentation is useful as shown by Erdal [20], it is time-consuming and error-prone, and U-Net's ability to segment major features in images makes it more effective and robust for CAde systems.

V. CONCLUSION

In conclusion, the results of the study have shown that applying both segmentation and augmentation techniques can lead to better performance measures in terms of accuracy, AUC, recall, precision, and F1-score when classifying TB related CXR images compared to using only one or none of these techniques.

Augmentation has become a popular technique in recent years, and for good reason. It allows for the creation of a large and diverse training set without the need for additional data collection efforts, which can be time-consuming and costly. However, augmentation alone may not always be sufficient, particularly when dealing with complex images with intricate features or objects.

Furthermore, while manual segmentation can be effective in capturing important details in the images, it is not robust or flexible enough to be widely used in computer-aided detection systems. The rise of automatic segmentation methods such as U-Net has made segmentation more viable as a pre-processing technique for image analysis tasks. Overall, the findings suggest that researchers consider using both segmentation and augmentation techniques as pre-processing methods when developing CAde systems.

ACKNOWLEDGMENT

The authors would like to thank Ministry of Higher Education (MOHE), Malaysia and Universiti Teknologi PETRONAS for supporting this work through a Research Grant, FRGS, (Ref: FRGS/1/2022/TK07/UTP/02/4) and Cost Centre, (015MA0-159).

REFERENCES

- [1] Tuberculosis (TB) | American Lung Association. (n.d.). Retrieved December 8, 2019, from <https://www.lung.org/lung-health-and-diseases/lung-disease-lookup/tuberculosis/>
- [2] W. H. Organization et al., "Global tuberculosis report 2020", World Health Organization, 2020.
- [3] Jaeger, S., Karagyris, A., Candemir, S., Siegelman, J., Folio, L., Antani, S., & Thoma, G. (2013). Automatic screening for tuberculosis in chest radiographs: a survey. *Quantitative Imaging in Medicine and Surgery*, 3(2), 89–99. <https://doi.org/10.3978/j.issn.2223-4292.2013.04.03>.
- [4] Ryu, Y. J. (2015, April 1). Diagnosis of pulmonary tuberculosis: Recent advances and diagnostic algorithms. *Tuberculosis and Respiratory Diseases*, Vol. 78, pp. 64–71. <https://doi.org/10.4046/trd.2015.78.2.64>.
- [5] Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. (2018) Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med* 15(11): e1002686.
- [6] Li, Q., & Nishikawa, R. M. (Eds.). (2015). *Computer-aided detection and diagnosis in medical imaging*. Taylor & Francis.
- [7] Afsaneh Jalalian, Syamsiah B.T. Mashohor et al., Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: A review, *Clinical Imaging*, 37, 3, 5 2013.
- [8] Dheeba, J., Singh, N. A., & Selvi, S. T. (2014). Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *Journal of biomedical informatics*, 49, 45-52.
- [9] Cascianelli, S., Scialpi, M., Amici, S., Forini, N., Minestrini, M., Luca Fravolini, M., ... & Palumbo, B. (2017). Role of artificial intelligence techniques (automatic classifiers) in molecular imaging modalities in neurodegenerative diseases. *Current Alzheimer Research*, 14(2), 198-207.
- [10] Sharon, H., Elamvazuthi, I., Lu, C.K., Parasuraman, S., Natarajan, E., 2019. Development of Rheumatoid Arthritis Classification from Electronic Image Sensor using Ensemble Method, *Sensors* 20, 167.
- [11] Timothy Ganesan, Pandian Vasant, Irraivan Elamvazuthi, *Advances in Metaheuristics, Applications in Engineering Systems*, CRC Press, 9 Dec, 2016.
- [12] Ku Abd. Rahim, K.N.; Elamvazuthi, I.; Izhar, L.I.; Capi, G. Classification of Human Daily Activities Using Ensemble Methods Based on Smartphone Inertial Sensors. *Sensors* 2018, 18, 4132.
- [13] Meraj, S. S., Yaakob, R., Azman, A., Rum, S. N. M., & Nazri, A. S. A. (2019). Artificial Intelligence in diagnosing tuberculosis: A review. *International Journal on Advanced Science, Engineering and Information Technology*, 9(1), 81–91.
- [14] Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 8, 53 (2021).

- [15] Antony, B., & Banu, N. (2017). Lung Tuberculosis Detection Using X-Ray Images. *International Journal of Applied Engineering Research*, 12, 15196–15201.
- [16] Muhathir, M., Sibarani, T. T. S., & Al-Khowarizmi, A.-K. (2020). Analysis K-Nearest Neighbors (KNN) in Identifying Tuberculosis Disease (Tb) By Utilizing Hog Feature Extraction. *Al'adzkiya International of Computer Science and Information Technology (AloCSIT) Journal*, 1(1).
- [17] Dunnmon, J. A., Yi, D., Langlotz, C. P., Ré, C., Rubin, D. L., & Lungren, M. P. (2019). Assessment of Convolutional Neural Networks for Automated Classification of Chest Radiographs. *Radiology*, 290(2), 537–544. <https://doi.org/10.1148/radiol.2018181422>.
- [18] Meraj, Syeda & Azman, Azreen & Yaakob, Razali & Nazri, Azree & Rum, Siti. (2019). Detection of Pulmonary Tuberculosis Manifestation in Chest X-Rays Using Different Convolutional Neural Network (CNN) Models. 10.35940/ijeat.A2632.109119.
- [19] Gordienko, Yu., Gang, P., Hui, J., Zeng, W., Kochura, Yu., Alienin, O., Rokovyi, O., & Stirenko, S. (2018). Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer. *Advances in Intelligent Systems and Computing*, 638–647.
- [20] Tasci, E. (2020). Pre-processing Effects of the Tuberculosis Chest X-Ray Images on Pre-trained CNNs: An Investigation. *Artificial Intelligence and Applied Mathematics in Engineering Problems*, 589–596.
- [21] Oloko-Oba, M., & Viriri, S. (2020). Diagnosing Tuberculosis Using Deep Convolutional Neural Network. *Lecture Notes in Computer Science*, 151–161.
- [22] Ahsan, M., Gomes, R., & Denton, A. (2019). Application of a Convolutional Neural Network using transfer learning for tuberculosis detection. 2019 IEEE International Conference on Electro Information Technology (EIT).
- [23] Colombo Filho, M. E., Mello Galliez, R., Andrade Bernardi, F., de Oliveira, L. L., Kritski, A., Koenigkam Santos, M., & Alves, D. (2020). Preliminary Results on Pulmonary Tuberculosis Detection in Chest X-Ray Using Convolutional Neural Networks. *Lecture Notes in Computer Science*, 563–576.
- [24] Showkatian, E., Salehi, M., Ghaffari, H., Reiazi, R., & Sadighi, N. (2022). Deep learning-based automatic detection of tuberculosis disease in chest X-ray images. *Polish Journal of Radiology*, 87(1), 118–124.
- [25] Jaeger, S., Candemir, S., Antani, S., Wang, Y. X. J., Lu, P.-X., & Thoma, G. (2014). Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative Imaging in Medicine and Surgery*, 4(6), 475–477.
- [26] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning*, 448-456.
- [27] Desjardins, G., Courville, A., & Bengio, Y. (2015). Statistics of natural image categories. *Advances in Neural Information Processing Systems*, 27, 962-970.
- [28] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [29] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *Proceedings of the International Conference on Learning Representations*.
- [30] Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1717-1724).
- [31] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations*.