

FAZ-BV: A DIABETIC MACULAR ISCHEMIA GRADING FRAMEWORK COMBINING FAZ ATTENTION NETWORK AND BLOOD VESSEL ENHANCEMENT FILTERS

Zailiang Chen*, Hailei Lan*, Yongan Meng†, Yuchen Xiong*, Jing Luo†, Hailan Shen**

* School of Computer Science and Engineering, Central South University, Changsha, China

†The Second Xiangya Hospital, Central South University, Changsha, China

ABSTRACT

Monitoring the progress of diabetic macular ischemia (DMI) is essential for providing timely and effective treatment plans and prognostic evaluations. Many approaches have recently been proposed for quantifying DMI based on optical coherence tomography angiography (OCTA) images. However, none of the existing methods can effectively segment the damaged foveal avascular zone (FAZ) and blood vessels (BV) of DMI patients. To avoid this disadvantage, this study proposes a DMI grading framework, i.e. FAZ-BV, combining accurate FAZ and vessel segmentation designed for DMI. For FAZ segmentation, we propose a FAZ attention network, namely FA-Net, coupled with residual fusion attention block (RFAB). For vessel segmentation, Frangi filter and multi-scale line detector can effectively highlight vessel pixels without annotations. The selection of FAZ and BV features follows the doctors' logic of diagnosing DMI. Hybrid features are fused to build a FAZ-BV grading model for DMI. We evaluate our proposed method on a newly collected dataset with 107 eyes. Experimental results show that FA-Net surpasses state-of-the-art segmentation methods with a 95.81% Dice score, an increase of 1.61% compared with U-Net. Our framework achieves encouraging grading performance with a 0.92 AUC, which indicates that the proposed framework is of potential clinical value in DMI grading.

Index Terms— Diabetic macular ischemia, FAZ segmentation, Attention block, Vessel segmentation, OCTA images

1. INTRODUCTION

Diabetic retinopathy (DR) is a common eye disease that causes visual impairment and blindness in working-age people, which seriously affects the vision and life of patients [1]. DR affects vision by destroying the blood perfusion state of the macular area, resulting in diabetic macular ischemia (DMI). In clinical practice, doctors usually judge the degree of diabetic macular ischemia by interpreting optical coherence tomography angiography (OCTA) images of the retinal vascular layer [2, 3]. OCTA produce high-resolution images in a fast and non-invasive manner [4], which contain abundant and accurate information about blood vessel. But interpreting OCTA images completely is time-consuming and requires doctors' extensive diagnostic experience. Therefore, accurate degree judgment of DMI from OCTA images via an automatic framework is urgently needed.

The shape and structure of the foveal avascular zone (FAZ) and retinal capillaries reflect the degree of DMI. The FAZ in the macular area is a biomarker for diagnosing DR, which is generally located in the center of the OCTA image. Babiuch *et al.* [2] used capillary

perfusion density measured by commercial software to grade DMI. Samra *et al.* [3] used OCTA software for FAZ segmentation and applied the threshold method for blood vessel segmentation. FAZ area, blood vessel area density and blood vessel length density are used for DMI quantitative analysis. Though their work promoted the development of DMI grading, their inaccurate FAZ segmentation and blood vessel segmentation influence the performance of grading due to technical limitations. Many studies applied deep learning to segment FAZ. Guo *et al.* [5] added the squeeze-and-excitation module [6] to the U-Net [7] architecture and used it to segment the FAZ of superficial OCTA image of healthy eyes. Afterwards, Li *et al.* [8] proposed lightweight U-Net using dilated convolution, which can quickly segment the OCTA images of healthy and diseased eyes. The main challenge of blood vessel segmentation is the lack of refined labels. Liang *et al.* [9] used the consistent style transfer loss to reconstruct blood vessel of the superficial OCTA image and FAZ segmentation was realized by U-Net with shared encoder. Compared to superficial OCTA images, we used the full-layer enface projection maps of OCTA. They contain more vessel information, but there are also more artifacts. In addition, the FAZ of patients with DMI is deformed and their non-perfusion area increases, which adds to the difficulty of segmentation.

In this study, we employ FAZ segmentation and blood vessel segmentation applicable to DMI and propose a novel DMI grading framework, namely FAZ-BV, to predict the grade of DMI. Our contributions include: (1) To the best of our knowledge, we first propose the DMI grading framework that combines automatic FAZ and vessel segmentation designed for damaged FAZ and vessel of DMI, as shown in Fig. 1. (2) The proposed FAZ attention network (FA-Net) with residual fusion attention block (RFAB) effectively facilitates the network to learn the complicated shape and position of the morbid FAZ. Frangi filter [10] and multi-scale line detector [11] are utilized to highlight vessel pixels without manual annotations significantly. (3) The proposed framework for DMI grading is automatic, highly accurate, and better interpretable. Particularly, the way hybrid features extracted from FAZ and blood vessel mimics the doctor's logic of diagnosing the grade of DMI.

2. MATERIALS AND METHODS

2.1. Data Acquisition and Annotation

On the approval of The Second Xiangya Hospital of Central South University, patients diagnosed with DMI were enrolled from April 2020 to April 2021. All the patients received standard eye examinations. The patients with other fundus diseases, e.g., age-related macular degeneration and retinal vein occlusion, were excluded. 75 eyes with DMI and 32 healthy eyes were included in this study. The OCTA images were obtained from Optovue RTVue XR Avanti (Op-

*This work was supported by the National Natural Science Foundation of China (No. 61972419 and 61672542), Natural Science Foundation of Hunan Province of China (No. 2020JJ4120, 2021JJ30865 and 2021JJ30879).

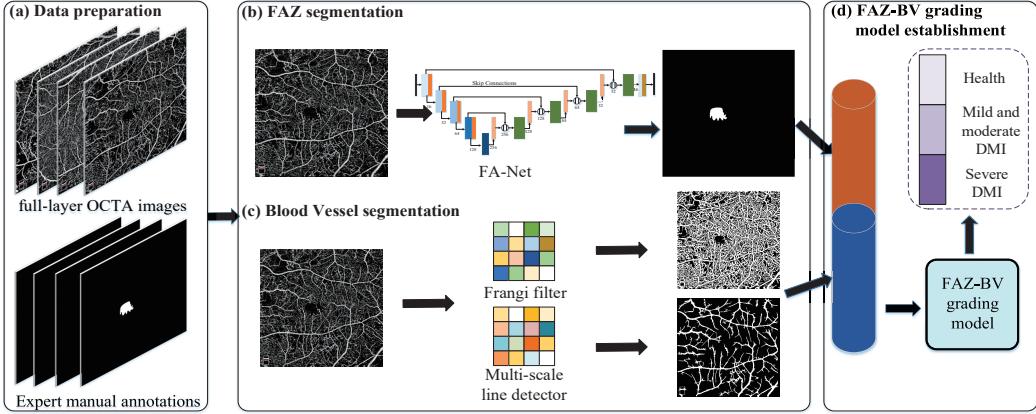


Fig. 1. Our FAZ-BV framework for DMI grading, including OCTA images acquisition and annotation, FAZ segmentation using FA-Net, blood vessel segmentation using Frangi filter and multi-scale line detector, and FAZ-BV grading model establishment.

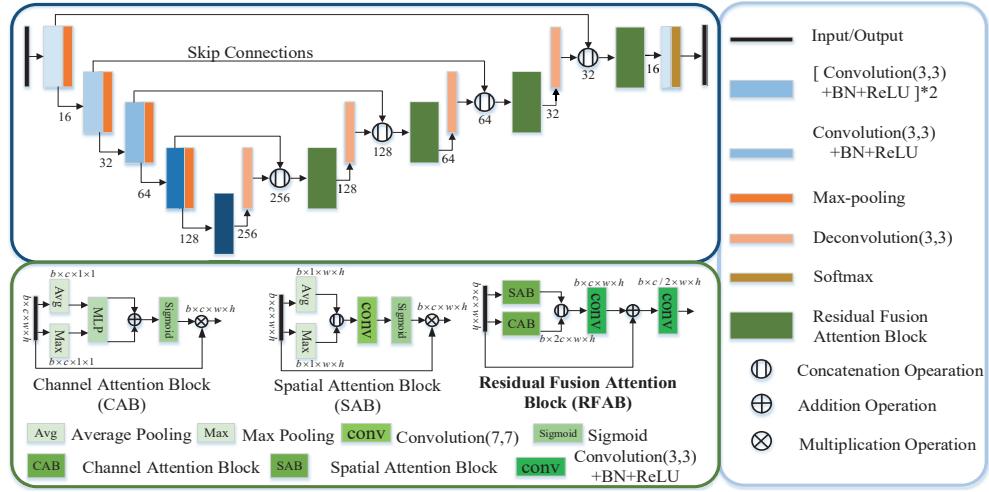


Fig. 2. Our FAZ Attention Network (FA-Net) architecture for FAZ segmentation. Based on U-Net, we utilize four residual fusion attention block (RFAB) after the concatenation of encoder and decoder for activating FAZ semantic information.

tovue, Inc, Fremont, CA) through the AngioVue OCTA system. The scanning range of OCTA images is 6×6 mm in the fovea and its scanning quality is greater than 6/10. A total of 107 full-layer retinal OCTA images were selected.

The equipment and operating environment affect the quality of collected OCTA images. We enhanced their quality by adjusting contrast, brightness, and histogram equalizing. Each FAZ was manually annotated by two doctors with extensive DMI diagnostic experience using ImageJ software [12]. To ensure the precision of all the annotations, they were reviewed by other two masked professional doctors. And the grade of DMI was acquired in the same way.

2.2. FAZ segmentation

For the OCTA images of DMI, the destruction of the shape of FAZ, the increasing non-perfusion area, and artifacts make it more difficult to segment FAZ and vessel precisely. Therefore, we proposed a FAZ attention network, called FA-Net, as shown in Fig. 2. Inspired by U-Net [7], FA-Net is mainly composed of encoders, decoders, skip connections between them, and residual fusion attention block

(RFAB). After the concatenation between the encoded low-level feature maps and the corresponding high-level ones from the decoder, the beneficial semantic information and redundant information for FAZ segmentation are simultaneously input to the next layer, possibly affecting the segmentation performance of the network. Effective feature selection strategies such as attention mechanism help solve this problem. It activates the meaningful semantic information of the feature map and suppresses redundant information. At present, the channel mechanism and the spatial attention mechanism are commonly used [13]. Our RFAB makes meaningful feature selection for FAZ from the channel and spatial dimensions.

The basic convolution module of FA-Net consists of a 3×3 convolution layer, a batch normalization layer, and a ReLU activation layer. In FA-Net, five decoder modules are utilized to obtain feature maps at different scales. The former four decoder modules are composed of two basic convolution modules and a max-pooling layer, and the last one contains a basic convolution module. The four decoders use the feature maps of the encoders to gradually restore the segmentation results. Each decoder contains a transposed convolution with 3×3 kernel size. To improve the ability of the network to

segment FAZ of patients with DMI, we add the RFAB after short-cut connection of the encoder and decoder with the same resolution. Finally, the segmentation maps are generated by a 3×3 convolution layer and a Softmax function.

The details of our proposed RFAB is as follows. On the assumption that the feature map $x \in R^{b \times c \times w \times h}$ after concatenation is input, the channel attention block(CAB) emphasizes the feature map that is significant for FAZ shape segmentation, and we obtain the channel weighted feature map $x_{CA} \in R^{b \times c \times w \times h}$. The spatial attention block(SAB) strengthens the position information of the FAZ in the feature map, then we obtain the spatially weighted feature map $x_{SA} \in R^{b \times c \times w \times h}$. The information extracted from each channel and feature map plays the same significant role in FAZ precise segmentation. Therefore, we concatenate them with the same weight and use one basic convolution module (*Conv1*) for fusion. A residual connection is utilized to promote information propagation during training. Finally, a basic convolution module (*Conv2*) is used to integrate all the information. Result $x_{RFAB} \in R^{b \times \frac{c}{2} \times w \times h}$ is obtained by the formula:

$$x_{RFAB} = \text{Conv2}(\text{Conv1}([x_{CA}, x_{SA}]) + x) \quad (1)$$

To make the network lighter, CAB and SAB are designed following Pang *et al.* [14]. Compared with conducting only one pooling operation, using both average-pooling and max-pooling can obtain more comprehensive information of each channel. Average-pooled feature maps and max-pooled ones are denoted as $x_{cavg} \in R^{b \times c \times 1 \times 1}$ and $x_{cmax} \in R^{b \times c \times 1 \times 1}$, respectively. Then, they are fed to the shared multiple layer perception (*MLP*) with two hidden layers separately. A *Sigmoid* is used to normalize the weight of each channel. Let \otimes denote element-wise multiplication. The channel attention map can be formulated as:

$$x_{CA} = \text{Sigmoid}(\text{MLP}(x_{cavg}) + \text{MLP}(x_{cmax})) \otimes x \quad (2)$$

Similarly, pooling operations in channel axes refine location information in all the feature maps. The obtained feature maps are separately denoted as $x_{savg} \in R^{b \times 1 \times w \times h}$ and $x_{smax} \in R^{b \times 1 \times w \times h}$. The 7×7 convolution layer (*Conv*) with a large receptive field is utilized to fuse and extract more semantic information. Its result is fed to a *Sigmoid* function to get the importance of each location. The spatial attention map can be obtained as:

$$x_{SA} = \text{Sigmoid}(\text{Conv}[x_{savg}, x_{smax}]) \otimes x \quad (3)$$

2.3. Blood Vessel Segmentation

To tackle with lack of refined mask, vessel segmentation, as shown in **Fig. 1 c**, is realized by blood vessel enhancing filters, which reduce the impact of superficial vessel projection and artifacts. We apply the 2D Frangi filter [10] to segment all the blood vessel in the OCTA images. Frangi filter is an enhanced filtering algorithm for vessel detection based on the eigenvalues of Hessian to obtain the measurement of "vesselness" [10]. We set the scale range at [0.3, 2] and the scale ratio at 0.15 for OCTA images. According to the calculated adaptive threshold, the binary operation is used to obtain the mask. Multi-scale line detection is designed for retinal vessel [11]. To segment the big blood vessel, we slightly adjusted its structure and parameters according to the characteristics of OCTA images.

2.4. FAZ-BV Model Establishment

2.4.1. FAZ Features and Blood Vessel Features Extraction

According to clinical experience, the FAZ of healthy eyes is almost a small circle, while DR patients' is deformed and enlarged. Four

features are utilized to describe its shape and size. Its actual size can be calculated based on the amount of segmentation results and the scale. To measure the similarity to a circle, Circularity [15], Aspect ratio [16], and Roundness are used to describe the FAZ shape. When FAZ is a complete circle, their values are 1. Circularity is the ratio of its area to the area with the same perimeter. Aspect ratio is defined as the ratio between the long axis and the short axis of the ellipse generated from FAZ [16]. Roundness is the ratio of its area to the area of the circle corresponding to the max axial length.

The blood vessel distribution of healthy people is tight, while there is non-perfusion area in DMI patients'. Referring to the binary masks of blood vessel and large blood vessel, the blood vessel density (BVD) and big blood vessel density (BBVD) is calculated by the ratio of the vessel pixels to all the pixels of OCTA images. The texture of blood vessels can reflect the state of DMI in the OCTA images. The gray-level co-occurrence matrix (GLCM) is calculated for refining texture information of blood vessel [17]. Average contrast and correlation in four angles [$45^\circ, 90^\circ, 135^\circ, 180^\circ$] are selected and extracted as GLCM eigenvalues for grading.

2.4.2. Hybrid Features and Grading Model Establishment

The hybrid features may promote the performance and robustness of classification [18]. In this study, we extract hybrid features from FAZ and blood vessel segmentation. Using these hybrid features, we build a FAZ-BV grading model based on tree-based ensemble methods (**Fig. 1 d**). Deep forest has better accuracy and faster training speed than other tree-based ensemble methods [19]. We employed it as a classifier to build a FAZ-BV grading model to predict the possibility of health, mild and moderate DMI, and severe DMI. The category with the highest probability is selected as the grading result.

3. EXPERIMENTAL RESULTS

3.1. Implementation and Evaluation Methods

We collected a dataset containing 107 OCTA full-layer images from patients with DMI and used it to evaluate the performance of the proposed method. After shuffling the dataset randomly, 75 images are assigned for training, 10 images for validation, and 22 images for testing. The data was augmented by multiple methods, including rotations of -30 degrees to 30 degrees with 0.8 possibilities, left-right flipping with 0.5 possibilities, up-down flipping with 0.5 possibilities, contrast normalization, and additive Gaussian noise. Every image has randomly applied the combination of 1 to 4 kinds of augmentation methods and the order of them was random in each batch. FA-Net was implemented in Pytorch¹ framework. We used Adam [20] to optimize Soft Dice loss. The learning rate was initialized to 5e-4 and weight decay was set to 10-8. The network was trained for 150 epochs with a batch size of 2. Training and testing were accelerated by the NVIDIA GeForce GTX 1080 Ti GPU. Five-fold cross-validation was utilized for fair performance evaluation. The probability threshold for FAZ pixels was set to 0.5. The extraction of the largest connected region and the hole filling were used to refine the FAZ segmentation results.

According to [5, 8, 9], Dice coefficient, Jaccard Index, Pearson correlation coefficient (Corr) [21] are commonly used metrics for FAZ segmentation. To measure the performance of grading, the area under the curve (AUC) value of receiver operating characteristic (ROC) curve, accuracy (Acc) and F1-score (F1) were selected.

¹<https://pytorch.org/>

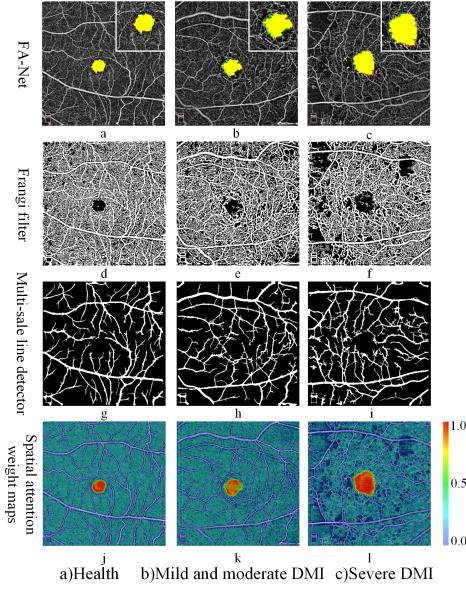


Fig. 3. Visualization of FAZ (a-c), all the blood vessel (d-f), big vessel (g-i) segmentation, and the fusion of spatial attention weight maps of the last RFAB of FA-Net and the original input images (j-l). For FAZ segmentation, the green, red and yellow regions represent the ground truth, the predicted area and the overlap of the two areas, respectively. For the spatial attention maps, the importance of pixels is relevant to their color according to the color bar.

Table 1. Performance comparison of FAZ segmentation.

Method	Dice(%)	Jaccard(%)	Corr(%)
Guo <i>et al.</i> [5]	93.15±3.59	88.37±5.12	99.44±0.36
Li <i>et al.</i> [8]	92.82±5.89	88.29±7.78	99.37±0.51
Gu <i>et al.</i> [13]	95.58±3.57	92.24±5.58	99.59±0.37
Baseline	94.20±3.98	90.09±5.97	99.48±0.40
FA-Net	95.81±3.40	92.59± 5.34	99.61± 0.35

3.2. FAZ Segmentation Results and Comparision with Other Algorithms

Guo *et al.* (2019) [5] and Li *et al.*(2020) [8] are the networks customized for FAZ segmentation recently. The comprehensive attention network (CA-Net,2021) [13] is composed of spatial attention, channel attention, and scale attention to achieve state-of-the-art performance on multiple medical datasets. Baseline is acquired by replacing the RFAB of FA-Net with a 3×3 convolution layer. In addition, the methods we compared in this study were implemented in accordance with the description or code of the original published paper. And we used the same learning rate of 5e-4 and training strategy for fair comparison. **Table 1** shows the performance of all methods for comparison on our dataset. Our FA-Net outperformed the others with 95.8% Dice, 92.6% Jaccard, and 99.6% Corr, while [5] and [8] only have Dice of 93.15% and 92.82%. Compared with the baseline, FA-Net increased Dice and Jaccard by 1.61% and 2.50% respectively, slightly surpassing the state-of-the-art CA-Net. Experimental results showed that RFAB can effectively extract important semantic information for segmentation when dealing with the complex shape and position of morbid FAZ. And our FA-Net can perform precise and robust segmentation of DMI patients' FAZ on a small dataset.

Table 2. Demographic characteristics, FAZ features, and BV features of each DMI grade.

Method	Health	Mild and Moderate	Severe
Count	32	45	30
Male:female	12:19	27:18	22:8
Age	33.86±11.82	49.82±10.35	54.71±7.95
Area(mm^2)	0.31±0.09	0.38±0.12	0.52±0.23
Circularity	0.70±0.10	0.49±0.17	0.34±0.16
Aspect ratio	1.12±0.06	1.28±0.17	1.30±0.22
Roundness	0.90±0.05	0.79±0.10	0.78±0.12
BVD(%)	56.06±3.25	54.48±3.12	50.11±3.53
BBVD(%)	15.12±0.49	16.1±0.69	16.6±0.63
Contrast	17.21±0.64	15.56±1.14	13.55±0.80
Correlation	0.28±0.03	0.36±0.05	0.44±0.03

Table 3. Grading performance of various methods.

Method	AUC	Acc(%)	F1(%)
FAZ	0.87±0.13	78.08±9.67	67.14± 14.35
BV	0.88±0.03	85.02±3.35	76.58± 5.39
FAZ-BV	0.92± 0.05	88.73± 4.97	84.03± 8.11

The **Fig. 3 a-c** shows visual results of FAZ segmentation. Despite complicated boundary and uncertain location of diseased FAZ, our model still maintains encouraging performance. As presented in **Fig. 3 j-l**, though the shape and size of FAZ changes, the FAZ area was highlighted when other areas were suppressed, which demonstrated that the network learned the definition of FAZ explainably.

3.3. Blood Vessel Segmentation Results

In **Fig. 3, d-f** and **g-i** show the visual results of all the vessel and big vessel segmentation, respectively. Though some artifacts are hard to remove cleanly, Frangi filter and multi-scale line detector effectively obtain the response of blood vessel in OCTA images.

3.4. Diagnostic validation and analysis

Table 2 shows the demographic features and the raw calculated feature values of each grade. We extracted a total of eight features, including four FAZ features and four blood vessel (BV) features. FAZ and BV grading model are established for comparison using their own features. As shown in **Table 3**, the BV model outperformed the FAZ model with 0.88 AUC. Our FAZ-BV model surpassed others with 0.92 AUC, 88.73% Acc, and 84.03% F1. Results indicate that the combination of FAZ and BV features boosts the grading performance to a great extent than utilizing single FAZ or BV features.

4. CONCLUSION

In this study, we have proposed the DMI grading framework, namely FAZ-BV, based on accurate FAZ and blood vessel segmentation referring to the doctors' diagnostic logic. Extensive experiments indicated that FA-Net is more suitable for FAZ segmentation of DMI patients and the hybrid FAZ-BV model significantly promote the performance of DMI grading. In the near future, we will evaluate our framework on OCTA images from other equipment and continue to increase the grading accuracy by designing better segmentation models and combining more explainable features.

5. REFERENCES

- [1] Charumathi Sabanayagam, Riswana Banu, Miao Li Chee, Ryan Lee, Ya Xing Wang, Gavin Tan, Jost B Jonas, Ecosse L Lamoureux, Ching-Yu Cheng, Barbara EK Klein, et al., “Incidence and progression of diabetic retinopathy: a systematic review,” *The lancet Diabetes & endocrinology*, vol. 7, no. 2, pp. 140–149, 2019.
- [2] Amy S Babiuch, Atsuro Uchida, Ming Hu, Mehnaz Khan, Sunil K Srivastava, Rishi P Singh, Peter K Kaiser, Aleksandra Rachtiskaya, Jamie L Reese, and Justis P Ehlers, “Use of octa capillary perfusion density measurements to detect and grade macular ischemia,” *Ophthalmic Surgery, Lasers and Imaging Retina*, vol. 51, no. 4, pp. S30–S36, 2020.
- [3] Wasim A Samara, Abtin Shahlaee, Murtaza K Adam, M Ali Khan, Allen Chiang, Joseph I Maguire, Jason Hsu, and Allen C Ho, “Quantification of diabetic macular ischemia using optical coherence tomography angiography and its relationship with visual acuity,” *Ophthalmology*, vol. 124, no. 2, pp. 235–244, 2017.
- [4] Richard F Spaide, James M Klancnik, and Michael J Cooney, “Retinal vascular layers imaged by fluorescein angiography and optical coherence tomography angiography,” *JAMA ophthalmology*, vol. 133, no. 1, pp. 45–50, 2015.
- [5] Menglin Guo, Mei Zhao, Allen MY Cheong, Houjiao Dai, Andrew KC Lam, and Yongjin Zhou, “Automatic quantification of superficial foveal avascular zone in optical coherence tomography angiography implemented with deep learning,” *Visual Computing for Industry, Biomedicine, and Art*, vol. 2, no. 1, pp. 1–9, 2019.
- [6] Jie Hu, Li Shen, and Gang Sun, “Squeeze-and-excitation networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.
- [7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [8] Mingchao Li, Yuexuan Wang, Zexuan Ji, Wen Fan, Songtao Yuan, and Qiang Chen, “Fast and robust fovea detection framework for oct images based on foveal avascular zone segmentation,” *OSA Continuum*, vol. 3, no. 3, pp. 528–541, 2020.
- [9] Zhijin Liang, Junkang Zhang, and Cheolhong An, “Foveal avascular zone segmentation of octa images using deep learning approach with unsupervised vessel segmentation,” in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1200–1204.
- [10] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever, “Multiscale vessel enhancement filtering,” in *International conference on medical image computing and computer-assisted intervention*. Springer, 1998, pp. 130–137.
- [11] Uyen TV Nguyen, Alauddin Bhuiyan, Laurence AF Park, and Kotagiri Ramamohanarao, “An effective retinal blood vessel segmentation method using multi-scale line detection,” *Pattern recognition*, vol. 46, no. 3, pp. 703–715, 2013.
- [12] Caroline A Schneider, Wayne S Rasband, and Kevin W Elceiri, “Nih image to imagej: 25 years of image analysis,” *Nature methods*, vol. 9, no. 7, pp. 671–675, 2012.
- [13] Ran Gu, Guotai Wang, Tao Song, Rui Huang, Michael Aertsen, Jan Deprest, Sébastien Ourselin, Tom Vercauteren, and Shaoting Zhang, “Ca-net: Comprehensive attention convolutional neural networks for explainable medical image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 2, pp. 699–711, 2021.
- [14] Shuchao Pang, Anan Du, Mehmet A Orgun, Yunyun Wang, and Zhenmei Yu, “Tumor attention networks: Better feature selection, better tumor segmentation,” *Neural Networks*, vol. 140, pp. 203–222, 2021.
- [15] Johnny Tam, Kavita P Dhamdhere, Pavan Tiruveedhula, Silvestre Manzanera, Shirin Barez, Marcus A Bearse, Anthony J Adams, and Austin Roorda, “Disruption of the retinal parafoveal capillary network in type 2 diabetes before the onset of diabetic retinopathy,” *Investigative ophthalmology & visual science*, vol. 52, no. 12, pp. 9257–9266, 2011.
- [16] Ming-Kuei Hu, “Visual pattern recognition by moment invariants,” *IRE transactions on information theory*, vol. 8, no. 2, pp. 179–187, 1962.
- [17] Alaa Eleyan and Hasan Demirel, “Co-occurrence based statistical approach for face recognition,” in *2009 24th international symposium on computer and information sciences*. IEEE, 2009, pp. 611–615.
- [18] Jianhong Cheng, Jin Liu, Meilin Jiang, Hailin Yue, Lin Wu, and Jianxin Wang, “Prediction of egfr mutation status in lung adenocarcinoma using multi-source feature representations,” in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1350–1354.
- [19] Zhi-Hua Zhou and Ji Feng, “Deep forest,” *National Science Review*, vol. 6, no. 1, pp. 74–86, 2019.
- [20] Diederik P Kingma and Jimmy Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [21] Mavuto M Mukaka, “A guide to appropriate use of correlation coefficient in medical research,” *Malawi medical journal*, vol. 24, no. 3, pp. 69–71, 2012.