BEB Conference on Biomedical Engineering and Biotechnology

BEB6129

Liver vessel segmentation based on inter-scale V-Net

Jinzhu Yang[1,*], Meihan Fu[2] 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

Abstract

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.

Methods

- *Improvement of V-Net network framework*

Although the downsampling used by the V-Net network 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. A dropout layer is added at the end of the residual unit of each layer to prevent the network from overfitting.

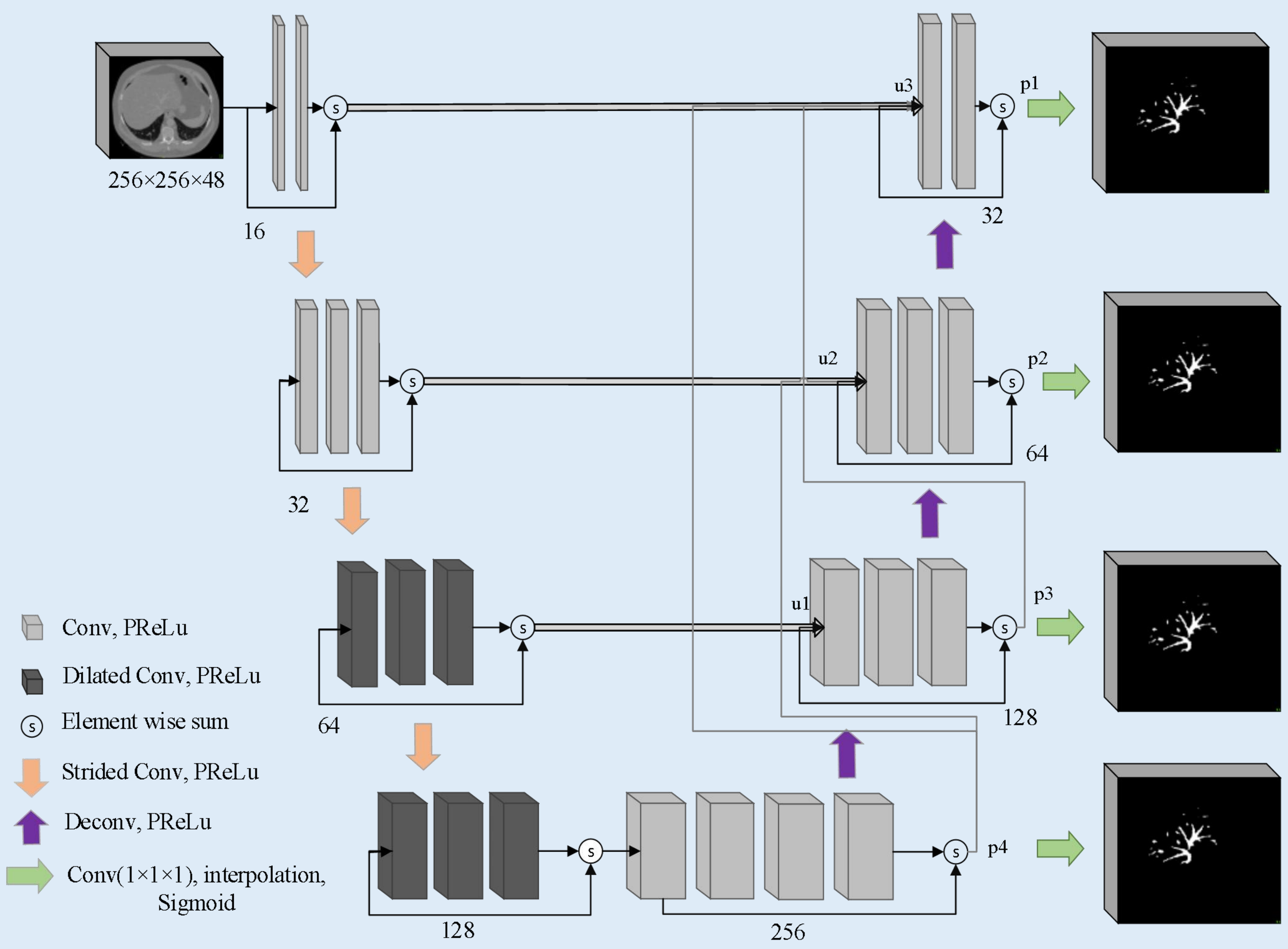

Figure 1. Schematic diagram of the overall structure of the improved V-Net network.

- *Inter-scale dense connections*

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→u3, p3→u3, p4→u2 is passed through a connection block (The connection block includes using trilinear interpolation for up-sampling and using 1 × 1 × 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 tion caused by operations such as up-sampling and multiple convolutions.

The inter-scale dense connections formula is as follows:

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

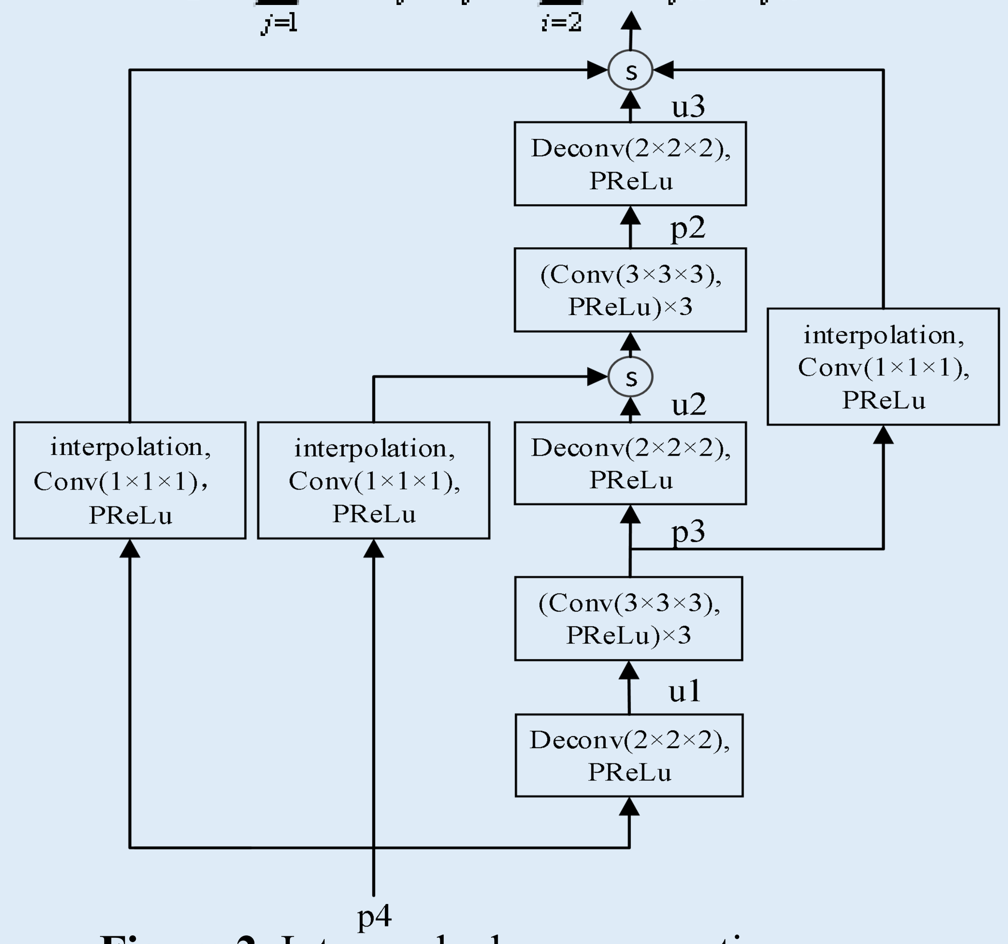

Figure 2. Inter-scale dense connections.

- *Loss function*

The combined loss function formula is as follows, where α is a weighting factor.

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

The binary cross-entropy loss function is as follows:

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

The dice loss function is as follows:

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

when α is 0.7, the performance of the network is good.

Evaluation and comparison

Table 1. 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

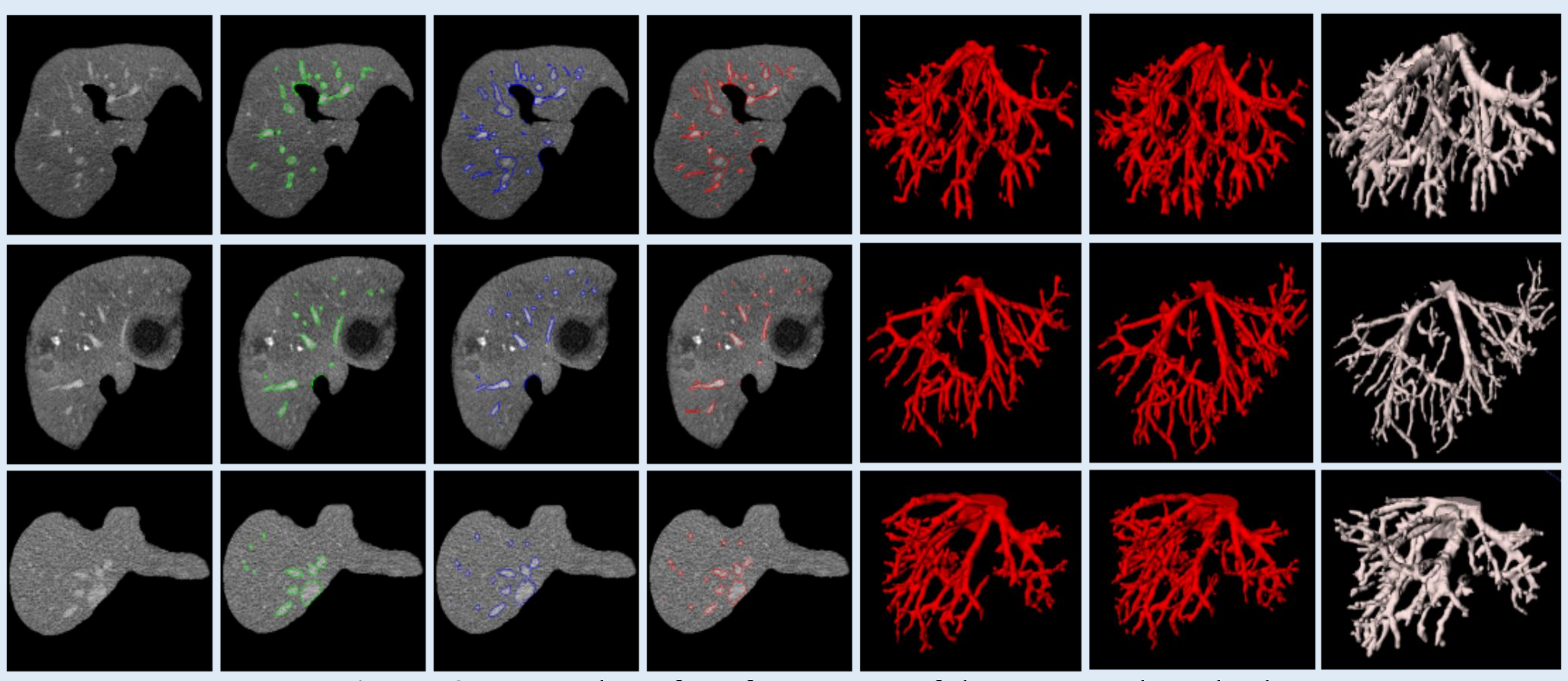

Figure 3. Examples of performances of the proposed method.

Conclusions

This paper proposes a method for automatically segmenting liver vessels from CT images based on an improved V-Net network. 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.