PROCEEDINGS OF SPIE

Group-wise attention fusion network for choroid segmentation in OCT images

Cheng, Xuena, Chen, Xinjian, Feng, Shuanglang, Zhu, Weifang, Xiang, Dehui, et al.

Group-wise Attention Fusion Network for Choroid Segmentation in OCT images

Xuena Cheng[1], Xinjian Chen[1,2], Shuanglang Feng[1], Weifang Zhu[1], Dehui Xiang[1], Qiuying Chen[3], Xun Xu[3], Ying Fan[3], and Fei Shi[1,*]

[1]School of Electronics and Information Engineering, Soochow University, Suzhou, 215006, China
[2]State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, 215123, China
[3]Shanghai General Hospital, Shanghai, 200080, China

ABSTRACT

The choroid is an important structure of the eye and choroid thickness distribution estimated from optical coherence tomography (OCT) images plays a vital role in analysis of many retinal diseases. This paper proposes a novel group-wise attention fusion network (referred to as GAF-Net) to segment the choroid layer, which can effectively work for both normal and pathological myopia retina. Currently, most networks perform unified processing of all feature maps in the same layer, which leads to not satisfactory choroid segmentation results. In order to improve this , GAF-Net proposes a group-wise channel module (GCM) and a group-wise spatial module (GSM) to fuse group-wise information. The GCM uses channel information to guide the fusion of group-wise context information, while the GSM uses spatial information to guide the fusion of group-wise context information. Furthermore, we adopt a joint loss to solve the problem of data imbalance and the uneven choroid target area. Experimental evaluations on a dataset composed of 1650 clinically obtained B-scans show that the proposed GAF-Net can achieve a Dice similarity coefficient of $95.21 \pm 0.73\%$.

KEYWORDS: Choroid segmentation, optical coherence tomography, group-wise attention fusion network, joint loss

1. INTRODUCTION

The choroid is a highly vascular structure of the eye. It provides metabolic support to the retinal pigment epithelium (RPE), and performs critical physiological functions [1-4]. The choroid plays a vital role in the pathophysiology of many conditions, such as age-related macular degeneration (AMD) [5], central serous chorioretinopathy (CSC) [6], Vogt-Koyanagi-Harada syndrome [7] and choroiditis [8]. With the advance of OCT techniques, the deeper structures of the eye become visible, and the whole choroid region can be visualized [9]. Accurate quantification of the choroid thickness from OCT is of great significance for the study of diseases associated with choroid.

Many existing traditional algorithms divide choroid segmentation into two tasks, first detecting the Bruch's membrane (BM) as its upper boundary and then detecting the choroidal-scleral interface (CSI) as the lower boundary [10-12]. Because the BM is normally a strong edge, various gradient-based approaches have proven to be effective [13-14]. However, the CSI is essentially a notional division between the choroidal granularity and the scleral uniformity, which is not defined by marked variation in brightness, and often open to subjective interpretation. Hence, the detection of CSI tends to be more challenging. Moreover, these traditional algorithms have other limitations. Some were only applicable to the normal retinas, or only to macula-centered OCT scans, or had high requirements for the quality of Bscans . For pathological myopia retina, low signal intensity, folding of the retina caused by the long axial length, and possible existence of pathological regions make detection of the retinal and choroidal structure more difficult.

In this paper, we propose an automatic method to segment the choroid region for either normal or pathological myopia retina, from wide-view swept source OCT scans including both macula and optic nerve head (ONH) area, as shown in Fig. 1.

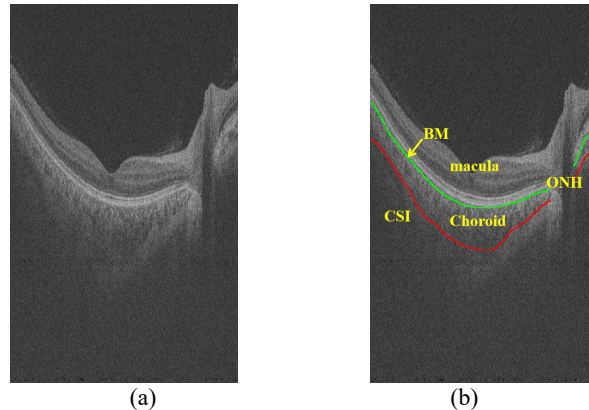*Corresponding author: E-mail: shifei@suda.edu.cn

Fig. 1. B-scans from wide-view volumetric OCT scans. (a) a B-scan of an eye. (b) the same B-scan in (a) with several structure manually labeled.

2. METHODS

In this section, we first introduce the group-wise channel module (GCM) and group-wise spatial module (GSM) in detail. And then the overall network structure and the joint loss are described.

2.1 Group-wise channel module

Fig.2(a) shows the detailed internal structure of the proposed group-wise channel module (GCM). The purpose of this module is to maximize the utilization of group-wise information by automatically fusing group-wise information under the guidance of channel information. As can be seen from the internal structure of GCM, the input feature maps are divided into four groups, and go through different convolution operations to obtain four new groups of feature maps. At the same time, the input feature maps go through global average pooling, convolution and activation operation, and finally return four groups of channel weights through group-wise softmax regression. This enables the network to automatically calculate the importance of each group of channels. Then the obtained channel weights are multiplied to the previously obtained four groups of feature maps. Finally, concatenation and residual operation are used to obtain the output feature maps. Therefore, the module tries to maximize the utilization of channel information by the idea of grouping and channel attention.

2.2 Group-wise spatial module

Fig.2(b) shows the detailed internal structure of the proposed group-wise spatial module (GSM). Same as the GCM, GSM aims to maximizes the utilization of group-wise information, but based on the spatial attention mechanism. As can be seen from the internal structure of GSM, the input feature maps are divided into three groups, and go through different convolution operations to obtained three new groups of feature maps. At the same time, the input feature maps go through 2×2 downsampling, followed by upsampling and softmax operation to obtain three feature maps with spatial weights. The purpose of the downsampling is to focus more on global information, and the upsampling can restore the size of the feature map. Then the obtained spatial weights are multiplied to the previously obtained three groups of feature maps. Finally, concatenation operation and residual operation are used to obtain the output feature maps.
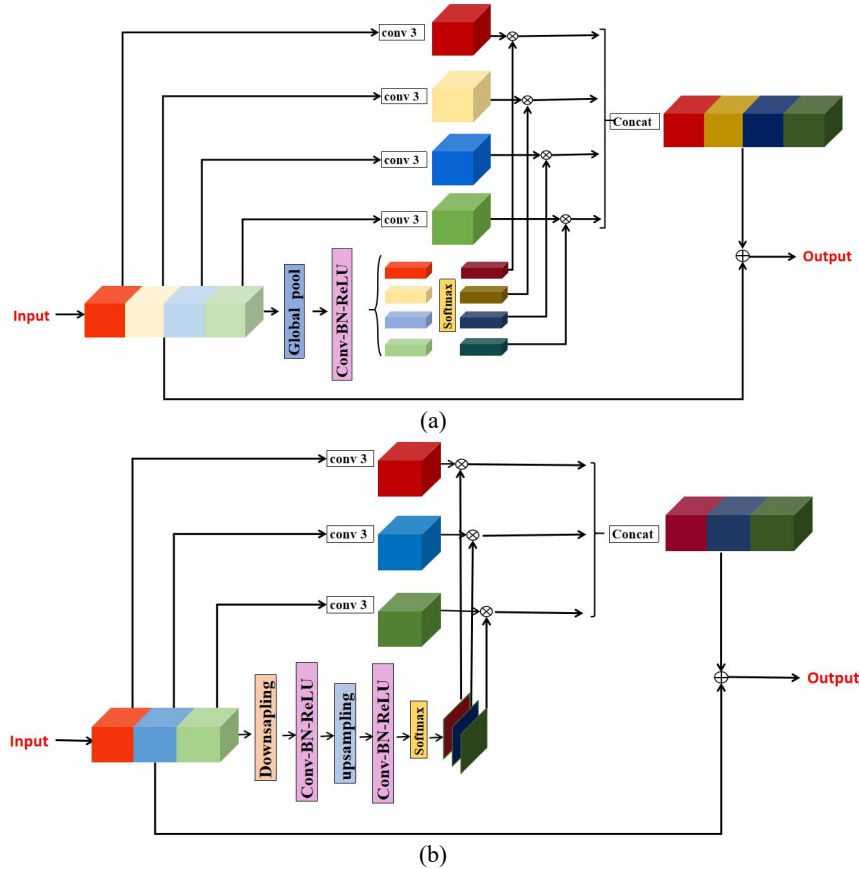
(a)

(b)

Fig. 2. Components of the group-wise channel module (GCM) and the group-wise spatial module (GSM)

2.3 Network Architecture

Fig.3 is the network architecture of the proposed group-wise attention fusion network for choroid segmentation. As can be seen, we use an encoder-decoder architecture with four layers as our baseline. GCMs are embedded in the skip-connection, while GSMs are added in the decoder path. The purpose of placing the GSM in the skip-connection is to optimize the feature maps transmitted from the encoder to the decoder under the guidance of channel attention. The purpose of placing the GSM in the decoder path is to automatically fuse group-wise information and to compensate for the loss of global information caused by upsampling in the decoding process. Compare to many other deep learning models, the proposed network is of compact size, with only 6.7M parameters.

2.4 Loss function

We use the combination of the the binary cross-entropy (BCE) and the Dice loss as the joint loss to solve the problem of data imbalance and the diverse target area in the task of choroid segmentation. The specific equations of Dice loss and BCE loss are as follows:

$$L_{Dice} = 1 - \frac{2\sum_{i}^{N} p_i g_i + \varepsilon}{\sum_{i}^{N} p_i + \sum_{i}^{N} g_i + \varepsilon} \tag{1}$$

$$L_{BCE} = -\frac{1}{N}\sum_{i}^{N}\left(g_i \log(p_i) + (1 - g_i)\log(1 - p_i)\right) \tag{2}$$

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

where N is the total size of the prediction map, $p_i \in [0,1]$ and $g_i \in \{0,1\}$ denote the predicted probability and ground truth label, and ε is a small smoothing factor.
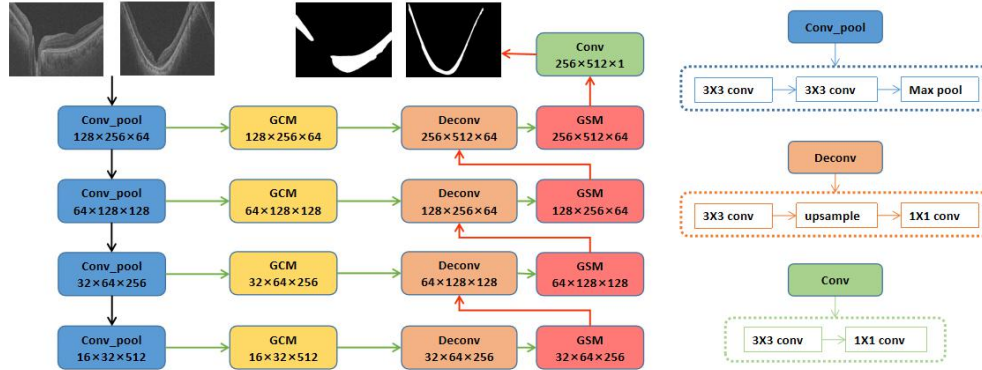

Fig. 3. An overview of the group-wise attention fusion network (GAF-Net). GCM: group-wise channel module, GSM: group-wise spatial module.

3. RESULTS

3.1 Datasets

The datasets used in this paper contain Bscans from three-dimensional OCT volumetric scans acquired at Shanghai General Hospital, China, by a Topcon Atlantis DRI-1 swept source OCT scanner (Topcon Corp., Tokyo, Japan) with center wavelength at 1050 nm, 20μm lateral resolution and 8μm axial resolution. The scan range included the macular center and the ONH region. The image size is 512(width of B-scans) × 256(number of B-scans) × 992(depth of B-scans), which corresponds to a 12×9×2.6 mm³ volume. The collection and analysis of image data were approved by the Institutional Review Board of Shanghai General Hospital and adhered to the tenets of the Declaration of Helsinki. An informed consent was obtained from each subject. The dataset included a total of 1650 OCT B-scans with ground truth labeled by a clinical professional. 1150 B-scans were from normal eyes and 500 B-scans were from pathological myopia eyes.

3.2 Parameter settings

We divided the data into five folds according to subjects and conducted cross-validation. In the training process for each fold, we used the stochastic gradient descent (SGD) algorithm with an initial learning rate of 0.01, momentum of 0.9 and weight decay of 0.0001 to optimize the network. The batch size was set to 8, and the iteration number was set to 60.

3.3 Evaluation metrics

We adopted seven evaluation metrics to quantitatively evaluate the performance of ours model: intersection-over-union (IoU), Dice similarity coefficient (DSC), sensitivity (Sen), specificity (Spe), absolute boundary difference (ABD) of BM and CSI, thickness difference (TD) of the choroid. The specific formulas of IoU and DCS are as follows:

$$IoU = \frac{TP}{TP + FP + FN} \tag{4}$$

$$DSC = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{5}$$

where TP, TN, FP, FN represents the number of true positive, true negative, false positive, and false negative predictions, respectively.

3.4 Results

3.4.1 Ablation study

In order to prove that the two proposed modules contribute to the network, we did ablation experiments between modles. As can be seen from Table 1, compared with the baseline, the GCM module improves the IoU by 0.47% and the GSM module improves the IoU by 0.54%. Adding both GCM and GSM to the baseline improves the IoU by 0.95 %. This shows both the GCM and GSM contribute to the final performance.

Table 1. Results of ablation experiments

Methods	IoU(%)	DSC(%)	Sen(%)	Spe(%)
Baseline	90.38±0.80	94.67±0.59	95.36±0.95	94.31±1.27
Baseline+GCM	90.85±0.88	94.86±0.52	95.20±0.56	94.75±1.47
Baseline+GSM	90.92±1.78	94.96±0.59	**95.70±0.91**	94.24±0.41
GAF-Net	**91.33±1.08**	**95.21±0.73**	95.33±0.95	**95.54±0.45**

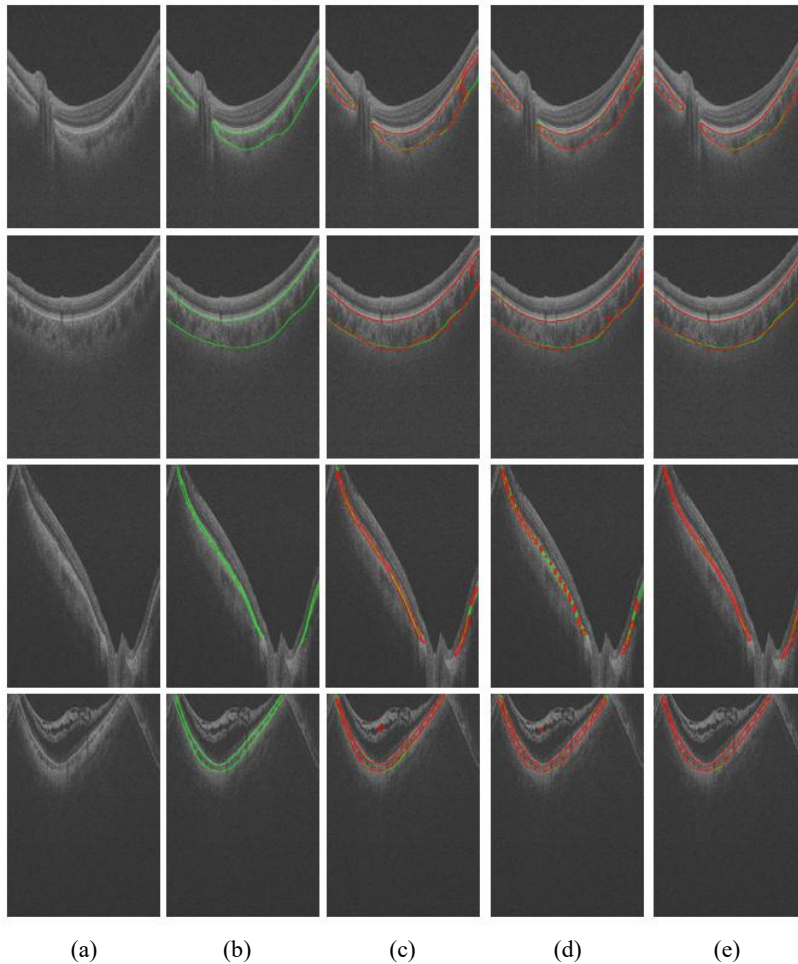|				
(a)	(b)	(c)	(d)	(e)

Fig. 4. Example B-scans with choroid segmentation results overlaid, the first two rows are the results of normal retina, the last two rows are the results of pathological myopia retina. The green and red outline represents the ground truth and the prediction, respectively. (a) Original images. (b) Ground truth. (c) Results of FCN. (d) Results of Deeplab v3. (e) Results of the proposed GAF-Net.

3.4.2 Comparison to other methods

We compared the proposed GAF-Net with two state-of-the-art algorithms: FCN [15] and Deeplab [16]. The performance comparisons are summarized in Table 2, where the average indices calculated on the normal eyes, pathological myopia (PM) eyes, and the total dataset, are listed respectively. Compared with the normal eyes, the choroid thickness of the PM eye is thinner, andpathological regions can lead to destruction of the retinal structure. This makes choroidal segmentation of highly myopic eyes more difficult. Therefore the indices on the PM dataset are worse than those on the normal dataset. However, the proposed method achieves the best results in most cases. In average, our method achieves the highest indices, and reaches 91.33%, 95.21%, 95.33%, and 95.54% in average IoU, DSC, Sen, and Spe, respectively. At the same time, the mean ABD_BM, ABD_CSI, and TD are the lowest, which show that our approach is competitive. Fig.4 shows some segmentation results. It can be seen that the results of the proposed method are more accurate in terms of boundaries and details when compared with other methods, whether for normal or pathological myopia retina.

Table 2. Results compared with other methods

metrics		FCN	Deeplab	GAF-Net
IoU(%)	total	87.88±0.60	88.15±1.22	**91.33±1.08**
	normal	91.75±0.99	92.70±0.93	**93.73±0.74**
	PM	78.97±1.82	77.70±3.44	**85.81±2.54**
DSC(%)	total	93.10±0.42	93.13±0.90	**95.21±0.73**
	normal	95.59±0.57	96.12±0.56	**96.70±0.44**
	PM	87.36±1.47	86.24±2.67	**91.78±1.81**
Sen(%)	total	94.55±1.48	94.18±0.86	**95.33±0.95**
	normal	96.44±0.97	**96.52±0.65**	96.51±0.59
	PM	90.20±3.12	88.80±2.29	**92.61±2.10**
Spe(%)	total	92.36±1.74	92.72±1.05	**95.54±0.45**
	normal	95.06±1.60	95.80±0.68	**96.99±0.47**
	PM	86.15±2.57	85.64±2.92	**92.22±1.28**
ABD_BM(μm)	total	4.54±0.41	4.23±0.28	**2.76±0.27**
	normal	3.54±0.34	3.01±0.14	**2.16 ± 0.08**
	PM	6.83±0.73	7.04±0.76	**4.12±0.77**
ABD_CSI(μm)	total	10.24±1.09	9.39±1.34	**8.25±1.26**
	normal	10.48±1.64	9.44 ±1.78	**8.86 ± 1.67**
	PM	9.68±0.72	9.28±0.79	**6.84 ± 0.78**
TD(μm)	total	11.37±1.29	10.90±1.44	**8.90±1.34**
	normal	11.52±1.77	10.35±1.83	**9.42 ± 1.70**
	PM	11.04±1.27	12.17±1.15	**7.71 ± 0.98**
#params		134M	42M	6.7M

4. CONCLUSIONS

In this paper, we design a novel lightweight group-wise attention fusion network for choroid segmentation in OCT B-scans. Two modules, the GCM and the GSM, are designed to automatically fuse group-wise information in two different

ways. GCM inserted in the skip-connection of the GAF-Net automatically fuses group-wise information under the guidance of the channel information, while GSM placed in the decoding path of the GAF-Net automatically fuses group-wise information under the guidance of the spatial information. The experimental results show that the proposed method can accurately segment choroid of either normal or high myopia retinas and outperforms some other state-of-the-art deep networks for segmentation.

5. ACKNOWLEDGEMENTS

6. REFERENCE

[1]D. L. Nickla, J. Wallman, and E. Research, "The multifunctional choroid," Progress in Retinal & Eye Research., vol. 29, no. 2, pp. 144-168, 2010.

[2]A. Bill, G. Sperber, and K. Ujiie, "Physiology of the choroidal vascular bed," International Ophthalmology., vol. 6, no. 2, pp. 101-107, 1983.

[3]L. M. Parver, "Temperature modulating action of choroidal blood flow," Eye., vol. 5 (Pt 2), no. 2, p. 181, 1991.

[4]A. Alm and S. F. Nilsson, "Uveoscleral outflow--a review," Experimental Eye Research., vol. 88, no. 4, pp. 760-768, 2009.

[5]V. Manjunath, J. Goren, J. G. Fujimoto, and J. S. Duker, "Analysis of choroidal thickness in age-related macular degeneration using spectral-domain optical coherence tomography," American Journal of Ophthalmology., vol. 152, no. 4, pp. 663-668, 2011.

[6]Y. Imamura, T. Fujiwara, R. Margolis, D. Xu and R. F. Spaider, "Enhanced depth imaging optical coherence tomography of the choroid in central serous chorioretinopathy," American Journal of Ophthalmology., vol. 29, no. 3, pp. 445-450, 2009.

[7] I. Maruko, T. Iida, Y. Sugano, H. Oyamada, T. Sekiryu, T. Fujiwara, and R. F. Spaide, "Subfoveal choroidal thickness after treatment of Vogt-Koyanagi-Harada disease," Retina 31(3), 510–517 (2011)

[8]D. A. Sim, P. A. Keane, H. Mehta, S. Fung, J. Zarranz, M. Fruttiger, P. J. Patel, and C.A. Egan, "Repeatability and reproducibility of choroidal vessel layer measurements in diabetic retinopathy using enhanced depth optical coherence tomography," Investigative Ophthalmology & Visual Science., vol. 54, no. 4, pp. 2893-2901, 2013.

[9]Y. Yasuno, Y. Hong, S. Makita, M. Yamanari, M. Akiba, and M. Miura, "In vivo high-contrast imaging of deep posterior eye by 1-microm swept source optical coherence tomography and scattering optical coherence angiography," Optics Express., vol. 15, no. 10, pp. 6121-39, 2007.

[10]F. Shi, B. Tian, W. F. Zhu, D. H. Xiang, L. Zhou, H. B. Xu, and X. J. Chen, "Automated choroid segmentation in three-dimensional 1-μm wide-view OCT images with gradient and regional costs," Journal of Biomedical Optics., vol. 21, no. 12, p. 126017, 2016.

[11]Z. Li, K. Lee, M.Niemeijer, R. F. Mullins, and M. Sonka, "Automated segmentation of the choroid from clinical SD-OCT," Investigative Opthalmology & Visual Science., vol. 53, no. 12, p. 7510, 2012.

[12]D. Alonso-Caneiro, S. A. Read, and M. J. Collins, "Automatic segmentation of choroidal thickness in optical coherence tomography," Biomedical Optics Express., vol. 4, no. 12, pp. 2795-2812, 2013.

[13]Z. Li, K. Lee, M.Niemeijer, R. F. Mullins, and M. Sonka, "Automated segmentation of the choroid from clinical SD-OCT," Investigative Opthalmology & Visual Science., vol. 53, no. 12, p. 7510, 2012.

[14]D. Alonso-Caneiro, S. A. Read, and M. J. Collins, "Automatic segmentation of choroidal thickness in optical coherence tomography," Biomedical Optics Express., vol. 4, no. 12, pp. 2795-2812, 2013.

[15]J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014.

[16]L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, and M. Intelligence, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," IEEE Transactions on Pattern Analysis & Machine Intelligence., vol. 40, no. 4, pp. 834-848, 2018.