Segmentation Trachea and Bronchial Branches in Chest Computed Tomography Image by Deep Learning - preliminary results -

Hoang Nhut Huynh[1,2,*], My Duyen Nguyen[1,2], Thai Hong Truong[1,2], Quoc Tuan Nguyen Diep[3], Anh Tu Tran[1,2] and Trung Nghia Tran[1,2,*]

[1] Department of Biomedical Engineering Physics, Faculty of Applied Sciences, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, ward 14, District 10, Ho Chi Minh City, Vietnam
[2] Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam
3 Dr Quoc Co. Ltd., Hung Long Ward, Binh Chanh District, Ho Chi Minh City, Vietnam
nhut.huynhvlkt@hcmut.edu.vn, ttnghia@hcmut.edu.vn

Abstract

Segmentation is one of the most common methods for analyzing and processing medical images, assisting doctors in making accurate diagnoses by providing detailed information about the required body part. However, segmenting medical images presents a number of challenges, including the need for medical professionals to be trained, the fact that it is time-consuming and prone to errors. As a result, it appears that an automated medical image segmentation system is required. Deep learning algorithms have recently demonstrated superior performance for segmentation tasks, particularly semantic segmentation networks that provide a pixel-level understanding of images. U-Net for image segmentation is one of the modern complex networks in the field of medical imaging; several segmentation networks have been built on its foundation with the advancements of Recurrent Residual convolutional units and the construction of recurrent residual convolutional neural network based on U-Net (R2U-Net). R2U-Net is used to perform trachea and bronchial segmentation on a dataset of 36,000 images. With a variety of experiments, the proposed segmentation resulted in a dice-coefficient of 0.8394 on the test dataset. Finally, a number of research issues are raised, indicating the need for future improvements.

Keywords: medical image, trachea, bronchial, segmentation, deep learning, R2U-Net

---

* Corresponding author

1 Introduction

Medical image segmentation is useful for evaluating and processing medical images since it divides the image into multiple useful parts based on pixel properties including intensity, color, and structure. Segmentation can be done by using a variety of image processing methodologies and procedures. Manual segmentation, carried out by specialists, is the most common practice. However, it is time-consuming and strongly reliant on the implementer's experience.

With the development of artificial intelligence, the convolutional neural networks (CNN) have shown remarkable success in a variety of computer vision applications, including image classification and segmentation referring to pixel-level insight images by labeling layers for each pixel [1]. The U-Net is a fully convolution neural network (FCN) that was constructed with an encoder and decoder network employing concatenations that allowed for greater segmentation accuracy [2 - 5]. It immediately became a leading approach for segmenting by splitting pixels into numerous categories based on their intensity value. If the pixel intensity is much larger than a threshold value, it is regarded as a foreground or part of the subject. Otherwise, pixels are considered background. However, the threshold method is sensitive to noise due to the inability to capture the spatial characteristics of the image. In this paper, a modified U-Net model is proposed for segmenting the trachea and bronchial branches from the Computed Tomography (CT) images to segment with the preliminary results. The segmented images can aid experts in accurately identifying locations, detecting abnormalities and monitoring the disease's progression.

2 Materials and Methods

2.1 Datasets

With the view toward the identification and location of the trachea and bronchial branches, a training dataset is used with 36,000 training DICOM images, including 12,000 chest CT images, 12,000 mask images for the trachea region, and 12,000 images for the bronchial branch area. The test dataset included 6,000 test images, including 2,000 chest CT images, 2,000 trachea mask images, and 2,000 bronchial branch mask images. The intensity value of the mask in relation to the trachea area and bronchial branches is set to 0, while the others are set to 1. Each image has a size of 1,024× 1,024 pixels and is on the grayscale. The distribution of training and testing datasets is shown in Table 1. The dataset was randomly divided into 75 percent training and 25 percent validation.

Attribute	Training	Validation	Testing
Chest CT	9,000	3,000	2,000
Trachea mask	9,000	3,000	2,000
Bronchial branch mask	9,000	3,000	2,000

Table 1: An overview of the datasets

2.2   Pre-processing and data argument

For improving data quality and reducing the image size to fit the current deep learning model, the preprocessing step was performed. The DICOM pixel arrays were extracted and then converted into three red, green, and blue channels. Then, the image was resized to 512×512 pixels. Figure 1 illustrates the chest CT image, trachea mask, and bronchial branch mask. Data enhancement techniques were used to avoid overfitting by rotating, shifting, adjusting images.
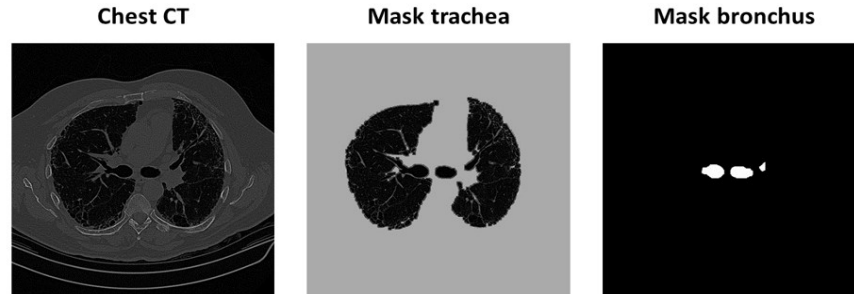

Figure 1: Examples of chest CT photos (left), trachea mask (center) and bronchial branch mask (right).

2.3   Architecture of Segmented Networks

In this study, a recurrent residual convolutional neural network based on U-Net (R2U-Net) is proposed with the U-Net network advancements and recurrent residual convolutional units (RRCU) [6]. Furthermore, the U-Net network has concatenation connections, which allow information to be transferred from encoder to decoder at corresponding levels to obtain more accurate segment images. The recurrent residual convolutional units are used with the feedback connection for storing information over time and the contextual information, which means that as the time steps increase, the more information the network takes advantage of in the vicinity.

The network's input is a data folder containing CT images and masks corresponding to the object to be segmented, and the desired output is the image as shown in Figure 2 containing the location information of the object to be partitioned. The trachea and bronchial Zoning were trained in two separate pieces of training in this study. The CT image and trachea mask are the input data pair for trachea zoning, and the CT image and trachea mask are the input data pair for bronchial branch zoning. Both of these pieces of training are trained with the same parameters for 100 epochs with Stochastic Gradient Descent optimizers (SGD) with momentous values of 0.9 [7], [8]. The learning rate is set to 1e-3 and gradually reduced to 1e-5, with a batch size of 5 to 10 images.
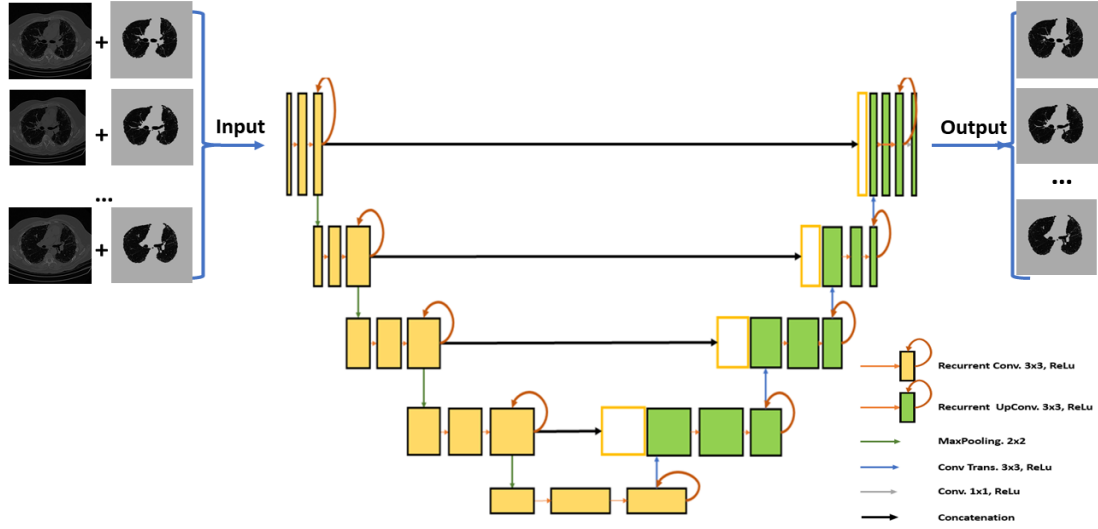
Figure 2: R2U-Net network architecture and training process

2.4 Experiments Setup

The proposed R2U-Net model was built in Python and use well-known libraries such as Keras (https://keras.io/) and Tensorflow (https://www.tensorflow.org/) as an alternative program. In addition, several Python packages were used including OpenCV (https://opencv.org/), pydicom for processing DICOM images [9]. All training was done on a workstation with an NVIDIA GeForce GTX 1050Ti GPU and 32GB RAM. The parameters for training experiments are shown in Table 2.

Experiment	Optimizer	LR schedule	Batch size	Epoch
EXP1	Adam	1e-4 to 1e-6	10	60
EXP2	SGD	1e-3 to 1e-5	6	80
EXP3	Adam	1e-4 to 1e-6	6	80

Table 2: Training parameters for experiments

2.5 Metrics

Segment performance is evaluated using Dice-coefficient and Dice-coefficient loss [10]. They are the loss jaws for the training model. The definition of Dice Coefficient is:

$$Dice - coef = \frac{2\sum_i^N g_i p_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{1}$$

$$Dice - coef\ loss = 1 - \frac{2\sum_i^N g_i p_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{2}$$

where the sums run over the N voxels, of the predicted binary segmentation volume $pi \in P$ and the ground truth binary volume $gi \in G$.

3 Result and Discussion

These two trachea area objects and two bronchial branches are segmented and located from two datasets.

Experiment	Train step time (s)	Train time/epoch (s)	Dice-coef.
EXP1	6.7	2010	0.7795
EXP2	5.8	1740	0.7535
EXP3	6.8	2040	0.8045

Table 3: Training time per epoch, step-by-step training time, and Dice-coef. points of confirming data of experiments.

Table 3 summarizes the model's components of training time, training time per epoch, and Dice-coef for each experiment. For the first experiment, each training step about 6.7 seconds, each training time per epoch about 2,010 seconds, and the Dice-coef value was 0.7795. In the second experiment, each training step about 5.8 seconds, each training time per epoch about 1,740 seconds, and the Dice-coef value was 0.7535. In the third experiment, each training step about 6.8 seconds, each training time per epoch about 2,040 seconds, and the Dice-coef value was 0.8045.

Experiment	Train step time (s)	Train time/epoch (s)	Dice-coef
EXP1	6.1	1830	0.8007
EXP2	5.7	1710	0.8030
EXP3	6.3	1890	0.8394

Table 4: Training time per epoch, step-by-step training time, and Dice-coef points of confirming data for experiments

Table 4 summarizes the components of the model of training time, training time per era, and Dice-coef for each experiment. For the first experiment, each training step about 6.1 seconds, each training time per epoch about 1,830 seconds and reached a Dice-coef value of 0.8007. For the second experiment, each training step took about 5.7 seconds, each training time per epoch about 1,710 seconds and reached a Dice-coef value of 0.8030. For the third experiment, each training step about 6.3 seconds, each training time per epoch about 1,890 seconds and reached a Dice-coef value of 0.8394.

The best-performing experiment from validation was selected. Tables 3 and 4 demonstrate how the model's performance varies depending on the condition. Figure 3 shows segmentation results using a chest CT image.
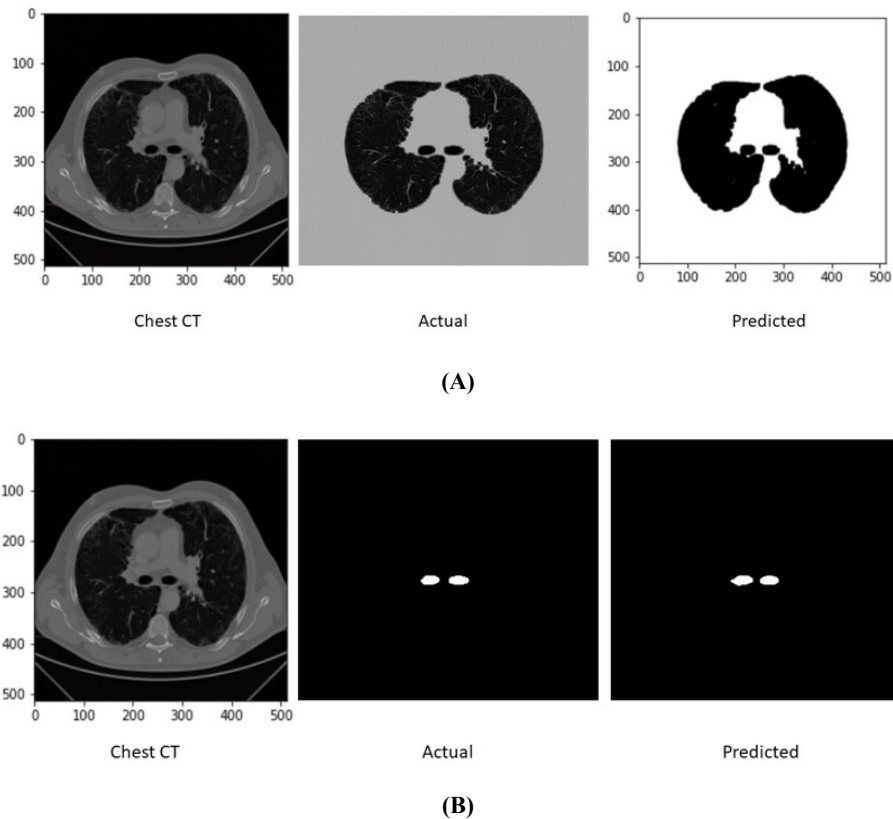
(A)

(B)

Figure 9: Segmentation results using a chest CT image (left), a real mask (middle), and a predictive mask (right): (A) for trachea and (B) for bronchial branch.

4 Conclusion

This paper proposed a method for segmenting the trachea region and bronchial branch using the R2U-Net model, with experiments performed under various conditions. The best Dice-coefficient values obtained were 0.8045 and 0.8394 for two cases of trachea and bronchial branch segmentation. The data enhancement method was used to generalize the model and obtain the Dice-coefficient value without significant differences between the data set and the validation dataset, thereby avoiding the current overfitting. Due to limited computing resources, all training images are resized to 512×512 pixels. For future works, training data will be processed in full size to avoid losing information. All of the study's code and data can be found at the following address in Kaggle: https://www.kaggle.com/nhthunhhong/g3-0805-aunet/edit/run/75804871

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgment

References

[1]. Arora R, Basu A, Mianjy P, Mukherjee A. 2016. Understanding deep neural networks with
rectified linear units. Available at https://arxiv.org/abs/1611.01491.

[2]. Jonathan Long, Evan Shelhamer, Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. Computer Vision and Pattern Recognition. Available at: https://arxiv.org/abs/1411.4038

[3]. Senthilkumaran N, Vaithegi S. 2016. Image segmentation by using thresholding techniques for medical images. Computer Science & Engineering: An International Journal 6(1):1–13.

[4]. Wu L, Songde M, Hanqing L. 1998. An effective entropic thresholding for ultrasonic images. In: Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No. 98EX170). Vol. 2. Piscataway: IEEE, 1552–1554.

[5]. Murthy TD, Sadashivappa G. 2014. Brain tumor segmentation using thresholding, morphological operations and extraction of features of tumor. In: 2014 International Conference on Advances in Electronics Computers and Communications. Piscataway: IEEE, 1–6.

[6]. Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. Available at: https://arxiv.org/abs/1802.06955

[7]. Sebastian Ruder. An overview of gradient descent optimization algorithms. Machine Learning. Available at: https://arxiv.org/abs/1609.04747

[8]. Diederik P. Kingma, Jimmy Ba. Adam: A Method for Stochastic Optimization. Machine Learning. Available at: https://arxiv.org/abs/1412.6980

[9]. Mason D. 2011. SU-E-T-33: pydicom: an open source dicom library. Medical Physics 38(6Part10):3493. DOI 10.1118/1.3611983.

[10]. Fausto Milletari, Nassir Navab, Seyed-Ahmad Ahmadi. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. Computer Vision and Pattern Recognition. Available at: https://arxiv.org/abs/1606.04797