FARGO: A Joint Framework for FAZ and RV Segmentation from OCTA Images

Linkai Peng[1], Li Lin[1,2], Pujin Cheng[1], Zhonghua Wang[1], and Xiaoying Tang[1(✉)]

[1] Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China
tangxy@sustech.edu.cn

[2] Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR, China

Abstract. Optical coherence tomography angiography (OCTA) is a recent advance in ophthalmic imaging, which provides detailed visualization of two important anatomical landmarks, namely foveal avascular zone (FAZ) and retinal vessels (RV). Studies have shown that both FAZ and RV play significant roles in the diagnoses of various eye-related diseases. Therefore, accurate segmentation of FAZ and RV from OCTA images is highly in need. However, due to complicated microstructures and inhomogeneous image quality, there is still room for improvement in existing methods. In this paper, we propose a novel and efficient deep learning framework containing two subnetworks for simultaneously segmenting FAZ and RV from *en-face* OCTA images, named FARGO. For FAZ, we use RV segmentation as an auxiliary task, which may provide supplementary information especially for low-contrast and low-quality OCTA images. A ResNeSt based encoder with split attention and ImageNet pretraining is employed for FAZ segmentation. For RV, we introduce a coarse-to-fine cascaded network composed of a main segmentation model and several small ones for progressive refining. Spatial attention and channel attention modules are utilized for adaptively integrating local features with global dependencies. Through extensive experiments, FARGO is found to yield outstanding segmentation results for both FAZ and RV on the OCTA-500 dataset, performing even better than methods that utilize 3D OCTA volume as an extra input.

Keywords: Foveal avascular zone · Retinal vessels · Joint segmentation · Coarse-to-fine · OCTA · ResNeSt.

1 Introduction

Optical coherence tomography angiography (OCTA) is a recent advance in ophthalmic imaging. It is a non-invasive imaging modality that does not require intravenously administering fluorescent dyes [2], which is much safer than prevenient

L. Peng and L. Lin contributed equally to this work.

H. Fu et al. (Eds.): OMIA 2021, LNCS 12970, pp. 1–10, 2021.
https://doi.org/10.1007/978-3-030-87000-3_5

forms of imaging. It provides details of vascular structures within the retina as well as images of blood flow in the retina and choroid [21]. Due to the high resolution of OCTA images, detailed microvascular structures can be displayed, which is beneficial for accurate extractions of foveal avascular zone (FAZ) and retinal vessels (RV) in the retina. These two anatomical landmarks play significant roles in the diagnoses of various eye diseases and eye-related systemic diseases [1]. For instance, the morphology and contour irregularity of FAZ is related to the condition of age-related macular degeneration [11,15] and the severity of glaucoma [4,22] and morphological changes in the RV tortuosity and caliber can reflect the progression of diabetic retinopathy [10,19]. Therefore, an efficient and accurate method for simultaneous FAZ and RV segmentation utilizing OCTA images is in need.

In the research direction of FAZ and RV segmentation from OCTA images, there exist a small number of related works [8,24,26]. For example, Díaz et al. [6] created a FAZ segmentation pipeline based on morphological processing and transformation methods. Eladawi et al. [7] proposed a joint Markov-Gibbs random field model to segment RV and used a Generalized Gauss-Markov random field model for denoising. In addition to traditional methods, a variety of deep learning based methods have been developed to tackle the medical image segmentation task in recent years. Ronneberger et al. [20] proposed the classical U-Net, concatenating the outputs of the encoders and the inputs of the decoders. Several U-Net variants such as U-Net++ [27] have also been proposed. For FAZ and RV segmentation, there have also been ongoing research efforts in this direction. For example, Ma et al. [16] introduced a split-based coarse-to-fine vessel segmentation network containing a coarse segmentation module and a refined segmentation module. Deng et al. [5] developed a U-Net based approach to segment and classify avascular, hypovascular, and capillary-dense areas. Li et al. [13] designed a 3D-to-2D image projection network which utilizes 3D OCTA and OCT volumes as the input to mitigate erroneous retina layer segmentation and to better segment FAZ and RV.

However, these methods still have limitations. For example, [16] and [13] require multi-stage training or a large quantity of model parameters. Also, most of them suffer from inhomogeneous OCTA image quality and complicated structures of FAZ and RV. For example, inferior image quality and erroneous layer projection may cause a network of interest to confuse FAZ with interfering structures, leading to imprecise boundaries and inevitable outliers. Retinal vessels are multi-scaled and it is difficult for fine vessels or terminal branches in noisy images to be segmented accurately. It will be even more complicated if there exist lesions, e.g., microaneurysms and non-perfusion. Such lesions may result in mis-segmentation of vessels and make this task even more challenging. Furthermore, existing works for simultaneously and fully-automatically segmenting FAZ and RV from OCTA images are relatively rare.

In such context, we propose a novel joint FAZ and RV segmentation framework, which mitigates the aforementioned issues and exhibits superior performance. This framework contains two subnetworks. For FAZ, we build a ResNeSt based U-Net with split attention and ImageNet pretraining for better segmentation performance. More importantly, we employ RV segmentation as an auxiliary task to

provide complementary edge and position information, which further improves the FAZ segmentation accuracy, especially for images wherein the contrast between FAZ and its neighborhood is low or the FAZ contours are blurry. For RV, we design a coarse-to-fine cascaded network flow to refine the output of each subnetwork step by step. Spatial attention and channel attention modules are also applied to promote inter-class discrimination and intra-class aggregation [9,17,18]. The final predictions are obtained by merging the outputs of the FAZ subnetwork and the RV subnetwork.

The main contributions of this paper are four-fold: (1) We propose an innovative joint framework for simultaneous FAZ and RV segmentation from OCTA images, achieving superior performance over representative state-of-the-art (SOTA) methods on the publicly-available OCTA-500 dataset [14]. (2) For FAZ segmentation, we make use of RV segmentation as an auxiliary task and a ResNeSt based U-Net structure with split attention and pretrained parameters for initialization. (3) For RV segmentation, we employ a spatial-and-channel dual-attention mechanism in a cascaded coarse-to-fine fashion to gradually refine the outputs. (4) Extensive comparison experiments are conducted, both quantitatively and qualitatively. The source code is available at https://github.com/lkpengcs/FARGO.

2 Methodology

2.1 The Proposed Architecture

The proposed framework is shown in Fig. 1, which consists of a joint FAZ segmentation subnetwork and a coarse-to-fine RV segmentation subnetwork.

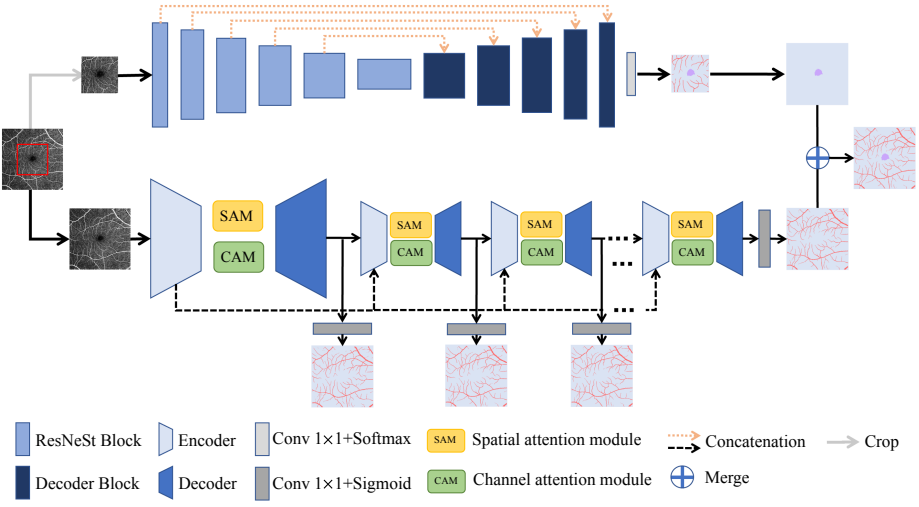

Fig. 1. Schematic demonstration of the architecture of our proposed framework FARGO. The upper part represents the FAZ subnetwork with region of interest extraction and ResNeSt blocks. The lower part represents the RV subnetwork including a main segmentation model and $\mathcal{N}$ small refining models. Spatial attention and channel attention modules are applied in the RV subnetwork.

FAZ Segmentation. Disturbing structures often interfere with the segmentation of FAZ, resulting in outlier predictions and inaccurate contours. Clinically, FAZ is defined as the area without any retinal vessel in the fovea, and thus it is natural to use retinal vessels as prior knowledge to constrain the position of FAZ and assist in distinguishing FAZ boundaries from interfering surroundings. Therefore, we build a joint segmentation model with an encoder-decoder structure to segment FAZ. Instead of treating it as a single-class segmentation task, we train our network to segment FAZ and RV together. It is desirable for such a joint segmentation model to learn the relationship between the distribution of RV and the location and shape of FAZ. Here, we define a weighted loss for FAZ segmentation, with $\lambda_1, \lambda_2, \lambda_3$ being trade-off parameters

$$L_{weighted} = \lambda_1 \mathcal{L}_{faz} + \lambda_2 \mathcal{L}_{rv} + \lambda_3 \mathcal{L}_{bg}, \tag{1}$$

where $\mathcal{L}_{faz}, \mathcal{L}_{rv}, \mathcal{L}_{bg}$ respectively denote the Dice loss for FAZ, RV and the background. We use ResNeSt50 [23] backbone with split attention and ImageNet pretraining initialization as the encoder to speed up network training and converging, resulting in better segmentation performance. Besides, appropriate region of interest (ROI) extraction is performed to focus on regions where FAZ is most likely to appear and reduce the influence of background noise.

RV Segmentation. Given retinal vessels are multi-scaled and are likely to be interfered by similar structures, resulting in disconnected vessels or outliers, we design a coarse-to-fine framework with an iterative correction mechanism. The refining framework consists of a main segmentation model and $\mathcal{N}$ small refining models. Compared to FAZ, retinal vessels are typically slender, and thus we employ shallower networks with fewer downsampling operations. Refining is simpler than segmentation, so we utilize four encoder blocks in the main model and three in the refining models. Each encoder block comprises two layers of 3 $\times$ 3 filters, batch normalization (BN), ReLU and maxpooling. RV segmentation from the input OCTA image is accomplished by the main model and the refining models mainly focus on removing potential errors and generating refined results. The input to the first refining model is the feature maps from the penultimate layer of the main model, and other refining models follow similarly [12]. Features obtained from the first layer of the main model are also concatenated with features from the first layer of each refining model. This allows each refining model to receive different inputs and to encounter differences and changes produced by preceding models. In this way, the main model can generate coarse vessel predictions from OCTA images and the refining models can polish those predictions by identifying potentially-missed vessels, eliminating outliers and recovering continuous details.

Furthermore, spatial and channel attention modules are employed in both the main model and each refining model. The features from the encoder are fed into two parallel attention blocks - a channel attention module (CAM) and a spatial attention module (SAM) as shown in Fig. 2. CAM utilizes features from each channel of the feature maps and improves feature representation by adaptively

assigning weights to channels. SAM obtains spatial information and selectively aggregates contexts by generating a spatial attention map [18]. For training the RV subnetwork, cross entropy is used for outputs from each model and then they are summed up with certain weights, defined as:

$$\mathcal{L}_i = -y_i \log(p_i) - (1 - y_i) \log(1 - p_i) \tag{2}$$

$$\mathcal{L}_{rv} = \sum_i w_i L_i \tag{3}$$

where y_i and p_i respectively represent the groundtruth and prediction of pixel i, and all w_i are set to be 1 since we do not treat any output specially.

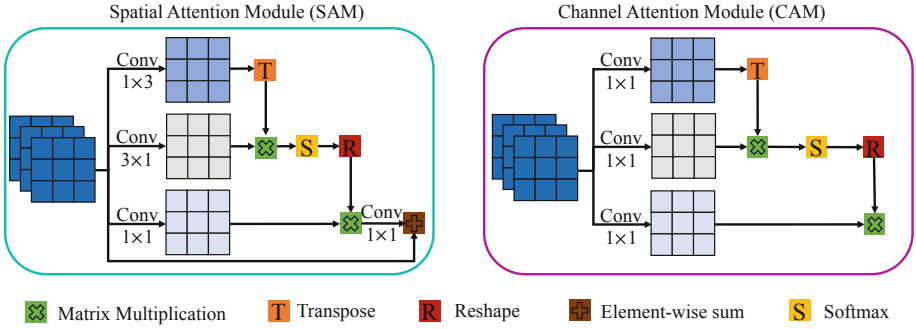

Fig. 2. Illustration of Spatial Attention Module (SAM) and Channel Attention Module (CAM).

3 Experiments

3.1 Dataset and Image Preprocessing

The OCTA-500 dataset are divided into two subsets according to the field of view (FOV) type. One subset contains 300 samples with 6 mm × 6 mm FOV, named OCTA_6M. The other subset contains 200 samples with 3 mm × 3 mm FOV, named OCTA_3M [14]. Please refer to the original paper for relevant biostatistic information [14]. We utilize *en-face* OCTA images generated by maximum-projection between internal limiting membrane layer and outer plexiform layer. Manual annotations for both FAZ and RV are provided. To generalize the entire pipeline, we employ the following data enhancement methods: random horizontal & vertical flipping and random rotation. In addition, for the two different subsets of OCTA-500, namely OCTA_6M and OCTA_3M, we respectively pad the images from 400×400 and 304×304 to 416×416 and 320×320 with reflect padding to meet the downsampling requirement that the input resolution of the network should be a power of 2.

3.2 Experimental Setting

We train and evaluate our pipeline on both OCTA_6M and OCTA_3M. For fair comparison, the training set, validation set and test set are divided according to [14]. All compared methods and the proposed one are implemented with Pytorch using NVIDIA TITAN RTX GPUs. We use the Adam optimizer with a learning rate of 1×10^{-4} with no learning rate policy and train the network for a total of 1000 epochs. The coefficients $\lambda_1, \lambda_2, \lambda_3$ in the weighted loss for FAZ segmentation are respectively set to be 0.6, 0.2, 0.2. Empirically, we choose $\mathcal{N} = 3$, i.e., three refining models for the RV segmentation task.

3.3 Results

All methods are evaluated using four metrics, i.e., Dice[%], Jaccard[%], 95% Hausdorff Distance (HD[px]) and Average Symmetric Surface Distance (ASSD[px]), the results of which are tabulated in Table 1. We compare our proposed framework with several SOTA segmentation models such as Deeplabv3+ [3] and PSPNet [25] for natural image segmentation as well as U-Net and U-Net++ for medical image segmentation. Apparently, our proposed framework FARGO achieves the best performance on both FAZ and RV segmentation on OCTA-500 among all methods that only use *en face* OCTA as the input. Compared with the most advanced IPN V2+ [14], our framework achieves overwhelming FAZ segmentation performance and competitive RV segmentation performance (slightly inferior), with much less computational complexity and memory consumption. To be noted, it is reasonable for IPN V2+ to achieve comparable or even better performance on some subsets or tasks due to its utilizations of 3D volume information and more complicated network structures.

Table 1. Quantitative evaluations of different methods for FAZ and RV segmentation. ‡ indicates the value is directly obtained from the cited paper.

FAZ	OCTA_6M				OCTA_3M			
Method	Dice↑	Jaccard↑	HD↓	ASSD↓	Dice↑	Jaccard↑	HD↓	ASSD↓
Deeplabv3+ [3]	87.75± 12.36	79.71 ± 14.44	8.05 ± 12.49	1.92 ± 3.07	95.43 ± 3.17	91.42 ± 5.51	**3.51 ± 2.56**	0.71 ± 0.27
Pspnet [25]	84.93 ± 13.09	75.56 ± 15.79	10.77 ± 21.84	2.01 ± 2.97	94.04 ± 3.87	88.99 ± 6.41	5.32 ± 7.20	0.95 ± 0.57
U-Net [20]	89.27 ± 12.51	82.27 ± 15.05	**6.40 ± 8.91**	**1.31 ± 1.94**	95.41 ± 4.27	91.51 ± 7.06	4.85 ± 4.39	0.75 ± 0.46
U-Net++ [27]	88.55 ± 15.90	81.90 ± 17.67	10.13 ± 24.54	2.15 ± 5.00	96.91 ± 1.59	94.06 ± 2.93	4.24 ± 3.26	**0.63 ± 0.35**
IPN V2+‡ [14]	**90.54 ± 10.05**	**84.00 ± 14.16**	–	–	**97.42 ± 2.16**	**95.04 ± 3.88**	–	–
Ours	**92.72 ± 6.74**	**87.01 ± 10.60**	**7.76 ± 12.17**	**1.58 ± 3.07**	**98.39 ± 0.92**	**96.84 ± 1.76**	**3.11 ± 2.29**	**0.40 ± 0.27**
RV	OCTA_6M				OCTA_3M			
Method	Dice↑	Jaccard↑	HD↓	ASSD↓	Dice↑	Jaccard↑	HD↓	ASSD↓
Deeplabv3+ [3]	78.82 ± 2.99	65.14 ± 3.93	5.51 ± 4.06	0.90 ± 0.31	78.74 ± 2.84	65.02 ± 3.79	11.94 ± 8.00	1.37 ± 0.71
Pspnet [25]	67.35 ± 2.84	50.84 ± 3.23	12.13 ± 2.48	1.88 ± 0.32	58.02 ± 2.99	40.93 ± 3.02	9.57 ± 4.79	1.93 ± 0.37
U-Net [20]	88.40 ± 2.69	79.32 ± 4.10	4.98 ± 4.29	0.64 ± 0.34	90.40 ± 2.04	82.54 ± 3.27	**3.20 ± 2.61**	**0.54 ± 0.25**
U-Net++ [27]	88.49 ± 2.55	79.44 ± 3.94	**4.44 ± 3.34**	**0.62 ± 0.26**	90.66 ± 2.01	82.97 ± 3.26	3.26 ± 2.40	0.56 ± 0.29
IPN V2+‡ [14]	**89.41 ± 2.74**	**80.95 ± 4.32**	–	–	**92.74 ± 3.95**	**86.67 ± 5.88**	–	–
Ours	**89.15 ± 2.39**	**80.50 ± 3.75**	**4.47 ± 3.73**	**0.61 ± 0.28**	**91.68 ± 2.05**	**84.70 ± 3.34**	**2.78 ± 2.67**	**0.44 ± 0.25**

Representative visualization results from FARGO on OCTA_6M and OCTA_3M are shown in Fig. 3. Clearly, the segmentation results of both FAZ and RV produced by FARGO are more precise and more accurate than those produced by other compared methods. Representative segmentation results for both high-quality images (row 2 and row 4) and low-quality images (row 1 and row 3) are presented in that figure. Our proposed framework has the highest Dice scores in all cases.

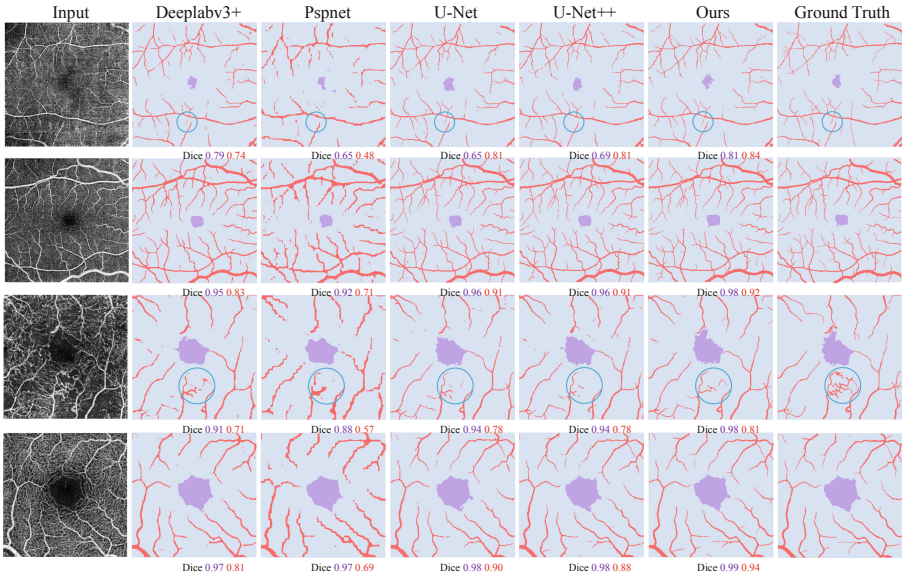

Fig. 3. Representative visualization results from OCTA_6M and OCTA_3M. The first two rows are results from OCTA_6M and the last two rows are results from OCTA_3M. Purple areas represent FAZ and the purple numbers represent their corresponding Dice scores. Red areas represent RV and the red numbers represent their corresponding Dice scores. Areas highlighted by blue circles reflect the ability of our framework to segment fine retinal vessels, even when the input image is of relatively low quality. (Color figure online)

In order to evaluate the effectiveness of several key components in FARGO, we also conduct several ablation studies. For FAZ segmentation, we compare with the proposed FAZ subnetwork without segmenting RV as an auxiliary task, without center ROI extraction, without ResNeSt as the encoder (using the original encoder of U-Net with BN), without pretrained parameters from ImageNet. As shown in Table 2, the Dice score of the original FAZ subnetwork is higher than that of the same network w/o the RV auxiliary task, that w/o center ROI extraction, that w/o ResNeSt as the encoder and that w/o pretrained parameters by 0.60%, 0.69%, 0.84% and 3.31% on OCTA_6M and by 0.17% and 0.27%, 0.54% and 1.25% on OCTA_3M, demonstrating the effectiveness of each component in the FAZ segmentation subnetwork.

Table 2. Ablation analysis results for FAZ segmentation.

RV	ROI	ResNeSt	Pretrain	OCTA_6M				OCTA_3M			
				Dice	Jaccard	HD	ASSD	Dice	Jaccard	HD	ASSD
	✓	✓	✓	92.12	86.69	9.53	1.76	98.22	96.35	4.41	0.49
✓		✓	✓	92.03	86.21	8.12	2.03	98.12	96.35	3.72	0.54
✓	✓		✓	91.88	86.14	10.26	1.80	97.85	95.82	3.87	0.58
✓	✓	✓		89.41	82.57	12.03	2.64	97.14	94.49	5.92	0.82
✓	✓	✓	✓	**92.72**	**87.01**	**7.76**	**1.58**	**98.39**	**96.84**	**3.11**	**0.40**

For RV segmentation, we demonstrate the significance of the two attention modules via ablation studies as well, the results of which are tabulated in Table 3. It is evident that the performance degrades when removing either SAM or CAM. Specifically, the Dice score decreases by 0.24% and 0.15% on OCTA_6M and decreases by 0.12% and 0.33% on OCTA_3M when removing CAM and SAM.

Table 3. The importance of the two attention modules for RV segmentation.

Method	OCTA_6M				OCTA_3M			
	Dice	Jaccard	HD	ASSD	Dice	Jaccard	HD	ASSD
w/o SAM&CAM	89.09 ± 2.57	80.43 ± 3.98	5.25 ± 3.94	0.68 ±0.30	87.65 ± 3.99	78.21 ± 5.75	8.58 ± 7.99	0.81 ± 0.50
w/o CAM	88.91 ± 2.59	80.13 ± 3.98	5.54 ± 4.08	0.70 ± 0.34	91.56 ± 2.25	84.50 ± 3.63	2.91 ± 2.76	0.47 ± 0.28
w/o SAM	89.00 ± 2.67	80.29 ± 4.10	5.42 ± 4.62	0.70 ± 0.35	91.35 ± 2.21	84.15 ± 3.56	3.44 ± 4.25	0.52 ± 0.33
Ours (w SAM&CAM)	**89.15 ± 2.39**	**80.50 ± 3.75**	**4.47 ± 3.73**	**0.61 ± 0.28**	**91.68 ± 2.05**	**84.70 ± 3.34**	**2.78 ± 2.67**	**0.44 ± 0.25**

The influence of the number of the refining models is also explored. The corresponding result is shown in Table 4. We observe that the one with 3 refining models achieves relatively good performance and has reasonable model complexity. The one with 2 refining models has lots of false positives or false negatives to be corrected and that with 4 refining models has a serious overfitting issue by misclassifying many interference structures as RV.

Table 4. The effect of the number of refining models for RV segmentation.

RV	OCTA_6M				OCTA_3M			
Method	Dice	Jaccard	HD	ASSD	Dice	Jaccard	HD	ASSD
2 refining models	88.96 ± 2.60	80.21 ± 4.02	5.37 ± 3.84	0.69 ± 0.30	91.26 ± 2.20	83.99 ± 3.55	3.49 ± 3.61	0.49 ± 0.29
3 refining models	**89.15 ± 2.39**	**80.50 ± 3.75**	**4.47 ± 3.73**	**0.61 ± 0.28**	**91.68 ± 2.05**	**84.70 ± 3.34**	2.78 ± 2.67	**0.44 ± 0.25**
4 refining models	88.99 ± 2.72	80.26 ± 4.16	5.72 ± 4.55	0.71 ± 0.37	91.53 ± 2.21	84.46 ± 3.56	**2.76 ± 2.98**	0.45 ± 0.30

4 Conclusion

In this work, we proposed and validated a novel framework for simultaneously segmenting FAZ and RV from OCTA images. For FAZ, we proposed a joint segmentation network with RV segmentation as an auxiliary task and ResNeSt with split attention as the encoder. Center ROI extraction and ImageNet pretraining were also employed for further improvements. For RV, we designed a coarse-to-fine flow with two attention modules to refine the RV segmentation results step by step. Based on extensive quantitative and qualitative experiments, the proposed method was found to be competitive or even better than representative SOTA segmentation methods, including the methods from the authors of the OCTA-500 dataset utilizing both 3D OCTA volumes and 2D OCTA images, particularly for the FAZ segmentation.

References

1. Abràmoff, M.D., Garvin, M.K., Sonka, M.: Retinal imaging and image analysis. IEEE Rev. Biomed. Eng. **3**, 169–208 (2010)
2. Ang, M., et al.: Optical coherence tomography angiography: a review of current and future clinical applications. Graefes Arch. Clin. Exp. Ophthalmol. **256**(2), 237–245 (2018). https://doi.org/10.1007/s00417-017-3896-2
3. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 833–851. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_49
4. Cheng, K.K., et al.: Macular vessel density, branching complexity and foveal avascular zone size in normal tension glaucoma. Sci. Rep. **11**(1), 1–9 (2021)
5. Deng, W., Tamplin, M.R., Grumbach, I.M., Kardon, R.H., Garvin, M.K.: Region-based segmentation of capillary density in optical coherence tomography angiography. In: Fu, H., Garvin, M.K., MacGillivray, T., Xu, Y., Zheng, Y. (eds.) OMIA 2019. LNCS, vol. 11855, pp. 18–25. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32956-3_3
6. Díaz, M., Novo, J., Cutrín, P., Gómez-Ulla, F., Penedo, M.G., Ortega, M.: Automatic segmentation of the foveal avascular zone in ophthalmological OCT-A images. PLoS ONE **14**(2), e0212364 (2019)
7. Eladawi, N., et al.: Automatic blood vessels segmentation based on different retinal maps from OCTA scans. Comput. Biol. Med. **89**, 150–161 (2017)
8. Fu, H., Xu, Y., Lin, S., Kee Wong, D.W., Liu, J.: DeepVessel: retinal vessel segmentation via deep learning and conditional random field. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 132–139. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_16
9. Fu, J., et al.: Dual attention network for scene segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3146–3154 (2019)
10. Klein, R., Myers, C.E., Lee, K.E., Gangnon, R., Klein, B.E.: Changes in retinal vessel diameter and incidence and progression of diabetic retinopathy. Arch. Ophthalmol. **130**(6), 749–755 (2012)

11. Koskosas, A., Muldrew, K., Patton, W., Topouzis, F., Chakravarthy, U.: Foveal avascular zone (FAZ) area in aging and age related macular degeneration (AMD). Investig. Ophthalmol. Vis. Sci. **50**(13), 948 (2009)
12. Li, L., Verma, M., Nakashima, Y., Nagahara, H., Kawasaki, R.: IterNet: retinal image segmentation utilizing structural redundancy in vessel networks. In: The IEEE Winter Conference on Applications of Computer Vision (WACV), March 2020
13. Li, M., et al.: Image projection network: 3d to 2d image segmentation in OCTA images. IEEE Trans. Med. Imaging **39**(11), 3343–3354 (2020)
14. Li, M., et al.: IPN-V2 and OCTA-500: methodology and dataset for retinal image segmentation. arXiv preprint arXiv:2012.07261 (2020)
15. Liu, H., Wong, D.W.K., Fu, H., Xu, Y., Liu, J.: DeepAMD: detect early age-related macular degeneration by applying deep learning in a multiple instance learning framework. In: Jawahar, C.V., Li, H., Mori, G., Schindler, K. (eds.) ACCV 2018. LNCS, vol. 11365, pp. 625–640. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20873-8_40
16. Ma, Y., et al.: ROSE: a retinal OCT-angiography vessel segmentation dataset and new model. IEEE Trans. Med. Imaging **40**(3), 928–939 (2020)
17. Mou, L., et al.: CS-Net: channel and spatial attention network for curvilinear structure segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 721–730. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_80
18. Mou, L., et al.: CS2-Net: deep learning segmentation of curvilinear structures in medical imaging. Med. Image Anal. **67**, 101874 (2021)
19. Pratt, H., Coenen, F., Broadbent, D.M., Harding, S.P., Zheng, Y.: Convolutional neural networks for diabetic retinopathy. Proc. Comput. Sci. **90**, 200–205 (2016)
20. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
21. Spaide, R.F., Fujimoto, J.G., Waheed, N.K., Sadda, S.R., Staurenghi, G.: Optical coherence tomography angiography. Prog. Retin. Eye Res. **64**, 1–55 (2018)
22. Yip, V.C., et al.: Optical coherence tomography angiography of optic disc and macula vessel density in glaucoma and healthy eyes. J. Glaucoma **28**(1), 80–87 (2019)
23. Zhang, H., et al.: RESNest: split-attention networks. arXiv preprint arXiv:2004.08955 (2020)
24. Zhang, S., et al.: Attention guided network for retinal image segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 797–805. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_88
25. Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.: Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2881–2890 (2017)
26. Zheng, Y., Gandhi, J.S., Stangos, A.N., Campa, C., Broadbent, D.M., Harding, S.P.: Automated segmentation of foveal avascular zone in fundus fluorescein angiography. Invest. Ophthalmol. Vis. Sci. **51**(7), 3653–3659 (2010)
27. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested U-Net architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS 2018. LNCS, vol. 11045, pp. 3–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_1