Biomedical Signal Processing and Control 80 (2023) 104407

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

A multi-scale interactive U-Net for pulmonary vessel segmentation method based on transfer learning

Rencheng Wu, Yu Xin *, Jiangbo Qian, Yihong Dong

Ningbo University, 818 Fenghua Road, Ningbo 315211, China

ARTICLE INFO

Keywords:
Pulmonary vessel segmentation
Deep learning
Multi-scale information interaction
Transfer learning

ABSTRACT

Pulmonary vessel segmentation is the key application of AI in lung disease diagnosis and surgical planning. Compared with manual labeling, automatic labeling of pulmonary vessels using an AI-based medical image segmentation method has the advantages of low cost, high accuracy, and efficiency, which is the development trend of medical images. In terms of pulmonary vessel segmentation, FCN and U-Net are the most widely used pulmonary vessel segmentation methods based on deep learning. However, the precision of pulmonary vessel segmentation, especially the small vessels, tends to be poor by such methods. Therefore, to solve the above problem, Multi-Scale Interactive U-Net (MSI-U-Net) is proposed. In MSI-U-Net, three decoder branches are used to extract small-scale, middle-scale and large-scale vessels respectively, which can improve the accuracy of small vessel segmentation by enhancing the representational ability of small vessels. In addition, to solve the problem of small vessel information loss caused by down-sampling, we introduce the attention mechanism into the skip-layer connection and propose a cross-layer aggregation module (CLA). Among the three decoder branches, a multi-scale information interaction strategy (MSIIS) is proposed based on transfer learning, which can effectively enhance the correlation of multi-scale vessels in lung CT images. In the training stage, we propose a scale-induced supervision strategy (SISS). This strategy uses the idea of fusion first and then supervision, which effectively solves the problem of inconsistency in multi-scale vessels classification, thereby reducing the segmentation errors. Finally, we use feature transmission instead of convolution parameter sharing to realize the multi-scale information interaction strategy, and propose an extension scheme called Multi-Level Cascade Interactive U-Net (MLCI-U-Net). The experimental results indicate that our MSI-U-Net and MLCI-U-Net have better performance than other state-of-the-art methods on pulmonary vessel segmentation. Specifically, the best Dice similarity coefficient (DSC), Sensitivity and Precision are obtained by the proposed methods to segment pulmonary vessels.

1. Introduction

With the advancement of computer technology, medical image processing technologies have been greatly developed, which mainly include image enhancement, segmentation and 3D visualization. At present, medical image processing technologies have become an important part of medical research and clinical practice. Compared with traditional methods, medical image processing technologies based on Artificial Intelligence (AI) have the advantages of low cost, high accuracy and efficiency, which is the trend of medical image processing.

Pulmonary vessel segmentation is an important part of medical image processing technologies. It can provide auxiliary diagnostic functions such as surgical planning, postoperative effect prediction in the diagnosis and treatment of pulmonary nodules, pulmonary embolism and lung tumors. In traditional methods, the pulmonary vessel labeling usually needs to be completed manually by professional doctors, which is time-consuming and error-prone. Therefore, it is of great significance for medical diagnosis and treatment to use AI-based image segmentation technology for automatic labeling and 3D reconstruction of pulmonary vessels.

In recent years, with the improvement of the performance of computer hardware equipment, many researchers have proposed a large number of AI-based pulmonary vessel segmentation methods. These methods are mainly divided into machine learning based and deep learning based vessel segmentation methods. Among them, machine learning based vessel segmentation methods mainly include Support Vector Machine [1,2], Clustering algorithm [3,4] and AdaBoost algorithm [5,6]. However, due to the problems of low contrast with surrounding tissues, blurred boundary and high noise in lung CT data, such methods will greatly reduce the accuracy and efficiency of pulmonary vessel segmentation.

* Corresponding author.
E-mail address: xinyu@nbu.edu.cn (Y. Xin).

https://doi.org/10.1016/j.bspc.2022.104407
Received 18 July 2022; Received in revised form 12 October 2022; Accepted 7 November 2022
Available online 18 November 2022

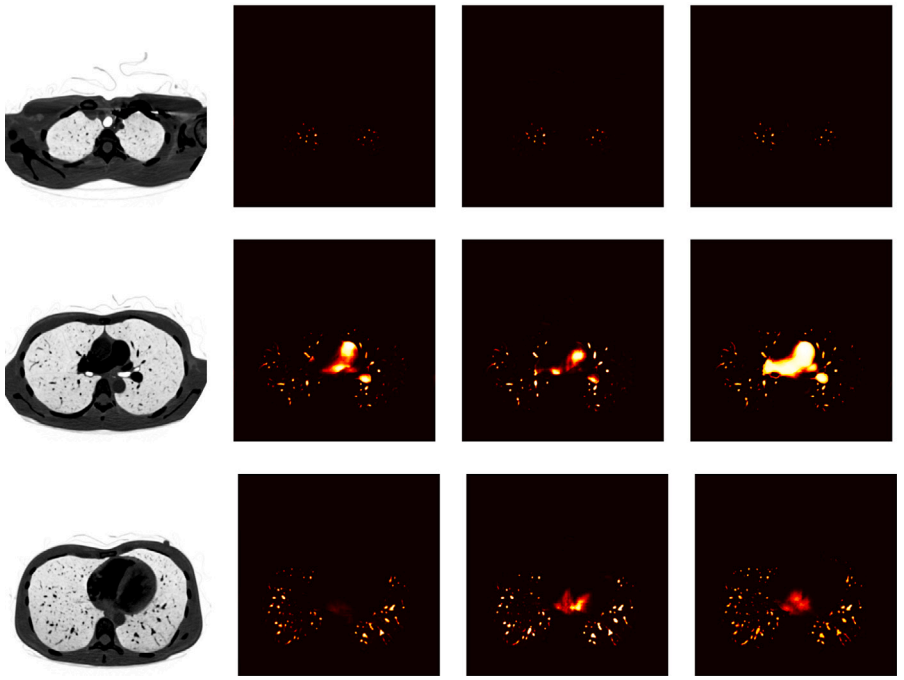


Fig. 1. Comparisons of the ability of different methods to extract small vessel features.

With the development of AI technology, deep learning has replaced traditional machine learning and become the cutting-edge technology of AI. At present, the image segmentation methods based on deep learning have surpassed the traditional segmentation methods in segmentation speed and accuracy, and have been widely used in the field of medical image segmentation. Among them, based on FCN [7], Ronneberger et al. [8] proposed U-Net, which is an encoder–decoder structure as a whole, and realizes multi-scale feature fusion by skip-layer connection. Based on the above characteristics, U-Net is more suitable for pulmonary vessel segmentation than FCN. In 2018, Huang et al. [9] implemented liver vessel segmentation using 3D U-Net, which introduced 3D convolution into the field of vessel segmentation. Compared with 2D CNN, 3D CNN can greatly improve the accuracy of 3D CT data segmentation.

Since 3D U-Net makes full use of the 3D characteristics of CT volume data, the accuracy of vessel segmentation is effectively improved. Therefore, more and more researchers have proposed a large number of vessel segmentation methods based on 3D U-Net. Currently, such methods usually only use a single network to segment multi-scale vessels simultaneously. However, as shown in Fig. 1, due to the serious imbalance in the proportion of small and large vessels, the features of small vessels are easily drowned in the features of large vessels, resulting in low accuracy of small vessel segmentation. In addition, such methods need to conduct down-sampling and pooling in feature extraction process. The drawback of down-sampling and pooling is that images would lose key information.

To this end, based on 3D U-Net, the Multi-Scale Interactive U-Net (MSI-U-Net) as shown in Fig. 3 is proposed, which consists of shared encoder and three separate decoder branches. Among them, shared encoder can supplement information during the training process to improve the segmentation performance and generalization ability of MSI-U-Net. In addition, three decoder branches are used to separately extract the vessel features of small-scale, middle-scale and large-scale, so that small vessel features will not be drowned in the vessel features of the other two scales.

In MSI-U-Net, to recover the information lost by down-sampling, a cross-layer aggregation module (CLA) is proposed in this paper. In CLA, attention mechanism is introduced to filter the encoder features which adapt to the three decoder branches. Moreover, based on the idea of transfer learning, we propose a multi-scale information interaction strategy (MSIIS), which can enhance the correlation between multi-scale vessel features and effectively reduce the amount of parameters and training time of MSI-U-Net. This strategy mainly realizes the information interaction of multi-scale vessels through convolution parameter sharing. The multi-scale model parameter sharing method proposed in this paper mainly studies the feature transfer learning between models of different sizes of the same class of objects. In training stage, we put forward a scale-induced supervision strategy (SISS). This strategy would solve the problem of inconsistency in the classification of vessels at three scales and effectively reduce the segmentation errors.

In summary, there are three main contributions of this work:

(1) We propose a Multi-Scale Interactive U-Net (MSI-U-Net) for pulmonary vessel segmentation, which can improve the accuracy of small vessel segmentation by enhancing the representational ability of small vessels.

(2) In this paper, a cross-layer aggregation module (CLA) is proposed for the skip-layer connection. This module enhances the encoder features adapted to the three decoder branches by introducing attention mechanism, so as to suppress invalid features.

(3) To enhance the information interaction between the three-scale vessels, we propose a multi-scale information interaction strategy (MSIIS) based on transfer learning.

(4) In the training stage, in order to solve the problem of inconsistency in the classification of three-scale vessels, we propose a scale-induced supervision strategy (SISS), which effectively reduces the errors of pulmonary vessel segmentation.

2. Related work

Recently, increasing efforts have been invested to explore how to segment the pulmonary vessels. According to the principle of segmentation methods, the existing vessel segmentation methods can be roughly divided into the following categories: traditional segmentation methods, machine learning based segmentation methods and deep learning based segmentation methods.

2.1. Traditional segmentation methods

In terms of traditional segmentation methods, the main methods include thresholding methods [10,11], region-growing methods [12–15] and graph-cut methods [16–18]. The basic idea of the thresholding methods is to separate the target from the background by the difference between gray values. Lassen et al. [19] used a single threshold to convert the raw CT images into binary images, and then classified each voxel which belongs to foreground or background. However, the gray values of pulmonary vessels have greatly differences, and it is difficult to select a fixed threshold to distinguish vessel voxels from other voxels, which also leads to the low accuracy of this method for vessel segmentation. The region-growing methods need to manually specify a single voxel or a small region as the growth starting point, and then merge sub-voxels or regions according to the predefined growth criteria and continue to grow outward until the complete segmentation target is extracted. Orkisz et al. [20] proposed a variational region-growing algorithm, which uses the eigenvalues of Hessian matrix to constrain the region growing. The region-growing methods are iterative methods, so its space and time costs are large. Furthermore these methods are sensitive to noise, especially when the vessels and other organs are highly similar, the segmentation results are poor. The graph-cut methods [16–18] map an image into a weighted-graph. The vertices represent the voxels, the edges represent the adjacency relationship between adjacent voxels and the weights of the edges represent the non-negative similarity between adjacent voxels in terms of grayscale, color or texture. Such methods realize image segmentation by defining the similarity criterion and optimizing the energy function. However, the graph-cut methods would fall into local optimal solution, and the segmentation performance will be greatly reduced when applied to low contrast images.

2.2. Machine learning based segmentation methods

The image segmentation methods based on machine learning take each pixel as center to obtain a fixed-size patch. Different feature extraction strategies are used to obtain the feature representation from each patch, and the classifiers are used to predict the category of the central pixel according to the feature representation of the patch. The exploration of the combination of feature extraction strategies and classifiers is of great significance for pulmonary vessel segmentation. Zhao et al. [21] extracted features by sparse autoencoders, and then used random forest to segment vessels. Ochs et al. [22] used Hessian features and Adaboost classifier to segment vessels. Goceri et al. [23] used the k-means clustering method to obtain the coarse segmentation results of vessels, and then used the iterative refinement method based on morphological operations to further refine the segmentation results. Zeng et al. [24] proposed a vessel segmentation method based on Extreme Learning Machines (ELM). This method first uses anisotropic filter to remove noise, then combines three classical vessel filters to extract vessel features based on the prior knowledge of shape and geometric structure. Finally, ELM is used to segment vessels. All the above methods need to manually design the feature models of the segmentation target. The designed feature models have many prior parameters in idealized conditions. Practically, these feature models will lead to a serious decline on image segmentation performance when the images exist large differences.

2.3. Deep learning based segmentation methods

Recently, deep learning has made remarkable progress in the field of medical image segmentation. Compared with machine learning based methods, without any handcraft features, deep learning based methods automatically learn the vessel features from the original image. Therefore, more and more researchers have begun to use deep learning techniques to solve the problem of pulmonary vessel segmentation. The medical image segmentation methods based on deep learning achieve pixel-level classification through FCN [7]. Based on FCN [7], many researchers have proposed excellent 2D fully convolutional neural networks such as U-Net [8], SegNet [25], DeepLab [26] and PSPNet [27], which have achieved good results in the field of image segmentation. Among them, U-Net [8] is used in medical image segmentation tasks with small datasets. It is a U-shaped structure as a whole, and adopts skip-layer connection combined with multi-scale features of encoder–decoder, greatly improving segmentation accuracy. Due to the wide application and better performance of U-Net in the field of medical image segmentation, Livne et al. [28] reduced the number of channels in each layer of classical U-Net by half, and proposed a half U-Net to achieve brain vessel segmentation. Meng et al. [29] proposed a Multiscale Dense Convolutional Neural Network (MDCNN) based on encoder–decoder to achieve brain vessel segmentation on Digital Subtraction Angiography (DSA) images. However, since medical images are usually 3D images and there is correlation between slices, it is easy to lose relevant information between slices by using only 2D segmentation network, resulting in low segmentation accuracy. Therefore, more and more researchers began to propose 3D segmentation networks for medical image segmentation, such as 3D U-Net [30], V-Net [31], and VoxResNet [32]. Huang et al. [9] adopted 3D U-Net with data augmentation techniques to successfully achieve liver vessel segmentation. At the same time, in order to solve the serious imbalance of voxels between vessels and other tissues, a weighted Dice loss function is proposed to increase the penalty for misclassified voxels. Sanches et al. [33] proposed a Uception through combining 3D U-Net and Inception for brain vessel segmentation. Gu et al. [34] used a cascaded convolutional neural network. The convolutional neural network at the first level segmented high-intensity structures including vessels and nodules, and the second level was used to distinguish vessels and non-vessel tissues. Xu et al. [35] adopted a stacked convolutional neural network to extract vessels, and then used the region-growing method to solve the discontinuity and false positive problems of vessel segmentation. However, such methods suffer from two major limitations. First, in lung CT images, there is a serious imbalance in the proportion of small and large vessels, which leads to the features of small vessels being easily drowned in the features of large vessels, resulting in low accuracy of small vessel segmentation. Second, during down-sampling, the features of small vessels would be lost, but using the skip-layer connection directly would introduce a lot of noise.

In conclusion, the U-Net-based segmentation methods have a good effect in medical image segmentation. In addition, since CT data are 3D volume data, the correlation between slices needs to be considered in feature extraction process. Therefore, we propose a Multi-Scale Interactive U-Net (MSI-U-Net) based on 3D U-Net [30], which can improve the accuracy of small vessel segmentation.

3. Method

3.1. Network overview

In terms of improving the accuracy of small vessel segmentation, the U-Net-based medical image segmentation methods preserve the small vessel information by the skip-layer connection. However, in lung CT images, the area of small vessels is relatively small, and the vessel features are also few, which make the representational ability of small vessels is insufficient. On the contrary, for large vessels, due to the large area of space, the extracted features of large vessels would be more than those of small vessels, resulting in a serious imbalance in the proportions of small and large vessels in lung CT images. Therefore, if a network segments multi-scale vessels at the same time, the features of small vessels are easily drowned, which would lead to insufficient training of small vessel feature extraction and low accuracy of small vessel segmentation.

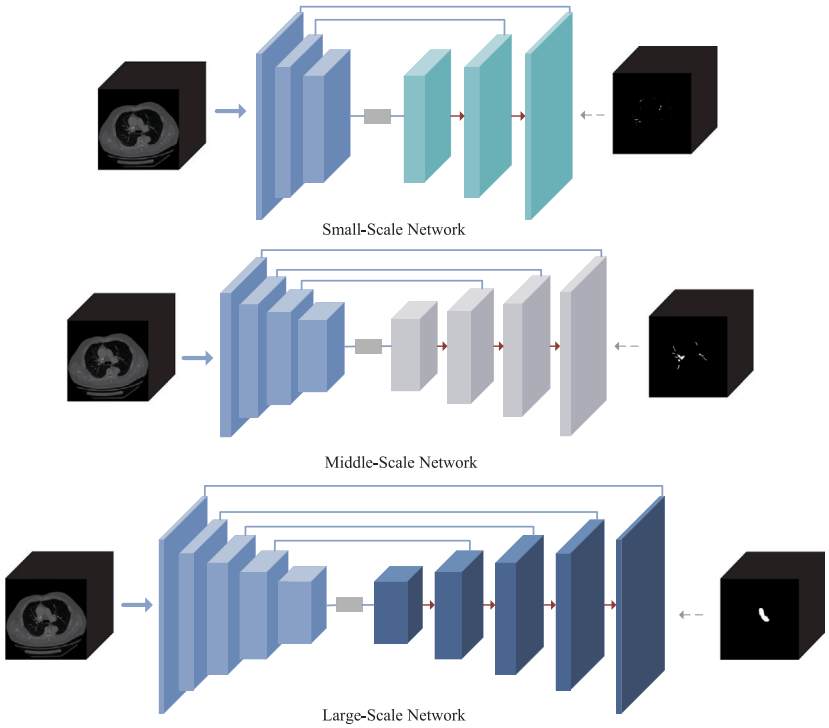


Fig. 2. Overall architecture of the three separate networks. The three networks extract small-scale, middle-scale and large-scale vessels respectively.

To solve the above problems, we design three networks with non-shared parameters according to the characteristics of vessels at different scales. As shown in Fig. 2, the three networks segment the small-scale, middle-scale and large-scale vessels separately. As the number of down-sampling increases, the loss of vessel information becomes more serious, so more small-scale vessel information can be preserved only in the shallow encoder. In this regard, we select a three-layer encoder to extract the features of small-scale vessels. In addition, as the number of encoder layers deepens, the features of small-scale vessels are gradually filtered out, leaving only the vessel features of the other two scales. Therefore, in this paper, we select a four-layer encoder to extract middle-scale vessel features, and a five-layer encoder to extract large-scale vessel features. The above solution can effectively solve the problem of insufficient representational ability of small vessels, and improve the accuracy of pulmonary vessel segmentation.

The three different networks in Fig. 2 need to be trained separately, which greatly increases both training and inference time. In addition, the three-scale vessels are trained using three separate networks. Therefore, due to the lack of information interaction among the three networks in Fig. 2, when multi-scale vessel feature fusion is conducted on the three networks, a large amount of noise would be introduced. From what has been discussed above, this method will reduce the final segmentation accuracy. In order to solve the two problems mentioned above, we propose a Multi-Scale Interactive U-Net (MSI-U-Net) as shown in Fig. 3, which consists of a shared encoder and three separate decoder branches.

When lung CT images are fed into the MSI-U-Net, the shared encoder extracts the pulmonary vessel features of three scales, which are sent into three separate decoder branches respectively. Finally, the MSI-U-Net is trained with the given scale-induced ground-truth maps. Among them, the shared encoder is used to extract the multi-scale features of pulmonary vessels. Since we adopt the convolutional neural network with shared parameters in this part, when training the MSI-U-Net, the encoder can learn the interrelation among the vessels of three different scales, and improve the vessel feature extraction ability of the encoder. Compared with using three separate networks, the shared encoder can supplement information in the training process, and improve the segmentation performance and generalization ability of the MSI-U-Net. In addition, the three decoder branches share a common encoder, thus the encoder features would contain multi-scale vessel features at the same time. Consequently, if the encoder features are directly used as the input of each branch, it would result in redundant features among the three branches. Therefore, in order to reduce the coupling between the three branches, we use the three convolutional layers with non-shared parameters as shown in Fig. 3 to extract features that adapt to different branches. And then these features are used as input for each branch.

In this paper, the shared encoder has five layers, which can be expressed as E^l, where $l \in \{1,2,3,4,5\}$. Since down-sampling will lose vessel detail features, we use shallow features $\left(\mu_1\left(E^3\right)\right)$ which contain more features of small vessels as the input of the first branch (small-scale branch), the fourth layer features $\left(\mu_2\left(E^4\right)\right)$ as the input of the third branch (middle-scale branch), and the fifth layer features $\left(\mu_3\left(E^5\right)\right)$ as the input of the second branch (large-scale branch). μ_1, μ_2, μ_3 are the convolutional layers with non-shared parameters. Since the inputs of the three decoder branches are different, the three decoder branches have different structures. The three decoder branches are three-layer, four-layer and five-layer convolutional neural network structure respectively, which are respectively represented as S_D^{sl}, M_D^{ml}, L_D^{ll}, where $sl \in \{1,2,3\}$, $ml \in \{1,2,3,4\}$, $ll \in \{1,2,3,4,5\}$.

3.2. Cross-layer aggregation module

As shown in Fig. 3, all the three branches take encoder–decoder as basic structure, so there is a problem of key information loss caused by down-sampling, and such lost information is difficult to recover only by decoder. As shown in Fig. 2, in U-Net, the spatial detail information can be preserved by directly overlaying the left and right feature maps with the same resolution using the skip-layer connection. However, this fusion method does not make full use of features, and will introduce a lot of noise, impairing the feature representational ability of the MSI-U-Net.

In response to the above problems, we consider that each decoder branch needs to fuse the encoder features that adapt to it. As shown

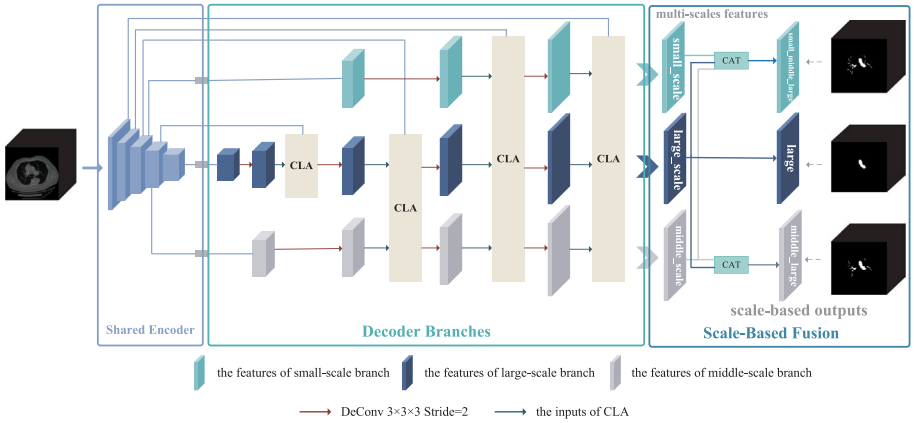
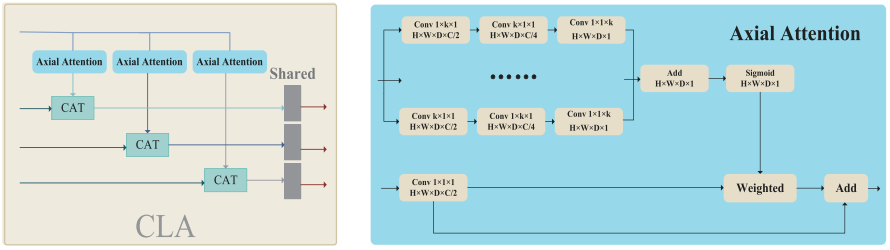


Fig. 3. Overall architecture of the proposed MSI-U-Net. Based on ResUNet, the MSI-U-Net consists of a shared encoder and three decoder branches.



(a) CLA (b) Axial Attention

Fig. 4. Overview of cross-layer aggregation module. In CLA, the Axial Attention module (blue) uses separable convolution instead of classical convolution to implement the attention mechanism, where k is the size of convolution kernel, which is set to 9 in the experiments. In addition, the output features are obtained after three convolutional layers (gray) with shared parameters.

in Fig. 4, we design a cross-layer aggregation module (CLA). In CLA, the attention mechanism is introduced to fuse the encoder features to recover the lost information. In Fig. 4, the encoder features will go through the three attention modules with non-shared parameters. The feature maps after the three attention modules are shown in Fig. 5, which have following advantages. First, the feature maps would filter out a lot of noise and keep the major features of the vessel regions. Second, the generated three feature maps show higher response values in small-scale, middle-scale and large-scale vessel regions respectively. In other words, the generated three feature maps are adapted to the three decoder branches respectively. Finally, the three feature maps enhanced by the attention modules are fused with the features of the corresponding decoder branch to recover the lost spatial detail information.

Due to the different layers of the three decoder branches, the inputs and outputs of CLA module are different. In this paper, the definitions of the operations in CLA are as follows:

when $l = 4$,

$$E\_f_{l\_l} = Cat\left(A_{2\_l}\left(E^l\right), L\_D^l\right) \tag{1}$$

when $l = 3$,

$$\begin{aligned} E\_f_{l\_l} &= Cat\left(A_{2\_l}\left(E^l\right), L\_D^l\right), \\ E\_f_{m\_l} &= Cat\left(A_{3\_l}\left(E^l\right), M\_D^l\right) \end{aligned} \tag{2}$$

when $l = 1$ or $l = 2$,

$$\begin{aligned} E\_f_{l\_l} &= Cat\left(A_{2\_l}\left(E^l\right), L\_D^l\right), \\ E\_f_{m\_l} &= Cat\left(A_{3\_l}\left(E^l\right), M\_D^l\right), \\ E\_f_{s\_l} &= Cat\left(A_{1\_l}\left(E^l\right), S\_D^l\right) \end{aligned} \tag{3}$$

where $E\_f_{l\_l}$, $E\_f_{m\_l}$, $E\_f_{s\_l}$ represent the large-scale, middle-scale and small-scale vessel features after the fusion of the lth layer respectively, E^l represents the encoder features of the lth layer, L_D^l, M_D^l, S_D^l represent the features of large-scale, middle-scale and small-scale branch of the lth layer respectively, A_{1_l}, A_{2_l}, A_{3_l} represent the three attention modules with non-shared parameters of the lth layer, Cat indicates the concatenation.

3.3. Multi-scale information interaction strategy

As shown in Fig. 3, Multi-Scale Interactive U-Net has three separate decoder branches. In general, the three branches require independent training. This strategy ignores the multi-scale vessel feature correlation, and will greatly increase the network parameters and training time. In [36], Oquab M et al. have demonstrated the transfer learning properties of convolutional neural networks. A model trained on the ImageNet for image classification shows good object detection performance with only fine-tuning on the PASCAL. The transfer learning method of neural networks is to transfer the feature expression ability learned from one task or dataset to another task or dataset. The reason for its effectiveness is that the neural networks can learn common features such as edge and texture features between objects, so such features can be shared. The source and target data used in this paper are the same class of objects (pulmonary vessels), but the sizes of the samples they contain are different. In this regard, we propose a multi-scale information interaction strategy. The strategy uses a multi-scale model parameter sharing method to share the convolution weights of convolutional neural network models at different scales, so as to realize the convolution feature sharing under different scale models. This method is essentially a transfer learning method of neural networks [37].

The multi-scale model parameter sharing method proposed in this paper mainly studies the feature transfer learning between models of different sizes of the same class of objects. That is, we share kernel parameters of convolutional neural network at different scales in the same layer. As shown in Fig. 4, there are three convolutional layers on

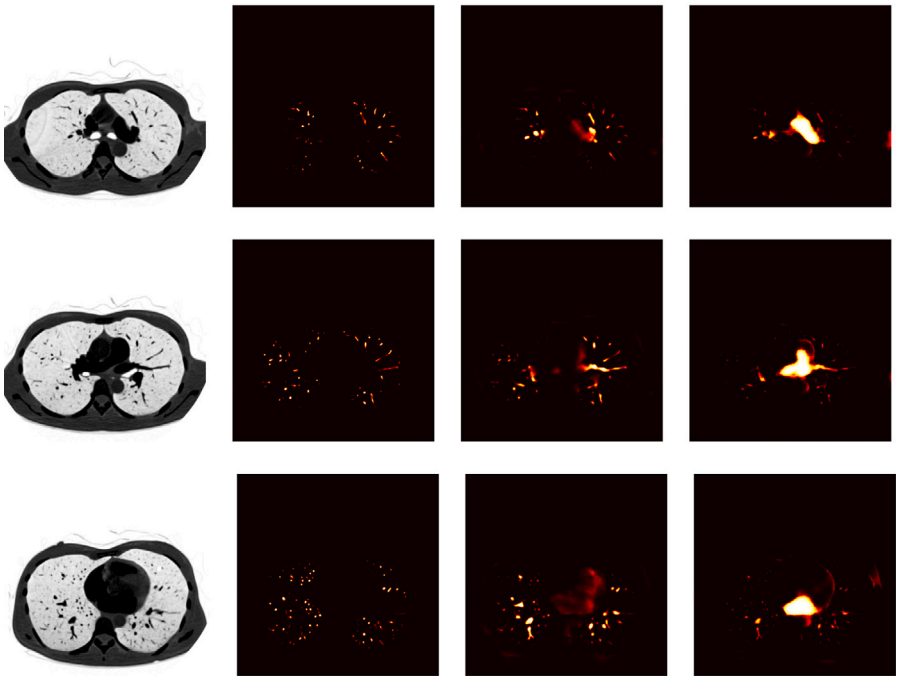


Fig. 5. Visualization results of outputs of three attention modules. (a) The input image. (b) The outputs of small-scale vessel attention module. (c) The outputs of middle-scale vessel attention module. (d) The outputs of large-scale vessel attention module.

the right in the CLA whose parameter settings are exactly the same. Therefore, the multi-scale information interaction strategy designed in this paper can not only reduce the computing cost and improve the segmentation efficiency, but also enhance the multi-scale vessel correlation in lung CT images, so as to improve the accuracy of pulmonary vessel segmentation, especially small vessels.

3.4. Scale-induced supervision

In Fig. 3, $\mu_i\left(E^j\right)(i \in \{1,2,3\}, j \in \{3,4,5\})$ pass through three separate decoder branches, and the multi-scale vessel feature maps can be obtained. The obtained features maps have the following characteristics. First, three decoder branches can respectively extract small-scale, middle-scale and large-scale vessel features in lung CT images. Second, since the same CT sequence is divided into three scales and each scale is separately encoded, there will be inconsistency phenomenon in vessel classification at the three scales. Aiming at the feature fusion of the three decoder branches, we design a multi-scale feature fusion module as shown in Fig. 3, and use the scale-induced supervision strategy to train the MSI-U-Net.

In Fig. 3, the extracted vessel features by the three decoder branches can be represented as F_S, F_M and F_L respectively. Due to the inconsistency in the classification of small-scale, middle-scale and large-scale vessels, it is easy to increase the segmentation errors if the segmentation results of the three-scale vessels are directly supervised by the corresponding ground-truth maps at this time. Therefore, the small-scale, middle-scale and large-scale vessel features are first fused ($Cat(F_S, F_M, F_L)$) and then the middle-scale and large-scale vessel features are fused ($Cat(F_M, F_L)$) to obtain two new feature maps (F_{S+M+L} and F_{M+L}). After that, the corresponding scale segmentation results (S_{S+M+L}, S_{M+L} and S_L) as shown in Fig. 6 are generated using $1 \times 1 \times 1$ convolutional layer for F_{S+M+L}, F_{M+L} and F_L respectively. Finally, the corresponding scale-induced ground-truth maps are used for supervised training.

The above supervision strategy adopts the idea of fusion first and then supervision. Since the fused feature maps contain multi-scale vessel features, compared with using single-scale vessel features, using the fused multi-scale vessel features can effectively solve the problem of inconsistency in the classification of three-scale vessels. And this strategy can reduce the segmentation errors, thereby improving the accuracy of multi-scale vessel segmentation.

In this paper, the Dice loss function is used to train the MSI-U-Net, and the total loss is calculated as follows:

$$\begin{aligned} L &= L_1 + \alpha \times (L_2 + L_3) \\ L_1 &= L_{Dice}(S_{S+M+L}, G_{S+M+L}) \\ L_2 &= L_{Dice}(S_{M+L}, G_{M+L}) \\ L_3 &= L_{Dice}(S_L, G_L) \end{aligned} \tag{4}$$

$$\begin{aligned} G_{S+M+L} &= G_S + G_M + G_L \\ G_{M+L} &= G_M + G_L \end{aligned} \tag{5}$$

$$L_{Dice} = 1 - \frac{2\sum_{i=1}^{n} S_i G_i}{\sum_{i=1}^{n} S_i^2 + \sum_{i=1}^{n} G_i^2} \tag{6}$$

where G_S, G_M and G_L represent the labels of small-scale, middle-scale and large-scale vessels respectively, N represents the number of voxels in lung CT images.

4. Model extension

In this paper, MSI-U-Net is proposed to solve the problem of serious imbalance in the proportion of small and large vessels in lung CT images. In addition, in MSI-U-Net, the convolution parameter sharing method is used to enhance the relationship between multi-scale vessels, emphasizing the feature consistency of multi-scale vessels, and effectively improving the accuracy of pulmonary vessel segmentation. In this paper, a new extension scheme for multi-scale information interaction strategy is proposed, that is, multi-scale vessel information interaction is realized by Multi-Level Cascade Interactive U-Net (MLCI-U-Net). This scheme emphasizes the spatial context of multi-scale vessels with cascade structure.

As shown in Fig. 7, we adopt a scale-induced supervision strategy for supervised training on the three branches. In other words, the

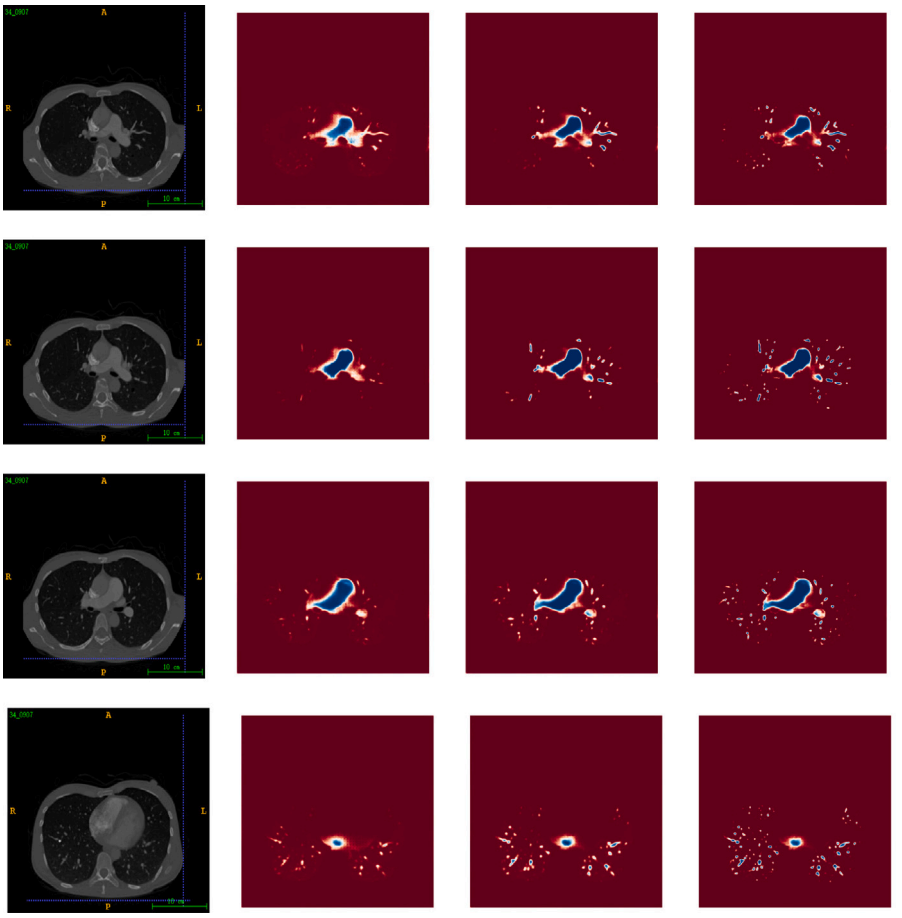


Fig. 6. Examples of scale masks generated by our model. (a) The input image. (b) The scale mask captures large-scale vessels. (c) The scale mask concentrates on middle and large scale vessels. (d) The scale mask catches the small, middle and large scale vessels.

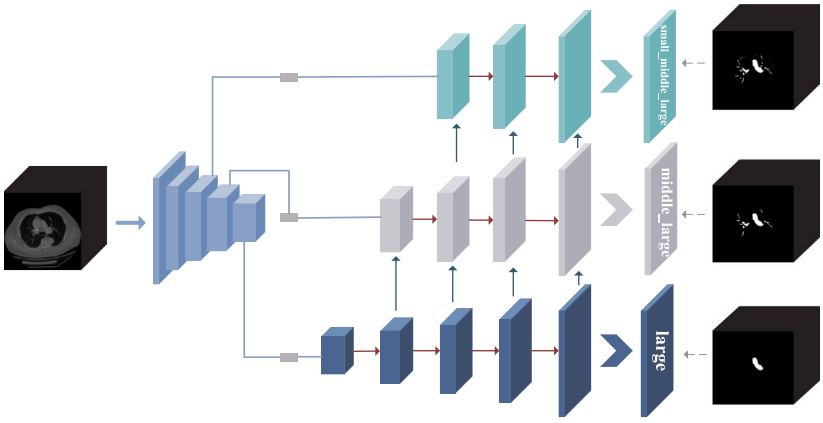


Fig. 7. Overall architecture of the proposed MLCI-U-Net. The MLCI-U-Net uses feature fusion instead of convolution parameter sharing to achieve information interaction between multi-scale vessels.

first branch contains small-scale, middle-scale and large-scale vessel features, the second branch contains middle-scale and large-scale vessel features, and the third branch contains only large-scale vessel features. Therefore, the target regions of the three branches exist in a contained relationship. In addition, as large vessels account for a relatively large area in lung CT images, while the area of small vessels is relatively small, it is hard to segment small vessels among the three scales. Therefore, the segmentation difficulty of the three branches is increasing from bottom to top.

To solve the above problems, we propose a Multi-Level Cascade Interactive U-Net (MLCI-U-Net) for multi-scale pulmonary vessel segmentation. In MLCI-U-Net, we cascade three convolutional neural networks in multi-layer to perform the multi-scale pulmonary vessel segmentation task at one time. As shown in Fig. 7, the cascade structure is composed of three decoder branches, so that the former branch can provide multi-scale supplementary features for the latter branch, as the spatial constraint information of the subsequent branches. The three branches generate segmentation results of three scales respectively

(from bottom to top are large-scale, large-scale + middle-scale, large-scale + middle-scale + small-scale). In the three decoder branches, the internal structure of the same layer is consistent. Among them, the third and the second decoder branches perform feature fusion by overlaying the features of the same layer. Since the third decoder branch contains the location information of large-scale vessels, it can provide the spatial information of large-scale vessels for the second branch. Similarly, the second decoder branch is connected in the same way as the first decoder branch, passing the spatial information of large-scale and middle-scale vessels to the first decoder branch. In the three decoder branches, since the large-scale vessels are easier to segment than the small-scale and middle-scale vessels, the third decoder branch can also obtain location information of large-scale vessels well without additional constraints.

In the above extension scheme, we realize a multi-scale information interaction strategy through feature transmission. Due to the multi-layer cascade feature transmission method, the former branch can transfer multi-scale prior knowledge to the latter branch, which can softly constrain the focus range of the latter branch. In addition, in the process of feature transmission, this strategy also adds auxiliary features for small-scale and middle-scale vessel segmentation, which alleviates the problem of sample imbalance to a certain extent and can effectively improve the accuracy of small-scale vessel segmentation.

5. Experiments

5.1. Datasets and pre-processing

In the experiments, we use the datasets that all from a hospital in Zhejiang Province. The datasets contain 143 cases, and each case is a real clinical CT sequence. In summary, the number of slices in each case is in the range of 135 to 513, and the 143 cases contain a total of 29,056 CT images with a size of 512*512. All the data used in this paper are manually-labeled one by one under the guidance of medical experts.

The original CT data obtained from the hospital are all in dicom format, whose voxel values are represented by the Hounsfield Unit (HU) values from -1000 to more than $+3000$. In order to make the lung and vessel region more obvious, the voxel values are truncated in the range of $[-950, +250]$, and then the voxel values are linearized and normalized to the range of $[0, 1]$. In addition, we use the linear gray level transformation to enhance the contrast between vessels and other background regions. The CT images after the linear gray level transformation can enhance the vessels, while the other background regions can be better suppressed. This method lays a good foundation for pulmonary vessel segmentation. In this paper, the experimental data is pre-processed, where 70% is used as the training set, 30% as the validation set.

5.2. Implementation details

In this paper, PyTorch is used to implement Multi-Scale Interactive U-Net. All experiments of our proposed method are performed on a workstation with Intel Core i7-10700KF CPU @ 3.80 GHz, 32G RAM, NVIDIA GeForce RTX 3090 GPU. In the training stage, the Adam optimizer with the initial learning rate of 0.0001 is employed to perform the gradient descent algorithm. And then the learning rate is divided by 10 every 40 epochs. If the loss of validation set does not decrease for 30 consecutive epochs, or the maximum number of training epochs has been reached, the training is stopped. The model with the smallest validation set loss during training is preserved and evaluated on the testing set. In terms of data augmentation, a random cropping strategy is used to limit the size of the network inputs, where the number of randomly cropped slices is 16. In the testing stage, we use a sliding window strategy to obtain the final segmentation results, and the inference time for each CT sequence is about 8 to 43 s.

Table 1
The settings of supervision for decoder branches.

Supervision	Sen	Pre	DSC
Full	0.6716	0.7202	0.6767
Scale-induced without fusion	0.6787	0.7563	0.6941
Scale-induced with fusion	**0.7234**	**0.7893**	**0.7168**

5.3. Evaluation metrics

In order to compare the performance of different methods, we use Dice similarity coefficient (DSC), Sensitivity (Sen) and Precision (Pre) as evaluation metrics in experiments. Among them, DSC is an indicator to measure the degree of overlap between segmentation results and labels, where 1 means complete overlap and 0 means no overlap at all. Sensitivity is used to evaluate the ratio of correctly segmented vessels to labeled vessels, that is, the ratio of correctly predicted samples to the total positive samples. Precision is used to evaluate the ratio of correctly segmented vessels to the segmented vessels, that is, the ratio of correctly predicted samples to predicted positive samples. The definitions of DSC, Sen and Pre are as follows:

$$\begin{aligned} DSC &= \frac{2TP}{2TP + FP + FN} \\ Sen &= \frac{TP}{TP + FN} \\ Pre &= \frac{TP}{TP + FP} \end{aligned} \tag{7}$$

where TP, TN, FP and FN represent the number of true positives, true negatives, false positives and false negatives respectively.

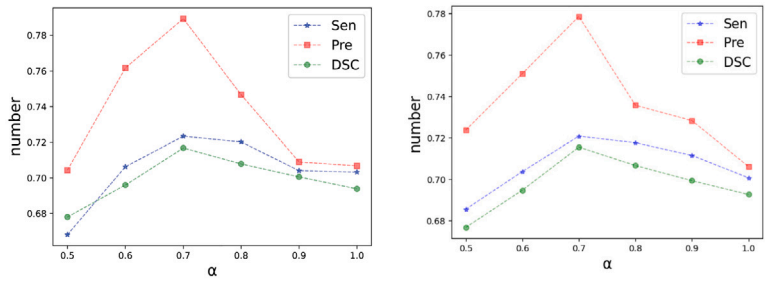
5.4. Comparison with other supervision strategies

In order to validate the effectiveness of the scale-induced supervision strategy, we conduct comparative experiments with different settings on the training targets of the three decoder branches. In Table 1, full supervision means that complete labels including small-scale, middle-scale and large-scale vessels are used for supervision training of the three branches. The scale-induced without fusion refers to the use of small-scale, middle-scale and large-scale vessel labels to separately train three branches, namely, as shown in Fig. 3, the first branch only extracts small-scale vessels, the second branch only extracts large-scale vessels and the third branch only extracts middle-scale vessels. The scale-induced with fusion represents the training of Multi-Scale Interactive U-Net using the scale-induced supervision proposed in this paper.

As shown in Fig. 3, the inputs of the three decoder branches are the encoder features of the third layer, the fourth layer and the fifth layer respectively, that is, the inputs of the three decoder branches already contain the prior knowledge of multi-scale vessels. Therefore, as shown in Table 1, if the three decoder branches are directly trained with complete vessel labels at this time, the training difficulty of the MSI-U-Net will be increased. This strategy would greatly reduce the accuracy of pulmonary vessel segmentation. In addition, due to the problem of inconsistency in the classification discussed in Section 3.4, if the segmentation results of the three-scale vessels are directly supervised by the corresponding ground-truth maps at this time, the segmentation errors tend to be increased.

As shown in Table 1, the idea of fusion first and then supervision, namely, the scale-induced supervision strategy proposed in this paper is used to train MSI-U-Net, which can effectively solve the problem of the inconsistency in the classification of three-scale vessels. This strategy can reduce the segmentation errors, so as to improve the accuracy of pulmonary vessel segmentation.

Since the vessel regions of the three feature maps (F_{S+M+L}, F_{M+L} and F_L) generated by MSI-U-Net and MLCI-U-Net all have the contained relationship, we use the hyperparameter α to control the weight



(a) MSI-U-Net (b) MLCI-U-Net

Fig. 8. Evaluation of the proposed network with different α.

of S_{M+L} and S_L generated by F_{M+L} and F_L in the total loss. In this section, we conduct comparative experiments on the choice of α. As shown in Fig. 8, since the segmentation difficulties of the three segmentation results (S_{S+M+L}, S_{M+L} and S_L) are successively decreasing, the optimal S_{M+L} and S_L can be obtained only by selecting a smaller α. However, as shown in Fig. 8, if the selected α is too small, the model training will also be insufficient. In conclusion, as shown in Fig. 8, when α is set to 0.7 in the experiment, both MSI-U-Net and MLCI-U-Net can obtain the optimal pulmonary vessel segmentation results.

5.5. Ablation studies

In order to validate the effectiveness of each module in the MSI-U-Net, we conduct ablation studies based on ResUNet architecture, and the results are shown in Table 2.

Decoder Branches: Based on ResUNet, we verify the effectiveness of the three decoder branches. ResUNet segments small-scale, middle-scale and large-scale vessels in the same neural network. However, at this time, the features of small vessels are easily drowned in the features of large vessels. As shown in Fig. 3, the MSI-U-Net extracts vessels of three scales through three decoder branches, which effectively improves the representational ability of small vessels. As shown in Table 2, in terms of segmentation performance, the decoder branches proposed in this paper have a significant improvement compared to ResUNet.

Cross-Layer Aggregation Module: ResUNet uses the skip-layer connections to preserve the information lost by down-sampling. However, there are three decoder branches in MSI-U-Net and the encoder features contain multi-scale vessel features at the same time. Therefore, if encoder features are directly fused with decoder branches features, a lot of noise would be introduced, which greatly reduces the segmentation accuracy. In this regard, we introduce an attention mechanism (Att) into the skip-layer connections, and filter out the encoder features which adapt to each decoder branch for aggregation, so as to suppress invalid features. As shown in Table 2, the cross-layer aggregation module proposed in this paper can effectively reduce the interference of noise and improve the segmentation accuracy to a certain extent.

Multi-Scale Information Interaction Strategy: In the three decoder branches of MSI-U-Net, we conduct ablation studies on whether to use the convolution kernel with parameter sharing. As shown in Table 2, the experimental results show that using three completely separate decoder branches would lead to the neglect of the correlation among multi-scale vessel features in the training process. On the contrary, the introduction of multi-scale information interaction strategy between decoder branches can effectively enhance the correlation between small-scale, middle-scale and large-scale vessels in lung CT images.

Table 2
Performance of the network with different modules.

Settings	Sen	Pre	DSC
Baseline	0.6977	0.7574	0.6928
Baseline + Branches	0.7005	0.7632	0.7002
Baseline + Branches + CLA without Att	0.7078	0.7689	0.7056
Baseline + Branches + CLA with Att	0.7164	0.7723	0.7078
Baseline + Branches + CLA with Att + Interactive	**0.7234**	**0.7893**	**0.7168**

Table 3
Comparative evaluation with some classic methods.

Methods	Sen	Pre	DSC	Inference time (s)
3D U-Net [30]	0.6868	0.7525	0.6852	3~26
3D Attention U-Net [38]	0.7127	0.7388	0.7075	4~28
3D ResUNet	0.6977	0.7574	0.6928	3~26
V-Net [31]	0.6799	0.7334	0.6701	3~26
3D SegNet [25]	0.5571	0.6385	0.5989	3~25
3D DeeplabV3+ [26]	0.7068	0.7632	0.7057	3~27
KiU-Net [39]	0.7121	0.7730	0.7159	5~29
3D UNet 3+ [40]	0.7102	0.7677	0.7134	8~59
Ours(MLCI-U-Net)	0.7207	0.7785	0.7154	7~50
Ours(MSI-U-Net)	**0.7234**	**0.7893**	**0.7168**	7~48

Table 4
Quantitative comparison results on the small vessels.

Methods	Sen	Pre	DSC
3D U-Net [30]	0.5086	0.6214	0.5910
3D Attention U-Net [38]	0.5426	0.6913	0.6114
3D ResUNet	0.5360	0.6686	0.5997
V-Net [31]	0.5295	0.6450	0.5974
3D SegNet [25]	0.4982	0.6164	0.5530
3D DeeplabV3+ [26]	0.5391	0.6834	0.6012
KiU-Net [39]	0.5455	0.6893	0.6215
3D UNet 3+ [40]	0.5436	0.6840	0.6190
Ours (MSI-U-Net)	**0.5568**	0.7056	**0.6224**
Ours (MLCI-U-Net)	0.5559	**0.7116**	0.6217

5.6. Comparison with other segmentation methods

In this subsection, we compare our method with other state-of-the-art methods. It turns out that the results of our model are consistently better than those of baselines. The experimental results are summarized in Table 3. In addition, in order to show the details of vessel segmentation, we intercept some segmentation results for visualization, as shown in Fig. 9.

In Fig. 9, from left to right are the original images, the corresponding labeling results, the segmentation results of ResUNet and the segmentation results of MSI-U-Net. It can also be seen from the visualization results that the MSI-U-Net proposed in this paper shows good segmentation performance in the pulmonary vessel segmentation task. Especially in most of the small vessel regions in the figures, it has a significant improvement compared with other segmentation methods.

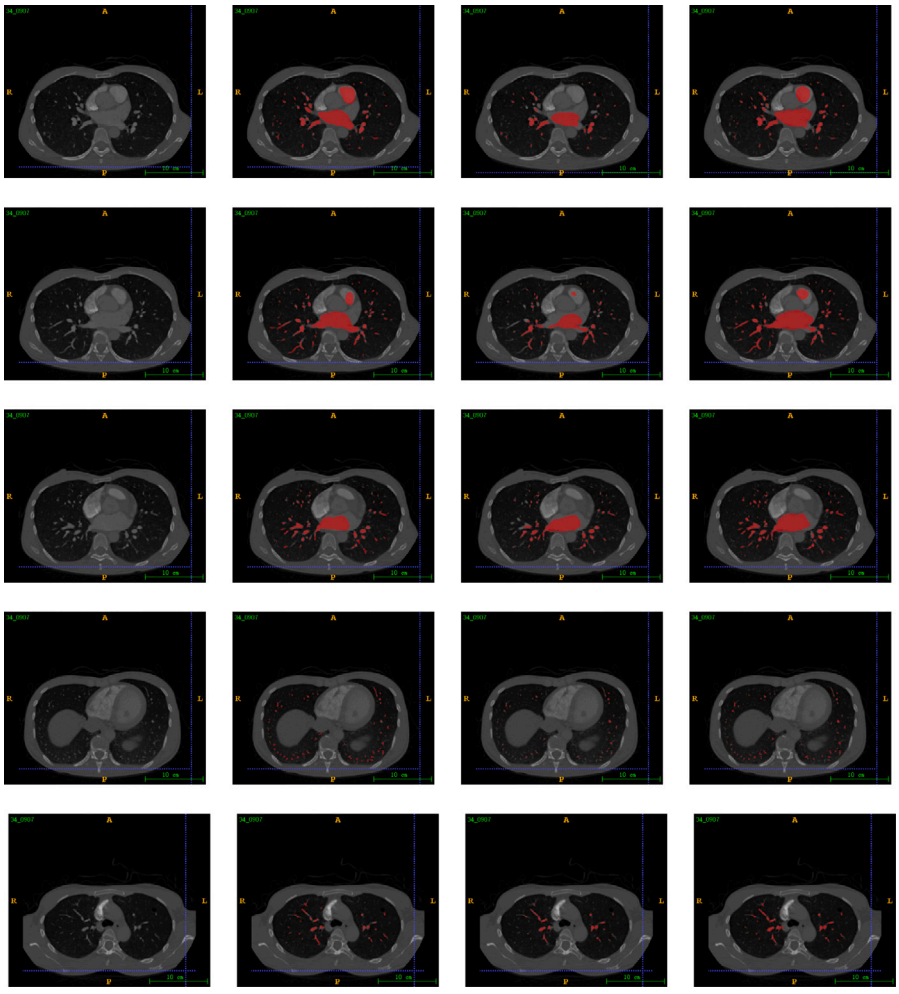


Fig. 9. Examples of performances of the MSI-U-Net and other methods on slices.

In addition, as shown in Fig. 10, we also show some 3D visualization results. In Fig. 10, from left to right are the labeled results, the segmentation results of ResUNet and the segmentation results of MSI-U-Net.

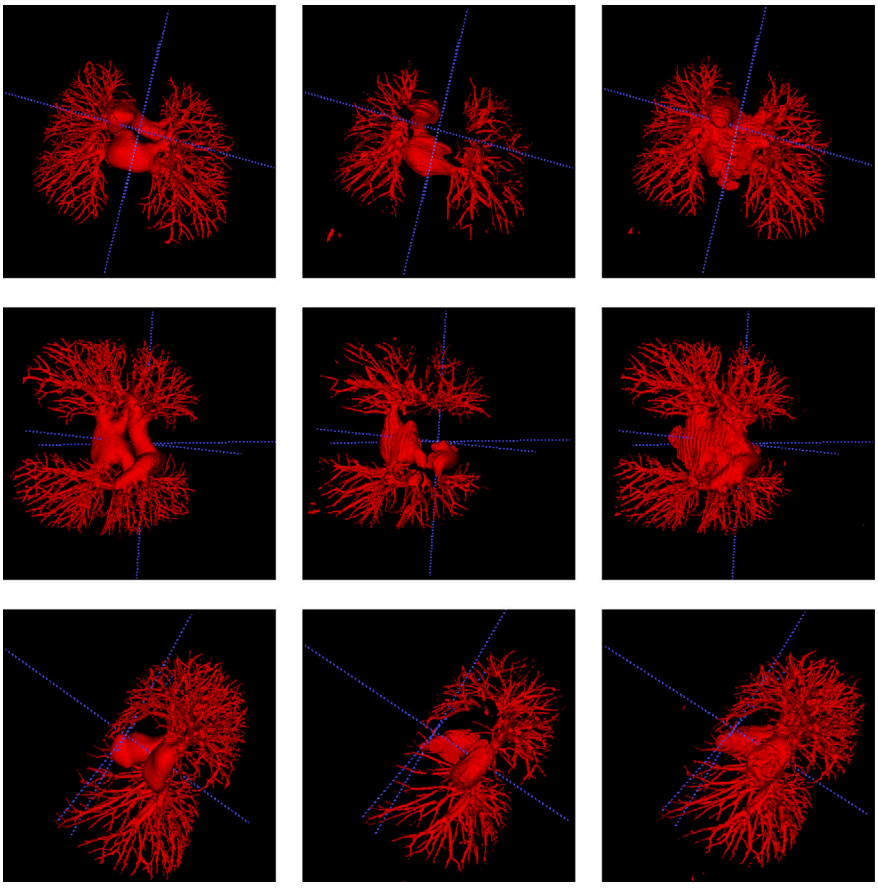
It can be seen from the 3D results that there are many misclassified voxels and some messy fragments in 3D visualization results of ResUNet, resulting in vessel structure disorder, while the vessels extracted by MSI-U-Net are closer to the real vessels. Compared with other methods, it can obtain more complete pulmonary vessel structure.

Small Vessels: Although the main target of pulmonary vessel segmentation is to extract all vessels in the lung, most existing methods can only extract large-scale vessels. However, in clinical practice, doctors usually need to design surgical plans based on the tendencies of small vessels. Therefore, we also validate the ability of MSI-U-Net to segment small vessels. As shown in Table 4, since the proposed MSI-U-Net can enhance the representational ability of small vessels and effectively reduce the information loss of small vessels, the proposed method is superior to other methods on the metrics. In addition, the extension scheme MLCI-U-Net proposed in this paper uses feature fusion instead of convolution parameter sharing to achieve a multi-scale information interaction strategy. In MLCI-U-Net, adding auxiliary features for small vessel segmentation can also alleviate the problem of sample imbalance and effectively improve the segmentation effect. Therefore, in the two schemes proposed in this paper, if it is necessary to extract small vessels for separate analysis and diagnosis, it is better to use MLCI-U-Net. MSI-U-Net has a better effect if multi-scale pulmonary vessels need to be extracted simultaneously for auxiliary diagnosis of pulmonary embolism and lung nodule.

6. Conclusion

In this paper, a Multi-Scale Interactive U-Net (MSI-U-Net) is proposed. In MSI-U-Net, there are three decoder branches to extract small-scale, middle-scale and large-scale vessels respectively, which greatly alleviates the problem of insufficient representational ability of small vessels. In addition, in order to solve the problem of information loss caused by down-sampling, we use a cross-layer aggregation module to select the encoder features which adapt to each decoder branch for aggregation, so as to suppress invalid features. On this basis, we propose a multi-scale information interaction strategy, which implements feature transfer among multi-scale vessels in the way of convolution kernel parameters sharing. This strategy enhances the correlation between small-scale, middle-scale and large-scale vessels in lung CT images.

In this paper, we use a pulmonary vessel dataset finely labeled by medical experts for experiment. In the experiments, we use the scale-induced supervision strategy to train MSI-U-Net, which effectively alleviates the problem of inconsistency in multi-scale vessel classification and reduces the segmentation errors. In addition, we propose an extension scheme named Multi-Level Cascade Interactive U-Net (MLCI-U-Net) for multi-scale information interaction strategy, which effectively improves the precision of small vessel segmentation. Among them, the evaluation scores of DSC, Sensitivity and Precision



(a) Ground Truth (b) ResUNet (c) Ours

Fig. 10. 3D visualization examples of different methods.

are the maximum values, which are 0.7168, 0.7234 and 0.7893 respectively. Besides, in terms of small vessel segmentation, the precision of MLCI-U-Net reaches 0.7116.

The MSI-U-Net has great advantages in the segmentation of small vessels, but the segmentation efficiency of MSI-U-Net is lower than that of classical medical image segmentation methods. Therefore, the MSI-U-Net still has a lot of room for improvement in achieving lightweight. In terms of model design, the MSI-U-Net has modeled the pulmonary vessel segmentation. In the following research, the MSI-U-Net will be further optimized and adjusted to realize the application of lung CT auxiliary diagnoses such as lung tumor and lung nodule segmentation.

CRediT authorship contribution statement

Rencheng Wu: Conceptualization, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Yu Xin:** Conceptualization, Resources, Supervision, Funding acquisition. **Jiangbo Qian:** Validation, Formal analysis. **Yihong Dong:** Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

We acknowledge the support of the Natural Science Foundation of Zhejiang Province, China (Grant No. LY22F020001), the 3315 Plan Foundation of Ningbo (Grant No. 2019B-18-G).

References

[1] I. Lizarazo, SVM-based segmentation and classification of remotely sensed data, Int. J. Remote Sens. 29 (23–24) (2008) 7277–7283.
[2] S. Park, H.S. Lee, J. Kim, Seed growing for interactive image segmentation using svm classification with geodesic distance, Electron. Lett. 53 (1) (2016) 22–24.
[3] A. Alush, J. Goldberger, Hierarchical image segmentation using correlation clustering, IEEE Trans. Neural Netw. Learn. Syst. 27 (6) (2016) 1358.
[4] S.M. Aqil Burney, H. Tariq, K-means cluster analysis for image segmentation, Int. J. Comput. Appl. 96 (4) (2014) 1–8.
[5] B. He, C. Huang, G. Sharp, S. Zhou, Q. Hu, C. Fang, Y. Fan, F. Jia, Fast automatic 3D liver segmentation based on a three-level AdaBoost-guided active shape model, Med. Phys. (2016).
[6] C.A. Lupascu, D. Tegolo, E. Trucco, FABC: retinal vessel segmentation using AdaBoost., IEEE Trans. Inf. Technol. Biomed. Publ. IEEE Eng. Med. Biol. Soc. 14 (5) (2010) 1267–1274.
[7] Long, Jonathan, Shelhamer, Evan, Darrell, Trevor, Fully convolutional networks for semantic segmentation, IEEE Trans. Pattern Anal. Mach. Intell. (2017).
[8] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.
[9] Q. Huang, J. Sun, D. Hui, X. Wang, G. Wang, Robust liver vessel extraction using 3D U-net with variant dice loss function, Comput. Biol. Med. 101 (2018) S0010482518302385.
[10] M. Borlu, Accurate Segmentation of Dermoscopic Images by Image Thresholding Based on Type-2 Fuzzy Logic, IEEE Press, 2009.
[11] V. Osuna-Enciso, E. Cuevas, H. Sossa, A comparison of nature inspired algorithms for multi-threshold image segmentation, Expert Syst. Appl. 40 (4) (2013) 1213–1219.

[12] S. Li, R. Li, Development of preoperative liver and vascular system segmentation and modeling tool for image-guided surgery and surgical planning, Proc. SPIE - Int. Soc. Opt. Eng. 6918 (2008).
[13] A. Sboarina, R.I. Foroni, A. Minicozzi, L. Antiga, F. Lupidi, M. Longhi, M. Ganau, A. Nicolato, G.K. Ricciardi, A. Fenzi, Software for hepatic vessel classification: feasibility study for virtual surgery, Int. J. Comput. Assist. Radiol. Surg. 5 (1) (2010) 39.
[14] L. Wang, C. Hansen, S. Zidowitz, H.K. Hahn, Segmentation and Separation of Venous Vasculatures in Liver CT Images, International Society for Optics and Photonics, 2014.
[15] Y. Chi, J. Liu, S.K. Venkatesh, S. Huang, J. Zhou, Q. Tian, W.L. Nowinski, Segmentation of liver vasculature from contrast enhanced CT images using context-based voting, IEEE Trans. Biomed. Eng. 58 (8) (2011) 2144–2153.
[16] Automatic quantitative analysis of pulmonary vascular morphology in CT images, Med. Phys..
[17] S. Moccia, E.D. Momi, S.E. Hadji, L.S. Mattos, Blood vessel segmentation algorithms – review of methods, datasets and evaluation metrics, Comput. Methods Programs Biomed. 158 (2018) 71–91.
[18] Z. Zhai, M. Staring, B.C. Stoel, Lung vessel segmentation in CT images using graph-cuts, 2016.
[19] B. Lassen, E.V. Rikxoort, M. Schmidt, S. Kerkstra, B. Van Ginneken, J.M. Kuhnigk, Automatic segmentation of the pulmonary lobes from chest CT scans based on fissures, vessels, and bronchi, IEEE Trans. Med. Imaging 32 (2) (2013) 210–222.
[20] M. Orkisz, M.H. Hoyos, V.P. Romanello, C.P. Romanello, J.C. Prieto, C. Revol-Muller, Segmentation of the pulmonary vascular trees in 3D CT images using variational region-growing, Irbm 35 (1) (2014) 11–19.
[21] B. Zhao, Z. Cao, S. Wang, Lung vessel segmentation based on random forests, Electron. Lett. 53 (4) (2017) 220–222.
[22] R.A. Ochs, J.G. Goldin, F. Abtin, H.J. Kim, K. Brown, P. Batra, D. Roback, M.F. Mcnitt-Gray, M.S. Brown, Automated classification of lung bronchovascular anatomy in CT using AdaBoost, Med. Image Anal. 11 (3) (2007) 315–324.
[23] E. Goceri, Z.K. Shah, M.N. Gurcan, Vessel segmentation from abdominal magnetic resonance images: adaptive and reconstructive approach, Int. J. Numer. Methods Biomed. Eng. 33 (4) (2017).
[24] Y.Z. Zeng, Y.Q. Zhao, M. Liao, B.J. Zou, X.F. Wang, W. Wang, Liver vessel segmentation based on extreme learning machine, Phys. Med. 32 (5) (2016) 709–716.
[25] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. (2017) 1.
[26] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2018) 834–848.
[27] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid Scene Parsing Network, IEEE Computer Society, 2016.
[28] M. Livne, J. Rieger, O.U. Aydin, A.A. Taha, V.I. Madai, A U-net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease, Front. Neurosci. 13 (2019).
[29] C. Meng, K. Sun, S. Guan, Q. Wang, L. Liu, Multiscale dense convolutional neural network for DSA cerebrovascular segmentation, Neurocomputing 373 (11) (2019).
[30] z. iek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, Springer, Cham, 2016.
[31] F. Milletari, N. Navab, S.A. Ahmadi, V-Net: Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 Fourth International Conference on 3D Vision (3DV), 2016.
[32] H.C. A, Q.D. A, L.Y. A, J.Q. B, P.A.H.A. C, VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images - ScienceDirect, NeuroImage 170 (2018) 446–455.
[33] P. Sanches, C. Meyer, V. Vigon, B. Naegel, Cerebrovascular Network Segmentation on MRA Images with Deep Learning, IEEE, 2019.
[34] X. Gu, J. Wang, J. Zhao, Q. Li, Segmentation and suppression of pulmonary vessels in low-dose chest CT scans, Med. Phys. 46 (8) (2019).
[35] Y. Xu, Z. Mao, C. Liu, B. Wang, Pulmonary vessel segmentation via stage-wise convolutional networks with orientation-based region growing optimization, IEEE Access PP (99) (2018) 1.
[36] M. Oquab, L. Bottou, I. Laptev, J. Sivic, Learning and transferring mid-level image representations using convolutional neural networks, in: Computer Vision & Pattern Recognition, 2014.
[37] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How Transferable are Features in Deep Neural Networks? MIT Press, 2014.
[38] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. Mcdonagh, N.Y. Hammerla, B. Kainz, Attention U-net: Learning where to look for the pancreas, 2018.
[39] J. Valanarasu, V.A. Sindagi, I. Hacihaliloglu, V.M. Patel, KiU-Net: Overcomplete Convolutional Architectures for Biomedical Image and Volumetric Segmentation, Institute of Electrical and Electronics Engineers (IEEE), 2021.
[40] H. Huang, L. Lin, R. Tong, H. Hu, J. Wu, Unet 3+: A full-scale connected unet for medical image segmentation, in: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020.