BSDA-Net: A Boundary Shape and Distance Aware Joint Learning Framework for Segmenting and Classifying OCTA Images

Li Lin[1,2], Zhonghua Wang[1], Jiewei Wu[1,2], Yijin Huang[1], Junyan Lyu[1], Pujin Cheng[1], Jiong Wu[3], and Xiaoying Tang[1(✉)]

[1] Department of Electrical and Electronic Engineering,
Southern University of Science and Technology, Shenzhen, China
`tangxy@sustech.edu.cn`
[2] SEIT, Sun Yat-Sen University, Guangzhou, China
[3] School of Computer and Electrical Engineering,
Hunan University of Arts and Science, Hunan, China

Abstract. Optical coherence tomography angiography (OCTA) is a novel non-invasive imaging technique that allows visualizations of vasculature and foveal avascular zone (FAZ) across retinal layers. Clinical researches suggest that the morphology and contour irregularity of FAZ are important biomarkers of various ocular pathologies. Therefore, precise segmentation of FAZ has great clinical interest. Also, there is no existing research reporting that FAZ features can improve the performance of deep diagnostic classification networks. In this paper, we propose a novel multi-level boundary shape and distance aware joint learning framework, named BSDA-Net, for FAZ segmentation and diagnostic classification from OCTA images. Two auxiliary branches, namely boundary heatmap regression and signed distance map reconstruction branches, are constructed in addition to the segmentation branch to improve the segmentation performance, resulting in more accurate FAZ contours and fewer outliers. Moreover, both low-level and high-level features from the aforementioned three branches, including shape, size, boundary, and signed directional distance map of FAZ, are fused hierarchically with features from the diagnostic classifier. Through extensive experiments, the proposed BSDA-Net is found to yield state-of-the-art segmentation and classification results on the OCTA-500, OCTAGON, and FAZID datasets.

Keywords: Boundary shape and distance · FAZ · Segmentation · Classification · OCTA · Joint learning

---

L. Lin and Z. Wang—Contributed equally to this work.

---

1   Introduction

Optical coherence tomography angiography (OCTA) is an emerging non-invasive ophthalmological imaging technique with an ability to generate high-resolution volumetric images of retinal vasculature. It has been increasingly recognized as an invaluable imaging technique to visualize retinal vessels and foveal avascular zone (FAZ) [19]. OCTA *en face* maps are produced by projection over regions of selective depths, and can be basically divided into superficial images and deep ones with different fields of view (FOVs), e.g., $3\,mm \times 3\,mm$ and $6\,mm \times 6\,mm$, obtained from different scan modes. The former mode has a higher scan resolution than the latter one, and thus the FAZ and capillaries are depicted more clearly. In contrast, the latter mode covers a broader area and has a greater ability in detecting pathological features such as microaneurysms and non-perfusion [6,13]. Many retinal biomarkers are extracted from the OCTA *en face* maps (hereinafter collectively referred to as OCTA images), given that the flattened retinal structures in OCTA images are more informative and convenient for ophthalmologists to examine [14]. Existing evidence suggests that the morphology and contour irregularity of FAZ are highly relevant to various ocular pathologies such as diabetic retinopathy (DR), age-related macular degeneration (AMD), and so on [17,23,31]. For instance, patients with high myopia typically have reduced FAZ areas, whereas macular ischemia caused by diabetes have enlarged FAZ areas [4,21]. As such, precise FAZ segmentation is of great clinical significance. Moreover, OCTA as a new modality shows its potential in computer-aided eye disease and eye-related systemic disease diagnoses [19].

In the past few years, several automatic FAZ segmentation algorithms have been proposed, which can be mainly divided into two categories. The first category is unsupervised methods, typically including statistical segmentation methods and mathematical morphology methods. For instance, Haddouche et al. [9] employed a Markov random fields based method to detect FAZ. Pipelines based on combinations of morphology processing methods and transformation methods also yielded reasonable results [7,18,24]. The past few years have witnessed a rapid development of the second category of methods, the deep learning based methods, which achieved overwhelming performance in almost all computer vision fields. Several works based on UNet [22] and its variants for FAZ segmentation have been reported [8,14–16]. Despite their progress, these methods still have limitations, such as imprecise boundaries due to inferior image quality, confusing FAZ with interfering structures, generating inevitable outliers when there exists erroneous layer projection, and failing to segment low contrast FAZ from its surrounding region. Some representative images are shown on the left panel of Figs. 1 and 2. The main reason for these problems is that UNet based methods typically lack the ability to learn sufficiently strong prior knowledge via single task/loss constraint on small medical image datasets. As for OCTA-based automatic diagnostic classification, only a few papers have reported their attempts. For example, Minhaj et al. designed several hand-crafted features such as vessel tortuosity and FAZ area, and employed support vector machine for classification and staging [2]. Deep ImageNet-pretrained networks have also been proposed for classifying small-sampled OCTA datasets [3,12]. However, these

methods are still very preliminary given most of them only utilized pretrained models followed by fine-tuning OCTA images from the specific study of interest, and there is still room for improvement. Meanwhile, there is no research showing that explicit or implicit FAZ-related features can be effectively and interpretably utilized in deep classification networks.

In such context, we propose a novel hierarchical and multi-level boundary shape and distance aware joint learning framework for FAZ segmentation and eye-related disease classification utilizing OCTA images. Two auxiliary branches, namely boundary heatmap regression and signed distance map (SDM) reconstruction branches, are constructed in addition to the segmentation branch following a shared encoder to improve the segmentation performance. Also, both low-level and high-level features from the aforementioned three branches, including shape, boundary, and signed directional distance map of FAZ, are fused hierarchically with features from the diagnostic classifier.

The main contributions of this paper are four-fold: (1) We present the first joint learning framework, named BSDA-Net, for FAZ segmentation and multi-disease (e.g. DR, AMD, diabetes and myopia) classification from OCTA images. (2) We propose boundary heatmap regression and SDM reconstruction auxiliary tasks, which can effectively improve the performance of FAZ segmentation in both joint learning and single-task learning settings. (3) Via hierarchically fusing features from decoders of the three segmentation-related tasks and the classifier, we demonstrate the effectiveness of FAZ features in guiding and boosting deep classification networks interpretably. (4) We validate the effectiveness of our method on three publicly-accessible OCTA datasets, and our approach achieves state-of-the-art (SOTA) performance in both tasks, establishing new baselines for the community. We make our code available at https://github.com/llmir/MultitaskOCTA.

2   Methodology

The proposed joint learning framework of BSDA-Net is shown in Fig. 1, which is composed of a multi-branch segmentation network (segmentor) and a classifier. Each component will be described detailedly in the following subsections.

2.1   Segmentation Network

As illustrated in Fig. 1, the segmentor in BSDA-Net adopts the widely employed encoder-decoder architecture, which is composed of a shared encoder E and three different decoders, namely a segmentation branch S, a soft boundary regression branch B, and a SDM reconstruction branch D.

BSDA-Net is a general framework, and any segmentation network with an encoder-decoder architecture, such as UNet [22], UNet++ [32], PSPNet [30], and DeepLabv3+ [5], can fit in. In this paper, we employ an adapted UNet structure with ResNeSt50 being the encoder for illustration [29]. The three decoders have the same structure, and features from the penultimate layer of branch B are concatenated to features from the corresponding layer of branch S for better contour perception and preservation. Each decoder module comprises nearest
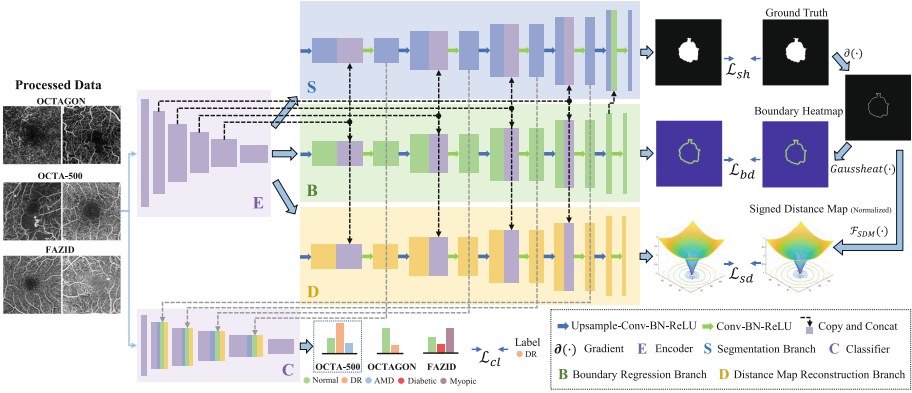
Fig. 1. Schematic representation of the architecture of our proposed framework.

upsampling with a scale factor of 2, followed by two layers of 3×3 filters, batch normalization (BN), and ReLU. In our setting, three small decoders are adopted, with the initial number of feature maps being 256 and getting halved after every upsampling layer. Under small dataset conditions, the edge region of a predicted mask may be inaccurate and has a high probability of being over or under segmented. Also, interfering structures and erroneous layer projection in OCTA images typically lead to outliers. We therefore propose a novel multi-task combination by constructing two auxiliary tasks to reconstruct edges and SDM, which provide the encoder with more topological priors both explicitly and implicitly and make them collaborate with the primary segmentation task to obtain a more accurate target segmentation mask.

Given a target FAZ and a point x in the OCTA image, the SDM of the ground truth G is formulated as

$$G_{sd} = \mathcal{F}_{SDM}(\partial G) = \begin{cases} -\inf_{y \in \partial G} \|x - y\|_2, & x \in G_{in} \\ 0, & x \in \partial G \\ \inf_{y \in \partial G} \|x - y\|_2, & x \in G_{out} \end{cases}, \qquad (1)$$

where $\|x - y\|_2$, ∂G, G_{in}, and G_{out} respectively denote the Euclidean distance between pixels x and y, the boundary, the inside and outside of the FAZ. Compared to the fixed distance map (DM) suggested in [14] which may produce wrong guidance and fail when the foveal center deviates from the image center and the normal DM which only calculates distance of either foreground or background [26,27], SDM has two main advantages. It considers the distance transformation information of both foreground and background, which also characterizes the diameter of FAZ and the fovea position. Moreover, since the target area is typically smaller than the background, through respectively normalizing the distance inside and outside FAZ to $[-1, 0]$ and $[0, 1]$, SDM naturally imposes more weights on the interior region and is beneficial for solving the class imbalance issue. As for the boundary regression branch, considering the subjectivity in

manual annotations, we generate a Gaussian kernel matrix $\mathcal{G}(\cdot)$ centered at each point c_n on the boundary ∂G and construct a soft label heatmap in the form of $Heatsum$ and treat it as a regression task using mean squared error (MSE) loss instead of treating it as a single pixel boundary segmentation problem:

$$\mathcal{G}(c_n) = \frac{1}{2\pi\sigma^2}e^{-\|x-c_n\|_2^2/2\sigma^2}, \tag{2}$$

$$Heatsum(\mathcal{G}(c_1), \mathcal{G}(c_2)) = 1 - (1 - \mathcal{G}(c_1)) \circ (1 - \mathcal{G}(c_2)), \tag{3}$$

$$G_{bd} = Gaussheat(\partial G) = Heatsum(\mathcal{G}(c_1), \mathcal{G}(c_2), ..., \mathcal{G}(c_n)), \forall c_n \in \partial G, \tag{4}$$

where $\circ$ denotes the Hadamard product. Also, G_{bd} is normalized to $[0, 1]$ with values less than 0.001 set to be 0 before normalization. For the segmentation branch S, we use Dice loss to evaluate the pixel-wise agreement between the prediction and the ground truth. Therefore, with trade-off parameters λ_1, λ_2, λ_3, the total objective function of the segmentation model is defined as

$$\begin{aligned}
\mathcal{L}_{seg} &= \lambda_1\mathcal{L}_{sh} + \lambda_2\mathcal{L}_{bd} + \lambda_3\mathcal{L}_{sd} \\
&= \lambda_1\mathcal{L}_{dice}(\mathcal{S}(p_i^s), G) + \lambda_2\mathcal{L}_{mse}(p_i^b, G_{bd}^n) + \lambda_3\mathcal{L}_{mse}(p_i^d, G_{sd}^n),
\end{aligned} \tag{5}$$

where p_i^s, p_i^b, and p_i^d respectively denote the predictions from branches S, B, and D given an input image i, and $\mathcal{S}$ represents the Sigmoid function. G_{bd}^n and G_{sd}^n are the normalized G_{bd} and G_{sd}.

2.2 Classification Network and Joint Learning Strategy

In the training phase, the segmentor is first trained via the aforementioned tasks, while the classifier C is initially frozen to avoid instability caused by inaccurate features from the segmentation network. We set a starting flag τ when the segmentation network almost converges to start joint learning. During joint learning, the multi-level features reconstructed by branches S, B, and D are sufficiently percepted by the unfrozen classifier in a hierarchical and interpretable way. For each feature concatenation, randomly initialized 1×1 convolutions are adopted for dimensionality reduction. Being consistent with the encoder E, we use ResNeSt50 partially (expect the above convolutions) initialized with weights pretrained on ImageNet as the classifier and employ the standard Cross-Entropy loss as $\mathcal{L}_{cl}$. So the final loss of BSDA-Net is defined as (with coefficient λ_0 to balance the two loss terms):

$$\mathcal{L}_{joint} = \begin{cases} \mathcal{L}_{seg}, & \text{epoch} \leqslant \tau \\ \mathcal{L}_{seg} + \lambda_0\mathcal{L}_{cl}, & \text{epoch} > \tau \end{cases}. \tag{6}$$

3 Experiments and Results

Dataset and Preprocessing. We evaluate our proposed BSDA-Net framework on three recently released OCTA datasets: OCTA-500, OCTAGON, and

Foveal Avascular Zone Image Database (FAZID), the details of which are listed in Table 1. In OCTA-500, we only use data from three categories (normal, DR, and AMD) each with a sample size greater than 20. For OCTAGON, the two categories are normal and DR, and FAZID includes normal, diabetic, and myopic. For each dataset, to enlarge the sample size and to better evaluate the classification performance, we unify and combine images with different spatial resolutions via resizing and center cropping. The final category distributions within the three datasets are listed in Tables A1 to A3 (in appendix). We adopt data augmentation methods that will not change the spatial resolution and FAZ shape, including random rotation, flipping, contrast adjustment, adding Gaussian noise, and random blurring.

Table 1. Details and preprocessing of the three datasets utilized in our experiments.

Dataset	OCTA-500 [14,15]		OCTAGON [7]		FAZID [1]
FOV [mm]	3×3	6×6	3×3	6×6	6×6
Original resolution [px]	304×304	400×400	320×320		420×420
Retinal depth	Superficial		Both superficial and deep		superficial
Number	195	169	108	105	304
Preprocessing	Resize	Crop + resize	Resize	Crop	Crop
Processed size [px]	192×192		160×160		224×224

Implementation Details. All compared models and the proposed BSDA-Net framework are implemented with Pytorch using NVIDIA Tesla A100 GPUs. We adopt ResNeSt50 as the encoder of both the segmentation network and the classifier in this work. We use the Adam optimizer with a learning rate of 1 $\times 10^{-4}$ with no learning rate policy for the segmentation network and another separate Adam optimizer with a learning rate of 2×10^{-5} for the classifier. Empirically, we set the starting point of the joint learning τ as 20 and train the network for a total of 200 epochs. Trade-off coefficients λ_0, λ_1, λ_2, λ_3, and σ for soft contour are respectively set to be 1, 3, 1, 1, and 2. For internal validation, we split each dataset into 70%, 20%, and 10% for training, testing, and validation. Five-fold cross-validation is used for fair comparison in all settings.

Evaluation of FAZ Segmentation. All methods are evaluated using four metrics, i.e., Dice[%], Jaccard[%], Average Symmetric Surface Distance (ASD[px]), and 95% Hausdorff Distance (95HD[px]) [11]. Table 2 tabulates the segmentation results on the three OCTA datasets. We compare BSDA-Net with the baseline ResNeSt-UNet (based on our implementation) and several ablation models (to verify the impact of each auxiliary task), as well as several SOTA segmentation models, e.g., Deeplabv3+, PSPNet for natural image segmentation and UNet, UNet++ for medical image segmentation. By compared with the sixth row, the seventh and eighth rows in Table 2 indicate that the soft boundary regression constraint and

the SDM reconstruction are effective in enhancing the FAZ segmentation performance in terms of every evaluation metric for all three datasets. The last two rows of that table indicate that when jointly learning classification and segmentation, though subsequent results identify the effectiveness of such joint learning strategy for boosting the classification performance, BSDA-Net may not achieve the best segmentation results. In other words, the joint learning framework slightly sacrifices the segmentation performance (without significant difference in any metric of any dataset, p-value > 0.05), with a reward of a much greater degree of improvement in the classification performance, as we will show later. Results shown in the penultimate row of Table 2 can be treated as the upper bound of our FAZ segmentation. Figure 2 displays representative segmentation results of the proposed method and four compared models. We also present two sets of intermediate outputs from BSDA-Net, which clearly show the effectiveness of the regressed soft boundary in assessing model's uncertainty and the reconstructed SDM in extracting fovea.
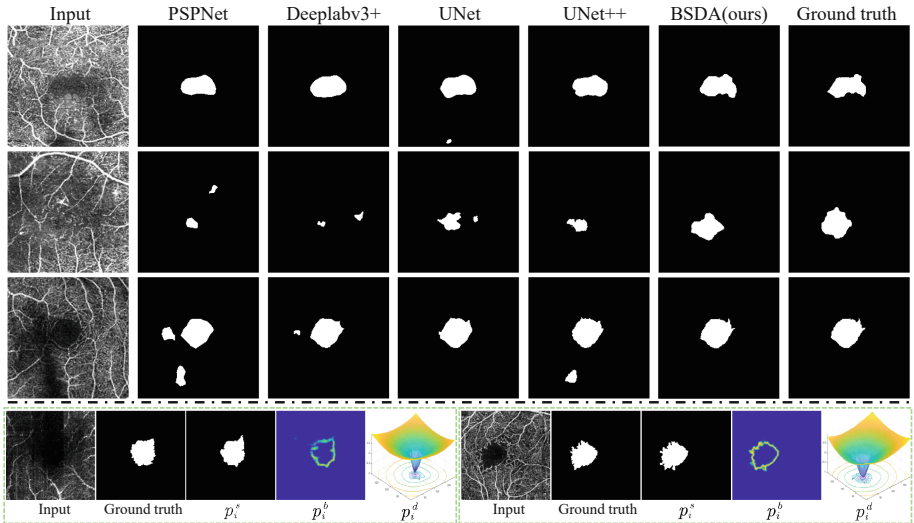

Fig. 2. Visualization results. The upper panel displays segmentation predictions of compared models and the proposed BSDA-Net. The bottom panel shows representative segmentation predictions p_x^s, regressed soft boundary p_x^b, and reconstructed SDM p_x^d. Zoom-in for details.

Evaluation of OCTA Classification. As for classification, accuracy and Cohen's kappa are adopted to evaluate all models. The compared methods include VGG16 (employed in previously proposed OCTA classification works), ResNet50, ResNeXt50, and the SOTA ResNeSt50 (with and without ImageNet pretrained). Moreover, we compare with the YNet [20] of a shared-encoder structure for joint classification and segmentation (based on our reimplementation using ResNeSt50 as the encoder), and the ablation structure without branches B and D. The quantitative results are shown in Table 3, which demonstrates

Table 2. Quantitative evaluations of different networks for FAZ segmentation. Joint learning slightly sacrifices the segmentation performance with no significant difference in any metric.

Method	OCTA-500				OCTAGON				FAZID			
	Dice↑	Jaccard↑	95HD↓	ASD↓	Dice↑	Jaccard↑	95HD↓	ASD↓	Dice↑	Jaccard↑	95HD↓	ASD↓
UNet [22] (w BN)	94.51 ± 6.71	90.16 ± 9.23	5.35 ± 9.67	1.02 ± 1.81	87.71 ± 7.09	78.73 ± 9.95	7.63 ± 12.93	2.28 ± 3.01	89.29 ± 8.09	81.40 ± 10.53	7.65 ± 11.48	2.34 ± 2.60
UNet++ [32]	94.71 ± 6.61	90.53 ± 9.42	5.33 ± 8.61	0.99 ± 1.81	87.86 ± 6.75	78.94 ± 9.90	6.35 ± 9.58	2.05 ± 2.33	88.73 ± 7.94	80.51 ± 10.95	8.26 ± 11.92	2.50 ± 2.81
PSPNet [30]	91.48 ± 6.89	84.87 ± 9.29	6.91 ± 11.01	1.63 ± 1.98	86.87 ± 6.69	77.34 ± 9.49	6.89 ± 10.92	2.18 ± 1.98	87.58 ± 9.40	78.91 ± 12.17	10.04 ± 16.94	2.96 ± 3.57
DeepLabv3+ [5]	92.35 ± 7.33	86.44 ± 9.71	6.06 ± 7.01	1.34 ± 1.49	87.28 ± 6.82	78.00 ± 9.50	6.65 ± 10.39	2.17 ± 2.38	88.22 ± 8.98	79.82 ± 11.57	8.02 ± 13.57	2.74 ± 4.12
Baseline ($E + S$)	95.24 ± 7.00	91.49 ± 9.05	4.52 ± 7.20	0.89 ± 2.25	87.73 ± 7.82	78.84 ± 10.24	6.95 ± 11.61	2.19 ± 2.81	89.89 ± 7.16	82.27 ± 10.02	7.05 ± 10.36	2.20 ± 2.57
BSDA (w/o D, C)	95.90 ± 4.20	92.40 ± 6.71	4.37 ± 7.26	0.76 ± 1.48	88.30 ± 6.32	79.56 ± 9.04	6.23 ± 9.81	2.00 ± 2.02	90.77 ± 5.87	83.57 ± 8.63	6.52 ± 10.15	1.98 ± 2.31
BSDA (w/o B, C)	95.84 ± 4.40	92.30 ± 6.92	4.08 ± 5.81	0.72 ± 1.20	88.14 ± 6.45	79.31 ± 9.10	5.94 ± 9.63	1.97 ± 2.18	90.60 ± 7.12	83.43 ± 9.57	6.19 ± 7.42	1.90 ± 2.26
BSDA (w/o C)	**96.21 ± 3.78**	**92.92 ± 6.14**	**3.61 ± 5.68**	**0.63 ± 1.23**	**88.64 ± 5.89**	**80.00 ± 8.66**	**5.43 ± 1.26**	**1.80 ± 1.26**	**91.03 ± 5.13**	**83.91 ± 7.92**	5.85 ± 6.05	**1.81 ± 1.23**
BSDA (ours)	96.07 ± 4.28	92.72 ± 6.82	3.90 ± 6.03	0.68 ± 1.28	88.37 ± 6.03	79.64 ± 8.75	5.79 ± 8.95	1.92v ± 1.96	90.98 ± 5.19	83.84 ± 8.03	**5.67 ± 5.53**	1.82 ± 1.23

Table 3. Classification performance of all methods on the three OCTA datasets.

Method	Pretrained (ImageNet)	OCTA-500		OCTAGON		FAZID	
		Acc	Kappa	Acc	Kappa	Acc	Kappa
VGG16 [3,12,25]	–	84.62	64.23	95.31	89.36	66.78	49.35
	✓	86.53	72.40	96.24	91.55	70.72	55.63
ResNet50 [10]	–	81.59	60.01	90.61	78.23	70.72	55.68
	✓	89.84	77.88	**97.65**	**94.58**	74.67	61.70
ResNeXt50 [28]	–	82.97	62.16	87.79	71.92	70.72	55.78
	✓	90.11	78.38	96.24	91.29	72.37	58.10
ResNeSt50 [29]	–	89.01	76.62	95.77	90.32	**75.00**	**62.08**
	✓	**90.93**	**80.32**	96.71	92.53	**75.00**	**62.08**
YNet (ResNeSt50)[20]	✓	91.76	82.21	96.71	92.53	74.34	61.24
BSDA (w/o B, D)	✓	92.03	82.67	97.65	94.58	78.62	67.81
BSDA (ours)	✓	**94.23**	**87.68**	**99.53**	**98.92**	**82.57**	**73.67**

that our model achieves the best performance and identifies the effectiveness of FAZ features and the proposed hierarchical feature perception strategy in boosting deep classification networks. Comparing results listed in the last two rows of Table 3, there is no doubt that our proposed joint learning strategy benefits the classification significantly, with only a very mild decrease in the segmentation performance (Table 2). In addition, we display detailed classification reports, including the precision, recall, F1-score of each disease and the macro avg, weight avg of each dataset, in Tables A1 to A3 (from our appendix) for the three datasets. These results establish new SOTA classification baselines for OCTA images. Across all three datasets, BSDA-Net is found to yield the best classification results on DR. Even in FAZID, samples misclassified in the diabetic category are mainly diabetic non-retinopathy samples. The classification

performance on myopia in the FAZID dataset is relatively inferior (79.68%), which partially agrees with existing evidence that FAZ features are relatively indistinguishable in low-moderate myopia [1,4].

4   Conclusion

This paper presents a novel hierarchical and multi-level boundary shape and distance aware joint learning (BSDA-Net) framework for FAZ segmentation and diagnostic classification. Specifically, by constructing a boundary heatmap regression branch and a SDM reconstruction branch, we essentially propose a soft contour and directional signed distance aware segmentation loss, which is found to predict more accurate FAZ boundaries and suppress outliers. We also design a hierarchical and interpretable joint learning strategy to fuse FAZ features with those from the classifier. Extensive experiments on three publicly-accessible OCTA datasets show that our BSDA-Net achieves significantly better performance than SOTA methods on both segmentation and classification. Collectively, our results demonstrate the potential of OCTA for automated ophthalmological and systemic disease assessments and the effectiveness of FAZ features in boosting deep learning classifiers' performance.

References

1. Agarwal, A., Balaji, J., Raman, R., Lakshminarayanan, V.: The Foveal avascular zone image database (FAZID). In: Applications of Digital Image Processing XLIII, vol. 11510, pp. 1151027. International Society for Optics and Photonics (2020)
2. Alam, M., Le, D., Lim, J., Chan, R., Yao, X.: Supervised machine learning based multi-task artificial intelligence classification of retinopathies. J. Clin. Med. **8**(6), 872 (2019)
3. Andreeva, R., Fontanella, A., Giarratano, Y., Bernabeu, M.: DR detection using optical coherence tomography angiography (OCTA): a transfer learning approach with robustness analysis. In: Fu, H., Garvin, M.K., MacGillivray, T., Xu, Y., Zheng, Y. (eds.) International Workshop on Ophthalmic Medical Image Analysis, vol. 12069, pp. 11–20. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-63419-3_2
4. Balaji, J., Agarwal, A., Raman, R., Lakshminarayanan, V.: Comparison of foveal avascular zone in diabetic retinopathy, high myopia, and normal fundus images. In: Ophthalmic Technologies XXX, vol. 11218, pp. 1121810. International Society for Optics and Photonics (2020)
5. Chen, L., 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.) Proceedings of the European Conference on Computer Vision (ECCV). LNCS, vol. 11211, pp. 801–818 (2018)
6. De Carlo, T., Romano, A., Waheed, N., Duker, J.: A review of optical coherence tomography angiography (OCTA). Int. J. Retin. Vitr. **1**(1), 5 (2015)
7. Díaz, M., Novo, J., Cutrín, P., Gómez-Ulla, F., Penedo, M., Ortega, M.: Automatic segmentation of the Foveal avascular zone in ophthalmological OCT-A images. PLoS One **14**(2), e0212364 (2019)

8. Guo, M., Zhao, M., Cheong, A., Dai, H., Lam, A., Zhou, Y.: Automatic quantification of superficial foveal avascular zone in optical coherence tomography angiography implemented with deep learning. Vis. Comput. Ind. Biomed. Art **2**(1), 1–9 (2019)
9. Haddouche, A., Adel, M., Rasigni, M., Conrath, J., Bourennane, S.: Detection of the foveal avascular zone on retinal angiograms using Markov random fields. Digital Sig. Process. **20**(1), 149–154 (2010)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
11. Heimann, T., et al.: Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans. Med. Imag. **28**(8), 1251–1265 (2009)
12. Le, D., et al.: Transfer learning for automated OCTA detection of diabetic retinopathy. Transl. Vis. Sci. Technol. **9**(2), 35 (2020)
13. Leitgeb, R.: En face optical coherence tomography: a technology review. Biomed. Opt. Express **10**(5), 2177–2201 (2019)
14. Li, M., et al.: Image projection network: 3D to 2D image segmentation in octa images. IEEE Trans. Med. Imag. **39**(11), 3343–3354 (2020)
15. Li, M., et al.: IPN-V2 and OCTA-500: methodology and dataset for retinal image segmentation. arXiv preprint arXiv:2012.07261 (2020)
16. Li, M., Wang, Y., Ji, Z., Fan, W., Yuan, S., Chen, Q.: Fast and robust fovea detection framework for OCT images based on Foveal avascular zone segmentation. OSA Continuum **3**(3), 528–541 (2020)
17. Linderman, R., Salmon, A., Strampe, M., Russillo, M., Khan, J., Carroll, J.: Assessing the accuracy of Foveal avascular zone measurements using optical coherence tomography angiography: segmentation and scaling. Transl. Vis. Sci. Technol. **6**(3), 16 (2017)
18. Lu, Y., et al.: Evaluation of automatically quantified foveal avascular zone metrics for diagnosis of diabetic retinopathy using optical coherence tomography angiography. Investig. Ophthalmol. Vis. Sci. **59**(6), 2212–2221 (2018)
19. Ma, Y., et al.: ROSE: a retinal OCT-angiography vessel segmentation dataset and new model. IEEE Trans. Med. Imag. (2020, in Press)
20. Mehta, S., et al.: Y-Net: joint segmentation and classification for diagnosis of breast biopsy images. In: In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. MICCAI 2018. LNCS, vol. 11071, pp. 893–901. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_99
21. Ometto, G., Montesano, G., Chakravarthy, U., Kee, F., Hogg, R., Crabb, D.: Fast 3-dimensional estimation of the foveal avascular zone from OCTA. arXiv preprint arXiv:2012.09945 (2020)
22. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Maier-Hein, G.E.B., Fritzsche, K., Deserno, G.E.B., Lehmann, T., Handels, H., Tolxdorff, T. (eds.) International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
23. Salles, M., Kvanta, A., Amrén, U., Epstein, D.: Optical coherence tomography angiography in central retinal vein occlusion: correlation between the foveal avascular zone and visual acuity. Investig. Ophthalmol. Vis. Sci. **57**(9), OCT242–OCT246 (2016)
24. Silva, A., et al.: Segmentation of Foveal avascular zone of the retina based on morphological alternating sequential filtering. In: Proceedings of the IEEE 28th International Symposium on Computer-Based Medical Systems, pp. 38–43 (2015)

25. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
26. Tan, C., et al.: Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In: IEEE International Symposium on Biomedical Imaging, pp. 1221–1224 (2018)
27. Wijnen, K., et al.: Automated Lesion detection by regressing intensity-based distance with a neural network. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, pp. 234–242. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32251-9_26
28. Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K.: Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1492–1500 (2017)
29. Zhang, H., et al.: ResNeSt: split-attention networks. arXiv preprint arXiv:2004.08955 (2020)
30. 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)
31. Zheng, Y., Gandhi, J., Stangos, A., Campa, C., Broadbent, D., Harding, S.: Automated segmentation of foveal avascular zone in Fundus Fluorescein angiography. Investig. Ophthalmol. Vis. Sci. **51**(7), 3653–3659 (2010)
32. Zhou, Z., Siddiquee, M., Tajbakhsh, N., Liang, J.: UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans. Med. Imag. **39**(6), 1856–1867 (2019)