The I0th International Conference on Control, Automation and Information Sciences
Xi'an, China - October 14-17, 2021

A Novel Lesion Segmentation Algorithm based on U-Net Network for Tuberculosis CT Image

Shaoyue Wen[1], Jing Liu[2], Wenge Xu[3]
[1]Department of Computing, Xi'an Jiaotong-Liverpool University, Suzhou, China
[2]Department of Radiology, The Fifth People's Hospital of Suzhou, Suzhou, China
[3]DMT Lab, Birmingham City University, Birmingham, United Kingdom

2021 International Conference on Control, Automation and Information Sciences (ICCAIS) | 978-1-6654-4029-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCAIS52680.2021.9624633

***Abstract*—Lung CT images provide several essential information for lung disease diagnosis and lung surgery. However, the traditional detection method through manual segmentation is laborious and time-consuming. This paper presents automatic tuberculosis (TB) lesion segmentation method based on U-Net neural network for detecting TB. In addition, we combined an edge detection algorithm called canny edge detector with this network to get a more accurate TB lesion boundary. This method is trained on two split databases with 3576 lung CT images obtained by data enhancement on 447 discontinuous lung CT images. The results show that the proposed approach is validated for complex TB lesions with a high dice coefficient (91.2%).**

***Index Terms*—Image processing, lesions segmentation, Unet, tuberculosis**

I. INTRODUCTION

Tuberculosis (TB) is an infectious bacterial disease caused by Mycobacterium tuberculosis (Mtb), which is transmitted between humans through the respiratory route and most commonly affects the lungs but can damage any tissue [1]. TB remains the world's leading cause of death from infectious agents, and it is one of the three highest mortality rate diseases worldwide [2, 3]. Early diagnosis and treatment have important clinical significance for treating tuberculosis, ensuring patients' physical health, and prevent TB from proliferating to the public.

CT image examination is the primary aiding method to diagnosis TB. However, to clarify the disease progression and give a precise diagnosis, millions of medical images from infected patients are accumulated to be examined, which would take a significant amount of time and effort from clinicians. What's more, manual examinations are prone to mistakes and have the potential for bias [4]. Therefore, it is essential to provide a computer-based framework that can automatically evaluate TB radiological manifestations in a short time with high accuracy.

Over the past decade, several segmentation algorithms have been proposed. There are many systems explored by researchers, including threshold method [5], clustering method [6], and regional growth method has been proposed and evaluated by researchers. However, these methods might reduce the gray level of the image and lead to details of the lungs being blurred [7]. Deep learning algorithms based on conventional neural networks (CNN) have recently become famous for visual recognition tasks [8]. The CNN-based networks like Alexnet, VGGNet, GoogleNet have proved their performance for visual recognition tasks [9]. However, there is little work has been done using CNN-based U-Net for medical image segmentation tasks. This paper aims to combine the U-Net and edge detection algorithms for segment TB lesions from CT images and evaluate its performance for medical image segmentation.

*Corresponding author. E-mail address: Liu8728@126.com

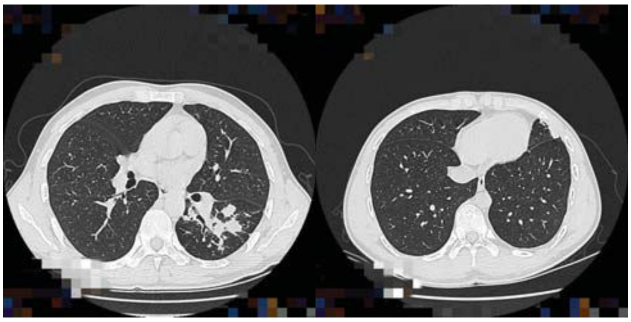
II. RELATED RESEARCH

Tuberculosis recognition has been a hit subject in medical image recognition research. Researchers used to implement traditional image processing methods to extract features in the past [5]. With the rapid improvement of the deep learning (DL) field, several high efficient deep learning algorithms and systems have been proposed.

Pham et al. [10], discussed the critical evaluation of the current period of a semi-automated and automated algorithm for medical image separation. Throttling, region growth and artificial neural networks are currently available. The combination of local and global classifications increased the ROC value to 84 %. Arzhaeva et al. [11] have presented a multivalued method to label chest X-ray images by combining essential image features with global image features and solving the problem of lacking accuracy. Based on X-ray images, J Jen Hong Tan, U. Rajendra Acharya and Collin Tan [12] have developed an artificially intervening serpentine algorithm based on X-ray images and extracted required regions. Then, detect TB by applying on C4.5 algorithm decision tree system, achieving an accuracy of 89.9 % in normal-tuberculosis classification. Wu et al. [13] a CT-based Tuberculosis Small Point Recognition Algorithm by using image masking and Gaussian filtering in the pre-processing. Then, they used the naive Bayesian algorithm for classification, and their proposed algorithm has reached 83.4% TPR and 14.4% FPR. Li et al. [14] have proposed a novel system called AE-CNN. In this method, they combined Conv and Auto Encoder, which emphasized unsupervised features. The AE-CNN process can do abnormal TB segmentation and classification. Yang et al. [15] have constructed an improved system of combing two DL networks U-Net and Faster-RCNN. They have optimized the Res block and found that this module has improved the sensitivity rate to 84%.

Although the above algorithms have been verified in many applications, there is a strong need for a higher efficiency

978-1-6654-4029-5/21/$31.00 ©2021 IEEE



(a) (b)

Fig. 1. Sample of lung CT images where (a) shows a lung CT with TB, (b) is normal lung CT image

algorithm. This is because these methods require a higher amount of data sample for accurate detection. Through a more robust connection between layers, upsampling and down convolution, full extraction of features can be realized by U-net, fewer training samples are required to obtain accurate segmentation.

III. METHODOLOGY

In this section, we will discuss our proposed algorithm in detail. The first step in this method is image pre-processing removing noise from the image. Then neural networks algorithm is applied to segment the TB lesion of the lung. Finally, the edges can be extracted according to the Canny edge detector.

A. CT Image Pre-Processing

Firstly, we randomly selected 40 patients who were hospitalized in the the fifth people's hospital of Suzhou. For each patient, we collected more than 260 CT images. In total, more than 10000 CT images were collected in the initial stage. Then, we removed repetitive images and images where TB lesions would not exist in these CT image slices. Figure 1 shows sample CT images used in our dataset. The normal Lung CT Image is posted in Fig. 1(a). Fig. 1(b) demonstrates the CT images infected with TB.

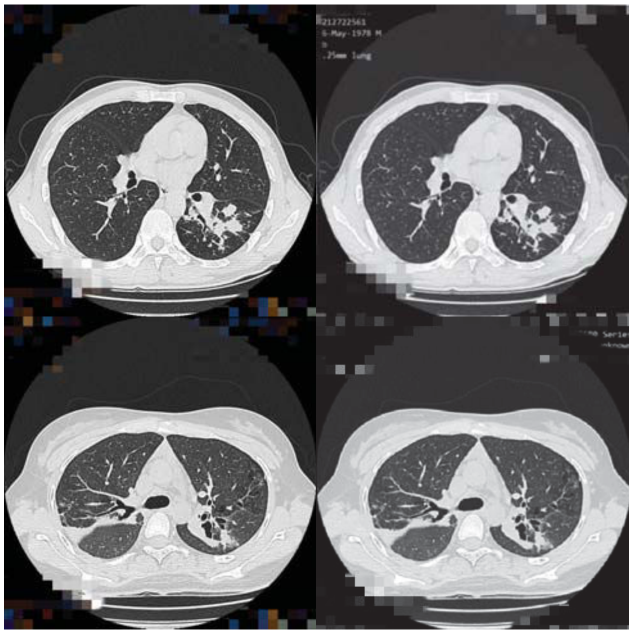
Second, we apply a Gaussian filter, a smooth linear filter, to process the noise of these medical images. Gaussian filter could can help to weaken sharp changes in the image gray and then reduce the interference caused by lung bubbles and blood vessels (see Fig. 2). The Gaussian formula is given by:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

B. U-Net

1) U-Net Architecture: U-Net is a popular segmentation DL networks proposed by Ronneberger et al. [16] in 2015, which mainly contains encoder-decoder neural architecture and skip connections. The U-Net system adopted in this paper is built under the segmentation models of Keras[1] and

[1]https://keras.io/



(a) (b)

Fig. 2. The Results of Gaussian filter where (a) shows original image, (b) is the result image of Gaussian Filter

Tensorflow[2]. The concept map of the architecture of the U-net is illustrated in Fig 3.

In the network structure, the propagation process pointed by a purple arrow is composed of 3 × 3 convolution layer and ReLu activation layer. The red arrows refer to the 2 × 2 pooling layer. The processes including bright green arrow consist of 2 × 2 convolution layer and Sigmoid activation function.

U-Net is famous for its symmetrical structure: a shrinking path on the left applied to extract image features and an expanding path used for precise positioning. On the right. Each layer in the shrinking path comprises two 2 × 2 convolution layers and non-linear function ReLU (Linear Unit) used as the activation function.

In each step of the sub-sampled, the number of feature channels is accumulated. During the expansion path, the obtained feature maps are up-sampled. After that, 2 × 2 deconvolution operations are performed to reduce feature channel numbers. Then, combine feature maps gained from sub-sampled and up-sampled procedures in a cascaded way. In addition, 3 × 3 convolution operation is implemented and Re-Lu worked as the activation function. The last layer of the network performs the 1 × 1 convolution operation.

2) Adam optimization algorithm: Adaptive Moment Estimation Adam algorithm is a traditional gradient descent optimization algorithm. Adam algorithm dynamically adjusts the learning rates by calculating the first-order matrix and second-order matrix values of the gradient, so that different parameters of the model have adaptive learning rates during the training process. Apart from this, the Adam algorithm

[2]https://www.tensorflow.org/

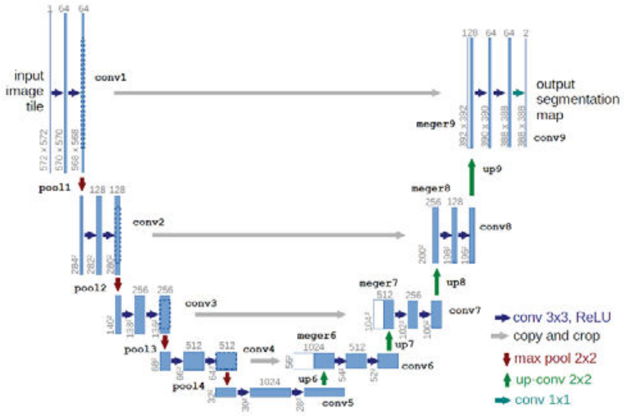


Fig. 3. Proposed U-Net network framework. Each blue box corresponds to a mufti-channel feature map

has little requirement for storage and equips with the characteristics of gradient focus scaling without deformation. Owing to its high efficiency in the calculation, it has become one of the most widely used optimization algorithms in deep learning at present.

In the iterative process, gradient calculation formula is used as show above:

$$\frac{1}{m}\nabla_{\theta}\sum_{i} L(f(x^{(i)},\theta),y^{(i)}) \rightarrow g \tag{2}$$

Update partial first order matrix estimators:

$$\begin{gathered} \beta_1 s+(1-\beta_1)g \rightarrow s \\ \beta_2 s+(1-\beta_2)g^2 \rightarrow r \end{gathered} \tag{3}$$

Correction of first order matrix estimation bias:

$$\frac{s}{1-\beta_1^t} \rightarrow \hat{s} \tag{4}$$

Update

$$\begin{gathered} \Delta\theta = -\alpha\frac{\hat{s}}{\sqrt{\hat{r}}+\delta} \\ \theta \leftarrow \theta+\Delta\theta \end{gathered} \tag{5}$$

C. Edge detection and boundary extraction

Edge is the most crucial feature for focal disease in CT images, and this feature can be used in target recognition and segmentation. Training models in the U-Net network to identify the lesion accurately and predict the locations of the lesions successfully can only obtain a rough segmentation. Numerous Slices of lung CT images are filled with capillaries and other substances. The edge detection algorithm is used in conjunction with the U-Net network to provide better segment lesions performance.

Compared to Sobel and Prewitt algorithm, canny is a further refinement with more accurate positioning. Hence, we choose the canny algorithm to extract boundaries from the target images. The results of boundary extraction are shown in Fig. 4. It can be seen that, after applying the contour extraction method, the boundaries of regions of foreground pixels which are typically bright in the image, are obtained and output the results with Sigmoid function.

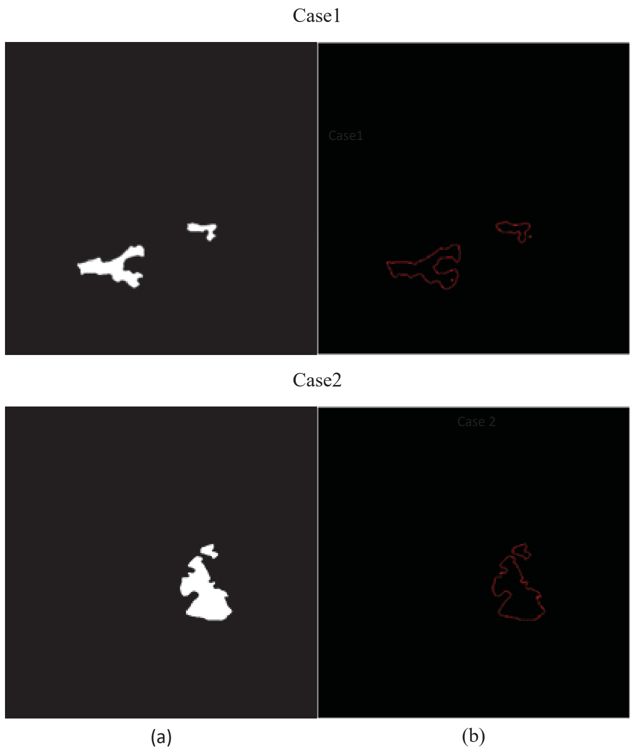


Fig. 4. The results of boundary detection where (a) shows the lesions of lung CT image, (b) is the boundary of lesions

TABLE I
AMOUNT OF DATA USED FOR TRAINING, VALIDATION, TEST BY SPLIT 1 AND SPLIT 2

Data	Split 1	Split 2
Training	355	302
Validation	56	45
Test	36	100
Total	447	447

IV. EXPERIMENT

We demonstrated the segmentation results on challenging datasets containing unpredictable factors. The network in Section 3.2 was implemented using the open-source Keras and Tensorflow framework. Model training was performed on an NVIDIA 11GB RTX 3090ti graphic card. All examples are implemented on a 3.2 GHz PC with an i9 CPU and 8GB RAM.

A. Initialization of Parameters

1) Training, validation, and testing sets: In order to obtain robustness of the results, two different splits of training, validation, and testing datasets were choose (Table 1). In Split 1, data were arranged with a relatively large number of training (validation) dataset and a small number of only 36 testing-set patients were selected. On the contrary, Split 2 had more cases in testing dataset with 100 and 347 in training and validation datasets. The datasets in this study were collected from the fifth people's hospital of Suzhou.

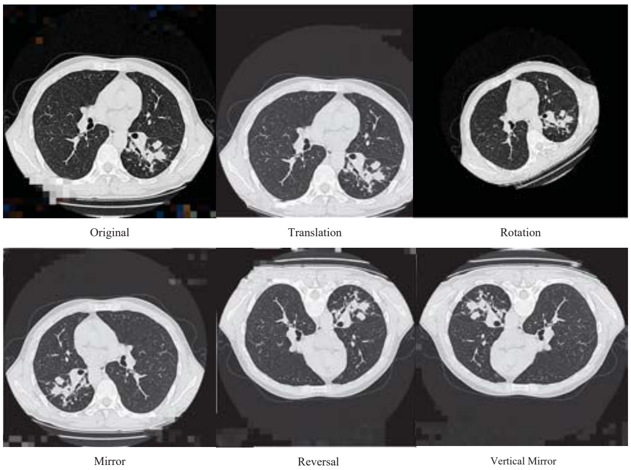


Fig. 5. Image Enhancement

2) Image Enhancement : To expand the training dataset, simple image enhancement was performed on the original images and label images such as rotation, flip, translation, and so on. Fig. 5 shows the results of CT image enhancement. The first image is the original image, and the subsequent images are the results of translation, rotation mirror, and flip. The expanded training data set was input into the model for training later.

3) Image Labeled: Before the training, the image should be labeled, and software "Labelme" was used in this experiment. The labeled CT images are shown in Fig. 6.

4) Parameters setting: The batch size of training was 32 and the epoch of training was 10. The parameter values of the Adam optimization algorithm were $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, δ=1×10^{-8}, respectively.

B. Experiment Results

Figure 7 and 8 illustrates the loss curve and accuracy curve during training and validation for Split 1 and Split 2 scenarios. Both loss curves are converging at last, which suggests that our training step is successful. The sharp growth and decline in Split 2 (see Fig. 8) might be due to the inadequate training set since Split 2 used a fewer training dataset. Therefore, Split 2 tends to have fewer kinds of TB lesions which could cause a transient change in Val loss.

As the training cycle accumulated, the error on the training set and the validation set were in continuous reduction in Split 1 and Split 2. At the same time, the accuracy of segmentation was continuing to improve. After 8 echoes, Split 1 training and validation accuracy reached 98.73%, Split 2 is 98.24%. Figure 9 illustrates the segmented results.

C. Evaluation metrics

The standard evaluation metrics that are commonly used in medial field are precision, recall, and dice coefficient [15]. Precision and Recall are calculated based on the following equations:

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

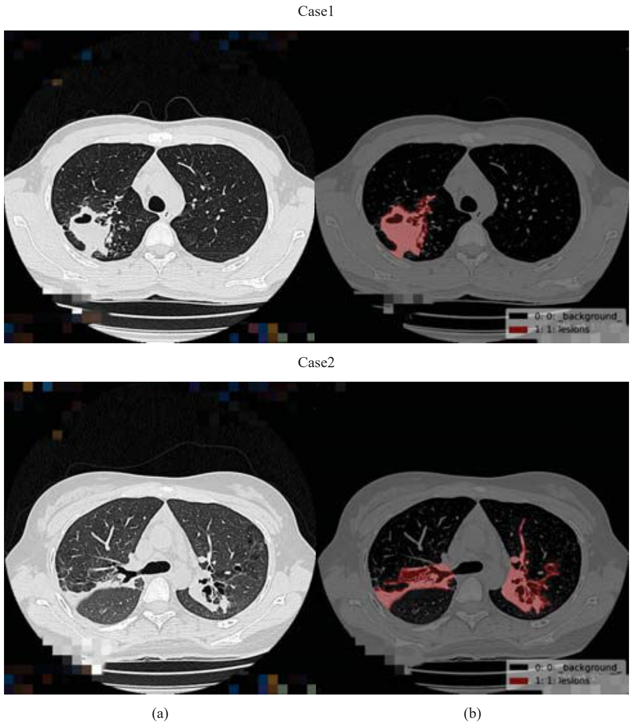


Fig. 6. FThe labeled CT images where (a) shows the original CT images, (b) is the labeled images

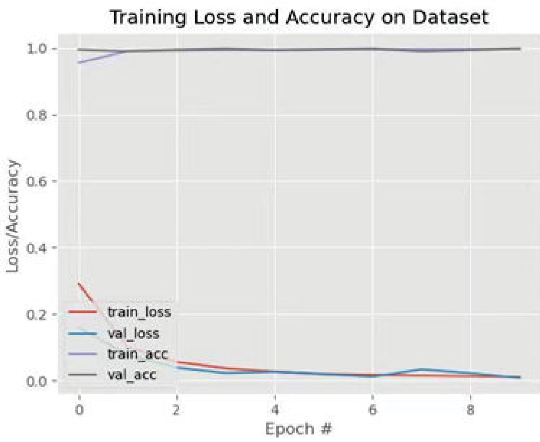


Fig. 7. The loss curve of Split 1

Recall is calculated based on the following equation:

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

Where TP refers to the true positive, FP refers to the false positive, FN refers to the false negatives which is the number of pixels mistakenly labeled as non-TB.

Dice coefficient (DSC) is another important measurement for segmenting multi-class image, it can be calculated by:

$$DSC = \frac{2\,|A \cap B|}{|A| + |B|} \tag{8}$$

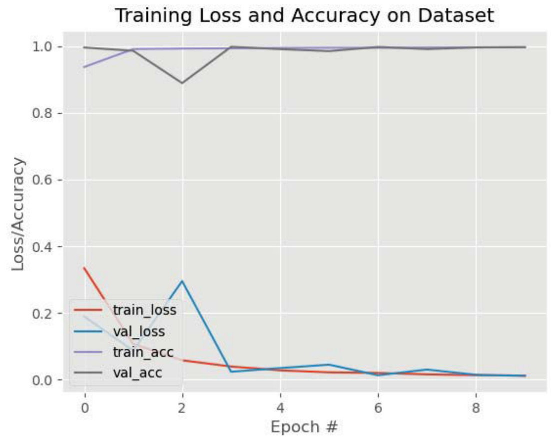


Fig. 8. The loss curve of Split 2

TABLE II
THE PERFORMANCE OF OUR PROPOSED ALGORITHM AND THE SELECTED BENCHMARK

Method	Precision (%)	Recall (%)	DSC (%)
U-Net (Split 1)	90.9	89.3	91.2
U-Net (Split 2)	81.5	88	86.7
YOLOv4 [17]	80	58	63.6

Where A and B denote the predicted and ground-truth masks.

Table 2 shows the evaluation results. It can be seen from the table that both Split 1 and Split 2 outperformed the benchmark YOLOv4 [17], suggesting U-Net architecture could achieve better performance to segment TB lesions. We also found that Split 1 had better performance (in all evaluation metrics) than Split 2. As expected, a higher number of training datasets could improve the performance of predicting.

D. Discussion, Limitation, Future Work

Our results suggest that using U-Net with a canny algorithm could achieve outstanding results in detecting CT images with TB. However, we also observed a few limitations of this research. First, we only used lesions. Other features/signs such as pericardium, bones, and lymph nodes may also help diagnose TB. Second, due to the lack of relevant training samples, other pulmonary lesions, such as infectious diseases (bacteria, fungi, viruses, and so on) and non-infectious diseases (tumors and vasculitis, and so on) could not be correctly identified and could be misjudged as a certain type of TB. Future work aims to solve the problems in Fig. 10, where lesions are stick to the alveolar. In addition, we aim to overcome the issue presented in Fig. 11 where multiple lesions appear in one area but the network could only classify it as a single lesion.

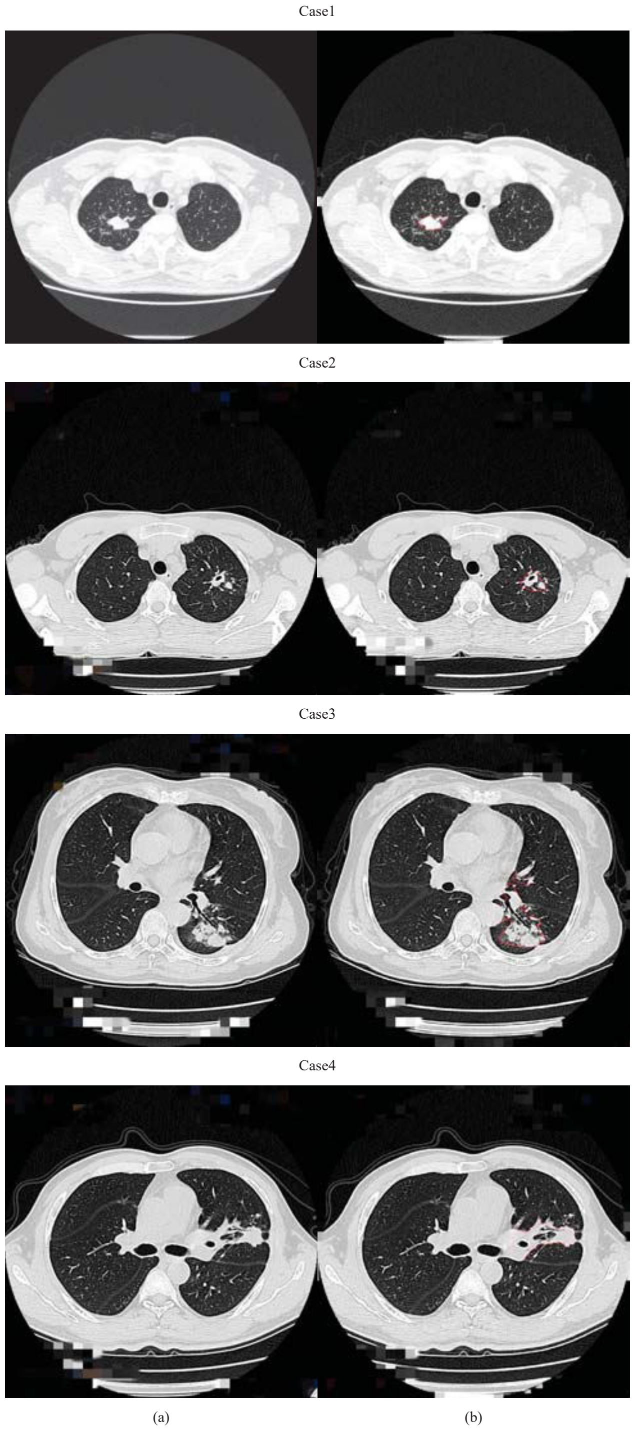


Fig. 9. The segmentation results of TB lesion CT image where (a) is original images, (b) shows the U-net segmented images

V. CONCLUSION

This paper presents a novel algorithm that combines U-Net with a canny algorithm for tuberculosis (TB) detection. Our experimental results show that the proposed segmentation method is effective, achieving a precision, recall, dice coefficient of 90.9%, 89.3%, and 91.2%, respectively.

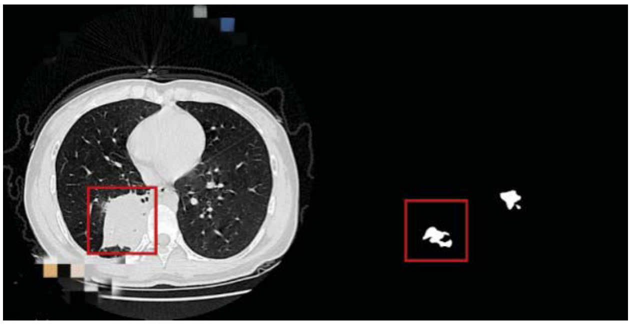


Fig. 10. Lesions are stick to alveolar

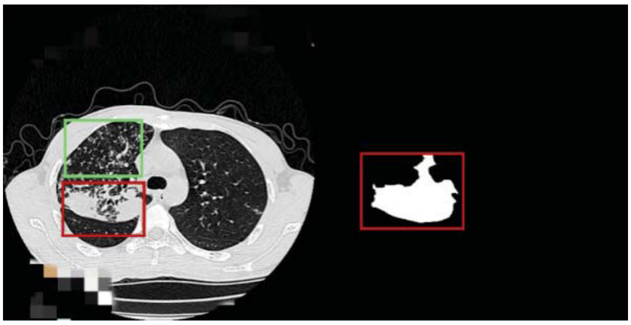


Fig. 11. Complex/Multiple TB lesions appear in one area

References

[1] T. W. Bank, "report of a mission organized by the international bank for reconstruction and development and the food and agriculture organization of the united nations at the request of the government of spain," 2017.

[2] W. H. Organization, "Global tuberculosis report 2018," 2018.

[3] B. R. Bloom, R. Atun, T. Cohen, C. Dye, H. Fraser, G. B. Gomez, G. Knight, M. Murray, E. Nardell, E. Rubin, J. Salomon, A. Vassall, G. Volchenkov, R. White, D. Wilson, and P. Yadav, "Tuberculosis," in *Major Infectious Diseases*, 3rd ed., K. K. Holmes, S. Bertozzi, B. R. Bloom, and P. Jha, Eds. Washington (DC): The International Bank for Reconstruction and Development / The World Bank, Nov. 2017.

[4] D. Núñez-Fernández, L. Ballan, G. Jiménez-Avalos, J. Coronel, P. Sheen, and M. Zimic, "Prediction of tuberculosis using u-net and segmentation techniques," 2021.

[5] A. Karthikeyan and M. Valliammai, "Lungs segmentation using multi-level thresholding in ct images," *International Journal of Electronics & Computer Science Engineering*, vol. 1, no. 3, 2012.

[6] "Lung segmentation on standard and mobile chest radiographs using oriented gaussian derivatives filter," *Biomedical Engineering Online*, vol. 14, no. 1, p. 20, 2015.

[7] P. Filho, P. C. Cortez, and V. Albuquerque, "3d segmentation and visualization of lung and its structures using ct images of the thorax," *Journal of Biomedical Science & Engineering*, vol. 6, no. 11, pp. 1099–1108, 2013.

[8] D. kumar and B. ganesh, "Segmenting the lung nodule image and classification by k-nn," *International Journal of Applied Engineering Research*, vol. 10, pp. 19 614–19 617, 01 2015.

[9] D. T. Kushnure and S. N. Talbar, "Ms-unet: A multiscale unet with feature recalibration approach for automatic liver and tumor segmentation in ct images," *Computerized Medical Imaging and Graphics*, vol. 89, p. 101885, 2021.

[10] Pham, Dzung, L., Chenyang, Prince, Jerry, and L., "Current methods in medical image segmentation." *Annual Review of Biomedical Engineering*, vol. 2, no. 1, pp. 315–315, 2000.

[11] Y. Arzhaeva, L. Hogeweg, P. A. de Jong, M. A. Viergever, and B. van Ginneken, "Global and local multi-valued dissimilarity-based classification: application to computer-aided detection of tuberculosis." *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 12, no. Pt 2, pp. 724–731, 2009, place: Germany.

[12] J. H. Tan, R. Acharya, C. Tan, T. Abraham, and C. M. Lim, "Computer-assisted diagnosis of tuberculosis: A first order statistical approach to chest radiograph," *Journal of Medical Systems*, vol. 36, no. 5, pp. 2751–2759, 2012.

[13] Y. Wu, H. Wang, and F. Wu, "Automatic classification of pulmonary tuberculosis and sarcoidosis based on random forest," in *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2017.

[14] L. Li, H. Huang, and X. Jin, "Ae-cnn classification of pulmonary tuberculosis based on ct images," in *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, 2018.

[15] A. Yang, X. Jin, and L. Li, "Ct images recognition of pulmonary tuberculosis based on improved faster rcnn and u-net," in *2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, 2019, pp. 93–97.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015.

[17] A. Bochkovskiy, C. Y. Wang, and H. Liao, "Yolov4: Optimal speed and accuracy of object detection," 2020.