



# HENet: Hierarchical Enhancement Network for Pulmonary Vessel Segmentation in Non-contrast CT Images

Wenqi Zhou<sup>1,2</sup>, Xiao Zhang<sup>1,3</sup>, Dongdong Gu<sup>2</sup>, Sheng Wang<sup>1,2</sup>, Jiayu Huo<sup>5</sup>, Rui Zhang<sup>4</sup>, Zhihao Jiang<sup>1</sup>, Feng Shi<sup>2</sup>, Zhong Xue<sup>2</sup>, Yiqiang Zhan<sup>2</sup>, Xi Ouyang<sup>2(✉)</sup>, and Dinggang Shen<sup>1,2(✉)</sup>

<sup>1</sup> ShanghaiTech University, Shanghai, China

[dgshen@shanghaitech.edu.cn](mailto:dgshen@shanghaitech.edu.cn)

<sup>2</sup> Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China

[xi.ouyang@uui-ai.com](mailto:xi.ouyang@uui-ai.com)

<sup>3</sup> School of Information Science and Technology, Northwest University, Xi'an, China

<sup>4</sup> Department of Pulmonary and Critical Care Medicine, West China Hospital, Sichuan University, Chengdu, China

<sup>5</sup> School of Biomedical Engineering and Imaging Sciences (BMEIS), King's College London, London, UK

**Abstract.** Pulmonary vessel segmentation in computerized tomography (CT) images is essential for pulmonary vascular disease and surgical navigation. However, the existing methods were generally designed for contrast-enhanced images, their performance is limited by the low contrast and the non-uniformity of Hounsfield Unit (HU) in non-contrast CT images, meanwhile, the varying size of the vessel structures are not well considered in current pulmonary vessel segmentation methods. To address this issue, we propose a hierarchical enhancement network (HENet) for better image- and feature-level vascular representation learning in the pulmonary vessel segmentation task. Specifically, we first design an Auto Contrast Enhancement (ACE) module to adjust the vessel contrast dynamically. Then, we propose a Cross-Scale Non-local Block (CSNB) to effectively fuse multi-scale features by utilizing both local and global semantic information. Experimental results show that our approach achieves better pulmonary vessel segmentation outcomes compared to other state-of-the-art methods, demonstrating the efficacy of the proposed ACE and CSNB module. Our code is available at <https://github.com/CODESofWenqi/HENet>.

**Keywords:** Pulmonary vessel segmentation · Non-contrast CT · Hierarchical enhancement

## 1 Introduction

Segmentation of the pulmonary vessels is the foundation for the clinical diagnosis of pulmonary vascular diseases such as pulmonary embolism (PE), pulmonary

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14222, pp. 551–560, 2023.

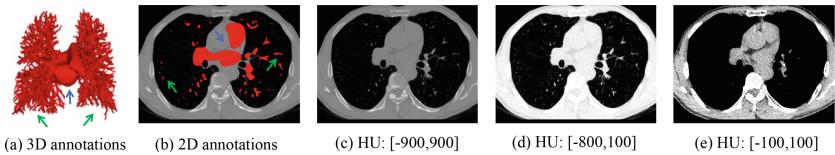
[https://doi.org/10.1007/978-3-031-43898-1\\_53](https://doi.org/10.1007/978-3-031-43898-1_53)

hypertension (PH) and lung cancer [9]. Accurate vascular quantitative analysis is crucial for physicians to study and apply in treatment planning, as well as making surgical plans. Although contrast-enhanced CT images have better contrast for pulmonary vessels compared to non-contrast CT images, the acquisition of contrast-enhanced CT images needs to inject a certain amount of contrast agent to the patients. Some patients have concerns about the possible risk of contrast media [2]. At the same time, non-contrast CT is the most widely used imaging modality for visualizing, diagnosing, and treating various lung diseases.

In the literature, several conventional methods [5, 16] have been proposed for the segmentation of pulmonary vessels in contrast-enhanced CT images. Most of these methods employed manual features to segment peripheral intrapulmonary vessels. In recent years, deep learning-based methods have emerged as promising approaches to solving challenging medical image analysis problems and have demonstrated exciting performance in segmenting various biological structures [10, 11, 15, 17]. However, for vessel segmentation, the widely used models, such as U-Net and its variants, limit their segmentation accuracy on low-contrast small vessels due to the loss of detailed information caused by the multiple down-sampling operations. Accordingly, Zhou *et al.* [17] proposed a nested structure UNet++ to redesign the skip connections for aggregating multi-scale features and improve the segmentation quality of varying-size objects. Also, some recent methods combine convolutional neural networks (CNNs) with transformer or non-local block to address this issue [3, 6, 13, 18]. Wang *et al.* [13] replaced the original skip connections with transformer blocks to better merge the multi-scale contextual information. For this task, Cui *et al.* [1] also proposed an orthogonal fused U-Net++ for pulmonary peripheral vessel segmentation. However, all these methods ignored the significant variability in HU values of pulmonary vessels at different regions.

To summarize, there exist several challenges for pulmonary vessel segmentation in non-contrast CT images: (1) The contrast between pulmonary vessels and background voxels is extremely low (Fig. 1(c)); (2) Pulmonary vessels have a complex structure and significant variability in vessel appearance, with different scales in different areas. The central extrapulmonary vessels near the heart have a large irregular ball-like shape, while the shape of the intrapulmonary vessels is delicate and tubular-like (Fig. 1(a) and (b)). Vessels become thinner as they get closer to the peripheral lung; (3) HU values of vessels in different regions vary significantly, ranging from -850 HU to 100 HU. Normally, central extrapulmonary vessels have higher HU values than peripheral intrapulmonary vessels. Thus, we set different ranges of HU values to better visualize the vessels in Fig. 1(d) and (e).

To address the above challenges, we propose a ***H*ierarchical *E*nhanement *N*etwork** (HENet) for pulmonary vessel segmentation in non-contrast CT images by enhancing the representation of vessels at both image- and feature-level. For the input CT images, we propose an Auto Contrast Enhancement (ACE) module to automatically adjust the range of HU values in different areas of CT images. It mimics the radiologist in setting the window level (WL) and window width



**Fig. 1.** The challenges of accurate pulmonary vessel segmentation. (a–b) The central extrapulmonary vessels (pointed by blue arrows) are large compared to tubular-like intrapulmonary vessels (pointed by green arrows), which become thinner as they get closer to the peripheral lung. (c) Hard to distinguish vessels in non-contrast CT images. (d–e) HU values of vessels in different regions vary significantly. (Color figure online)

(WW) to better enhance vessels from surrounding voxels, as shown in Fig. 1(d) and (e). Also, we propose a Cross-Scale Non-local Block (CSNB) to replace the skip connections in vanilla U-Net [11] structure for the aggregation of multi-scale feature maps. It helps to form local-to-global information connections to enhance vessel information at the feature-level, and address the complex scale variations of pulmonary vessels.

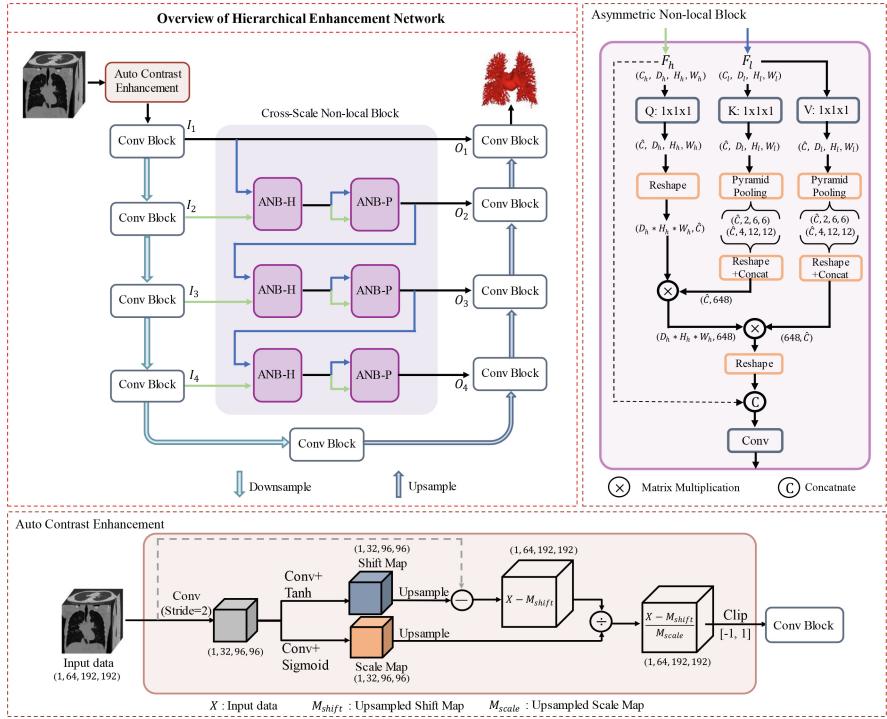
## 2 Method

The overview of the proposed method is illustrated in Fig. 2. Our proposed Hierarchical Enhancement Network (HENet) consists of two main modules: (1) Auto Contrast Enhancement (ACE) module, and (2) Cross-Scale Non-local Block (CSNB) as the skip connection bridge between encoders and decoders. In this section, we present the design of these proposed modules. First, the ACE module is developed to enhance the contrast of vessels in the original CT images for the following vessel segmentation network. After that, we introduce the CSNB module to make the network pay more attention to multi-scale vessel information in the latent feature space.

### 2.1 Auto Contrast Enhancement

In non-contrast CT images, the contrast between pulmonary vessels and the surrounding voxels is pretty low. Also, the HU values of vessels in different regions vary significantly as ranging from  $-850$  HU to  $100$  HU. Normally, radiologists have to manually set the suitable window level (WL) and window width (WW) for different regions in images to enhance vessels according to the HU value range of surrounding voxels, just as different settings to better visualize the extrapulmonary and intrapulmonary vessels (Fig. 1(d) and (e)). Instead of a fixed WL/WW as employed in existing methods, we address it by adding an ACE module to automatically enhance the contrast of vessels.

The ACE module leverages convolution operations to generate dynamic WL and WW for the input CT images according to the HU values covered by the kernel. Here we set the kernel size as  $15 \times 15 \times 15$ . First, we perform min-max



**Fig. 2.** Overview of the proposed HENet. It consists of two components: (1) Auto Contrast Enhancement module (bottom). (2) The Asymmetric Non-local Blocks in Cross-Scale Non-local Block (top-right corner).

normalization to linearly transform the HU values of the original image  $X$  to the range  $(-1, 1)$ . Then, it passes through a convolution layer to be downsampled into half-size of the original shape, which is utilized to derive the following shift map and scale map. Here, the learned shift map and scale map act as the window level and window width settings of the “width/level” scaling in CT images. We let values in the shift map be the WL, so the tanh activation function is used to limit them within  $(-1, 1)$ . The values in the scale map denote the half of WW, and we perform the sigmoid activation function to get the range  $(0, 1)$ . It matches the requirement of the positive integer for WW. After that, the shift map and scale map will be upsampled by the nearest neighbor interpolation into the original size of the input  $X$ . This operation can generate identical shift and scale values with each  $2 \times 2 \times 2$  window, avoiding sharp contrast changes in the neighboring voxels. The upsampled shift map and scale map are denoted as  $M_{shift}$  and  $M_{scale}$ , respectively, and then the contrast enhancement image  $X_{ACE}$  can be generated through:

$$X_{ACE} = \text{clip}\left(\frac{X - M_{shift}}{M_{scale}}\right). \quad (1)$$

It can be observed that the intensity values of input  $X$  are re-centered and re-scaled by  $M_{shift}$  and  $M_{scale}$  (Fig. 3(c)). The clip operation ( $\text{clip}(\cdot)$ ) truncates the final output into the range  $[-1, 1]$ , which sets the intensity value above 1 to 1, and below -1 to -1. In our experiments, we find that a large kernel size for learning of  $M_{shift}$  and  $M_{scale}$  could deliver better performance, which can capture more information on HU values from the CT images.

## 2.2 Cross-Scale Non-local Block

There are studies [14, 18] showing that non-local operations could capture long-range dependency to improve network performance. To segment pulmonary vessels with significant variability in scale and shape, we design a Cross-Scale Non-local Block (CSNB) to fuse the local features extracted by CNN backbone from different scales, and to accentuate the cross-scale dependency to address the complex scale variations of pulmonary vessels.

Inspired by [18], our CSNB incorporates 6 modified Asymmetric Non-local Blocks (ANBs), which integrate pyramid sampling modules into the non-local blocks to largely reduce the computation and memory consumption. As illustrated in Fig. 2, the CSNB works as the information bridge between encoders and decoders while also ensuring the feasibility of experiments involving large 3D data. Specifically, the  $I_1 \sim I_4$  are the inputs of CSNB, and  $O_1 \sim O_4$  are the outputs. Within the CSNB, there are three levels of modified ANBs, we denote them as ANB-H (ANB-Head) and ANB-P (ANB-Post). For the two ANBs in each level, the ANB-H has two input feature maps, and the lower-level feature maps (denoted as  $F_l$ ) contain more fine-grained information than the higher-level feature maps (denoted as  $F_h$ ). We use  $F_h$  to generate embedding  $Q$ , while embeddings  $K$  and  $V$  are derived from  $F_l$ . By doing this, CSNB can enhance the dependencies of cross-scale features. The specific computation of ANB proceeds as follows: First, three  $1 \times 1 \times 1$  convolutions (denoted as  $\text{Conv}(\cdot)$ ) are applied to transform  $F_h$  and  $F_l$  into different embeddings  $Q$ ,  $K$ , and  $V$ ; then, spatial pyramid pooling operations (denoted as  $P(\cdot)$ ) are implemented on  $K$  and  $V$ . The calculation can be expressed as:

$$Q = \text{Conv}(F_h), K_p = P(\text{Conv}(F_l)), V_p = P(\text{Conv}(F_l)). \quad (2)$$

Next, these three embeddings are reshaped to  $Q \in \mathbb{R}^{N \times \hat{C}}$ ,  $K_p \in \mathbb{R}^{S \times \hat{C}}$ ,  $V_p \in \mathbb{R}^{S \times \hat{C}}$ , where  $N$  represents the total count of the spatial locations, i.e.,  $N = D \times H \times W$  and  $S$  is equivalent to the concatenated output size after the spatial pyramid pooling, i.e., setting  $S = 648$ . The similarity matrix between  $Q$  and  $K_p$  is obtained through matrix multiplication and normalized by softmax function to get a unified similarity matrix. The attention output is acquired by:

$$O = \text{Softmax}(Q \times (K_p)^T) \times V_p, \quad (3)$$

where the output  $O \in \mathbb{R}^{N \times \hat{C}}$ . The final output of ANB is given as:

$$O_{final} = \text{Conv}(\text{cat}(O^T, F_h)), \quad (4)$$

where the final convolution is used as a weighting parameter to adjust the importance of this non-local operation and recover the channel dimension to  $C_h$ . ANB-P has the same structure as ANB-H, but the inputs  $F_h$  and  $F_l$  here are the same, which is the output of ANB-H at the same level. The ANB-P is developed to further enhance the intra-scale connection of the fused features in different regions, which is equivalent to the self-attention mechanism.

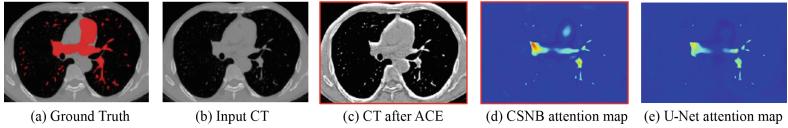
Note that,  $O_1$  is directly skipped from  $I_1$ . For the first level of CSNB, the input  $F_l$  of ANB-H is the  $I_1$ , while for the other levels, the input  $F_l$  is the output of ANB-P of the above level. That is, each level of CSNB can fuse the feature maps from its corresponding level with the fused feature maps from all of the above lower levels. Thereby, the response of multi-scale vessels can be enhanced.

### 3 Experiments and Results

**Dataset and Evaluation Metrics.** We use a total of 160 non-contrast CT images with the inplane size of  $512 \times 512$ , where the slice number varies from 217 to 622. The axial slices have the same spacing ranging from 0.58 to 0.86 mm, and the slice thickness varies from 0.7 to 1.0 mm. The annotations of pulmonary vessels are semi-automatically segmented in 3D by two radiologists using the 3D Slicer software. This study is approved by the ethical committee of West China Hospital of Sichuan University, China. These cases are randomly split into a training set (120 scans) and a testing set (40 scans). The quantitative results are reported by Dice Similarity Coefficient (Dice), mean Intersection over Union (mIoU), False Positive Rate (FPR), Average Surface Distance (ASD), and Hausdorff Distance (HD). For the significance test, we use the paired t-test.

**Implementation Details.** Our experiments are implemented using Pytorch framework and trained using a single NVIDIA-A100 GPU. We pre-process the data by truncating the HU value to the range of  $[-900, 900]$  and then linearly scaling it to  $[-1, 1]$ . In the training stage, we randomly crop sub-volumes with the size of  $192 \times 192 \times 64$  near the lung field, and then the cropped sub-volumes are augmented by random horizontal and vertical flipping with a probability of 0.5. In the testing phase, we perform the sliding window average prediction with strides of  $(64, 64, 32)$  to cover the entire CT images. For a fair comparison, we use the same hyper-parameter settings and Dice similarity coefficient loss across all experiments. In particular, we use the same data augmentation, no post-processing scheme, Adam optimizer with an initial learning rate of  $10^{-4}$ , and train for 800 epochs with a batch size of 4. In our experiments, we use a two-step optimization strategy: 1) first, train the ACE module with the basic U-Net; 2) Integrate the trained ACE module and a new CSNB module into the U-Net, and fix the parameters of ACE module when training this network.

**Ablation Study.** We conduct ablation studies to validate the efficacy of the proposed modules in our HENet by combining them with the baseline U-Net [11].



**Fig. 3.** The effectiveness of the proposed components, with the images in red circles generated from our method. (a–c) The contrast of vessels is significantly enhanced in the CT images processed by ACE module. (d–e) Compared to baseline, CSNB can enhance the ability to capture vascular features with widely variable size, shape, and location. (Color figure online)

The quantitative results are summarized in Table 1. Compared to the baseline, both ACE module and CSNB lead to better performance. With the two components, our HENet has significant improvements over baseline on all the metrics. For regional measures Dice and mIoU, it improves by 3.02% and 2.32% respectively. For surface-aware measures ASD and HD, it improves by 35% and 53%, respectively. Results demonstrate effectiveness of the proposed ACE module and CSNB.

To validate the efficacy of ACE module, we show the qualitative result in Fig. 3. As shown in Fig. 3(c), the ACE module effectively enhances the contrast of pulmonary vessels at the image-level. We also visualize the summation of feature maps from the final decoder in Fig. 3(d) and (e). As can be seen, the baseline U-Net can focus on local features for certain intrapulmonary vessels, but it fails to activate complete vascular regions of multiple scales.

**Table 1.** Quantitative results of ablation study. \* denotes significant improvement compared to the baseline U-Net ( $p < 0.05$ ).

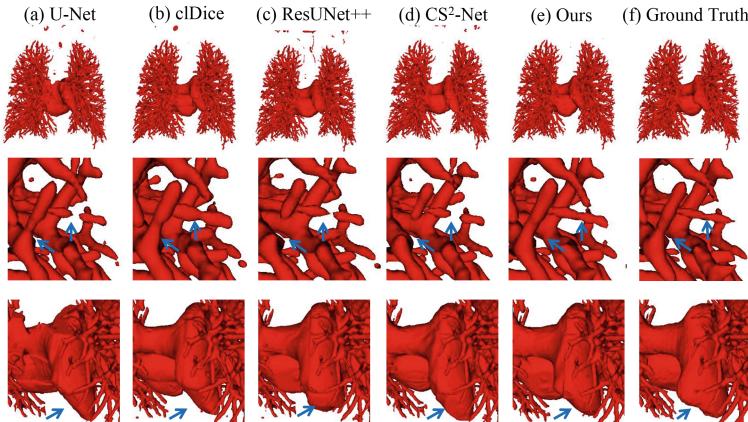
Method	Dice [%]↑	mIoU [%]↑	ASD [mm]↓	HD [mm]↓	FPR [%]↓
U-Net	$82.88 \pm 3.44$	$85.02 \pm 2.56$	$1.11 \pm 0.56$	$20.16 \pm 21.78$	$0.35 \pm 0.15$
U-Net+ACE	$84.71 \pm 2.69^*$	$86.40 \pm 2.09^*$	$0.88 \pm 0.46^*$	$10.80 \pm 11.29^*$	$0.33 \pm 0.15$
U-Net+CSNB	$85.22 \pm 2.79^*$	$86.80 \pm 2.18^*$	$0.85 \pm 0.43^*$	$14.00 \pm 16.64$	$0.30 \pm 0.14$
<b>Ours</b>	<b><math>85.90 \pm 2.92^*</math></b>	<b><math>87.34 \pm 2.29^*</math></b>	<b><math>0.72 \pm 0.41^*</math></b>	<b><math>9.43 \pm 10.91^*</math></b>	<b><math>0.27 \pm 0.14^*</math></b>

**Comparison with State-of-the-Art Methods.** Since we adopt U-Net as our baseline, we compare our method with several state-of-the-art encoder-decoder CNNs and the transformer-based method VT-UNet [8] within a considerable computational complexity. We also compare our method with state-of-the-art deep learning-based vessel segmentation methods, including cIDice [12], CS<sup>2</sup>-Net [7], and OF-Net [1]. The quantitative and qualitative results are presented in Table 2 and Fig. 4, respectively.

As shown in Table 2, our method outperforms the competing methods and achieves the best Dice, mIoU, ASD, and HD. CS<sup>2</sup>-Net performs best on FPR,

**Table 2.** Quantitative comparison with the state-of-the-art methods.

Method	Dice [%]↑	mIoU [%]↑	ASD [mm]↓	HD [mm]↓	FPR [%]↓
VT-UNet [8]	$77.52 \pm 3.33$	$81.14 \pm 2.27$	$1.96 \pm 2.05$	$30.02 \pm 31.92$	$0.49 \pm 0.17$
OF-Net [1]	$82.70 \pm 3.53$	$84.89 \pm 2.65$	$1.45 \pm 0.88$	$28.62 \pm 31.99$	$0.31 \pm 0.14$
U-Net [11]	$82.88 \pm 3.44$	$85.02 \pm 2.56$	$1.11 \pm 0.56$	$20.16 \pm 21.78$	$0.35 \pm 0.15$
clDice [12]	$85.08 \pm 2.65$	$86.67 \pm 2.04$	$0.82 \pm 0.41$	$10.38 \pm 10.59$	$0.43 \pm 0.17$
ResUNet++ [4]	$85.30 \pm 3.21$	$86.88 \pm 2.49$	$1.31 \pm 1.27$	$23.67 \pm 31.40$	$0.30 \pm 0.15$
CS <sup>2</sup> -Net [7]	$85.34 \pm 3.04$	$86.91 \pm 2.37$	$0.89 \pm 0.46$	$13.50 \pm 17.13$	<b><math>0.25 \pm 0.13</math></b>
<b>Ours</b>	<b><math>85.90 \pm 2.92</math></b>	<b><math>87.34 \pm 2.29</math></b>	<b><math>0.72 \pm 0.41</math></b>	<b><math>9.43 \pm 10.91</math></b>	$0.27 \pm 0.14$

**Fig. 4.** Qualitative segmentation results. The blue arrows are used to highlight the regions for visual presentation. (Color figure online)

but our method has better Dice and mIoU than CS<sup>2</sup>-Net (increasing 0.56% and 0.43%, respectively), indicating the under-segmentation of CS<sup>2</sup>-Net and more vessels being correctly segmented by our method. In the first row of qualitative results (Fig. 4), the competing methods can produce satisfactory results for the overall structure but generate many false positives. Furthermore, due to low contrast between small intrapulmonary vessels and the surrounding voxels, results of competing methods exist many discontinuities (the second row), while our method obtains more connective segmentation for these small vessels. Also, for the segmentation of large extrapulmonary vessels (the last row), our method can produce more accurate results. Note that, although clDice can also yield connective results for small vessels, their FPR is 0.16% higher than ours. This implies that clDice tends to over-segment vessels, and it cannot obtain precise segmentation for the large extrapulmonary vessels. Results proved the superiority of our method.

## 4 Conclusion

In this paper, we have proposed a hierarchical enhancement network to enhance the representation of vessels at both image- and feature-level for pulmonary vessel segmentation first time in non-contrast CT images. In the proposed HENet, an Auto Contrast Enhancement module is designed to enhance vessels in different regions of the input CT. And the Cross-Scale Non-local Block is further designed as the information bridge between encoders and decoders, to enhance the ability to capture and integrate vascular features of multiple scales. Experimental results show that our method outperforms the competing methods and demonstrates effectiveness of the proposed ACE module and CSNB. At the same time, it can be observed that the learning of  $M_{shift}$  and  $M_{scale}$  is only supervised by the final segmentation loss. One of our future research directions is to develop explicit constraints for the ACE module to better re-center and re-scale the CT images.

**Acknowledgments.** This work was supported in part by National Key Research and Development Program of China (2021ZD0111100), National Natural Science Foundation of China (62131015), and Science and Technology Commission of Shanghai Municipality (STCSM) (21010502600).

## References

1. Cui, H., Liu, X., Huang, N.: Pulmonary vessel segmentation based on orthogonal fused U-Net++ of chest CT images. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11769, pp. 293–300. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-32226-7\\_33](https://doi.org/10.1007/978-3-030-32226-7_33)
2. Hasebroock, K.M., Serkova, N.J.: Toxicity of MRI and CT contrast agents. Expert Opin. Drug Metab. Toxicol. **5**(4), 403–416 (2009)
3. Huang, H., et al.: ScaleFormer: revisiting the transformer-based backbones from a scale-wise perspective for medical image segmentation. arXiv preprint [arXiv:2207.14552](https://arxiv.org/abs/2207.14552) (2022)
4. Jha, D., et al.: ResUNet++: an advanced architecture for medical image segmentation. In: 2019 IEEE International Symposium on Multimedia (ISM), pp. 225–2255. IEEE (2019)
5. Kaftan, J.N., Kiraly, A.P., Bakai, A., Das, M., Novak, C.L., Aach, T.: Fuzzy pulmonary vessel segmentation in contrast enhanced CT data. In: Medical Imaging 2008: Image Processing, vol. 6914, pp. 585–596. SPIE (2008)
6. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV), pp. 565–571. IEEE (2016)
7. Mou, L., et al.: CS<sup>2</sup>-net: deep learning segmentation of curvilinear structures in medical imaging. Med. Image Anal. **67**, 101874 (2021)
8. Peiris, H., Hayat, M., Chen, Z., Egan, G., Harandi, M.: A robust volumetric transformer for accurate 3D tumor segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13435, pp. 162–172. Springer, Heidelberg (2022). [https://doi.org/10.1007/978-3-031-16443-9\\_16](https://doi.org/10.1007/978-3-031-16443-9_16)

9. Pu, J., et al.: Automated identification of pulmonary arteries and veins depicted in non-contrast chest CT scans. *Med. Image Anal.* **77**, 102367 (2022)
10. Qin, Y., Zheng, H., Gu, Y., Huang, X., Yang, J., Wang, L., Zhu, Y.-M.: Learning bronchiole-sensitive airway segmentation CNNs by feature recalibration and attention distillation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 221–231. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-59710-8\\_22](https://doi.org/10.1007/978-3-030-59710-8_22)
11. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
12. Shit, S., et al.: cIDice-a novel topology-preserving loss function for tubular structure segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 16560–16569 (2021)
13. Wang, H., Cao, P., Wang, J., Zaiane, O.R.: UCTransNet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, pp. 2441–2449 (2022)
14. Wang, X., Girshick, R., Gupta, A., He, K.: Non-local neural networks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7794–7803 (2018)
15. Zhang, X., et al.: Progressive deep segmentation of coronary artery via hierarchical topology learning. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13435, pp. 391–400. Springer, Heidelberg (2022). [https://doi.org/10.1007/978-3-031-16443-9\\_38](https://doi.org/10.1007/978-3-031-16443-9_38)
16. Zhou, C., et al.: Automatic multiscale enhancement and segmentation of pulmonary vessels in CT pulmonary angiography images for cad applications. *Med. Phys.* **34**(12), 4567–4577 (2007)
17. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested U-net architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 3–11. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1)
18. Zhu, Z., Xu, M., Bai, S., Huang, T., Bai, X.: Asymmetric non-local neural networks for semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 593–602 (2019)