Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.4, August 2025, pp. 1346~1353
e-ISSN: 2548-6861

Lung Segmentation in X-ray Images of Tuberculosis Patients Using U-Net with CLAHE Preprocessing

Ibnu Farid Mabina [1], Christy Atika Sari [2], Eko Hari Rachmawanto [3]
* Teknik Informatika, Universitas Dian Nuswantoro
111202214591@mhs.dinus.ac.id [1], christy.atika.sari@dsn.dinus.ac.id [2], eko.hari@dsn.dinus.ac.id [3]

Article Info

Article history:

Received 2025-06-18
Revised 2025-06-30
Accepted 2025-07-03

Keyword:

Tuberculosis,
Image Segmentation,
U-Net,
CLAHE,
X-ray Image,
Deep Learning.

ABSTRACT

Tuberculosis (TB) is an infectious disease that commonly affects the lungs and remains one of the leading causes of death from infectious diseases. Early detection is essential to prevent further spread and organ damage. Chest X-ray images are one of the main methods for diagnosing TB, but image quality is often affected by low contrast and noise. This study proposes the application of Contrast Limited Adaptive Histogram Equalization (CLAHE) method to improve X-ray image quality, combined with U-Net deep learning architecture for lung segmentation in X-ray images of tuberculosis patients. U-Net was chosen due to its excellent capability in medical image segmentation, thanks to its architectural structure that has encoder-decoder with skip connections, which allows the model to retain detailed information on high-resolution images, even on complex and noisy data. Experimental results using the Shenzhen and Montgomery datasets show that the U-Net model with CLAHE achieves Pixel Accuracy 97.96%, Recall 94.93%, Specificity 98.97%, Dice Coefficient 95.87%, and Jaccard Index (IoU) 92.07%.

I. INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by the bacterium Mycobacterium tuberculosis, which commonly affects the lungs [1]. The disease is spread through the air, usually when an infected person coughs or sneezes [2]. Although curable with proper treatment, TB remains one of the leading causes of death from infectious diseases in the world [3]. According to the Global Tuberculosis Report 2024, by 2023 it is estimated that there will be a total of 10.8 million global TB cases, of which 8.2 million will be new cases [4]. Indonesia alone accounts for about 10% of the 56% of pulmonary TB cases in the world from the five countries with the highest caseloads [4]. According to data from the Ministry of Health of the Republic of Indonesia, by 2022 there will be more than 969,000 cases of tuberculosis (TB) with approximately 150,000 deaths [5]. Then, from an estimated one million TB cases in Indonesia, case finding increased from 724,309 in 2022 to 821,200 in 2023 [5]. Early detection of tuberculosis is essential to minimize the risk of disease spread and prevent further organ damage [6]. One of the main methods for early detection of TB disease is through X-ray image analysis, which can provide a visual picture of the condition of the lungs and help in identifying early signs of infection [7].

In the medical world, imaging using tools such as X-rays and CT scans of the lungs play an important role in detecting and monitoring the progression of TB [8]. Medical images in the form of X-rays and CT scans provide a visual representation of the condition of body tissues, including infected areas. However, the quality of medical images is often affected by noise or low contrast, which can hinder image analysis and make it difficult to detect infected areas [9]. Therefore, image processing is necessary to improve the image quality, which allows the diagnosis process to be performed more accurately [10]. One effective image processing method is Contrast Limited Adaptive Histogram Equalization (CLAHE) because it can improve image contrast without adding significant noise [11]. CLAHE is chosen over other methods, such as standard Histogram Equalization (HE), Adaptive Gamma Correction (AGC), and Median Filtering, because it prevents the over-amplification of noise that often occurs with standard HE, while providing more adaptive contrast adjustment that is focused on important

areas without compromising details [12]. AGC is less effective in local adjustment, while Median Filtering tends to smooth out edges and important details [13]. With CLAHE's ability to enhance contrast without degrading image quality, this method is highly suitable for medical applications such as TB detection [14]. To further improve diagnosis in segmenting TB-infected lung images, deep learning technology can be one of the more effective diagnosis methods [15].

In medical image analysis, Convolutional Neural Network (CNN) technology has emerged as one of the most representative deep learning models in medical image analysis, thanks to its ability to recognize complex patterns and support various tasks such as classification, segmentation, and image detection [16]. Convolutional Neural Network is one of the algorithms in deep learning that can receive an image input and is used to identify and recognize an image that has the advantage of having high accuracy but has the disadvantage that the training process is quite long [17]. CNN are a further development of Artificial Neural Networks that consist of neural networks that perform weight, bias, and activation functions, in addition, the most common layers in CNN are convolution, pooling, and fully connected [18]. CNN work by applying convolution operations to extract local features from 2D input data such as images, using local filters and local connections to reduce the number of parameters, which speeds up training and enables automatic identification of relevant features without human supervision [19].

One of the effective Convolutional Neural Network architectures for medical image segmentation is U-Net. U-Net is specifically designed for medical image segmentation with an architecture consisting of an encoder and decoder [20]. The encoder, also known as the contracting layer, is used to extract important features from the image by reducing the size of the input matrix while increasing the feature map [21]. Meanwhile, the decoder (expanding layer) is in charge of taking the feature map from the encoder to produce an image that matches the original input or desired output, by returning the matrix to its original size through reducing the number of feature maps[21]. The main advantage of U-Net is its ability to provide accurate segmentation despite the limited data used [22]. As a Convolutional Neural Network (CNN) based model, U-Net suffers from the inherent locality of the convolution operation. This results in limitations in its capacity to explicitly grasp the interaction of semantic information that is global and remote in nature [23].

U-Net offers several advantages, especially in medical image segmentation, due to its simple yet effective architecture [24]. The encoder-decoder structure combined with skip connections allows the model to effectively capture both local and global features, which is crucial for segmenting fine structures such as lung tissues and lesions in tuberculosis X-ray images [25]. This makes U-Net highly efficient in scenarios where datasets are small, which is often the case in medical imaging. Compared to other models like U-Net++, which is more complex with additional skip connections and deep supervision, U-Net is lighter in terms of computation, allowing for faster training and deployment without sacrificing accuracy [26]. Models like DeepLabv3+, which use dilated convolutions and larger receptive fields, require more data to perform optimally, while U-Net performs well even with limited datasets [27]. Additionally, U-Net's flexibility allows for easy adaptation to various segmentation tasks, such as detecting organs, without major changes to the architecture [28]. These advantages make U-Net a more practical and efficient choice compared to more complex models, while still providing high performance in medical image segmentation.

Lung segmentation in medical images, especially in chest X-rays (CXR), is one of the rapidly growing areas in the field of medical image processing, especially with the advances in deep learning technology. This study uses a dataset of chest X-ray images of tuberculosis patients consisting of 1,408 X-ray images divided into two main files, namely, 704 chest X-ray images (images) and 704 lung segmentation masks (masks) that represent the ground truth of the lung area. These images are obtained from two sources, namely Montgomery County Chest X-ray Database (United States) and Shenzhen Chest X-ray Database (China). This rich data is the basis for various studies related to lung segmentation, one of which is Hashem and Kamil's research [29], which utilizes deep learning algorithms for lung segmentation in chest X-ray images using the U-Net architecture, known as one of the most significant semantic segmentation frameworks in convolutional neural networks (CNN). The study involved a dataset of 565 X-ray images divided into 500 training images and 65 validation images, with highly competitive results, achieving an accuracy rate of 91.47% in the training data and 89.18% in the validation data.

While in the research conducted by Sousa [30], developed a deep learning model that combines two architectures, U-Net and ResNet34, for lung segmentation in computed tomography (CT) images. The model was tested on a cross-cohort dataset and achieved an average Dice similarity coefficient (DSC) of more than 0.93 for the four cohorts tested. Evaluation by two experienced radiologists showed a slight decrease in performance in cases of consolidation and pneumocystis pneumonia, which underscores the importance of testing the model on a variety of pathological conditions to ensure the accuracy and effectiveness of the model in diverse clinical practices. Other research conducted by Alam [31], proposed AMRU++, an attention-based network with multi-residual blocks for lung segmentation in CXR images. This study introduces a data augmentation technique that aims to simulate the features and characteristics of CXR pathology, thus addressing the problem of limited annotated data. Experimental results show that AMRU++ successfully outperforms the U-Net model with a DSC value of 94.13% and a Jaccard Index (JI) of 89.16%.

Then the research conducted by Arvind [32], developed a modified U-Net architecture for lung organ segmentation in

chest X-ray images. The model is more lightweight with the addition of dropouts on all deconvolutional layers to overcome overfitting issues and provides higher accuracy on a combination of three different datasets. With a training accuracy of 92.71% and a generalization accuracy of 93.87%, this model shows significant potential to be used in lung segmentation with higher efficiency. Lastly, Rahman's research [33], proposed a two-stage segmentation system based on a modified U-Net architecture for lung field segmentation in CXR images. The method was tested on a dataset of 138 CXR images obtained from the Montgomery County Tuberculosis Control Program, with results demonstrating the effectiveness of the method in overcoming segmentation challenges on CXR images containing pathological abnormalities.

II. METHOD

This study conducts a systematic comparison between the segmentation method using U-Net and CLAHE preprocessing on chest X-ray images of tuberculosis patients. The methodology follows several stages that combine medical image segmentation protocols with specific adjustments to enhance lung image quality in X-ray scans. Figure 1 illustrates the research framework, outlining the steps from data preparation to model performance evaluation. This systematic approach allows for an objective and structured comparison of segmentation performance between U-Net applied with CLAHE preprocessing on the tuberculosis image dataset.

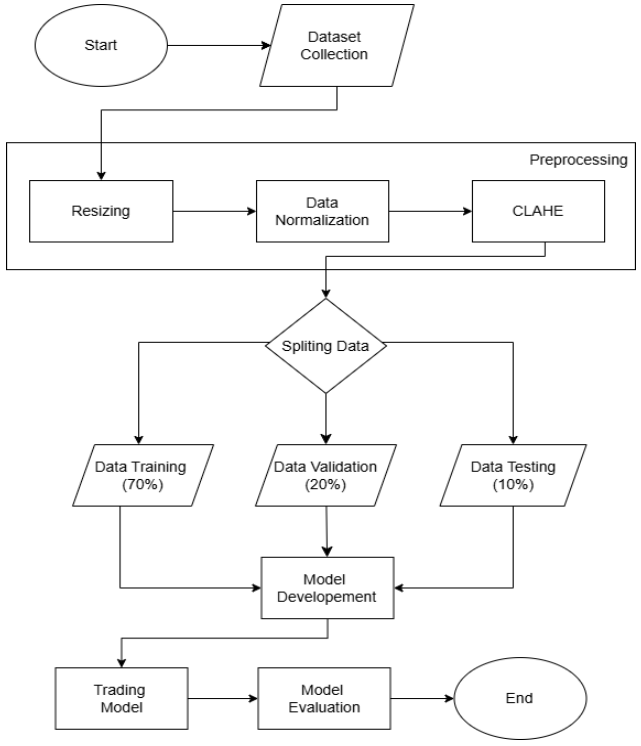


Figure 1. Flowchart Penelitian

A. Dataset Preparation

To give a clearer picture of the dataset used, here is a visual representation of the dataset structure:

Category	Subcategory	Count	Percentage
Gender	Male	872	51%
	Female	430	25%
	Other	292	24%
Age	28	86	5%
	25	74	4%
	Other	1454	92%
County	Shenxhen	1363	80%
	Montgomery	341	20%
Tuberculosis Status	Normal	730	44%
	Bilateral PTB	113	7%
	Other	860	49%
Remarks	Normal	730	44%
	Bilateral PTB	113	7%
	Other	860	49%

Figure 2. Dataset Data.csv

Based on Figure 2, the table shows the metadata information of the chest X-ray image dataset used in this study, which includes several important attributes such as id, gender, age, county, ptb, and remarks. The id attribute is a unique identification for each X-ray image. The gender column indicates the sex of the patient, where the majority are male (51%), followed by female (25%), and other (24%). Meanwhile, the age column represents the age of the patient with a diverse distribution, and the highest age is recorded at 28 years (5%) and 25 years (4%). The county column indicates the location of origin, with the largest proportion coming from Shenzhen (80%) and the rest from Montgomery (20%). The ptb column indicates the status of the patient as either having TB (1) or not (0), with the data split between normal and TB positive. Additional information can be seen in the remarks column, which describes the patient's specific condition, such as “normal”, “STB, ATB, tuberculosis pleurisy”, and “secondary PTB in the bilateral upper field.” The dataset used in this study is the Chest X-ray Dataset for Tuberculosis Segmentation, which can be accessed through the following link, https://www.kaggle.com/datasets/iamtapendu/chest-x-ray-lungs-segmentation

B. Data Preprocessing

In the preprocessing stage of the chest X-ray images, the first step involves resizing each image to a fixed dimension of 256×256 pixels, which corresponds to the standard input size required by the U-Net model architecture employed in this study. This resizing ensures uniformity across all images, allowing consistent processing by the convolutional neural

network. Following resizing, normalization is performed by scaling the pixel intensity values from the original range of [0, 255] to a normalized range of [0, 1]. This is achieved by dividing each pixel value by 255, with the purpose of accelerating model convergence and reducing the impact of feature scale disparities during the learning process.

Once the image has been resized and normalized, the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique is applied to enhance local contrast. CLAHE operates by dividing the image into small contextual regions (tiles) and applying limited histogram equalization within each tile, thus emphasizing fine details—particularly in lung regions affected by diseases such as tuberculosis—without amplifying noise excessively. In this study, CLAHE is implemented with specific parameters: clipLimit = 2.0 and tileGridSize = (8, 8). The clipLimit parameter constrains the contrast enhancement to prevent over-amplification, while tileGridSize defines the number of contextual regions over which the histogram equalization is performed.

C. Splitting Data

The division of the dataset follows a structured approach to meet the needs of medical segmentation model development. The process begins with splitting the dataset into three proportional parts with a ratio of 70:20:10, where the largest portion is allocated for model training. A total of 70% of the data is used as the training set for deep feature learning, while 20% is set aside as the validation set which plays a crucial role in performance monitoring and parameter tuning during training. The remaining 10% is reserved as a completely independent test set, serving as an objective final evaluation of the model's generalization capabilities. This division ensures that each subset retains the characteristics of the original data while maintaining the integrity of the image-mask pair. Validation was performed on 20% of the data allowing early detection of overfitting, while 10% of the test data provided a true picture of the model's performance on novel cases.

D. Model Development

The method used in this research is medical image segmentation using U-Net architecture. U-Net is one of the Convolutional Neural Network (CNN)-based deep learning models specifically designed for image segmentation, especially in the medical field. This architecture is able to effectively learn spatial representations and produce precise segmentation of target areas, such as lungs in chest X-ray images. The U-Net model is trained using a pre-processed image dataset to produce accurate segmentation predictions. The structure of the algorithm can be seen in Figure 3.

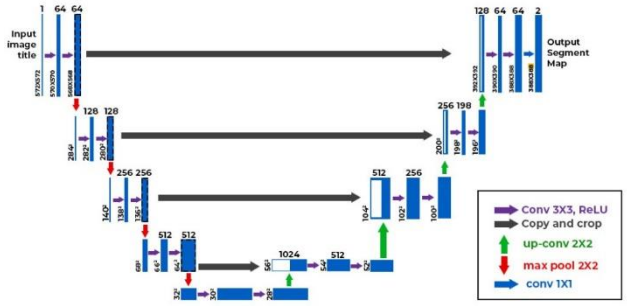


Figure3 U-Net Architecture [34]

1) *Input Layer:* The U-Net model is designed to accept as input a grayscale image of size 256×256×1. This size was chosen to match the standard input configuration commonly used in the U-Net architecture, ensuring compatibility and optimal performance for medical image segmentation tasks.

2) *Contracting Path (Encoder):* This path is the first part of the U-Net architecture that functions as a feature extractor. The encoder is composed of several convolution blocks. Each of these blocks consistently consists of two convolution layers with a 3×3 kernel, followed by a ReLU activation function to introduce non-linearity, and then a 2×2 max pooling operation that serves for downsampling. This downsampling process progressively reduces the spatial dimension of the input image. Along with each block, the number of filters is gradually increased, starting from 64 filters until it reaches 256 filters in later blocks. This increase in the number of filters allows the encoder to capture and represent increasingly complex and abstract features of the original image.

3) *Bottleneck:* This section occupies the deepest position in the U-Net architecture, serving as the bridge between the encoder and decoder. The bottleneck consists of two convolution layers, each using 512 filters, followed by a ReLU activation function. Its main role is to compress and capture the most dense and complex feature representation of the input image. This representation will then be passed to the decoder path for segmentation image reconstruction.

4) Expansive Path (Decoder): This path is the reverse of the encoder, responsible for restoring the image resolution to its original size and gradually generating a detailed segmentation map. Each block in the decoder begins with an upsampling operation (using 2×2 up-convolution) to increase the spatial dimension of the image. The result of this upsampling is then concatenated or merged with the relevant feature map from the encoder via skip connection. The use of skip connection is crucial as it allows the decoder to utilize fine spatial details that may be lost during the downsampling process at the encoder. Along with upsampling, the number of filters in each decoder block is gradually reduced, from 256 to 64 filters.

5) Output Layer: As the last layer of the U-Net architecture, the Output Layer is a 1×1 convolution layer. This layer is responsible for generating one output channel with the same size as the original input, which is 256×256. For binary segmentation tasks, such as separating the lung area from the background in an X-ray image, a sigmoid activation function is applied to the output layer. This function generates a probability value between 0 and 1 for each pixel, indicating the likelihood of that pixel belonging to the segmented class.

E. Training Model

Model training was conducted in two scenarios with a total of 17 epochs. The first scenario uses the standard U-Net architecture for image segmentation, where the model is trained with image data that has been labeled using the Adam optimizer and binary cross-entropy loss. Metrics used for evaluation included accuracy, Dice Coefficient, and Jaccard Index. The second scenario involves the use of U-Net combined with CLAHE (Contrast Limited Adaptive Histogram Equalization) technique, which is applied to the image before training to enhance contrast. CLAHE helps the model recognize features in low-light areas of the image, which is expected to result in more accurate segmentation. These two scenarios aim to compare the performance of the model with and without the CLAHE contrast enhancement technique, with training performed for 17 epocs for each scenario

F. Model Evaluation

The evaluation model used in this research is:

1) *Confusion Matrix*: Confusion Matrix is a summary table used to evaluate the performance of classification models, including in the context of image segmentation where each pixel is classified. This matrix presents the comparison between the actual class label (ground truth) and the class label predicted by the model.

TABLE I
CONFUSION MATRIX

Confusion Matrix		Predicted Class	
		Positive	Negative
Acctual Class	Positive	TP	FN
	Negative	FP	TN

Based on Table 2.1, there are four main categories: True Positives (TP), which is the amount of positive data that is correctly classified as positive; False Positives (FP), which is the amount of negative data that is incorrectly classified as positive; False Negatives (FN), which is the amount of positive data that is incorrectly classified as negative; and True Negatives (TN), which is the amount of negative data that is correctly classified as negative.

2) *Jaccard Index (Intersection Over Union – IoU):* Jaccard Index, also known as Intersection Over Union (IoU), is a metric that measures the overlap between two sets. In the context of image segmentation, it is calculated as the ratio between the intersection area of the predicted segmentation result and the ground truth, divided by the union area of the two. The Jaccard Index value ranges between 0 and 1, where a value of 1 indicates perfect overlap (prediction exactly matches the ground truth), while a value of 0 indicates no overlap at all. This metric is very sensitive to small differences between prediction and ground truth, making it a good choice for evaluating the spatial accuracy of segmentation and effective for dealing with cases of unbalanced classes. The Jaccard Index can be formulated by,

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

3) *Dice Coefficient*: Dice Coefficient, also known as F1-score or Sørensen-Dice coefficient, is another popular overlap metric, often used in medical image segmentation. It is calculated as twice the area of the intersection between the prediction and ground truth, divided by the total sum of the prediction and ground truth areas. Similar to the Jaccard Index, the Dice Coefficient value also ranges between 0 and 1, with 1 indicating perfect similarity. The Dice Coefficient tends to give higher weight to areas of overlap and is often considered slightly less restrictive than the Jaccard Index. This is because the denominator "recalculates" the True Positive (TP) twice, which makes it slightly more fault-tolerant than the Jaccard Index. The Dice Coefficient can be formulated by,

$$D = \frac{2 \times TP}{2 \times TP+FP+FN} \quad (2)$$

4) *Recall (Sensitifity): Re*call, also known as Sensitivity or True Positive Rate, is a metric that measures the model's ability to identify all positive cases that actually exist. In the context of segmentation, Recall calculates the proportion of object (foreground) pixels correctly detected by the model out of all object pixels actually present in the ground truth. A high Recall value is particularly important in medical applications, such as disease detection, where missing positive cases (False Negatives) can have serious consequences.

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

5) Specificity: Specificity, or True Negative Rate, is a metric that measures the model's ability to correctly identify negative cases. In segmentation, Specificity calculates the proportion of background pixels that are correctly predicted as background out of all background pixels actually present in the *ground* truth. A high Specificity value indicates that the model is effective in avoiding False Positives, i.e. predicting healthy areas as infected areas, which is also important in medical diagnosis to avoid unnecessary medical anxiety or action.

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

III. Results and Discussion

This research was conducted using a dataset with a training, validation, and testing ratio of 80:10:10. The U-Net algorithm will be optimized and combined with CLAHE in one scenario, and without CLAHE in the other. In both scenarios, there are hyperparameters and parameters in the Adam Optimizer optimization function. The values of the hyperparameters are learning rate (lr) = 0.0001, batch size = 16, and epoch = 17. The following in Figure 3 is a comparison graph of Loss and Pixel Accuracy from training on a dataset with an epoch of 17, showing the performance for both scenarios: with and without the CLAHE method.

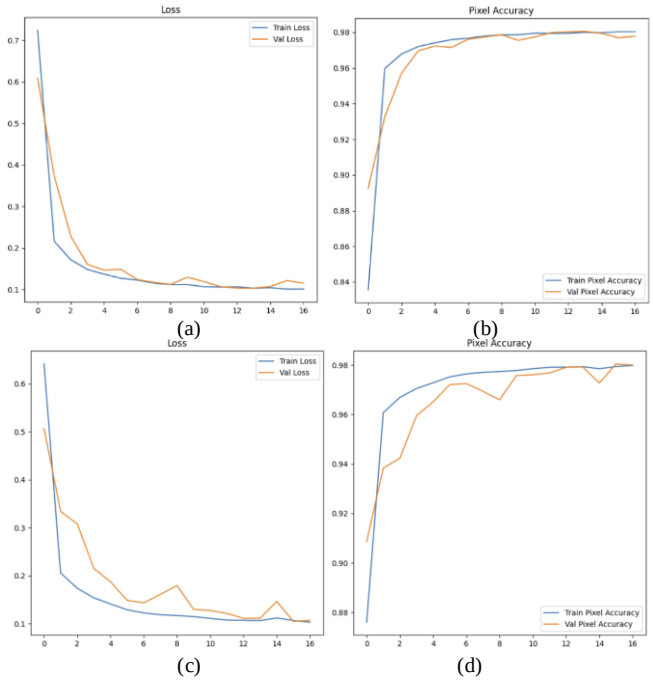
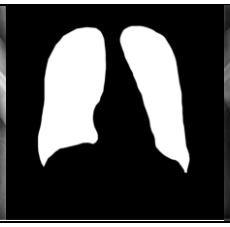
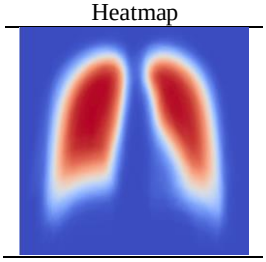


Figure 3. (a) Loss Result Graph with CLAHE (b) Pixel Accuracy Result Graph with CLAHE (c) Loss Result Graph without CLAHE (d) Pixel Accuracy Result Graph without CLAHE

Based on Figure 2, it can be concluded that the use of Contrast Limited Adaptive Histogram Equalization (CLAHE) has a positive impact on the model's performance compared to not using CLAHE. The graph in panel (a) shows that the loss with CLAHE decreases rapidly in the early stages of training and stabilizes at a low value, both for training and validation data. This indicates that the model adapts well without experiencing overfitting. In panel (b), the pixel accuracy for both training and validation data increases sharply, with train pixel accuracy approaching 100%, showing that the model can recognize objects effectively in both training and validation data. On the other hand, without CLAHE, the graph in panel (c) shows that although the loss also decreases, the validation loss remains higher than the training loss, suggesting a tendency for overfitting. In panel (d), while pixel accuracy increases, there is a significant difference between the train pixel accuracy, which is very high, and the validation pixel accuracy, which is lower, indicating that the model performs well on the training data but not as effectively on the validation data. Overall, the use of CLAHE improves the model's performance, reduces overfitting, and ensures that the model performs more consistently on the validation data compared to when CLAHE is not used.

TABLE II
Result of Image Segmentation

Original Image	Ground Truth (Mask)	Segmentation Result vs Ground Truth
		

Heatmap


Based on the provided image and the data from Table 2, it can be concluded that the segmentation results are highly accurate and closely match the ground truth. The original image, ground truth, and the segmentation result clearly show that the model successfully identifies and segments the lung regions, with a high degree of precision. The segmentation result vs. ground truth image demonstrates that the segmentation model correctly outlines the lungs in the X-ray, which is visualized in a highlighted form, showing that the model can efficiently segment the relevant areas of the chest X-ray.Moreover, the heatmap further reinforces this analysis by visually indicating areas with the highest confidence, represented by warmer colors. This heatmap serves as a visual interpretation of the model's confidence in its predictions, where the red areas highlight the regions the model is most confident about, which corresponds to the lung regions. In terms of performance metrics, the calculated True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP) values are highly favorable, indicating that the model has successfully segmented the lungs with minimal error. Specifically, the True Negative (TN) value of 3,455,781 and True Positive (TP) value of 1,102,335 suggest that the model correctly identified a large portion of both the background and lung regions. The False Negative (FN) of 58,860 and False Positive (FP) of 36,080 are relatively low, suggesting a minimal amount of incorrect predictions. From these values, the results are obtained as in Table 3. following:

TABLE III
MATRIX RESULTS OF THE RESEARCH

Matrix	Result With CLAHE	Result Without CLAHE
Pixel Accuracy	0.9796	0.9751
Recall	0.9493	0.9188
Specificity	0.9897	0.9939
Dice Coeficient	0.9587	0.9486
Jaccard Index	0.9207	0.9021

Based on Table 3, show that the use of CLAHE significantly enhances the model's performance across multiple evaluation metrics. Specifically, Pixel Accuracy improves from 0.9751 without CLAHE to 0.9796 with CLAHE, highlighting the model's increased overall segmentation accuracy when CLAHE is applied. The Recall also demonstrates a notable increase, rising from 0.9188 to 0.9493, indicating that CLAHE helps the model more accurately detect lung regions, thereby reducing false negatives. Additionally, the Dice Coefficient increases from 0.9486 to 0.9587, showing a stronger overlap between the predicted segmentation and the ground truth when CLAHE is used. The Jaccard Index improves from 0.9021 to 0.9207, further confirming that CLAHE enhances the similarity between the predicted and true masks. While there is a slight decrease in Specificity from 0.9939 without CLAHE to 0.9897 with CLAHE, the overall performance is still positively impacted by CLAHE, demonstrating its effectiveness in improving the segmentation quality and model accuracy for medical image analysis. Thus, CLAHE proves to be an important enhancement for improving the model's performance in lung segmentation tasks on chest X-ray images.

TABLE IV
COMPARISON OF RESULT WITH PREVIOUS RESEARCHERS

Method	Dataset	Accuracy	Loss	Jaccard Index	Dice Coeficient	Recall	Specificity
U-Net [24]	Shenzhen + Montgomery	91.47%	19.23	74.94	-	-	-
U-Net dan ResNet34 [25]	Shenzhen + Montgomery	-	-	-	93%	-	-
U-Net [26]	Shenzhen	-	-	85.19	91.68	-	-
AMRU++ [26]	Shenzhen	-	-	89.16	94.13	-	-
UNET [27]	JSRT + MC + SCHENZEN	93.87	-	91	91.87	98.33	-
UNET [28]	Montgomery	-	-	91.37	92.21	-	-
UNET + CLAHE (Phurpose Method	**Shenzhen + Montgomery**	**97.96**	**9.98%**	**92.07**	**95.87**	**94.93**	**98.97**

Based on Table 4, a comparison is shown between the results of several methods applied to the X-ray image dataset for segmentation. The Shenzhen dataset, consisting of 662 images (336 with tuberculosis manifestations and 326 normal cases), combined with the Montgomery dataset, consisting of 138 posterior-anterior X-ray images (80 normal and 58 abnormal with tuberculosis manifestations), was used for the U-Net model. This model achieved an Accuracy of 91.47%, IoU of 74.94%, and Loss of 19.23%, although data for Dice Coefficient, Recall, and Specificity are not available. The combination of U-Net and ResNet34 on the same dataset showed a Dice Coefficient of 93%, but no data were provided for Accuracy, Loss, IoU, Recall, and Specificity. Meanwhile, the U-Net model applied to the Shenzhen dataset alone (662 images) achieved an IoU of 85.19% and Recall of 91.68%, though data for Accuracy, Loss, and Dice Coefficient are missing. The AMRU++ model on the Shenzhen dataset reached an IoU of 89.16% and Recall of 94.13%, but other data were not provided. When U-Net was applied to the JSRT, MC, and SCHEZEN datasets, the model achieved an Accuracy of 93.87%, Dice Coefficient of 91%, and Specificity of 98.33%, though Loss, IoU, and Recall data were unavailable. The U-Net model applied to the Montgomery dataset yielded an Accuracy of 91.37%, Recall of 92.21%, and Dice Coefficient of 91%, while data for Loss, IoU, and Specificity were not reported. Finally, the U-Net + CLAHE method (Purpose Method) applied to the Shenzhen and Montgomery datasets demonstrated superior results compared to previous studies, achieving an Accuracy of 97.96%, IoU of 92.07%, Dice Coefficient of 95.87%, Recall of 94.93%, and Specificity of 98.97%.

IV. CONLUSION

Based on the results of the conducted study, it was found that the implementation of the U-Net architecture combined with the Contrast Limited Adaptive Histogram Equalization (CLAHE) method in the segmentation of lung images from tuberculosis patients achieved superior results compared to previous models, with an accuracy of 97.96%, sensitivity of 94.93%, specificity of 98.97%, Dice Coefficient of 95.87%, and Jaccard Index of 92.07%. This indicates that CLAHE helps to clarify the distinction between the lung objects and the background of the X-ray images of tuberculosis patients, allowing the U-Net model to predict pixels more accurately. Therefore, the combination of U-Net architecture and CLAHE shows great potential in improving the quality of medical image segmentation, particularly in the identification of tuberculosis in the lungs. For future research, it is recommended to develop a classification technique for lung abnormalities to expand the model's capability in detecting various abnormalities in X-ray images of tuberculosis patients.

REFERENCES

[1] N. Aini and H. Rahmania Hatta, "Expert System for Diagnosing Tuberculosis Disease," *Jurnal Informatika Mulawarman*, vol. 12, no. 1, p. 56, 2017, doi: 10.30872/jim.v12i1.224.

[2] K. Mar'iyah and Zulkarnain, "Pathophysiology of Tuberculosis Infectious Disease," *Prosiding Biologi Achieving The Sustainable Development Goals With Biodiversity In Confronting Climate Change*, vol. 7, no. 1, Nov. 2021, doi: 10.24252/psb.v7i1.23169.

[3] R. Ramadhan, E. Fitria, and R. Rosdiana, "Detection of Mycobacterium Tuberculosis by Microscopic Examination and PCR Technique in Patients with Pulmonary Tuberculosis at the Darul Imarah Health Center," *Sel Jurnal Penelitian Kesehatan*, vol. 4, pp. 73–80, Nov. 2017, doi: 10.22435/sel.v4i2.1463.

[4] T. A. Ghebreyesus and T. Kasaeva, *Global Tuberculosis Report 2024*. Geneva: World Health Organization, 2024. Accessed: Jun. 17, 2025. [Online]. Available: https://www.who.int/publications/i/item/9789240101531

[5] K. W. D. Nugraha, T. Seviana, and F. Sibuea, *Indonesia Health Profile 2022*. Jakarta: Kementerian Kesehatan Republik Indonesia.

[6] R. M. Natsir and F. Aipassa, "The Role of Anti-Tb Testing in Tuberculosis Early Detection Efforts in the Lower Taeno Hamlet Community," *Jurnal Kreativitas Pengabdian Kepada Masyarakat*, vol. 7, no. 12, pp. 5652–5660, Dec. 2024, doi: 10.33024/jkpm.v7i12.17622.

[7] A. A. Alimi, A. R. Adriansyah, and P. Prima, "Development of Tuberculosis Detection System on X-Ray Image Using Convolutional Neural Network (CNN) Method with Laravel Framework," *Jurnal Informatika Terpadu*, vol. 10, no. 2, pp. 165–171, Oct. 2024, doi: https://doi.org/10.54914/jit.v10i2.1437.

[8] I. Haq *et al.*, "Machine Vision Approach for Diagnosing Tuberculosis (TB) Based on Computerized Tomography (CT) Scan Images," *Symmetry (Basel)*, vol. 14, no. 10, Oct. 2022, doi: 10.3390/sym14101997.

[9] O. A. M. F. Alnaggar *et al.*, "Efficient Artificial Intelligence Approaches for Medical Image Processing in Healthcare: Comprehensive Review, Taxonomy, And Analysis," *Artif Intell Rev*, vol. 57, no. 8, p. 221, 2024, doi: 10.1007/s10462-024-10814-2.

[10] D. Priyawati, "Spatial-Domain Digital Image Processing Techniques for X-ray Image Enhancement," *Jurnal Komunikasi dan Teknologi Informasi*, vol. 2, no. 2, 2011.

[11] M. R. Pratama, S. Zahrah Hidayat, A. R. Nuruddin, H. W. Niamaputri, D. Fetty, and T. Anggraeny, "Image Contrast Enhancement Optimization Using Interval-Valued Intuitionistic Fuzzy Sets and Contrast Limited Adaptive Histogram Equalization (CLAHE)," *Prosiding Seminar Implementasi Teknologi Informasi dan Komunikasi*, vol. 4, no. 1, 2025, doi: 10.31284/p.semtik.2025-1.6884.

[12] E. Ijaz and I. Jr, "Quantitative Analysis of Image Enhancement Algorithms for Diverse Applications," vol. 5, pp. 694–707, Dec. 2023.

[13] P. Singh and A. K. Bhandari, "Globally and locally tuned filtering structure for high contrast intensity degradation," *Multimed Tools Appl*, vol. 83, no. 35, pp. 82471–82494, 2024, doi: 10.1007/s11042-024-18749-0.

[14] L. Li, Y. Si, and Z. Jia, "Medical Image Enhancement Based on CLAHE and Unsharp Masking in NSCT Domain," *J Med Imaging Health Inform*, vol. 8, pp. 431–438, Mar. 2018, doi: 10.1166/jmihi.2018.2328.

[15] D. Gunawan and H. Setiawan, "Convolutional Neural Network in Medical Images," *Konvergensi Teknologi dan Sistem Informasi*, vol. 2, May 2022, doi: 10.24002/konstelasi.v2i2.5367.

[16] H. Yu, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen, "Convolutional Neural Networks for Medical Image Analysis: State-of-the-art, Comparisons, Improvement and Perspectives," *Neurocomputing*, vol. 444, pp. 92–110, 2021, doi: https://doi.org/10.1016/j.neucom.2020.04.157.

[17] A. Susanto, Y. Kusumawati, E. Niagara, and A. Sari, "Convolutional Neural Network in Helmet Detection System for Motorcycle Riders," *Seminar Nasional Teknologi dan Multidisiplin Ilmu (SEMNASTEKMU)*, vol. 2, pp. 91–99, Dec. 2022, doi: 10.51903/semnastekmu.v2i1.158.

[18] M. M. I. Al-Ghiffary, C. A. Sari, E. H. Rachmawanto, N. R. D. Cahyo, N. M. Yaacob, and R. R. Ali, "Milkfish Freshness Classification Using Convolutional Neural Networks Based on Resnet50 Architecture," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 3, Aug. 2023, doi: 10.26877/asset.v5i3.17017.

[19] L. Alzubaidi *et al.*, "Review of deep learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions," *J Big Data*, vol. 8, no. 1, p. 53, 2021, doi: 10.1186/s40537-021-00444-8.

[20] S. Kordnoori, M. Sabeti, H. Mostafaei, and S. S. A. Banihashemi, "An Effective U-Net Model for Diagnosing Covid-19 Infection," *Intell Based Med*, vol. 10, p. 100156, 2024, doi: https://doi.org/10.1016/j.ibmed.2024.100156.

[21] A. Pravitasari *et al.*, "UNet-VGG16 with Transfer Learning for MRI-based Brain Tumor Segmentation," *Telecommunication Computing Electronics and Control*, vol. 18, p. 1310, Jun. 2020, doi: 10.12928/telkomnika.v18i3.14753.

[22] A. Misbullah, W. Mursyida, L. Farsiah, and K. Martiwi Sukiakhy, "Performance Analysis of Brain Tumor MRI Image Segmentation with U-Net and Res-UNet Architecture," *J-SIGN*, vol. 2, no. 2, pp. 83–95, 2024, doi: 10.24815/j-sign.v2i02.

[23] W. Liao, Y. Zhu, X. Wang, C. Pan, Y. Wang, and L. Ma, "LightM-UNet: Mamba Assists in Lightweight UNet for Medical Image Segmentation," *ArXiv*, Mar. 2024, doi: https://doi.org/10.48550/arXiv.2403.05246.

[24] J. D. Bodapati, R. Sajja, and V. Naralasetti, "An Efficient Approach for Semantic Segmentation of Salt Domes in Seismic Images Using Improved UNET Architecture," *Journal of The Institution of Engineers (India): Series B*, vol. 104, no. 3, pp. 569–578, 2023, doi: 10.1007/s40031-023-00875-2.

[25] A. Santone, R. De Vivo, L. Recchia, M. Cesarelli, and F. Mercaldo, "A Method for Retina Segmentation by Means of U-Net Network," *Electronics (Basel)*, vol. 13, no. 22, 2024, doi: 10.3390/electronics13224340.

[26] U. Tarik, F. Ali, and S. Dawwd, "U-Net Convolutional Networks Performance Based on Software-Hardware Cooperation Parameters: A Review," *International Journal of Computing and Digital Systems*, vol. 11, pp. 977–990, Mar. 2022, doi: 10.12785/ijcds/110180.

[27] Y. Wu and Q. Li, "ConvNeXt embedded U-Net for semantic segmentation in urban scenes of multi-scale targets," *Complex & Intelligent Systems*, vol. 11, no. 4, p. 181, 2025, doi: 10.1007/s40747-024-01735-2.

[28] I. Dimitrovski, V. Spasev, S. Loshkovska, and I. Kitanovski, "U-Net Ensemble for Enhanced Semantic Segmentation in Remote Sensing Imagery," *Remote Sens (Basel)*, vol. 16, no. 12, 2024, doi: 10.3390/rs16122077.

[29] S. Hashem and M. Kamil, "Segmentation of Chest X-Ray Images Using U-Net Model," *MENDEL*, vol. 28, pp. 49–53, Dec. 2022, doi: 10.13164/mendel.2022.2.049.

[30] J. Sousa *et al.*, "Lung Segmentation in CT Images: A Residual U-Net Approach on a Cross-Cohort Dataset," *Applied Sciences (Switzerland)*, vol. 12, no. 4, Feb. 2022, doi: 10.3390/app12041959.

[31] Md. S. Alam *et al.*, "Attention-Based Multi-Residual Network for Lung Segmentation in Diseased Lungs with Custom Data Augmentation," *Sci Rep*, vol. 14, no. 1, p. 28983, 2024, doi: 10.1038/s41598-024-79494-w.

[32] S. Arvind, J. V Tembhurne, T. Diwan, and P. Sahare, "Improvised light Weight Deep CNN based U-Net for The Semantic Segmentation of Lungs from Chest X-rays," *Results in Engineering*, vol. 17, p. 100929, 2023, doi: https://doi.org/10.1016/j.rineng.2023.100929.

[33] M. Rahman, B. Tseng, M. Pokojovy, W. Qian, B. Totada, and H. Xu, *An Automatic Approach to Lung Region Segmentation in Chest X-Ray Images Using Adapted U-Net Architecture*. 2021. doi: 10.1117/12.2581882.

[34] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241.