Journal of computer science and software applications
https://www.mfacademia.org/
ISSN:2377-0430
Vol. 2, No. 4, 2022

Attention-Unet: A Deep Learning Approach for Fast and Accurate Segmentation in Medical Imaging

Ziyi Zhu[1],Yafeng Yan[2],Ruilin Xu[3],Yun Zi[4],Jinyin Wang[5]

[1]New York University,New York,USA,zz1831@nyu.edu

[2]Stevens Institute of Technology,Hoboken,USA,yanyafeng0105@gmail.com

[3]The University of Chicago,Chicago,USA,harveytsui915@gmail.com

[4]Illinois Institute of Technology,Chicago,USA,yzi9@gatech.edu

[5]Stony Brook University,New York,USA,jinyinsbu@gmail.com

Abstract:

Accurate extraction of the bronchial tubes from lung computed tomography (CT) images is crucial for evaluating respiratory function and diagnosing diseases. Current bronchial segmentation methods often rely heavily on substantial human-computer interaction to improve segmentation accuracy. Although deep learning has been widely applied in medical image processing, especially in lung nodule detection and diagnosis of malignancy, its use in bronchial segmentation in lung CT images faces challenges such as image noise and partial volume effects, which lead to segmentation leakage and difficulties in identifying small bronchi. Additionally, original lung CT images contain non-relevant regions like bones and the patient bed, which increase data processing time and risk errors. By leveraging the anatomical structure of the bronchial tree, we propose a stepwise approach to bronchial segmentation and introduce an Attention-Unet-based method. Experimental results demonstrate that applying the deep learning-based Attention-Unet network to bronchial segmentation in lung CT images enhances segmentation speed and accuracy while effectively preventing leakage.

Keywords:

Convolutional neural network, Image classification, ResNet, ShuffleNet.

1. Introduction

The bronchial tubes serve as crucial biomarkers, yet effective segmentation of bronchial tubes in lung CT images remains unresolved. Developing three-dimensional images of the bronchial tree not only aids in diagnosing bronchial-related diseases but also accurately locates disease sites, assisting doctors in examinations and surgeries to avoid unnecessary damage to other areas. The lower bronchial branches have intact walls and clear boundaries, making them easier to segment. However, as the bronchi branch further, the walls thin out, and the distinction between the bronchi and lung parenchyma blurs, leading to segmentation leakage [1]. The region growing method is a prominent technique in bronchial segmentation, merging pixels or regions into a connected area based on predefined growth criteria [2.3]. Fabijanska [4] introduced a secondary region growing method to mitigate leakage in small bronchi. Ginneken et al. [5] proposed a multi-threshold region growing method, establishing different growth rules for the main bronchus and various bronchi, making the method suitable for bronchial trees in different regions. Pei Hongliang et al. [6] improved the region growing algorithm by using histogram equalization to highlight the boundaries of bronchial walls and extract bronchial tubes. Gong Huayao [7] employed a hysteresis threshold-based region growing method to segment the bronchial tree at an optimal threshold. Besides region growing, rule-based bronchial segmentation methods include morphological techniques. Aykac et al. [8] suggested a morphological grayscale reconstruction algorithm, using a four-neighborhood low-pass filter to remove noise and circular structural elements on 2D axial slices to capture potential bronchial regions, reconstructing the 3D bronchial tree based on regional coherence. However, traditional methods can only segment major bronchi and are less effective in identifying smaller bronchi

In recent years, deep learning-based bronchial segmentation methods, particularly those utilizing convolutional neural networks (CNNs), have become a focal point of research. Numerous studies have implemented deep neural networks for bronchial segmentation in CT images. Jin et al. [9] initially trained a 3D fully convolutional network to extract bronchi and then refined it using a graph-based approach. Zhao et al. [10] trained 3D and 2D neural networks to detect bronchi in horizontal and vertical orientations, respectively, and introduced a linear programming-based tracking method to optimally combine the two detection results. Zhang Lin [11] devised a self-supervised transfer learning-based bronchial segmentation algorithm to address the imbalance between foreground and background voxels in annotated tracheal images. Wang Jiwei et al. [12] proposed a bronchial image segmentation method based on an enhanced adversarial generative network model, enabling the segmentation of morphologically complex 3D bronchial structures without manual intervention. Li Kang [13] utilized deep learning to overcome the varying difficulties in detecting different bronchial bifurcation points. Cheng Liying et al. [14] introduced a lung parenchyma segmentation method for lung CT images based on the U-net neural network. While these methods have significantly advanced bronchial segmentation, the precision in segmenting smaller bronchi still requires enhancement.

This study examines the effectiveness of deep learning algorithms for bronchial segmentation in lung CT images, conducting experiments with the U-net network and its optimized variants, such as Unet++ and Attention-Unet. The experimental results indicate that the Attention-Unet network, based on deep learning, not only improves the speed and accuracy of bronchial segmentation but also effectively prevents leakage, thus improving the accuracy of lung disease diagnosis.

2. Model Introduction

2.1. U-Net Network Model

In 2015, the U-shaped network architecture emerged prominently in the field of image segmentation [15]. The U-net, extensively utilized in segmentation tasks, was developed based on fully convolutional networks (FCN), addressing the issue of FCNs' inability to retain and confirm contextual positional information. The U-net architecture was initially introduced by Hinton in 2006. The encoder-decoder structure holds significant importance, with its fundamental principle involving the encoding of input images through downsampling to reduce the original image into smaller feature maps, followed by decoding to reconstruct the original input image. Consequently, only the relevant features and the decoder need to be stored to preserve an image. The primary innovation of U-net is the integration of low-level feature maps with subsequent high-level feature maps, and the symmetrical structure facilitates more comprehensive feature fusion, leading to minimal information loss. Downsampling at low resolution is employed for object recognition, while upsampling at high resolution is utilized for segmentation and localization, with the fusion process compensating for information loss.

2.2. Network Model Optimized Based on U-net

2.2.1.Unet++ Network Segmentation Algorithm

Unet++[16] enhances the skip connections in U-net to prevent the semantic gap that arises when combining shallow encoder features with deep decoder features, resulting in a highly adaptable feature fusion strategy. As illustrated in Figure 1, Unet++ mitigates issues related to unknown network depth by effectively integrating U-nets of varying depths. The updated Unet++ exhibits several key characteristics. Firstly, U-net can share part of an encoder and engage in joint learning through deep supervision, which enhances segmentation performance while simplifying the model. Secondly, Unet++ overcomes the limitations of skip connections by redesigning them to fuse image features with the corresponding decoder nodes at various scales.

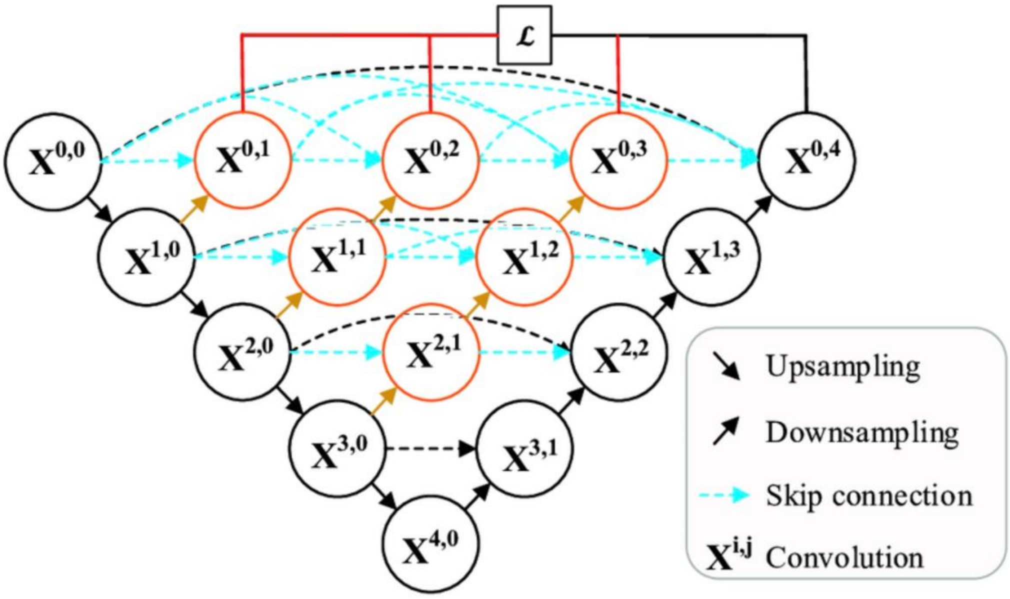

Figure 1. Unet++ architecture

2.3.Attention-Unet Network Segmentation Algorithm

Attention-Unet integrates an attention mechanism within the U-net framework. Prior to concatenating the encoder's features at each resolution with the corresponding decoder features, an attention module is employed to reconfigure the encoder's output features. This attention module produces a gating signal to manage the significance of features at various spatial locations, as illustrated by the AttentionGate in Figure 2. The primary distinction between Attention-Unet and U-net is found in the decoding phase; the encoding stage first applies the AttentionGate before proceeding to the decoder. Unlike multi-stage CNNs, the AttentionGate reduces the focus on irrelevant target regions while enhancing the attention on relevant areas. The AttentionGate demands fewer training models and no extra parameters, effectively suppressing the background noise and minimizing the need for manual region cropping. Additionally, soft attention usage helps to suppress responses in irrelevant regions and decreases redundant skip connections. Integrating the AttentionGate into the U-net architecture accentuates the crucial features through skip connections, addressing the problems of irrelevant and noisy responses in these connections.

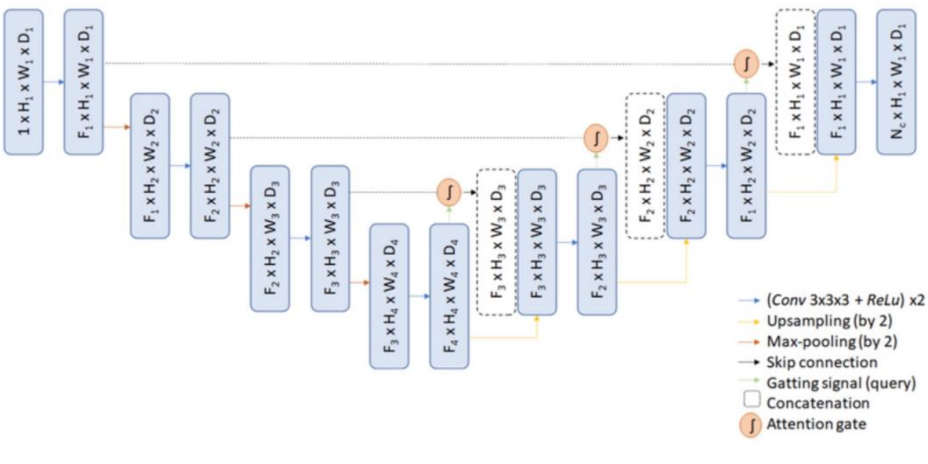

Figure 2. Attention-Unetarchitecture

3. The general segmentation process

By integrating the unique characteristics of the lung bronchial tree, this paper enhances the Attention-Unet network to design a specialized lung bronchial segmentation model for lung CT images. The overall structure is illustrated in Figure 3. Regarding lung CT image preprocessing, an efficient and straightforward scheme is employed. Morphological processing and smoothing operations reduce noise impacts on regions of interest in lung CT images, ensuring that the segmented lung regions retain complete lung texture. This is followed by the preliminary extraction of lung parenchyma tissue contours, which provides an effective scope for subsequent lung bronchial segmentation, minimizing the interference of redundant data. For the segmentation of the lung bronchi, pre trained results are fed into the Attention-Unet. The AttentionGate module within the model suppresses learning irrelevant parts and enhances the focus on task-related features, thereby effectively extracting the tracheal structure.

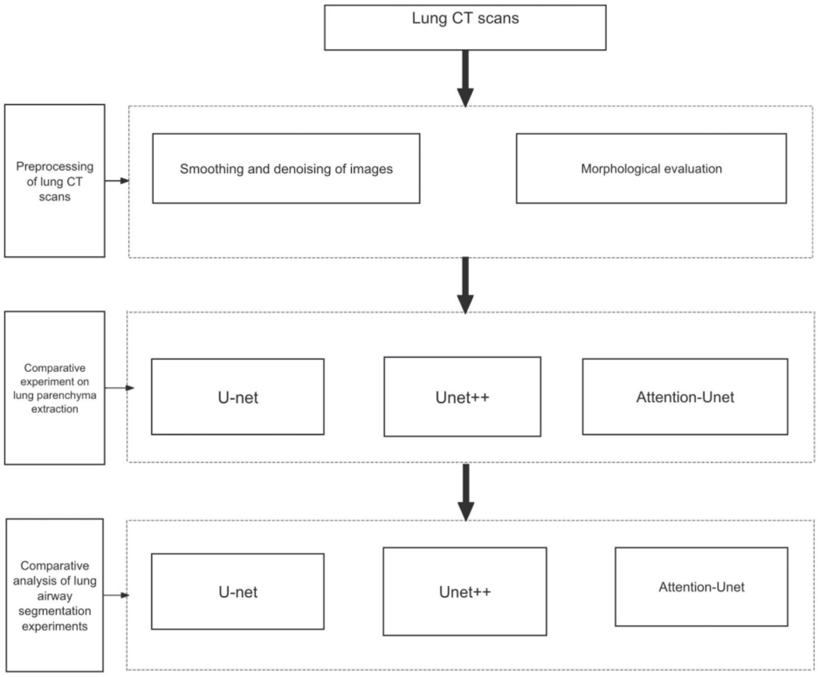

Figure 3. Flowchart of the lung bronchial segmentation process

3.1. Preprocessing of lung CT images

Original lung CT images can be influenced by patient movement or environmental interference during CT scanning, leading to various types of noise in the CT images. To mitigate noise interference, enhance the differentiation between lung parenchyma and surrounding tissues, and ensure the accuracy and reliability of the extracted lung parenchyma data, preprocessing operations are required before segmenting lung CT images.

This paper employs morphological operations and smoothing techniques for image preprocessing. The fundamental concept of morphology is to utilize a specific structuring element to measure or extract relevant shapes or features from the input image, aiding in further image analysis and target recognition. Image dilation, a common morphological processing technique, involves using a custom structuring element to perform a sliding operation analogous to "filtering" on the binary image. The corresponding pixels of the binary image are then compared with those of the structuring element, and the union obtained constitutes the pixels of the dilated image. This process may smooth the edges of the image region, increase the number of pixels in the region, and potentially connect previously unconnected parts, facilitating subsequent processing.

3.2.Extraction of Lung Parenchyma and Segmentation of Lung Bronchi

Given the significant differences in CT values between the lung parenchyma region and the surrounding areas, this paper first extracts the lung parenchyma region to reduce subsequent workload and enhance segmentation accuracy and efficiency (Figure 4).

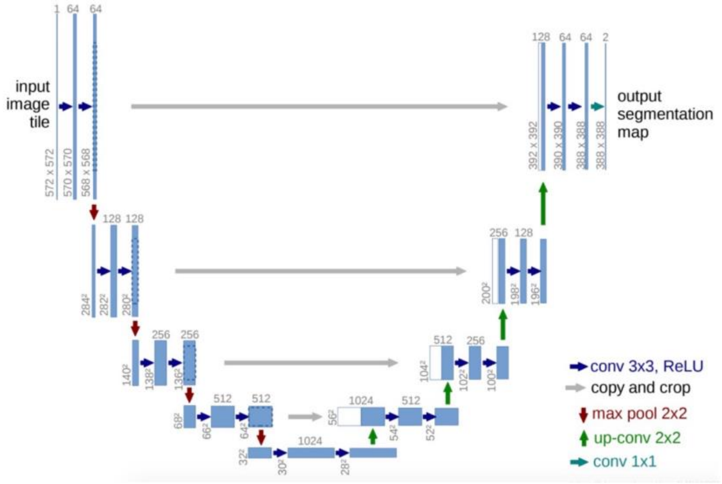

Figure 4. U-net segmentation flowchart

The extracted lung parenchyma data is fed into the Attention-Unet model for lung bronchial segmentation. The internal structure of the AttentionGate within the Attention-Unet is illustrated in Figure 5. In this structure, g represents the matrix from the decoding part, and xl represents the matrix from the encoding (left) part. After being multiplied by a coefficient (to apply Attention), x is concatenated with g and then enters the next decoding layer. Here, the Resampler resamples the feature map back to its original size. The mathematical formula is as follows:

$$q_{att}^{l} = \psi^{T}(\sigma_1(W_x^{T} x_i^{l} + W_g^{T} g_i + b_g)) + b_{\psi} \alpha_i^{l} = \sigma_2(q_{att}^{l}(x_i^{l}, g_i; \Theta_{att}))$$

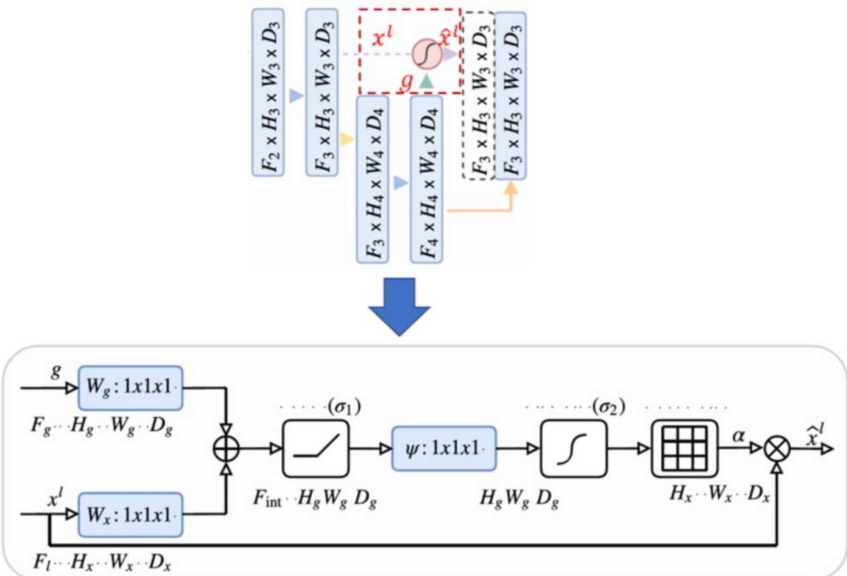

Figure 5. Diagram of the internal structure of the AttentionGate

4. Experimental Results and Analysis

For model training, Adam was employed as the optimizer, using default parameters to train U-net, Unet++, and Attention-Unet over 100 epochs. The neural network algorithms were primarily programmed in Python 3.6 and executed on the PyCharm Community Edition 2020.2.3 x64 platform. The experimental configuration utilized a computer running a 64-bit Windows 10 operating system, with an Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz, 1.80GHz processor.

4.1.Experimental Data and Evaluation Metrics

The experimental data utilized in this paper is sourced from the EXACT09 website, providing lung CT image sequences for competition purposes. These sequences are scanned by various teams worldwide. The CT image sequences are in DICOM format, with individual slice thicknesses ranging from 0.6 to 1.25mm, and the total sequence thickness varying between approximately 260 to 670 slices. Each slice measures 512×512 pixels. The dataset is divided into training and test sets in an 8:2 ratio, with experimental results

Journal of computer science and software applications
https://www.mfacademia.org/
ISSN:2377-0430
Vol. 2, No. 4, 2022

obtained through 5-fold cross-validation.

Three evaluation metrics are selected in this paper: Miou, Dice, and Aver_hd:

Miou (Mean Intersection over Union): This metric calculates the ratio of the intersection to the union of the ground truth and predicted values.

Dice (Dice Similarity Coefficient): This coefficient measures the ratio of twice the area of the intersection of the ground truth and predicted values to the total area.

Aver_hd (Average Hausdorff Distance): This metric is primarily used to assess the segmentation accuracy of boundaries, measuring the maximum mismatch degree between two sets of points.

4.2.Experimental Results of Lung Parenchyma Segmentation

To evaluate the effectiveness of different algorithms for lung parenchyma segmentation, this paper compares the results of U-net, Unet++, and Attention-Unet. Examples of segmentation results are illustrated in Figures 6 to 8, respectively.

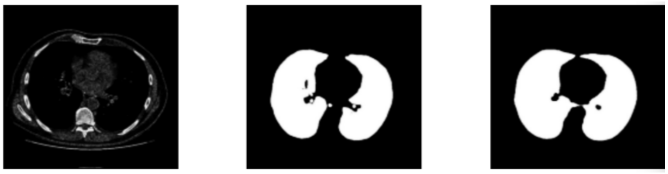

Figure 6. Segmentation results of lung parenchyma using U-net

Figure 7. Segmentation results of lung parenchyma using

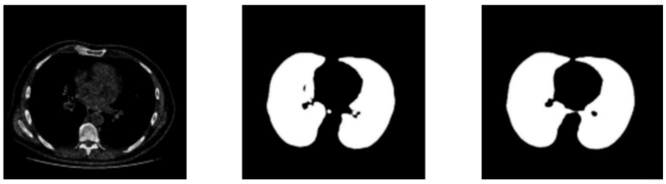

Figure 8. Segmentation results of lung parenchyma using Attention-Unet

From the results in Table 1, both U-net and Attention-Unet can effectively segment the lung parenchyma, with the Attention-Unet network showing superior performance due to the added attention mechanism. However, Unet++ demonstrates less satisfactory segmentation results for the lung parenchyma, attributed to its weaker edge recognition capability in the absence of preprocessing.

Table 1 Evaluation table of pulmonary parenchymal segmentation results

Evaluation Indicators	**U-net**	**Attention-Unet**	**Unet++**
Miou	0.930156	0.934629	0.459280
Dice	0.963423	0.965217	0.350792
Averhd	7.564404	7.310715	14.755167

4.3.Experimental Results of Lung Bronchial Segmentation

The segmentation results of Unet++ are presented in Figure 9, and those of Attention-Unet are shown in Figure 10. The evaluation comparison of lung bronchial segmentation in Table 2 clearly indicates that due to the small size of the bronchial segmentation area, the Unet++ network shows significant under-segmentation, with notable discrepancies between the results and the ground truth data, suggesting the need for further network improvement. However, the proposed Attention-Unet network shows superior performance in lung bronchial segmentation, effectively segmenting even small blood vessels and resolving the issue of bronchial leakage, demonstrating excellent segmentation performance.

Table 2 Evaluation table of pulmonary tracheal segmentation results

Metric	Unet++	Attention-Unet
Miou	0.367869	0.653435
Dice	0.471135	0.721395
Averhd	3.936440	3.453897

Figure 9. Visualization of Unet++ lung airway segmentation results

Figure 10. Visualization of Attention-Unet lung airway segmentation results

5. Conclusion

This paper starts by analyzing image segmentation techniques and proposes a lung bronchial segmentation method for lung CT images based on an improved Attention-Unet network. The segmentation results are compared and analyzed against U-net and its variants. The trained model can automatically segment the original images. Due to structural optimizations and the incorporation of the attention mechanism, the Attention-Unet network improves data adaptability, demonstrating superior performance compared to U-net and Unet++ on the same dataset. This algorithm is simple, effectively extracts complete lung bronchi, resolves the issue of bronchial leakage, and exhibits strong robustness.

References:

[1] Wang Jiwei, Wang Hongxuan, Huang Shaohui, et al. "Lung trachea image segmentation based on improved generative adversarial network model." Digital Medicine, 2021, 16(10): 93-97.

[2] Duan Huihong, Gong Jing, Wang Lijia, et al. "Research progress on tracheal tree segmentation technology in pulmonary CT images." Journal of Biomedical Engineering, 2018, 37(6): 739-748.

[3] Peng Shuang, Xiao Changyan. "CT image lung tracheal tree segmentation combining region growing and gray level reconstruction." Journal of Image and Graphics, 2014, 19(9): 1377-1383.

[4] Fabijanska, A. "Results of applying two-pass region growing algorithm for airway tree segmentation on MDCT chest scans from EXACT database." Proceedings of the Second International Workshop on Pulmonary Image Analysis, USA: CreateSpace, 2009: 251-260.

[5] van Ginneken, B., Baggerman, W., Rikxoort, E. M. "Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans." Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Berlin, Heidelberg: Springer, 2008: 219-226.

[6] Pei Hongliang, Jiang Yunju, Fan Qingwen, et al. "Extraction of the lung tracheal tree based on an improved region growing algorithm." Journal of Integrated Traditional and Western Medicine on Cardiovascular Disease, 2020, 8(21): 27-29.

[7] Gong Huayao. "Study on automatic segmentation algorithm of lung trachea CT images based on three-dimensional region growing method." Chengdu: Sichuan Normal University, 2019.

[8] Aykac, D., Hoffman, E. A., McLennan, G., et al. "Segmentation and analysis of the human airway tree from three-dimensional X-ray CT images." IEEE Transactions on Medical Imaging, 2003, 22(8): 940-950.

[9] Jin, D., Xu, Z., Harrison, A. P., et al. "3D convolutional neural networks with graph refinement for airway segmentation using incomplete data labels." Proceedings of the International Workshop on Machine Learning in Medical Imaging, Cham: Springer, 2017: 141-149.

[10] Zhao, T., Yin, Z., Wang, J., et al. "Bronchus segmentation and classification by neural networks and linear programming." Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Cham: Springer, 2019: 230-239.

[11] Zhang Lin. "Study on lung trachea and blood vessel segmentation algorithm based on self-supervised transfer learning." Hangzhou: Zhejiang University, 2021.

[12] Bian Zijian, Qin Wenjun, Liu Jiren, et al. "A review of anatomical structure segmentation methods in pulmonary CT images." Journal of Image and Graphics, 2018, 23(10): 1450-1471.

[13] Li Kang. "Detection of bifurcation points in CT images of lung tracheal tree based on deep learning." Wuhan: Huazhong University of Science and Technology, 2019.

[14] Cheng Liying, Gao Xuanshuang, Shen Hai, et al. "Pulmonary tissue segmentation based on U-Net network." Journal of Shenyang Normal University (Natural Science Edition), 2020, 38(3): 278-282.

[15] Ronneberger, O., Fischer, P., Brox, T. "U-net: Convolutional networks for biomedical image segmentation." Proceedings of the International Conference on Medical Image Computing and Computer-assisted Intervention, Cham: Springer, 2015: 234-241.

[16] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." IEEE Transactions on Medical Imaging, 2019, 9(6): 1856-1867.