MLU-Net: Efficient Segmentation for Retinal Layers In Optical Coherence Tomography Images

Xiangcong Xu1,3, Xuehua Wang2,3*, Dingan Han2,3, Xuedong Guo1,3, Jingyi Lin2,3, Ke Xiong4, Yixu Zheng4, and Yaguang Zeng2,3

1School of Mechatronic Engineering and Automation, Foshan University, Foshan, Guangdong, 528225, China
2School of Physics and Optoelectronics, Foshan University, Foshan, Guangdong, 528225, China
3Guangdong-Hong Kong-Macao Joint Laboratory for Intelligent Micro-Nano Optoelectronic Technology, Foshan University, Foshan, Guangdong, 528225, China
4Department of Ophthalmology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, 510515, China

*Corresponding author’s e-mail: wangxuehua10000@fosu.edu.cn

Abstract. The automatic segmentation of retinal images obtained by optical coherence tomography is increasingly important for ophthalmologists to diagnose and monitor many kinds of ophthalmic diseases. U-Net is the most widely used deep learning network in retinal segmentation, but the limited number of data-flow paths made it hard to capture complex features. We proposed here an optimized Mobile Ladder U-Net (MLU-Net), which consists of a Ladder Connection for increasing the network’s data-flow paths and a depthwise separable convolution for reducing the model’s parameters. Experiments on 100 B-scans from 10 human eyes demonstrated that the 9 retinal layer boundaries can be segmented accurately with the MLU-Net. In addition, compared with the original U-Net and LadderNet, the segmentation result of our method is closest to the expert label.

1. Introduction

Optical Coherence Tomography (OCT) is a widely used imaging technology that can image various biological tissue [1] . Because of its high sensitivity and noninvasive detection, it has important application value in the field of clinical medicine. In recent years, OCT has been applied in the field of fundus imaging to characterize retinal images and their respective layer structures in detail to examine and monitor ocular diseases such as Age-related Macular Degeneration (AMD), Diabetic Retinopathy (DR), and glaucoma [2] .

To date, a great deal of work has been done in developing segmentation algorithms for retinal multilayer structures in OCT images. Mujat et al [3] used anisotropic noise suppression operation and deformation spline to segment the NFL layer in SD-OCT images, but it is relatively sensitive to vascular artifacts in OCT images. Yazdanpanah et al [4] used Chan-Veses energy-minimized active contour to segment the retinal inner layer, but boundary constraints would lead to errors in the irregular layers. Novosel et al [5] segmented the three retinal layers of the SD-OCT retinal image using level set
Bayesian inference, but this method also sensitive to vascular shadows. Guo et al [6] proposed guided bi-directional graph search to retinal segmentation, but it is time-consuming.

Machine learning is also one of the research hotspots of retinal segmentation. Fang et al [7] predicted the class of the center pixel of a given image patch using the deep network, and then segmented the boundary using the graph search method. Hamwood et al [8] optimized the algorithm and compared the segmentation performance of normal retinal images with different patch sizes and networks. U-Net [9] is the most widely used structure in medical image analysis because the skip connection allows efficient information flow and performs well without sufficiently large datasets. However, the number of paths U-Net uses for information flow is limited, resulting in insufficient extraction of some complex feature information.

Therefore, we proposed Mobile Ladder U-Net (MLU-Net) which uses the Ladder Connection to connect two U-Nets to increase the data-flow paths of the network and enhance the network’s ability of capturing more complex features to improve the segmentation accuracy. In the contrast mutation region, the segmentation result is more accurate. In the network, we replace the standard convolution with the depthwise separable convolution, which can reduce the number of parameters of the model and training difficulty. Compared with expert labels, the proposed MLU-Net can segment the retinal layer boundary accurately.

2. Method

2.1. Neural network model
In this paper, we proposed Mobile Ladder U-Net (MLU-Net) for retinal segmentation, as shown in Fig. 1. It is a modified network based on U-Net, which has a limited number of paths for information flow. MLU-Net used the Ladder Connection to increase the network’s data-flow paths, and replaced the standard convolution with the depthwise separable convolution to reduce the number of parameters.

2.2. Ladder Connection
The Ladder Connection uses the skip connections to connect two U-Nets. Unlike U-Net, which concatenates features from the encoder branch with features from the decoder branch, we concatenate features from both networks, which adds more U-Nets to form a complex network structure. It allows the total number of network paths to grow exponentially with the number of codec pairs and the number of space levels, making it possible for the network to capture more complex features to produce higher accuracy.
2.3. Depthwise Separable Convolution
The Ladder Connection adds branches of the network which makes the number of parameters of the model larger and the training is more difficult. We used the depthwise separable convolution \[10\] to replace standard convolution to reduce the amount of parameters. Standard convolution is to filter the input and then combine it into a new output. Depthwise separable convolution is a method of decomposing convolution into depthwise convolution and point-by-point convolution. The former is used for filtering and the latter is used for splicing. The deep convolution uses a separate filter for each input, and the point-by-point convolution is to splice the output and the deep convolution with \(1 \times 1\) convolution. This decomposition method can greatly reduce the parameters of the model.

3. Experiment

3.1. Dataset
We collected retinal data from 10 samples under the Topcon fundus tomography system 3D OCT-1. Each sample contains 128 B-scans with a resolution of \(784 \times 1024\). Randomly select 30 B-scans from each sample as the data set, of which 200 are used as the training set and 100 are used as the prediction set. To increase the number of training samples, we cut each B-scan into 15 patches with a size of \(128 \times 128\) that only contain the retina area. Finally, 20% of the training samples were used as validation data and the rest as training data.

3.2. Training of the neural network
To solve the problem of limited data, we use data augmentation (e.g. horizontal flip, rotation, scale). In the training phase, we use cross-entropy loss and Adam optimizer. The learning rate was set to \(1 \times 10^{-5}\), the batchsize of training set was set to 8, the batchsize of validation set was set to 2, and the number of iterations was set to 100.

3.3. Evaluation strategy
In the experiment, we used two indexes to evaluate the performance of MLU-Net, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). \(L\) is the size of the image, \(pre\) is the prediction result, \(true\) is the ground label, the formula is as follows:

\[
\text{MAE} = \frac{1}{L} \sum_{i=1}^{L} |pre(i) - true(i)|
\]

\[
\text{RMSE} = \sqrt{\frac{1}{L} \sum_{i=1}^{L} (pre(i) - true(i))^2}
\]

4. Result
By comparing other different networks, the performance of the MLU-Net network in retinal segmentation is verified. The prediction results of U-Net [9], LadderNet [11], and MLU-Net are shown in Fig. 2. The background of Fig. 2(a), (c), (e) is the original B-scan, and the background of Fig. 2(b), (d), (f) is the expert label. We used the red dashed box to mark the obvious difference areas on the figure with the expert label as the background, which makes us see the contrast effect among different networks more clearly. Through the comparison, it is found that U-Net is prone to error in the contrast mutation region because of the insufficient characteristics of its learning. However, LadderNet and MLU-Net use the Ladder Connection to extract more features, which can enhance the network discriminative ability and make the predicted result more similar to the expert label.
Figure 2. Retinal segmentation map of U-net, LadderNet, and MLU-Net. (a) and (b) are the segmentation results of U-Net on the original B-scan and expert label respectively; (c) and (d) are the segmentation results of LadderNet on the original B-scan and expert label respectively; (e) and (f) are the segmentation results of MLU-Net on the original B-scan and expert label respectively.

Table 1 shows the segmentation results of different networks compared with the expert label. It can be seen that LadderNet and MLU-Net are significantly better than U-Net. We showed the parameters of different networks in Table 2. The parameters of MLU-Net are only half of LadderNet, because MLU-Net uses depthwise separable convolution instead of standard convolution, which greatly reduces model parameters and training difficulty, while maintaining the original segmentation results. As shown in Fig. 2 and Table 1, the segmentation results of MLU-Net are consistent with that of LadderNet.

Table 1. MAE and RMSE Values for U-Net, Ladder Net, MLU-Net Compared to Expert Labels

| Boundary | MAE          | RMSE         |
|----------|--------------|--------------|
|          | U-Net | LadderNet | MLU-Net | U-Net | LadderNet | MLU-Net |
| ILM      | 2.040  | 1.587     | 1.509    | 2.660  | 1.808     | 1.306   |
| RNFL-GCL | 2.530  | 1.703     | 1.636    | 3.230  | 2.028     | 2.070   |
| IPL-INL  | 1.790  | 1.112     | 1.052    | 2.410  | 1.884     | 1.685   |
| INL-OPL  | 1.840  | 0.777     | 0.726    | 2.250  | 1.386     | 1.006   |
| OPL-ONL  | 1.430  | 0.947     | 0.876    | 1.860  | 1.396     | 1.225   |
| ELM      | 0.810  | 0.771     | 0.698    | 1.060  | 0.905     | 0.848   |
| IS-OS    | 0.840  | 0.671     | 0.698    | 1.220  | 0.905     | 0.848   |
| OS-RPE   | 1.840  | 0.447     | 0.475    | 2.270  | 1.669     | 1.689   |
| BM       | 2.040  | 1.285     | 1.277    | 2.660  | 1.552     | 1.532   |

Table 2. Parameters of different models

| Model    | Params |
|----------|--------|
| U-Net    | 359.3M |
| LadderNet | 756.0M |
| MLU-Net  | 400.0M |
5. Conclusion

We proposed Mobile Ladder U-Net (MLU-Net) for the retinal segmentation. The Ladder Connection is used to connect two U-Nets to increase the network’s data-flow paths, thus enhancing the ability to capture more complex features and improving the accuracy. The model parameters and the training difficulty are reduced by replacing the standard convolution with the depthwise separable convolution. We tested MLU-Net on 100 B-scans from 10 human eyes, which attested that the 9 retinal layer boundaries can be segment accurately. Compared with U-Net and LadderNet, our proposed method has less model parameters and a segmentation result that is closest to the expert label especially in the contrast mutation region.

Reference

[1] Huang D, Swanson E A, Lin C P, et al. Optical coherence tomography [J]. science, 1991, 254(5035): 1178-1181.
[2] Bhagat N, Grigorian R A, Tutela A, et al. Diabetic macular edema: pathogenesis and treatment [J]. Survey of ophthalmology, 2009, 54(1): 1-32.
[3] Mujat M, Chan R C, Cense B, et al. Retinal nerve fiber layer thickness map determined from optical coherence tomography images [J]. Optics Express, 2005, 13(23): 9480-9491.
[4] Yazdanpanah A, Hamarneh G, Smith B R, et al. Segmentation of intra-retinal layers from optical coherence tomography images using an active contour approach [J]. IEEE transactions on medical imaging, 2010, 30(2): 484-496.
[5] Novosel J, Vermeer K A, Thepass G, et al. Loosely coupled level sets for retinal layer segmentation in optical coherence tomography [C]//2013 IEEE 10th International Symposium on Biomedical Imaging. IEEE, 2013: 1010-1013.
[6] Guo Y, Camino A, Zhang M, et al. Automated segmentation of retinal layer boundaries and capillary plexuses in wide-field optical coherence tomographic angiography [J]. Biomedical optics express, 2018, 9(9): 4429-4442.
[7] Fang L, Cunefare D, Wang C, et al. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search [J]. Biomedical optics express, 2017, 8(5): 2732-2744.
[8] Hamwood J, Alonso-Caneiro D, Read S A, et al. Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers [J]. Biomedical optics express, 2018, 9(7): 3049-3066.
[9] Hu J, Shen L, Sun G, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, PP(99).
[10] Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
[11] Zhuang J. Laddernet: Multi-path networks based on u-net for medical image segmentation[J]. arXiv preprint arXiv:1810.07810, 2018.