PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Attention multi-scale network for pigment epithelial detachment segmentation in OCT images

Bao, Dengsen, Cheng, Xuena, Zhu, Weifang, Shi, Fei, Chen, Xinjian

Dengsen Bao, Xuena Cheng, Weifang Zhu, Fei Shi, Xinjian Chen, "Attention multi-scale network for pigment epithelial detachment segmentation in OCT images," Proc. SPIE 11313, Medical Imaging 2020: Image Processing, 1131335 (10 March 2020); doi: 10.1117/12.2548959

SPIE.

Event: SPIE Medical Imaging, 2020, Houston, Texas, United States

Attention Multi-Scale Network for Pigment Epithelial Detachment Segmentation in OCT Images

Dengsen Bao[1], Xuena Cheng[1], Weifang Zhu[1,2], Fei Shi[1], Xinjian Chen[1,3,*]

[1]School of Electronics and Information Engineering, Soochow University, Suzhou, Jiangsu Province, 215006, China
[2]Collaborative Innovation Center of IoT Industrialization and Intelligent Production, Minjiang University, Fuzhou 350108, China
[3]State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou 215123, China

ABSTRACT

Accurate segmentation of pigment epithelial detachment (PED) in retinal optical coherence tomography (OCT) images can help doctors comprehensively analyze and diagnose chorioretinal diseases, such as age-related macular degeneration (AMD), central serous chorioretinopathy and polypoidal choroidal vasculopathy. Due to the serious uneven sizes of PED, some traditional algorithms or common deep networks do not perform well in PED segmentation. In this paper, we propose a novel attention multi-scale network (named as AM-Net) based on a U-shape network to segment PED in OCT images. Compared with the original U-Net, there are two main improvements in the proposed method: (1) Designing channel multi-scale module (CMM) to replace the skip-connection layer of the U-Net, which uses channel attention mechanism to obtain multi-scale information. (2) Designing spatial multi-scale module (SMM) based on dilated convolution, which is inserted in the decoder path to make the network pay more attention on the multi-scale spatial information. We evaluated the proposed AM-Net on 240 clinically obtained OCT B-scans with 4-fold cross validation. The mean and standard deviation of Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Sensitivity (Sen) and Specificity (Spe) are 72.12±9.60%, 79.17±8.25%, 93.05±1.72% and 79.93±5.77%, respectively.

KEYWORDS: Pigment epithelial detachment, optical coherence tomography, attention multi-scale network, channel multi-scale module, spatial multi-scale module

1. INTRODUCTION

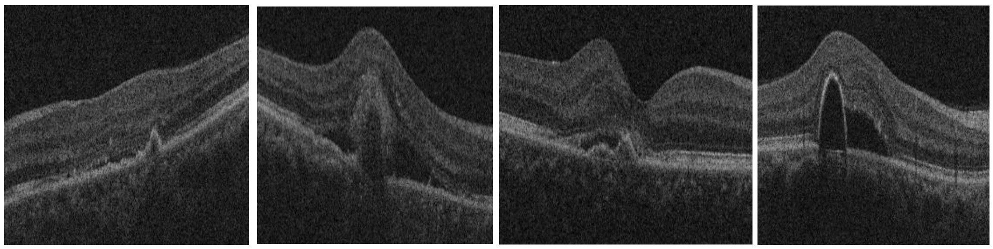
Pigment epithelial detachment (PED) is an important feature of several chorioretinal diseases, such as age-related macular degeneration (AMD), central serous chorioretinopathy and polypoidal choroidal vasculopathy[1,2]. Severe PED will cause damage to central vision finally[3]. Precise and efficient segmentation of PED is necessary for clinical diagnosis of these eye diseases.

In recent years, many methods based on convolutional neural networks (CNN) have been proposed for PED segmentation. Ronneberger et al. proposed a convolutional network named as U-Net for biomedical image segmentation[4]. U-Net and some of its variant networks have also been widely used in the target segmentation of ophthalmic imaging[5,6,7,8]. However, these networks may not be suitable for PED segmentation due to two main challenges as shown in Fig.1. First, the sizes of the PED vary in a wide range, which cause many segmentation algorithms do not perform well. Second, the segmentation performance is easily affected by the blurred boundaries, various shapes and random positions of PED. To overcome these problems, we propose a novel U-shape network to segment PED. We design two modules including channelmulti-scale

*Corresponding author: E-mail:xjchen@suda.edu.cn

Medical Imaging 2020: Image Processing, edited by Ivana Išgum, Bennett A. Landman,
Proc. of SPIE Vol. 11313, 1131335 · © 2020 SPIE · CCC
code: 1605-7422/20/$21 · doi: 10.1117/12.2548959

module (CMM) and spatial multi-scale module (SMM) to capture the multi-scale information in two different ways. Last but not least, we adopt data augmentation strategies, including random flips, random rotations, and Gaussian noise addition during training process, to increase the generalization of the model. Experiment evaluations show a great improvement compared with other networks.



(a) (b) (c) (d)

Fig. 1. Examples of PEDs in original OCT B-scans with variant shapes and sizes and blurred boundaries

2. METHODS

In this section, the overall network structure is firstly given with explanations on how the two modules work for our PED segmentation task. Subsequently, the channel multi-scale module (CMM) and spatial multi-scale module (SMM) are introduced in detail. Finally, the joint loss is described.

2.1 Overall structure of the AM-Net

With CMM and SMM, we propose our attention multi-scale network for PED segmentation as shown in Fig.2. The proposed method is based on U-Net, which is widely used in medical image segmentation[9]. We replace the skip-connection of the U-Net layer with CMM, which can use the shallow channel attention of the network to discard redundant information and obtain multi-scale information. At the same time, we insert SMM into the decoder path to make the network pay more attention on the multi-scale spatial information. With CMM and SMM, the proposed network can not only extract more useful multi-scale information from the encoding path, but also restore the PED segmentation more accurately from the decoder path.

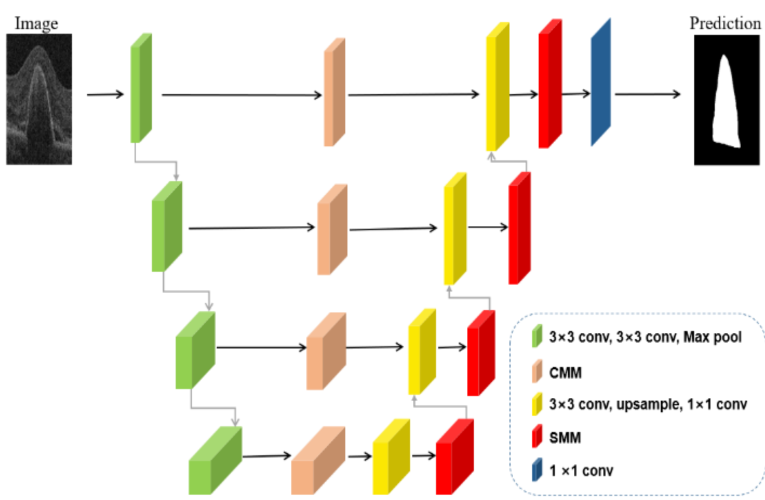


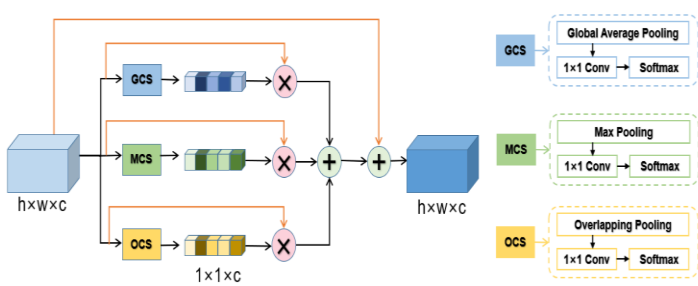
Fig. 2. An overview of the proposed attention multi-scale network (AM-Net).

2.2 Channel multi-scale module

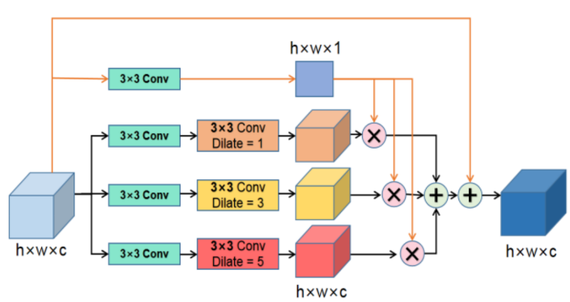
Fig.3 (a) shows the proposed channel multi-scale module (CMM). The input feature maps (h×w×c) go through a series of operations (pooling, 1×1 convolution and softmax excitation) three times to obtain 3 groups of channel information, which are taken as channel weights and multiplied by the input feature map respectively. The greater the channel weight is, the greater the feature map contributes to the final prediction result. The purpose of repeating this operation 3 times is to use multiple channels to increase the weight of the effective channel and reduce the weight of the invalid channel. To obtain multi-scale channel information, global average pooling, max pooling and overlapping pooling are applied during these three operations, respectively. The ablation experiment proves that corresponding improvement is obvious. Then, we make fusion on the three groups of the feature maps with channel information, and finally do the residual operation with the input feature map to obtain the final output of this module. The function of this module is to discard redundant information and guide the model to focus on the useful channel information. What's more, the module can obtain multi-scale information by three different pooling methods to solve the problem of the inconsistent sizes of PED.

2.3 Spatial multi-scale module

Fig.3 (b) shows the proposed spatial multi-scale module (SMM). The input feature maps (h×w×c) go through 3×3 convolution and dilated convolution with dilated ratio of 1, 3, 5, respectively, to obtain 3 groups of feature maps with multi-scale information. The dilated convolution here is another method to solve the problem of inconsistent sizes of target. The dilated convolution with dilated ratio of 1 focuses on the small target, the one with ratio of 5 focuses on the large target, and the medium area is identified by convolution with dilated ratio of 3. At the same time, the input feature maps go through a 3×3 convolution to obtain spatial information. Then the spatial feature map multiplies by the obtained three groups of feature maps respectively to further guide the optimization of multi-scale information. Finally, we do the residual operation with the input feature map to obtain the final output of this module. The function of this module is to use spatial information to guide the dilated convolution to obtain multi-scale receptive fields, thereby solving the problems of the inconsistent sizes of PED.



(a)



(b)

Fig. 3. Components of the two proposed modules. (a) Channel multi-scale module (CMM). (b) Spatial multi-scale module (SMM).

2.4 Loss function

To solve the problem of data imbalance and the various shapes and sizes of PED, we use the combination of the Dice loss and the Binary Cross-Entropy (BCE) loss as the joint loss, which is described as:

$$L_{Dice} = 1 - \frac{2\sum_{i}^{N} \bar{y}_{(k,i)} y_{(k,i)} + \varepsilon}{\sum_{i}^{N} \bar{y}_{(k,i)} + \sum_{i}^{N} y_{(k,i)} + \varepsilon} \tag{1}$$

$$L_{BCE} = -\frac{1}{N}\sum_{i}^{N}\left(y_{(k,i)}\log(\bar{y}_{(k,i)}) + (1-y_{(k,i)})\log(1-\bar{y}_{(k,i)})\right) \tag{2}$$

$$L_{Total} = L_{Dice} + L_{BCE} \tag{3}$$

Where N indicates the batch size, $\bar{y}_i \in [0,1]$ and $y_i \in \{0,1\}$ denote the predicted probability and ground truth label respectively. ε is a small smoothing factor.

3. RESULTS

3.1 Dataset

The evaluation dataset contains 240 retinal OCT B-scans from 60 patients (4 B-scans are randomly selected from each patient), which are labeled under the supervision of 2 senior ophthalmologists. The dataset is randomly divided into 4 folds according to subjects and 4-fold cross-validation strategy is adopted to evaluate the performance of the proposed method. To increase the generalization of the model, some data augmentation strategies, including random flips, random rotations, and Gaussian noise addition are applied during training process.

3.2 Results

To evaluate the performance of our method, we perform comparison experiments with other networks, including FCN[10] and U-Net. Fig.4 shows some PED segmentation results with different methods. It can be seen from Fig.4 that the segmentation results of the proposed method are the best in terms of boundaries and details no matter the size of PED is small or large. In order to further verify the effectiveness of CMM and SMM, we use U-Net as baseline and perform two ablation experiments. We first insert CMM into the baseline and the Intersection over Union (IoU) index decreases by 0.92% than the complete AM-Net, which implies that the capture of channel multi-scale information is necessary. Second, we insert SMM into the baseline which also achieves a performance reduction of 1.57% compared to the complete AM-Net and indicates that spatial multi-scale information is more conducive to PED segmentation. Table 1 shows objective evaluation metrics of experimental results, including the mean and standard deviation of IoU, Dice Similarity Coefficient (DSC), Sensitivity (Sen) and Specificity (Spe). As can be seen from Table 1, the proposed AM-Net performs better than FCN and U-Net in all evaluation metrics. The ablation experiments (U-Net + CMM, U-Net + SMM, and AM-Net) show the necessary and effectiveness of the proposed CMM and SMM modules.

Table 1. The performance of segmentation with different evaluation metrics.

Architecture	IoU (%)	DSC (%)	Sen (%)	Spe (%)
FCN	50.29±11.05	56.11±12.30	92.99±2.02	79.23±8.14
U-Net	67.75±9.13	75.29±8.97	91.27±1.60	77.03±6.81
U-Net + CMM	71.31±8.99	78.33±9.68	91.69±1.06	79.62±6.56
U-Net + SMM	70.55±9.24	77.84±11.36	92.66±1.30	78.74±4.72
Our Methods	**72.12±9.60**	**79.17±8.25**	**93.05±1.72**	**79.93±5.77**

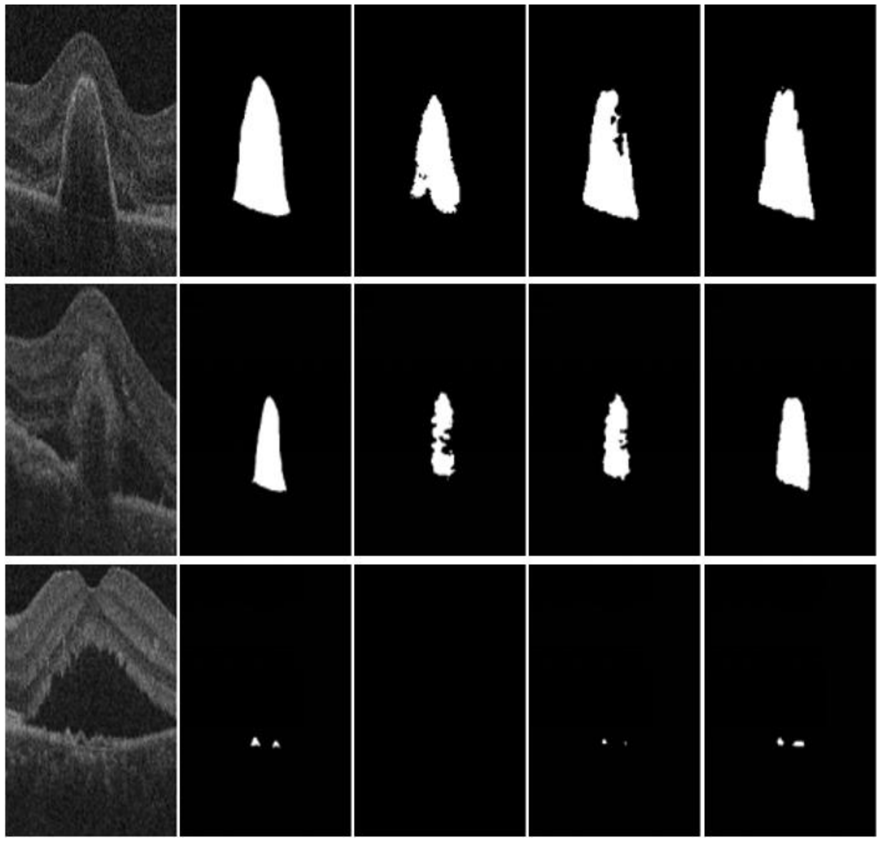


Fig. 4. PED segmentation results with different methods. (a) Original images. (b) Ground truth. (c) Results of FCN. (d) Results of U-Net. (e) Results of the proposed AM-Net.

4. CONCLUSIONS

In this paper, we propose a novel attention multi-scale network (AM-Net) for PED segmentation in retinal OCT B-scans. Two modules, the CMM and SMM, are designed to let the model pay attention on channel information and spatial information. The experiment result shows that the proposed method can accurately segment PED without pre-processing and post-processing.

5. ACKNOWLEDGEMENTS

This work was supported in part by the National Key R&D Program of China under Grant 2018YFA0701700, and in part by the National Basic Research Program of China (973 Program) under Grant 2014CB748600, the National Natural Science Foundation of China (NSFC) under Grant 61622114, 81401472, and in part by Collaborative Innovation Center of IoT Industrialization and Intelligent Production, Minjiang University under Grant IIC1702.

6. REFERENCE

[1] Zayit-Soudry, S., Moroz, I., & Loewenstein, A. (2007). Retinal pigment epithelial detachment. Survey of ophthalmology, 52(3), 227-243.

[2] Lommatzsch, A., Heimes, B., Gutfleisch, M., Spital, G., Zeimer, M., & Pauleikhoff, D. (2009). Serous pigment epithelial detachment in age-related macular degeneration: comparison of different treatments. Eye, 23(12), 2163.

[3] Mrejen, S. (2013). Multimodal imaging of pigment epithelial detachment: a guide to evaluation. Retina, 33(9), 1735-1762.

[4] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

[5] Al-Bander, B., Williams, B., Al-Nuaimy, W., Al-Taee, M., Pratt, H., & Zheng, Y. (2018). Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis. Symmetry, 10(4), 87.

[6] Tennakoon, R., Gostar, A. K., Hoseinnezhad, R., & Bab-Hadiashar, A. (2018, April). Retinal fluid segmentation in OCT images using adversarial loss based convolutional neural networks. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 1436-1440). IEEE.

[7] Roy, A. G., Conjeti, S., Karri, S. P. K., Sheet, D., Katouzian, A., Wachinger, C., & Navab, N. (2017). ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomedical optics express, 8(8), 3627-3642.

[8] Cheng, X., Chen, X., Ma, Y., Zhu, W., Fan, Y., & Shi, F. (2019, March). Choroid segmentation in OCT images based on improved U-net. In Medical Imaging 2019: Image Processing (Vol. 10949, p. 1094921). International Society for Optics and Photonics.

[9] Tennakoon, R., Gostar, A. K., Hoseinnezhad, R., & Bab-Hadiashar, A. (2018, April). Retinal fluid segmentation in OCT images using adversarial loss based convolutional neural networks. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 1436-1440). IEEE.

[10] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation.

In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).