PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Segmentation of common lesions in age-related macular degeneration based on boundary attention mechanism

Chengjian Hong, Yuanyuan Peng, Weifang Zhu, Dehui Xiang, Fei Shi, et al.

Chengjian Hong, Yuanyuan Peng, Weifang Zhu, Dehui Xiang, Fei Shi, Baoqing Nie, Xinjian Chen, "Segmentation of common lesions in age-related macular degeneration based on boundary attention mechanism," Proc. SPIE 12464, Medical Imaging 2023: Image Processing, 124643U (3 April 2023); doi: 10.1117/12.2653863

SPIE.

Event: SPIE Medical Imaging, 2023, San Diego, California, United States

Segmentation of Common Lesions in Age-related Macular Degeneration Based on Boundary Attention Mechanism

Chengjian Hong[1], Yuanyuan Peng[1], WeiFang Zhu[1], Dehui Xiang[1], Fei Shi[1], Baoqing Nie[1], Xinjian Chen[1,2,*]

[1]School of Electronics and Information Engineering, Soochow University, Suzhou, Jiangsu Province, 215006, China

[2]State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, 215123, China

*indicates the corresponding author, E-mail: xjchen@suda.edu.cn

ABSTRACT

Age-related macular degeneration (AMD) is a common ophthalmic disease, mainly occurring in the elderly. After the occurrence of pigment epithelial detachment (PED), neuroepithelial detachment and subretinal fluid (SRF) are further caused, and patients need follow-up treatment. Quantitative analysis of these two symptoms is very important for clinical diagnosis. Therefore, we propose a new joint segmentation network to accurately segment PED and SRF in this paper. Our main contributions are: (1) a new multi-scale information selection module is proposed. (2) based on the U-shape network, a novel decoder branch is proposed to obtain boundary information, which is critical to segmentation. The experimental results show that our method achieves 72.97% for the average dice (DSC), 79.92% for the average recall, and 67.11% for the average intersection over union (IOU).

KEYWORDS: AMD，PED&SRF, Neural Networks, Multi-scale, Boundary information, OCT, Segment ation.

1. INTRODUCTION

Age related macular degeneration is a very common ophthalmic disease, mostly in the elderly, and its occurrence is related to some environmental and genetic factors. The disease often occurs when the fluid in retinal capillaries leaks into the retina and accumulates in the retina. Once retinal neuroepithelial detachment leads to pigment epithelial detachment (PED) and subretinal fluid (SRF), treatment is needed.

In recent years, many related works have been proposed for AMD fluid segmentation. For example, Roy et al. proposed RelayNet[1] to segment retinal layer and fluid simultaneously on normal and abnormal OCT images, which takes the accurate segmentation of retinal layer as a priori knowledge to realize the accurate

Medical Imaging 2023: Image Processing, edited by Olivier Colliot, Ivana Išgum, Proc. of SPIE Vol. 12464, 124643U · © 2023 SPIE
1605-7422 · doi: 10.1117/12.2653863

segmentation of fluid. In addition, in the RETOUCH challenge of MICCAI2017, the RMIT[2] team has achieved good performance based on modified UNet by introducing the adversarial network. However, when carefully observing the characteristics of PED and SRF, we found that the boundary of these two fluids is not clear, and the size is relatively variable, as shown in Figure 1. Therefore, the above methods cannot solve these problems in essence. To alleviate the above problems, we design a branch of boundary detection and a multi-scale information selection （MSIS）module based on the U-shape network to achieve the accurate join segmentation of PED and SRF.

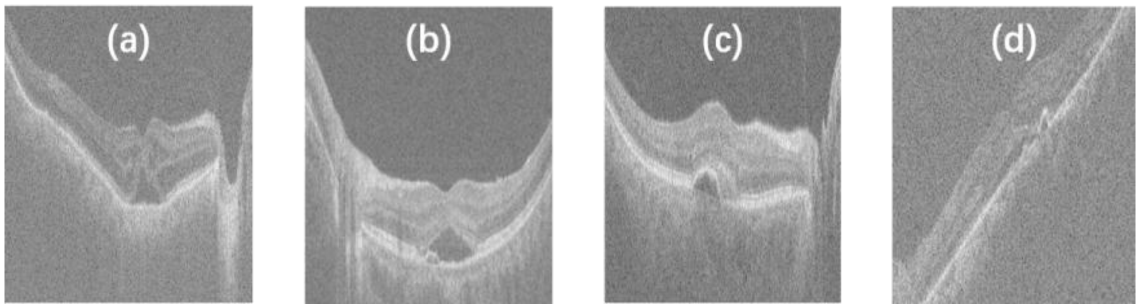

Figure 1. Four OCT images with PED and SRF

2. METHODS

2.1 Network structure

Inspired by reverse attention (RA)[3] shown in Figure 2, we design a novel boundary decoder branch to preserve the edge information as much as possible. On the one hand, we integrate the boundary information generated by the difference between the high-level features and the low-level features into a separate branch to generate the image boundaries, and realize the boundary supervision. On the other hand, we converge the boundary information into the main branches of segmentation to realize the fusion of boundary information. In addition, inspired by SKNet[4] and to preserve the multi-scale information of SRF and PED, we develop multi-scale information selection (MSIS) module to extract multi-scale information, which can focus on different multi-scale information and ensure the fusion of different scale information on each channel. Finally, the improved multi-scale semantic information is obtained by fusing the multi-scale spatial feature maps generated by a common concatenation operator. The proposed MSIS module and the overall architecture of our network are shown in Figure 3 and Figure 4 respectively, where the middle branch is the branch that generates the boundary information.

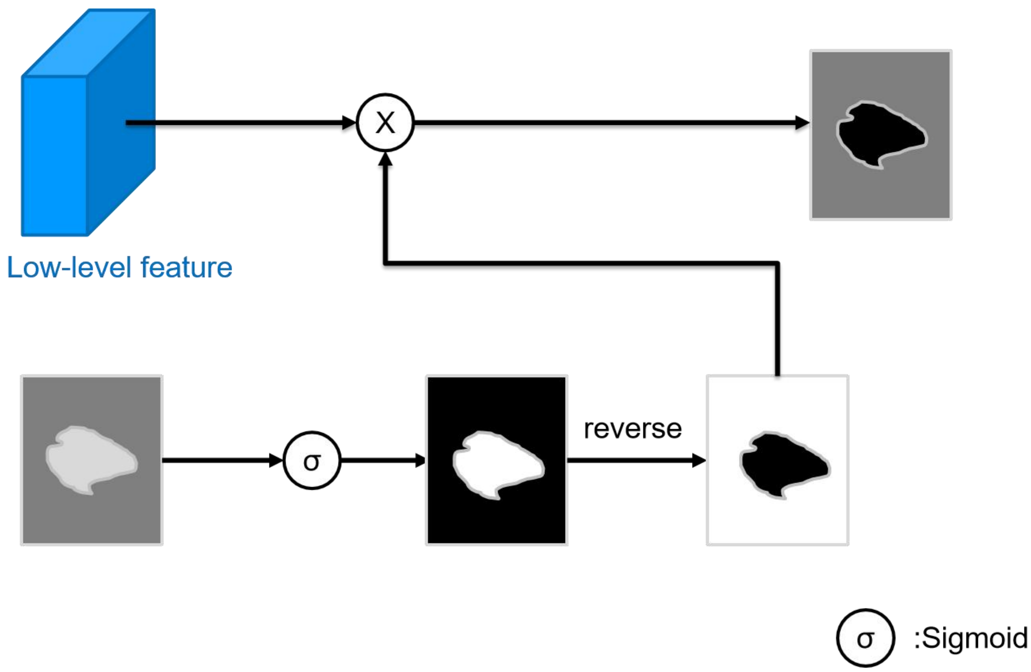

Figure 2. Reverse Attention (RA)

2.2 Training and testing

We train the model with back-propagation algorithm by minimizing the cross-entropy cost function and dice loss function. Adam is used as the optimizer, where initial learning rate is set to 0.0001. The batch size and epoch are set to 8 and 200, respectively. During training, all networks are trained with identical optimization schemes and we save the best model on validation set.

3. RESULTS

3.1 Preprocessing

the OCT B-scans were acquired using BV1000 from the BigVisionTech, Suzhou, China, which were labeled by two professional annotators within the company. All the annotation work was conducted by the two annotators independently, and only the data with relatively consistent results were used to evaluate the proposed method. After labeling, a total of 638 OCT B-scans (638×1024×512) from 219 human eyes were obtained, where we used 439 of them as the training set, 89 as the validation set, and 110 as the testing set. To save computational resources, all OCT images are resized to 512×256 by bilinear interpolation. In addition, to overcome the overfitting problem, we perform a series of data enhancement operations such as random rotation, Gaussian additive noise, and contrast enhancement.

3.2 Parameter settings

The implementation of the proposed network is based on the public platform PyTorch and a NVIDIA

RTX3090 GPU with 24GB memory. The model is trained with back-propagation algorithm by minimizing the joint loss function L_{joint}, which is an combination of the cross-entropy loss L_{ce} and Dice loss L_{dice}. The overall loss function is defined as follow:

$$L_{joint} = L_{ce} + L_{dice} \tag{1}$$

where,

$$L_{ce} = \frac{1}{N}\sum_i \sum_{c=1}^{M} y_{ic} \log p_{ic} \tag{2}$$

$$L_{dice} = 1 - \frac{1}{N}\sum_{i=1}^{N} \frac{2y_i p_i}{y_i^2 + p_i^2} \tag{3}$$

where the y_{ic} is a symbolic function indicating that the value of channel c is 1 if it is the i-th pixel, otherwise it is 0. p_{ic} is the output of the model prediction, which is normalized by softmax. N represents the total number of pixels in the output feature map. y_i represents the taregt pixel value of a particular lesion at the i-th location on that channel, and p_i represents the predicted value of the model.

3.3 Results

To verify the effectiveness of our method, we compare the proposed network with other excellent CNN based segmentation networks, such as PSPNet[5], DeepLabv3[6], and CE-Net[7]. For convenience, we call the original UNet[8] as Backbone. Table 1 shows the quantitative results of different methods.

As shown in Table 1, compared to Backbone, the proposed method improves the average DSC, Recall, and IOU by 5.94%, 2.94%, and 5.90%, respectively. Then, compared to other state-of-the-art segmentation networks, the performance of our method also gets an overall improvement in terms of all metrics. For example, compared with the network with suboptimal segmentation performance (DeepLabv3), the proposed method improves the segmentation performance of PED and SRF, and achieves 72.97% for the average DSC, 79.92% for the average Recall, and 67.11% for the averaged IOU.

In addition, to evaluate the effectiveness of the proposed MSIS module and the addition of boundary information extraction branch (BEB), a series of ablation studies are conducted. As shown in Table 1, compared to Backbone, Backbone with MSIS module (Backbone+MSIS) improves the main segmentation metrics DSC and IOU. Table 1 also shows the effectiveness of the addition of boundary information extraction branch. Compared to Backbone, Backbone with boundary information extraction branch (Backbone+BEB) improves the average DSC, Recall and IOU by 3.38%, 2.40% and 4.21%, respectively.

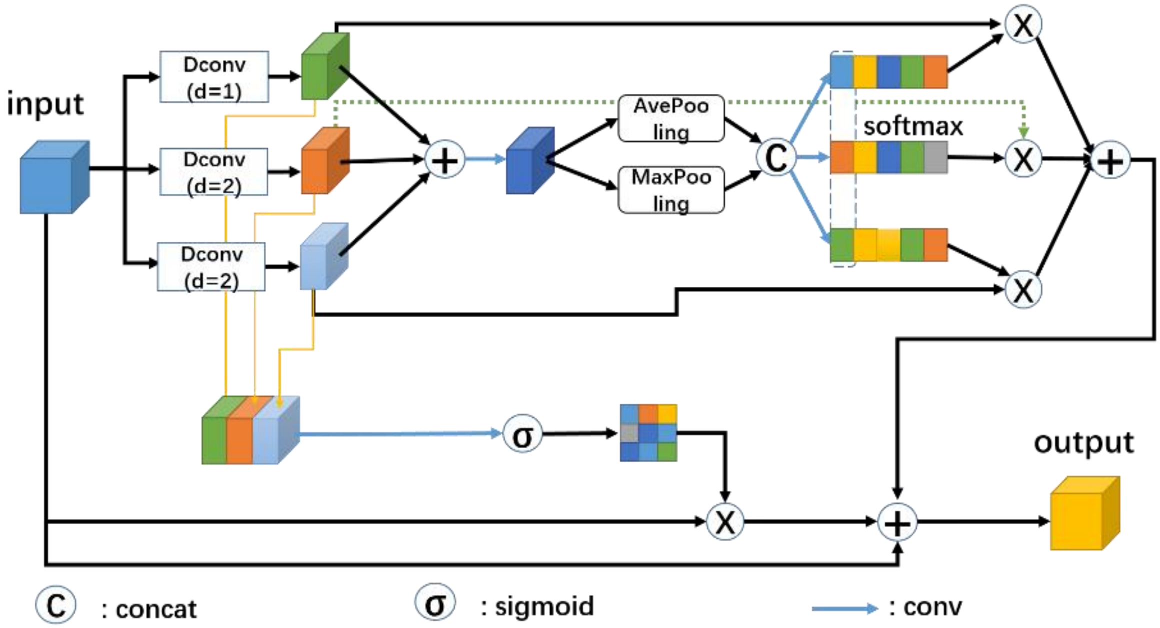

Figure 3. Multi-scale Information Selection Module (MSIS)

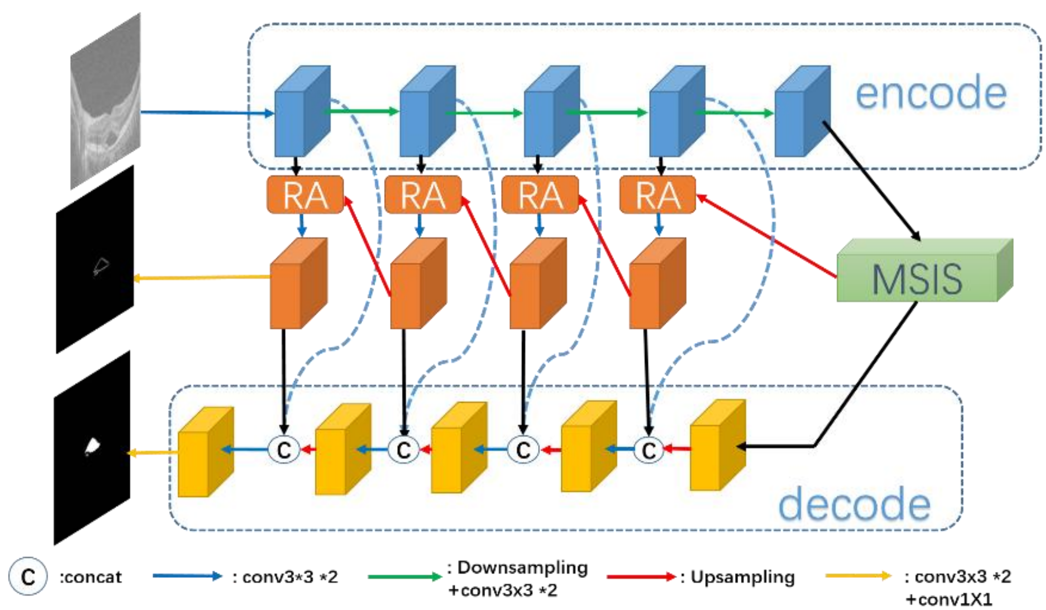

Figure 4. The whole architecture of our network.

Table 1 Segmentation Results

	DSC (%)			Recall (%)			IOU (%)		
Method	SRF	PED	Ave	SRF	PED	Ave	SRF	PED	Ave
CE-Net [7]	67.84	67.77	67.80	68.53	71.43	69.98	63.40	61.96	62.68
DeepLabv3[6]	69.24	67.36	68.30	71.77	72.70	72.23	63.37	61.47	62.42
PSPNet [5]	65.82	58.69	62.26	70.66	63.30	66.98	60.81	53.87	57.34
Backbone	71.43	66.33	68.88	73.46	**81.82**	77.64	66.82	59.92	63.37
Backbone+MSIS	74.35	67.70	71.03	**79.28**	68.98	74.13	68.27	63.00	65.63
Backbone+BEB	73.66	68.76	71.21	76.52	75.29	75.91	68.50	63.58	66.04
proposed	**75.79**	**70.16**	**72.97**	79.11	80.73	**79.92**	**70.45**	**63.76**	**67.11**

It is worth noting that we also show some segmentation results for qualitative analysis as shown in Figure 5. It can be observed from Figure 5 that compared with other segmentation networks, our network obtains the best results in preserving the boundary information and the segmentation of images with simultaneous appearance of large and small targets.

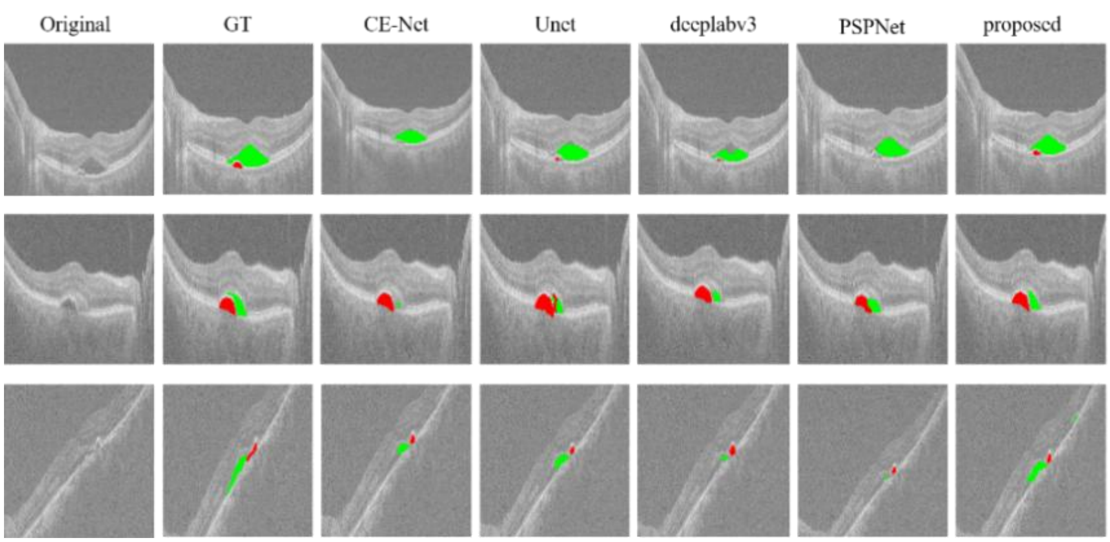

Figure 5. The segmentation results of different methods, where the green area and the red area indicate SRF and PED, respectively.

4. CONCLUSIONS

In this paper, we proposed the MSIS module and BEB to improve the UNet and achieve accurate PDE and SRF segmentation. On the one hand, the introduction of MSIS module in UNet for multi-scale filtering of features can ensure the maximum retention of features when the size of segmented objects changes unevenly. One the other hand, when PED and SRF frequently have uneven and unsmooth boundaries, the proposed BEB can help the network better learn additional boundary features, and integrate them into the final segmentation to achieve performance improvement.. The preliminary results demonstrate the effectiveness of the proposed method.

5. ACKNOWLEDGEMENTS

This study was supported in part by the National Key R&D Program of China (2018YFA0701700) and part by the National Nature Science Foundation of China (U20A20170，61622114，62271337 and 61971298).

6. REFERENCE

[1] Roy A G, Conjeti S, Karri S P K, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks[J]. Biomedical optics express, 2017, 8(8): 3627-3642.

[2] Tennakoon R, Gostar A K, Hoseinnezhad R, et al. Retinal fluid segmentation and classification in OCT images using adversarial loss based CNN[J]. Proc. MICCAI Retinal OCT Fluid Challenge (RETOUCH), 2017: 30-37.

[3] Fan D P, Ji G P, Zhou T, et al. Pranet: Parallel reverse attention network for polyp segmentation[C]//International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2020: 263-273.

[4] Li X, Wang W, Hu X, et al. Selective kernel networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 510-519.

[5] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2881-2890.

[6] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 801-818.

[7] Gu, Zaiwang, et al. "Ce-net: Context encoder network for 2d medical image segmentation." *IEEE transactions on medical imaging* 38.10 (2019): 2281-2292.

[8] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.