



Research article

Liver vessel segmentation based on inter-scale V-Net

Jinzu Yang^{1,*}, Meihan Fu² and Ying Hu^{2,*}

¹ Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education Northeastern University, Shenyang 110000, China

² College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116000, China

* **Correspondence:** Email: yangjinzhu@cse.neu.edu.cn, hawk_huy@sina.com.

Abstract: Segmentation and visualization of liver vessel is a key task in preoperative planning and computer-aided diagnosis of liver diseases. Due to the irregular structure of liver vessel, accurate liver vessel segmentation is difficult. This paper proposes a method of liver vessel segmentation based on an improved V-Net network. Firstly, a dilated convolution is introduced into the network to make the network can still enlarge the receptive field without reducing down-sampling and save detailed spatial information. Secondly, a 3D deep supervision mechanism is introduced into the network to speed up the convergence of the network and help the network learn semantic features better. Finally, inter-scale dense connections are designed in the decoder of the network to prevent the loss of high-level semantic information during the decoding process and effectively integrate multi-scale feature information. The public datasets 3Dircadb were used to perform liver vessel segmentation experiments. The average dice and sensitivity of the proposed method reached 71.6 and 75.4%, respectively, which are higher than those of the original network. The experimental results show that the improved V-Net network can automatically and accurately segment labeled or even other unlabeled liver vessels from the CT images.

Keywords: liver vessel; V-Net, dilated convolution; 3D deep supervision mechanism; inter-scale dense connections

1. Introduction

Liver cancer is the second most deadly cancer after lung cancer and one of the cancers with the fastest increasing morbidity and mortality in China [1]. At present, the computer-assisted treatment method for liver cancer is thermal ablation which is an effective treatment method to eliminate

malignant liver tumors in addition to surgical cutting and liver transplantation. Three-dimensional visualization of liver vessels is essential for the path planning and guidance for thermal ablation. The relative position of the liver vessels and liver tumors can determine the final ablation effect and affect the tumor recurrence rate [2]. Manually delineating liver vessels is time-consuming and laborious, and the delineation results between different experts are quite different. Due to the irregularity, uneven distribution, and the low contrast with surrounding organs of liver vessels, it is difficult to segment. Therefore, an accurate, fast, and efficient liver vessel segmentation method is needed for clinical applications.

The traditional methods of liver vessel segmentation mainly include the region growing method, image filtering and enhancement algorithm, tracking algorithm, and machine learning method [3]. The region-growing method is a semi-automatic segmentation method based on pixels or voxels, which relies on gray similarity and spatial proximity. Oliveira et al. [4] used the region-growing method to extract liver vessels in the segmented liver region, which can obtain the main branches of vessels, but it has a weak effect on small vessel segmentation and is sensitive to noise. Chi et al. [5] proposed a vascular context-based voting system to segment vessels using regional features, but they required manual labeling of seed points and took a long time. Image filtering enhancement method, based on the characteristics of the tree structure of the liver vessels as a whole, using the relationship between Hessian matrix eigenvalues to extract the tubular structure in the image, which is used for multi-scale enhancement of liver vessels [6,7]. However, this method cannot distinguish liver vessels correctly on images with high noise or uneven intensity, so region growth [8], graph cuts [9], and morphological operations are often used for post-processing. The tracking algorithm starts from a certain point of the vessel boundary on the image, searches for and tracks the adjacent boundary points to obtain the entire vessel boundary [10]. The above all algorithms are prone to produce segmentation errors and often require user interaction.

In machine learning algorithms, Zeng et al. [11] used K-means clustering combined with a hybrid active contour model to perform coarse vessel segmentation. Gaussian filters were utilized to promote 3D region growth for fine vessel segmentation. The final segmentation result combines these two segmentation results, but the segmented liver vessel surface was rough, and some thick vessels with lower intensity failed to be extracted.

Recently, deep learning has achieved breakthrough results in computer vision tasks. For example, convolutional neural network (CNN) is a classic deep learning model that can learn complex data features from numerous training samples, thereby avoiding the cumbersome feature extraction process and possessing generalization capabilities. Kitrungrotsakul et al. [12] used three deep convolutional networks (DNN) to learn liver vessel features from different planes of CT images, but the network failed to segment liver vessels that have diverse intensity range than the training datasets. The fully convolutional neural network U-Net makes a great contribution to medical image segmentation [13]. It introduces the encoder-decoder and skip connections, which is very suitable for small sample training. Since most medical images are volume data, networks suitable for 3D medical image segmentation have appeared, such as 3DU-Net [14] and V-Net [15]. Yu et al. [16] utilized 3D residual U-Net spatial contextual information to extract liver vessels from the CT images. Xu et al. [17] applied a 3DCNN network to extract liver vessel features. The network uses dilated convolutional layers instead of the down-sampling layers to capture multi-scale context information without reducing the size of the feature map. Huang et al. [18] also used the 3DU-Net network to extract all liver vessels.

Due to the irregular shape of liver vessels, the surrounding tissue has low contrast, and 3D medical training samples with annotations are minimal. Besides, image segmentation is essentially a classification problem at the pixel level. Suppose the foreground and background categories are not balanced. In that case, it will easily cause training to fall into the optimal local value, and the small foreground area will be lost or without detected. Although the spatial information of the decoder in the V-Net network is relatively rough, it has powerful semantic features and accurate resolution. However, de-convolution and convolution often result in the loss of high-level semantic information so that the contextual information cannot be propagated to a higher resolution layer. The above problems make it difficult to segment liver vessels with conventional deep convolutional networks.

V-Net is chosen as the basic network structure of liver vessel segmentation and improved. The main contributions are as follows: 1) The original network structure is optimized and a 3D deep supervision mechanism [19] is introduced into the network, which helps the network learn semantic features better, accelerate the convergence speed and improve prediction accuracy. 2) Inter-scale dense connections are designed in the decoder, aiming to reduce the loss of high-level semantic information during the decoding process and effectively fuse multi-scale feature information. 3) A loss function composed of binary cross-entropy and dice coefficient is utilized to ensure that the network can still effectively train in the case of category imbalance.

2. Methods

2.1. Preprocessing

The preprocessing steps are as follows: 1) CT values of the CT image are limited to $[-200,200]$ HU, which can filter out other organs in the image. 2) Due to the limitation of GPU memory, the original resolution is changed from 512×512 to 256×256 by down-sampling. 3) Because the thickness of most training dataset slices is 1.6 mm, the thickness of data slices less than 1.25 mm or greater than 2 mm is normalized to 1.6 mm by trilinear interpolation. A three-dimensional training data is taken multiple 48 slices continuously for training with a sliding step of 5. 4) Rotation and mirroring operations are used to augment the data.

2.2. Improvement of V-Net network framework

V-Net is a 5-layer symmetrical network architecture with an encoder that extracts spatial features from images, and a decoder that constructs segmentation graphs from encoded features, as well as the skip connection structure that combines the position information in the encoding path with the context information in the decoding path to make up for the missing edge features and spatial information during the decoding process. To mitigate the disappearance of the network gradient, the residual units are added to the network. The formula is as follows:

$$x_L = x_l + \sum_{i=l}^{L-1} \mathcal{F}(x_i, W_i) \quad (2.1)$$

where \mathcal{F} represents the residual function, x_i is the input feature, and W_i is a group of weights related to the residual units. Any deeper feature $x_L (L > l \geq 1)$ can be expressed as a shallow feature

x_i plus an accumulated residual function $\sum_{i=1}^{L-1} \mathcal{F}(x_i, W_i)$.

Because the original V-Net has many parameters, it is easy to cause the network to overfit. Therefore, $3 \times 3 \times 3$ convolution kernels are applied to replace the original $5 \times 5 \times 5$ convolution kernels in each layer of the network. A PReLU activation function adopted throughout the network. The down-sampling of the V-Net network adopts the convolution method. That is, the feature map is convolved using a $2 \times 2 \times 2$ convolution kernel with stride 2 in the encoder to reduce the resolution rate of the feature map. At the same time, the number of feature channels in each layer is doubled to learn in-depth features more accurately and fully.

Although down-sampling can increase the receiving field, it also reduces the spatial resolution. Therefore, the last layer of the network is changed. Only the number of feature channels is increased without change the feature map size (see Figure 1). Three dilated convolutions are introduced in the third and fourth layers of the encoder to avoid losing the resolution and still increase the receptive field. The third layer dilation rate are 1, 2, and 4, and the corresponding receptive fields are 3, 7, and 15, respectively. The fourth layer dilation rate 3, 4, and 5, and the corresponding receptive fields are 11, 15, and 19, respectively. Adjusting the dilation rate of the dilated convolution can extract the context information about different scales of the feature map. The network can locate the target more accurately due to the improvement of the resolution. Each layer in the decoding path uses a $2 \times 2 \times 2$ de-convolution with stride 2 for up-sampling. The number of feature channels is halved, followed by 3 padded convolutions (the last layer is 2 padded convolutions). Finally, in the output layer, a $1 \times 1 \times 1$ convolution is performed to adjust the number of channels of the characteristic map. Because the image resolution is reduced in the preprocessing stage, trilinear interpolation is performed to restore the feature map to the original image size. A sigmoid function is applied to obtain the final probability map. A dropout layer is added at the end of the residual unit of each layer to prevent the network from overfitting.

2.3. 3D deep supervision mechanism

A 3D deep supervision mechanism is introduced into the network to optimize the model, speed up the network learning speed and prevent information loss during the forward propagation. Because the parameters of each path are initialized randomly, this mechanism allows different paths to update the weights independently without interfering with each other so that the learning of the network will not stay in the same local minimum. Moreover, introducing a deep supervision mechanism allows the network to obtain more feedback information during the back propagation process than just using the last output layer for back propagation (see Figure 1). Three output layers are added to the decoder of the improved V-Net network. Each layer's characteristic map is up-sampling by trilinear interpolation, and then the loss value is calculated after a sigmoid function. The 3D deep supervision formula is as follows:

$$L = L(\mathcal{X}; W) + \sum_{d \in D} \eta_d L_d(\mathcal{X}; W_d, \hat{w}_d) + \lambda (\|W\|^2 + \sum_{d \in D} \|\hat{w}_d\|^2) \quad (2.2)$$

where $L(\mathcal{X}; W)$ is the loss value calculated by comparing the prediction result of the main network with the label of the ground truth, \mathcal{X} and W respectively represent the training database and main network weights, $L_d(\mathcal{X}; W_d, \hat{w}_d)$ is the auxiliary loss of all hidden layers, η_d is the balancing weight of

L_d . w_d represents the weight of the d -th layer in the main network and the third term is the weight attenuation regularizations and λ is the hyper-parameter of weight.

The 3D deep supervision mechanism can promote the expression of high-level features by hidden layers, thereby promoting the discrimination capability of the model. As these different loss components propagate backward, the equivalent training data expands, thereby effectively preventing the network overfitting and further boosting its generalization capability.

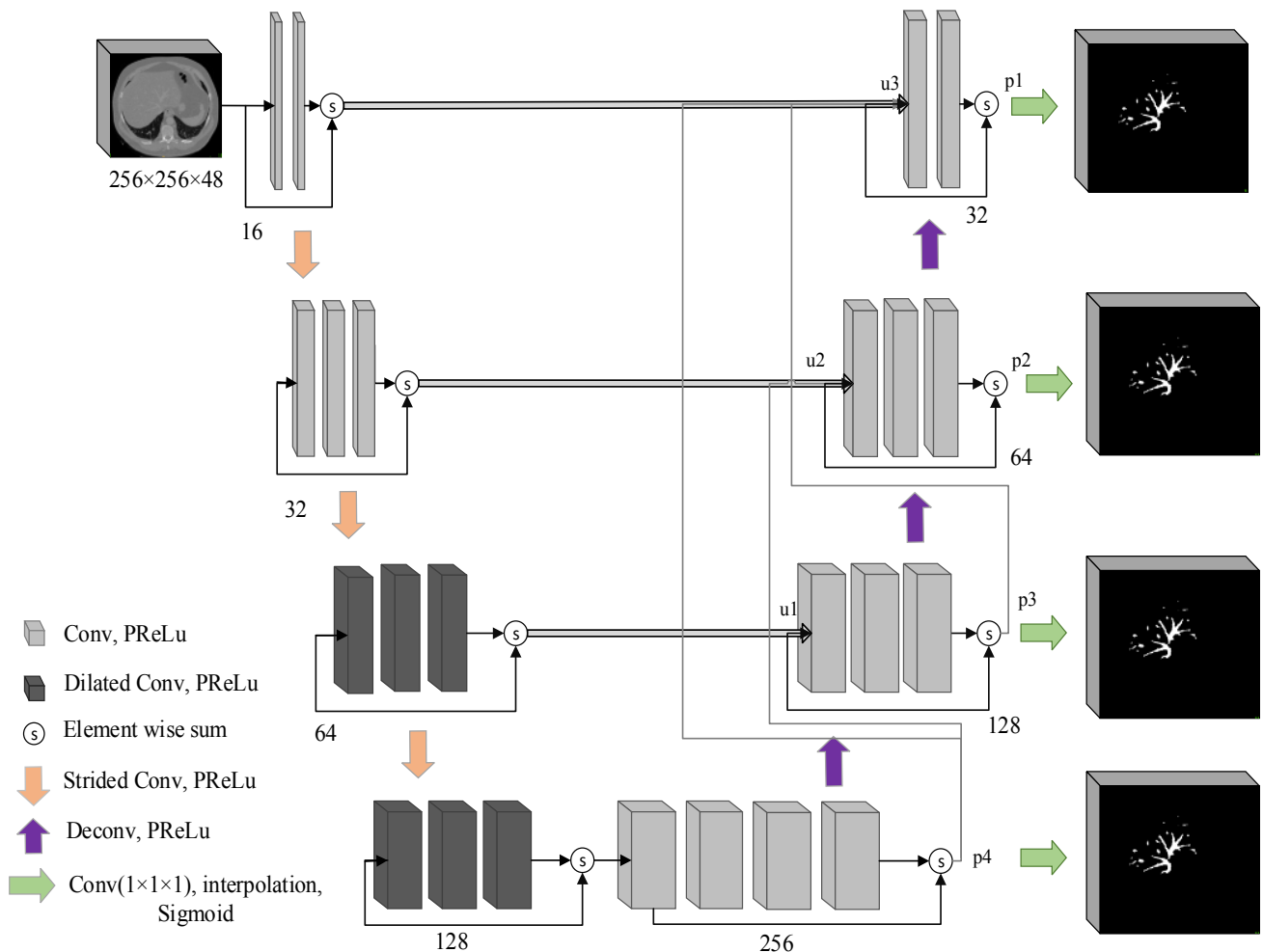


Figure 1. Schematic diagram of the overall structure of the improved V-Net network.

2.4. Inter-scale dense connections

Inter-scale dense connections are introduced in the decoder to further reduce the information loss during the decoding process. The network constructs encoders and decoders for top-down and bottom-up methods. Although the spatial information is coarse in the decoder, it has powerful semantic features and precise resolution. Due to the large semantic gap between the layers, these specific inter-scale dense connections can directly propagate the feature information from one scale stage to another scale stage so that it can fuse feature information of different scales to prevent high-level semantic information loss.

The improved V-Net network is a four-layer network structure, and we use the feature activation

of the residual block output of each stage from bottom to top in the decoder. We indicate that the output of residual block is $\{p1, p2, p3, p4\}$, and the up-convolution block is $\{u1, u2, u3\}$ (see Figure 1). To achieve inter-scale dense connections at the decoder (see Figure 2), p in $p4 \rightarrow u3$, $p3 \rightarrow u3$, $p4 \rightarrow u2$ is passed through a connection block (The connection block includes using trilinear interpolation for up-sampling and using $1 \times 1 \times 1$ convolution to reduce the number of channels.) is fused with the corresponding u . Then it is fused again with the feature maps propagated through the skip connection in the same layer to achieve multi-fusion. The inter-scale dense connections effectively avoid the loss of deep semantic information caused by operations such as up-sampling and multiple convolutions.

The inter-scale dense connections formula is as follows:

$$x_L = \sum_{j=1}^L \{ \Gamma(x_j, w_j) + \sum_{i=2}^{L-j} \Theta(x_{j+i}, w_{j+i}) \} \quad (2.3)$$

where L represents the number of network layers, Γ is the j -th layer de-convolution block in the encoder, x_j is the input feature of the de-convolution block, and w_j is a group of weights related to the de-convolution block. Θ is the connection block after the residual structure of $j+i$ layer. x_{j+i} is the input feature of the connection block, and w_{j+i} is a group of weights related to the connection block.

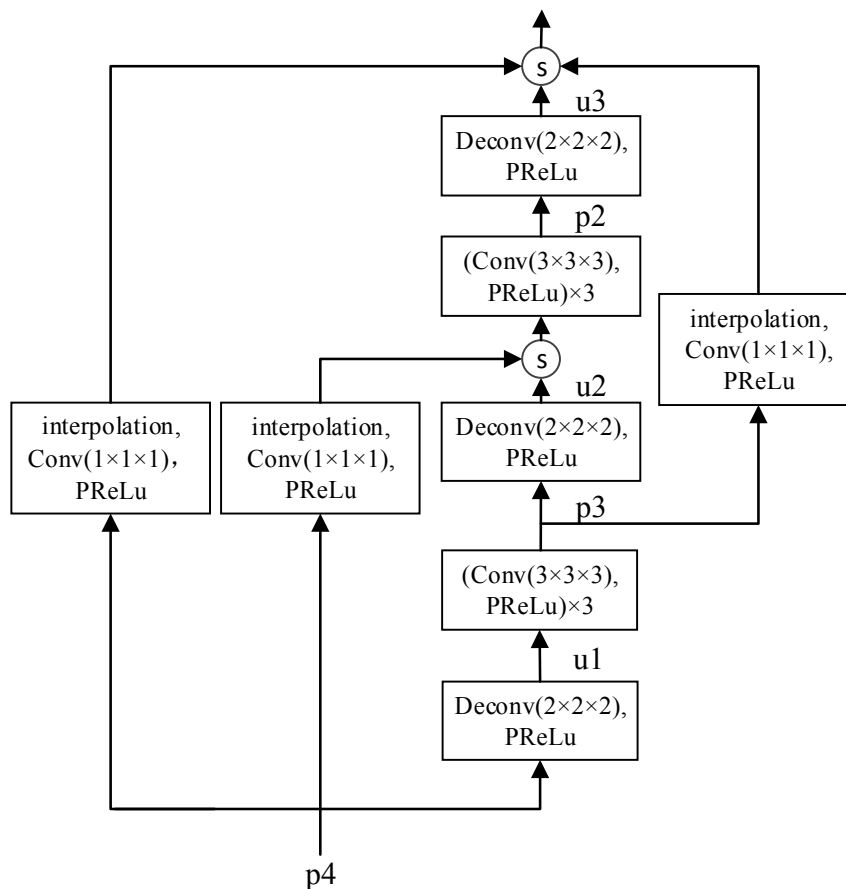


Figure 2. Inter-scale dense connections.

2.5. Loss function

A combined loss function, which is composed of binary cross-entropy loss function and dice loss function [20], enables the network to be effectively trained in imbalanced categories. The combined loss function formula is as follows, where α is a weighting factor.

$$L_{BD} = (1 - \alpha)L_{BCE} + \alpha(1 - L_{Dice}) \quad (2.4)$$

The binary cross-entropy loss function is as follows:

$$L_{BCE} = -\frac{1}{n} \sum_{i=1}^n (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (2.5)$$

The dice loss function is as follows:

$$L_{Dice} = \frac{2 \sum_i^n y_i \hat{y}_i}{\sum_i^n y_i^2 + \sum_i^n \hat{y}_i^2} \quad (2.6)$$

where \hat{y} represents the prediction result of the network, y represents the true label of the corresponding voxel.

2.6. Post-processing

In post-processing, the volume of each connected region is calculated. To prevent the predicted disconnected liver vessels from being removed, we remove the small area noise (less than 450 mm³) caused by classification errors through volume judgment, effectively reducing false positives in segmentation results (see Figure 3).

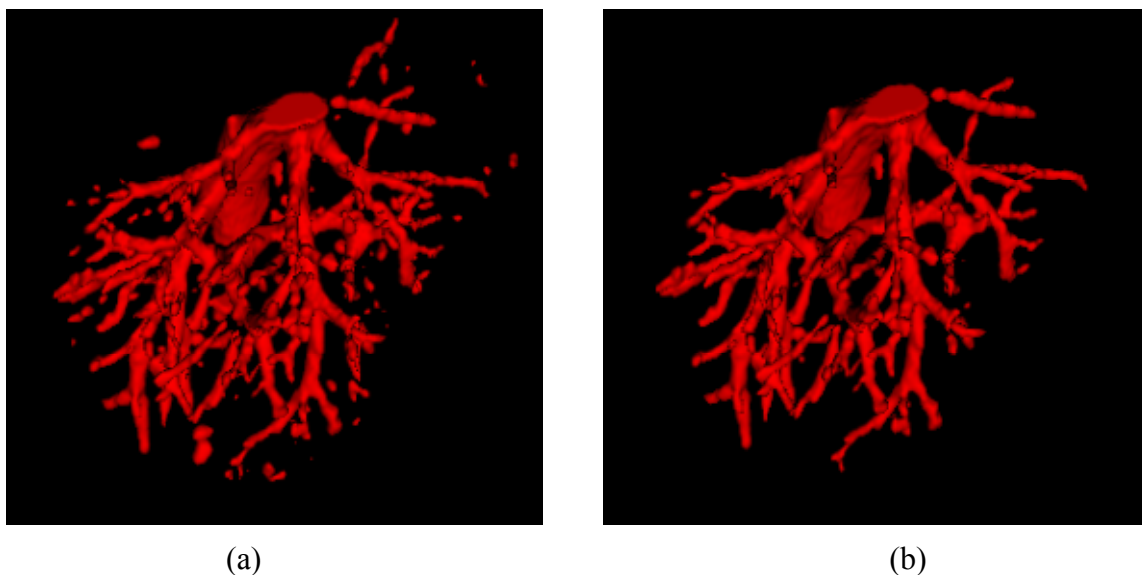


Figure 3. The process of post-processing. (a) (b) 3D visualization result before and after post-processing.

3. Experiments and results

3.1. Experimental environment and experimental data

The hardware configuration required for the experiment is Intel (R) Xeon (R) Silver 4110 CPU @ 2.10GHz and an NVIDIA Tesla T4 GPU (16 GB memory) and the development tools are Python3.7 and PyTorch.

The experimental data were selected from the public CT image datasets 3Dircadb provided by the Research Institute against Digestive Cancer. The datasets contain 20 three-dimensional images of enhanced portal venous phase with pixel spacing ranging from 0.56 to 0.86 mm, slice thickness ranging from 1mm to 4 mm, number of slices ranging from 64 to 502, and single-layer resolution of 512×512 , by manually selecting 12 cases for training and 8 cases for testing.

3.2. Parameter settings and training

The dropout parameter was set as 0.5 and the 3D deep supervision weight was initialized to 0.33, which decays as the training progress. A typical Adam optimizer was selected for network training, and the initial learning rate was 0.0001. Considering the computing resources, the batch size was set as 1, and the final number of the epoch was 35. Because the inter-scale dense connections were designed into the improved V-Net network, the probability map of the last output layer is used as the final segmentation result when making predictions. The training time of the model was about 10 h, the testing time of the 8 test datasets was 9.58–26.53 s, and the average testing time was 13.46 s.

3.3. Evaluation metrics

The following four evaluation metrics were selected, which include the dice coefficient (Dice), accuracy (Acc), sensitivity (Sen), and specificity (Spe). The formulas are as follows:

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (3.1)$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \quad (3.2)$$

$$Sen = \frac{TP}{FN + TP} \quad (3.3)$$

$$Spe = \frac{TN}{TN + FP} \quad (3.4)$$

where TP and TN are the numbers of voxels correctly divided into liver vessels and background, and FP and FN are the numbers of voxels incorrectly divided into liver vessels and background.

3.4. Selection of the weighting factor

The selection of weighting factor α in the combined loss function was analyzed, which was set to 0, 0.3, 0.5, 0.7, 0.9 and 1, respectively. Table 1 shows the effect of combined loss function on improved V-Net network performance under different weight factors. It can be seen from Table 1 that when α is 0.7, the performance of the network is good. In particular, when α is 0, the loss function is the binary cross-entropy loss function; when α is 1, the loss function is the dice loss function. Therefore, the weighting factor α in this experiment was set to 0.7.

Table 1. The impact of different weighting factors on improved V-Net performance.

Weighting factor α	Dice (%)	Sen (%)
0	66.9	67.4
0.3	67.0	69.1
0.5	67.9	72.7
0.7	68.7	73.4
0.9	68.3	72.8
1	66.7	71.4

As shown in Figure 4, using the combined loss function in the network can segment even smaller liver vessels than using the dice loss function. However, there are still disparities compared with the annotated data. Therefore, in the following experiments, the combined loss function is used to train the network.

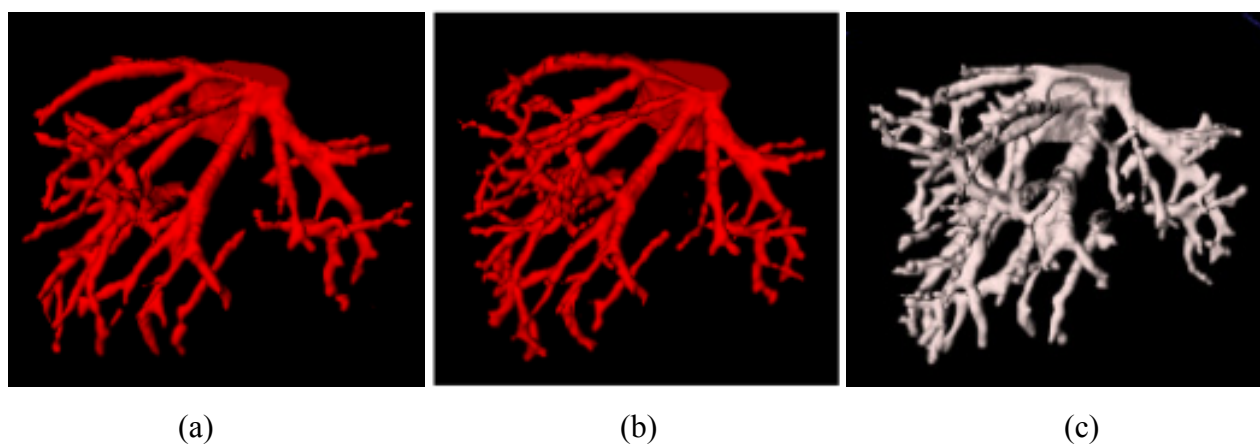


Figure 4. Performance comparison on different loss functions. (a) 3D visualization result using dice loss function; (b) 3D visualization result using combined loss function; (c) 3D visualization result of expert segmentation.

3.5. Evaluation and comparison

Each improved method was tested on 8 3Dircadb datasets, and the results were post-processing. As shown in Table 2, after introducing the 3D deep supervision mechanism into the improved V-Net network, the average dice, sensitivity, accuracy, and specificity are improved by 1.3, 0.7, 0.5,

and 0.2%, respectively. The 3D deep supervision mechanism can alleviate the gradient disappearance or explosion of the network during the training process, make the network update parameters from different paths without interference, and help the network learn discrimination features better. When inter-scale dense connections were introduced into the improved V-Net network, the average dice was 71.2%, sensitivity was 74.8%, accuracy and specificity were 98.4 and 99.4%, respectively. Compared with the improved V-Net network evaluation, the average dice value improved by 2.5%, and sensitivity improved by 1.4%, which shows that this method can effectively compensate for the loss of high-level semantic information due to multiple up-sampling and convolution and achieve inter-scale feature fusion. As shown in Figure 5, this method can extract more thin-walled small liver vessels and further optimize the segmentation results.

Table 2. Comparison of segmentation performance of the improved network on 3Dircadb test data.

Methods	Dice (%)	Sen (%)	Acc (%)	Spe (%)
3DU-Net	65.7	68.1	97.1	98.4
Improved V-Net + L_{BD}	68.7	73.4	97.6	99.2
Improved V-Net + L_{BD} + DS	70.0	74.1	98.1	99.4
Improved V-Net + L_{BD} + ISD	71.2	74.8	98.4	99.4
Improved V-Net + L_{BD} + DS + ISD	71.6	75.4	98.5	99.5
Improved V-Net + L_{BD} + DS + ISD (no pp)	71.5	75.5	98.4	99.5

(DS: 3D deep supervision mechanism; ISD: inter-scale dense connections; no pp: no post-processing)

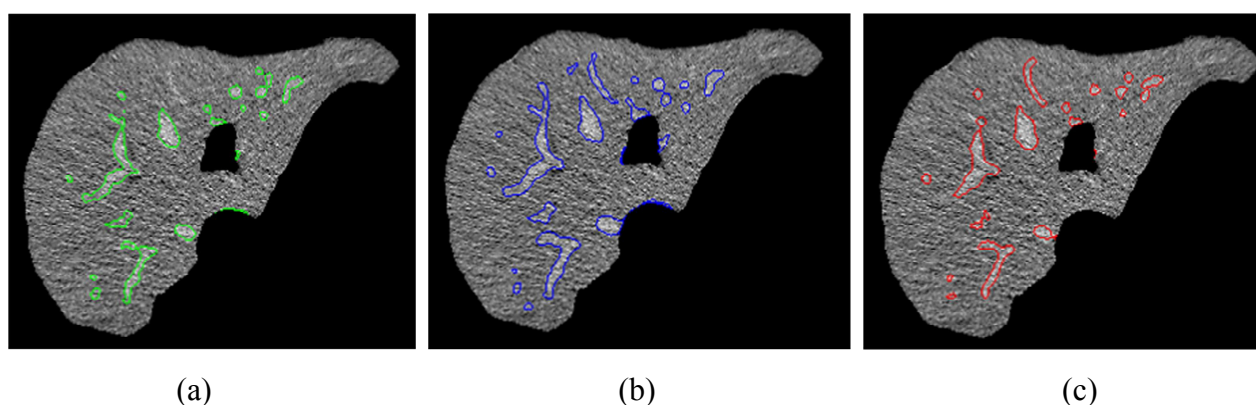


Figure 5. An example of performances of the proposed method. (a) liver vessel slice result using improved V-Net network; (b) liver vessel slice result using the inter-scale dense connections; (c) liver vessel slice result of expert segmentation.

Finally, the 3D deep supervision mechanism and the inter-scale dense connection were introduced to the network simultaneously. The final average dice, sensitivity, accuracy, and

specificity of the testing data were 71.6, 75.4, 98.5, and 99.5%, respectively. The average dice, sensitivity, accuracy and specificity of liver vessels without post-processing were 71.5, 75.5, 98.4 and 99.5%, respectively. As shown in Table 3, the average sensitivity of our proposed method is slightly lower than the method [9], but it belongs to a semi-automatic segmentation method, and other metrics are significantly higher than the comparison methods, which indicates that our proposed method has better segmentation performance. As shown in Figure 6, the narrow liver vessels segmented by our method are closer to the real liver vessel contour and have high accuracy and robustness for images with high noise, low contrast and varied intensity distribution.

Table 3. Comparison of segmentation performance between the proposed algorithm and other algorithms.

Methods	Dice (%)	Sen (%)	Acc (%)	Spe (%)
Method in [5]	–	70.0	98.0	99.0
Method in [9]	–	79.8	97.7	98.6
Method in [18]	67.5	74.3	97.1	98.3
Proposed method	71.6	75.4	98.5	99.5

(The bold value is the highest value of each metric)

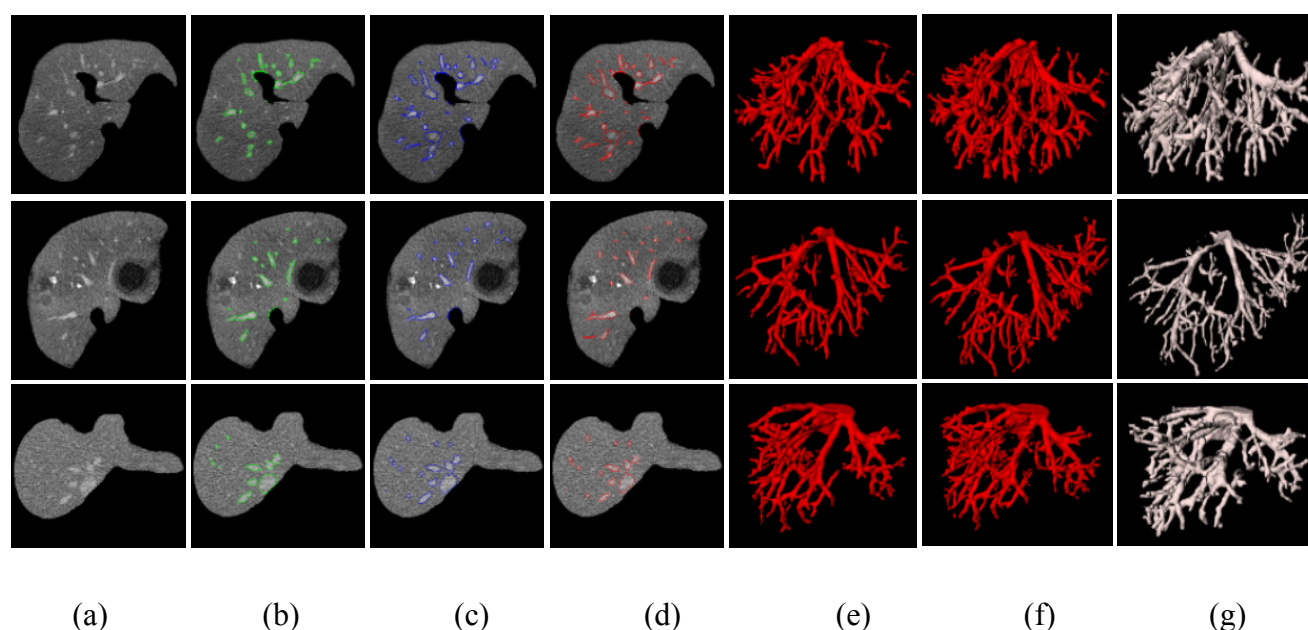


Figure 6. Examples of performances of the proposed method. (a) CT images; (b) (e) liver vessel slices and 3D visualization results using a combined loss function; (c) (f) liver vessel slices and 3D visualization results using the 3D deep supervision mechanism and the inter-scale dense connections; (d) (g) liver vessel slices and 3D visualization results of expert segmentation.

Kitrungrotsakul et al. [12] finally predicted the average dice value was 83%. Although the dice value is high, unlabeled liver vessels were not extracted in the results. Through experiments, we find that the proposed method can extract liver vessels unlabeled by experts, and these liver vessels have been recognized by experts, as shown in Figure 7. Therefore, our evaluation results are closer to the

clinical results rather than the comparison results based on incomplete annotations. The proposed method is proved to effectively and accurately extract liver vessels, which is used to replace the interactive segmentation of liver vessels in clinical practice and assist surgical planning through three-dimensional visualization. In the future, we will also verify the proposed method on more vascular datasets, such as aortic vessels [21].

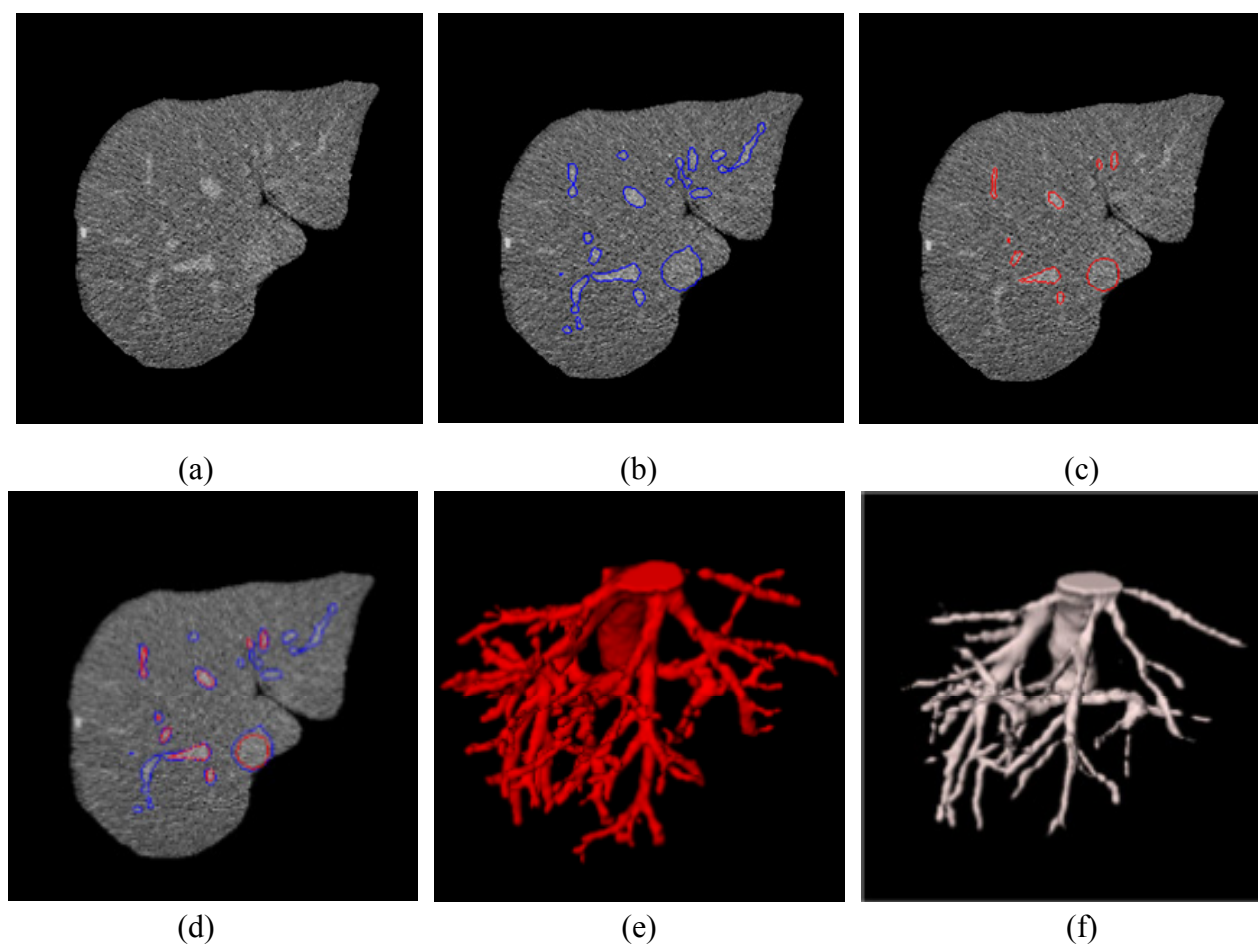


Figure 7. An example of unannotated liver vessel segmentation. (a) CT image; (b) (e) liver vessel slice and 3D visualization result using the proposed method; (d) combined liver vessel slice for (b) and (c); (c) (f) liver vessel slice and 3D visualization result of expert segmentation.

4. Conclusions

This paper proposes a method for automatically segmenting liver vessels from CT images based on an improved V-Net network. Rotation and mirroring operations are performed to augment the data. A combined loss function is utilized to improve the segmentation accuracy and sensitivity of liver vessels with unbalanced categories. The dilated convolution is introduced in the network encoder to increase the receptive field of the network in the case of reducing down-sampling. The 3D deep supervision mechanism is introduced into the network to speed up the network learning speed and improve the network's discrimination ability. Besides, inter-scale dense connections are designed into

the network, effectively integrating multi-scale feature information. The final experimental results show that all metrics have been significantly improved and have been recognized by experts. The algorithm can automatically and accurately segment liver vessels with complex structures and low contrast with surrounding tissues from CT images.

Acknowledgments

Supported by the Key Laboratory of the Ministry of Medical Imaging, Ministry of Education, China (80119008).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. W. Chen, R. Zheng, P. D. Baade, S. Zhang, H. Zeng, F. Bray, et al., Cancer statistics in China, 2015, *CA A Cancer J. Clin.*, **66** (2016), 115–132.
2. H. W. Huang, Influence of blood vessel on the thermal lesion formation during radiofrequency ablation for liver tumors, *Med. Phys.*, **40** (2013), 073303.
3. 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.
4. D. A. Oliveira, R. Q. Feitosa, M. M. Correia, Segmentation of liver, its vessels and lesions from CT images for surgical planning, *Biomed. Eng. Online*, **10** (2011), 30.
5. Y. Chi, J. Liu, S. K. Venkatesh, S. Huang, J. Zhou, Q. Tian, et al., Segmentation of liver vasculature from contrast enhanced CT images using context-based voting, *IEEE Trans. Biomed. Eng.*, **58** (2011), 2144–2153.
6. A. Foruzan, R. Zoroofi, Y. Sato, M. Hori, A Hessian-based filter for vascular segmentation of noisy hepatic CT scans, *Int. J. Comput. Assisted Radiol. Surg.*, **7** (2012), 199–205.
7. J. Li, M. Zhang, Y. Gao, Vessel segmentation of liver CT images by hessian-based enhancement, in *International Conference on Image and Graphics*, (2019), 442–445.
8. H. Zhang, P. Bai, X. Min, Q. Liu, Y. Ren, H. Li, et al., Hepatic vessel segmentation based on an improved 3D region growing algorithm, *J. Phys.*, **1486** (2020), 032038.
9. Y. Z. Zeng, Y. Q. Zhao, P. Tang, M. Liao, Y. X. Liang, S. H. Liao, et al., Liver vessel segmentation and identification based on oriented flux symmetry and graph cuts, *Comput. Methods Programs Biomed.*, **150** (2017), 31–39.
10. S. Cetin, G. Unal, A higher-order tensor vessel tractography for segmentation of vascular structures, *IEEE Trans. Med. Imaging*, **34** (2015), 2172–2185.
11. Y. Z. Zeng, S. H. Liao, P. Tang, Y. Q. Zhao, M. Liao, Y. Chen, et al., Automatic liver vessel segmentation using 3D region growing and hybrid active contour model, *Comput. Biol. Med.*, **97** (2018), 63–73.

12. T. Kitrungrotsakul, X. H. Han, Y. Iwamoto, A. H. Foruzan, L. Lin, Y. W. Chen, Robust hepatic vessel segmentation using multi deep convolution network, in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging. International Society for Optics and Photonics*, (2017), 1013711.
13. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, (2015), 234–241.
14. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: Learning dense volumetric segmentation from sparse annotation, in *Medical Image Computing and Computer-assisted Intervention–MICCAI 2016: 19th International Conference*, (2016), 424–432.
15. F. Milletari, N. Navab, S. A. Ahmadi, V-Net: Fully convolutional neural networks for volumetric medical image segmentation, in *Fourth International Conference on 3D Vision (3DV)*, (2016), 565–571.
16. W. Yu, B. Fang, Y. Liu, M. Gao, S. Zheng, Y. Wang, Liver vessels segmentation based on 3d residual U-NET, in *International Conference on Image Processing (ICIP)*, (2019), 250–254.
17. M. Xu, Y. Wang, Y. Chi, X. Hua, Training liver vessel segmentation deep neural networks on noisy labels from contrast CT imaging, in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, (2020), 1552–1555.
18. Q. Huang, J. Sun, H. Ding, X. Wang, G. Wang, Robust liver vessel extraction using 3D U-Net with variant dice loss function, *Comput. Biol. Med.*, **101** (2018), 153–162.
19. Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, P. A. Heng, 3D deeply supervised network for automatic liver segmentation from CT volumes, in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016-19th International Conference*, (2016), 149–157.
20. F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S. Kohl, et al., nnU-Net: Self-adapting framework for U-Net-based medical image segmentation, preprint, arXiv: 1809.10486.
21. A. Pepe, J. Li, M. R. Pissarczyk, C. Gsaxner, C. Xiaojun, G. A. Holzapfel, et al., Detection, segmentation, simulation and visualization of aortic dissections: A review, *Med. Image Anal.*, **65** (2020), 101773.



AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)