A Hyperreflective Foci Segmentation Network for OCT Images with Multi-dimensional Semantic Enhancement

Xingguo Wang[1], Yuhui Ma[1*], Xinyu Guo[1], Yalin Zheng[2], Jiong Zhang[1], Yonghuai Liu[4], and Yitian Zhao[1,*]

[1] Laboratory of Advanced Theranostic Materials and Technology, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo, China
{yitian.zhao; mayuhui}@nimte.ac.cn
[2] Department of Eye and Vision Science, University of Liverpool, Liverpool, UK
[3] Department of Computer Science, Edge Hill University, Ormskirk, UK

Abstract. Diabetic macular edema (DME) is a leading cause of vision loss worldwide. Optical Coherence Tomography (OCT) serves as a widely accepted imaging tool for diagnosing DME due to its non-invasiveness and high resolution cross-sectional view. Clinical evaluation of Hyperreflective Foci (HRF) in OCT contributes to understanding the origins of DME and predicting disease progression or treatment efficacy. However, limited information and a significant imbalance between foreground and background in HRF present challenges for its precise segmentation in OCT images. In this study, we propose an attention mechanism-based **MU**lti-dimensional **S**emantic **E**nhancement **Net**work (MUSE-Net) for HRF segmentation to address these challenges. Specifically, our MUSE-Net comprises attention-based multi-dimensional semantic information enhancement modules and class-imbalance-insensitive joint loss. The adaptive region guidance module softly allocates regional importance in slice, enriching the single-slice semantic information. The adjacent slice guidance module exploits the remote information across consecutive slices, enriching the multi-dimensional semantic information. Class-imbalance-insensitive joint loss combines pixel-level perception optimization with image-level considerations, alleviating the gradient dominance of the background during model training. Our experimental results demonstrate that MUSE-Net outperforms existing methods over two datasets respectively. To further promote the reproducible research, we made the code and these two datasets online available.

Keywords: Hyperreflective foci · OCT · Attention · Segmentation

1 Introduction

Diabetic macular edema (DME) is the leading cause of vision impairment in patients with diabetes [8, 2, 17]. Clinical research on Hyperreflective Foci (HRF) provides new insights into understanding the pathogenesis of DME, and offers

directions for predicting the prognosis of DME or efficacy of treatments. Optical Coherence Tomography (OCT), an imaging tool for the retina, improves visualization of retinal microstructure, aiding in accurate disease diagnosis [9]. HRF in OCT images are well-defined lesions with high reflectivity [1, 5, 3]. Consequently, automated segmentation of HRF based on OCT images can facilitate quantitative analysis, and improve the diagnostic performance of DME. However, the limited information provided and the significant imbalance between foreground and background pose challenges to its accurate segmentation. Therefore, there is an urgent need for reliable HRF segmentation to assist ophthalmologists in making informed clinical decisions.

Several methods have been proposed to segment HRF from OCT images, utilizing either traditional machine learning or deep learning techniques. Okuwobi *et al.* [10] introduced an automatic segmentation method based on the component tree, identifying regions of interest in the retina through clustering algorithms. However, traditional machine learning methods rely on retinal layer segmentation, which can be particularly challenging in the presence of ocular pathologies. Wei *et al.* [14] proposed a lightweight network involving image preprocessing and patch classification. Goel *et al.* [6] employed a patch-based strategy, ensuring that the regions of interest are composed of pixels from the retinal region, and utilized U-Net for HRF segmentation. However, these patch-based methods may lose semantic information across the entire image, and come with a high computational cost. Therefore, there is a necessity to explore dedicated methods to capture more comprehensive information in HRF to address existing challenges.

In clinical practice, ophthalmologists also need to leverage global and local perspectives while combining consecutive OCT images for improved lesion localization. Therefore, a well-designed attention mechanism may enable networks to incorporate this clinical knowledge, potentially improving HRF segmentation. In this work, we propose an attention mechanism-based HRF segmentation network (MUSE-Net) that integrates multi-dimensional semantic information enhancement. Our approach introduces an adaptive region-guided single-slice enhancement module (SEM) for balancing local and global perspectives within individual OCT images, and an adjacent slice-guided multi-dimensional enhancement module (MEM) for combining information across multiple OCT images. MUSE-Net takes multiple adjacent OCT slices as the inputs and produces a segmentation map for the target OCT slice image. To address the foreground-background imbalance issues, we use a joint loss function that combines pixel-level perceptual optimisation with image-level considerations.

The main contributions are summarized as follows:

(1) We propose an attention-based multi-dimensional semantic information enhancement network for accurate HRF segmentation. The proposed network contains the adaptive region-guidance and adjacent slice-guidance modules, which aim at enhancing intra-slice and inter-slice semantic information, respectively.

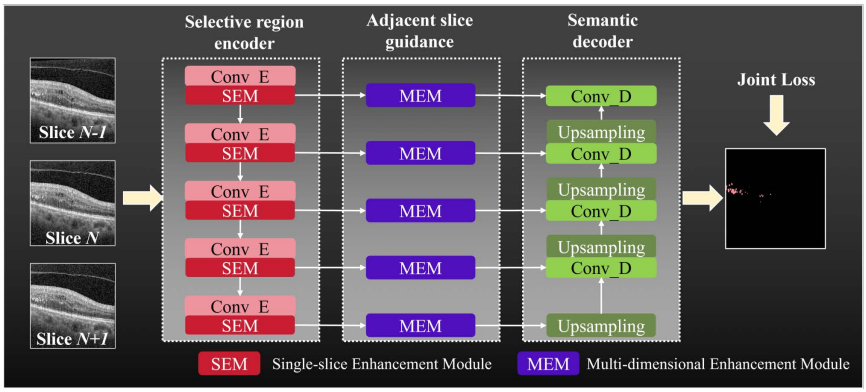

Fig. 1. The overall architecture of the proposed MUSE-Net.

(2) We introduce a joint loss function that combines pixel-level perceptual optimization with constraints on image-level similarity, to improve the segmentation performance on class-imbalance samples.

(3) To promote relevant research in the field of 2D/3D HRF analysis, for the first time, we establish two OCT datasets with HRF regions manually annotated, and they are online available at https://github.com/iMED-Lab/MUSEnet-Pytorch

2 Methodology

The proposed MUSE-Net comprises a selective region encoder, adjacent slice guidance module and a semantic decoder, as shown in Fig. 1. The selective region encoder consists of five layers with convolution followed by SEM, while the semantic decoder has the five symmetric layers with upsampling followed by convolution. The adjacent slice guidance module contains five MEMs located at skip connections between the encoder and the decoder. A joint loss function based on pixel-level perceptual optimization with image-level considerations is applied for network training. These are detailed as follows.

2.1 Multi-dimensional Semantic Information Enhancement

Single-slice Enhancement Module (SEM): The small size and limited information of HRFs in OCT images pose challenges for accurate segmentation. To extract richer semantic information, we introduce an SEM module with region sensitivity. This module softly allocates regional importance based on the distribution of HRF across different areas of the image, thereby enhancing semantic information in single images.

Specifically, SEM comprises two main components: regional representation module (RRM) and regional selection module (RSM), as depicted in Fig. 2. RRM

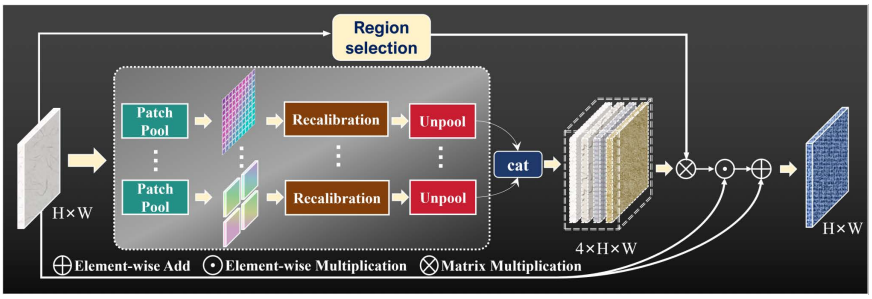

Fig. 2. Schematic diagram of the proposed SEM module.

consists of four paralleled patch-wise squeeze excitation adapters, which aims at capturing regional representations with different sizes. The input feature map $X \in R^{H\times W\times C}$ of each patch-wise squeeze excitation adapter is first squeezed to $S \times S \times C$ by adaptive average pooling, where S are set as 2, 5, 10 and 20 respectively for the four branches. To obtain the significance weight of each patch, the squeezed feature map is then activated through recalibration convolution, which adopts two cascaded convolutional layers. At the end of each branch, unpooling is applied to the activation map to reconstruct to the dimensions of the input feature map. Finally, the outputs of the four branches are concatenated to produce multi-scale regional representation space $P \in R^{C\times 4\times H\times W}$. In order to achieve prediction fusion in a data-driven manner, RSM first employs global average pooling to squeeze the input feature map, then 1×1 convolution with softmax is applied to obtain 4-channel adaptive region-sensitive weights:

$$W = \text{Softmax}(\text{Conv1} \times 1(\text{GAP}(X))), \tag{1}$$

where $W \in R^{4\times 1}$. The final region-sensitive channel attention weights are achieved by applying sigmoid after matrix multiplication of the multi-scale regional representation space P and region-sensitive weights W. After region-sensitive channel attention weighting, the weighted image is combined with the input image via residual connections, which can be expressed as: $X' = X + (P \cdot W) \odot X$.

Multi-dimensional Enhancement Module (MEM): To exploit the multi-dimensional contextual information across consecutive slices, we introduce the MEM inspired by [13]. It comprises two components: channel attention module, and spatial attention module, as shown in Fig. 3. We define remote feature as the combined refined feature maps from adjacent slices, improved by SEM within the same encoder layer, denoted as $X_{rem} = [X'_{N-1}, X'_N, X'_{N+1}] \in R^{3\times C\times H\times W}$, and target feature as $X_{tar} = X'_N \in R^{C\times H\times W}$. Here, N is the sequential number of the target OCT image in the volumetric data. In the channel attention module, patch pool modules squeeze the remote and target features, followed by recalibrated convolutions to activate the pooled features. This is implemented as:

$$F(X) = \text{Conv3} \times 3(\text{RELU}(\text{Conv3} \times 3(\text{GAP}_{\text{S}\times\text{S}}(X)))), \tag{2}$$

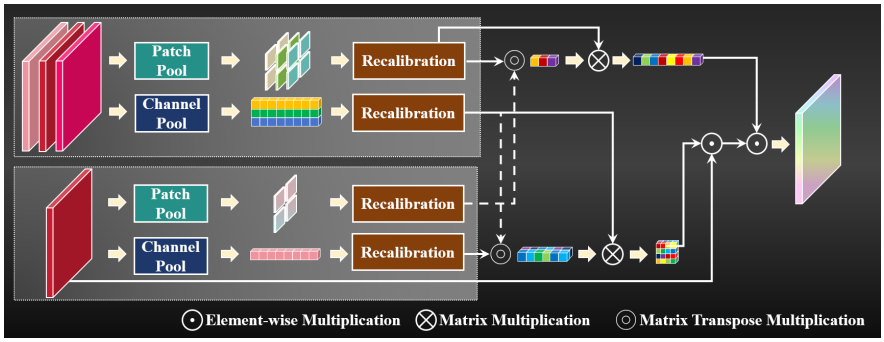

Fig. 3. Schematic diagram of the proposed MEM module.

Here, S is empirically set to 10. Matrix multiplication and softmax operation are employed to obtain inter-slice attention signals:

$$S_{att}^{s} = \text{softmax}(F(X_{tar}) \cdot F(X_{rem})), \tag{3}$$

The inter-slice attention signal is applied to long-range features $F(X_{rem})$, according to $S_{att}^{s} \cdot F(X_{rem})$. An activation function, a 3×3 convolution, unpooling, and normalization are then applied to capture nonlinear interactions among channels.

Similar to channel attention block, spatial attention block adopts channel pool, recalibrated convolution, matrix multiplication, and softmax operations to generate the inter-slice attention signal. Finally, the input undergoes pixel-wise multiplication sequentially with the channel attention map and spatial attention map, yielding a refined feature map enriched with multi-dimensional contextual information.

2.2 Joint loss

In medical image segmentation tasks, imbalance between foreground and background can lead to background regions of large portion dominating the gradient during training. To address this challenge, we propose a joint loss function aimed at mitigating easy sample's gradient dominance during model training.

Pixel-level perceptual optimization loss: We introduce a perceptual optimization loss ($\mathcal{L}_{PO}$) function, using $(1-p)^{\gamma}$ to estimate weights for challenging samples, where p is the predicted probability, and γ is the adjustment factor [18]. Normalizing the loss distribution without altering the sum is achieved with $\mathrm{Z} = \frac{\sum_i (1-p_i)^{\gamma} l(p_i, y_i)}{\sum_i l(p_i, y_i)}$. Here, p_i is the predicted probability, y_i is the label, and $l(p_i, y_i)$ is the cross-entropy loss for class i. To address early inaccurate prediction issues, the loss function is dynamically weighted based on an annealing function:

$$\mathcal{L}_{PO} = \sum_{i} \left[\frac{1}{Z} \cdot (1-p_i)^{\gamma} + \varphi(t) \cdot \left(1 - \frac{1}{Z} \cdot (1-p_i)^{\gamma} \right) \right] l(p_i, y_i), \tag{4}$$

where $\varphi(\cdot)$ represents the annealing function relative to the current training step $t \in [0, 10000]$, and $\varphi(t)$ is a monotonic decreasing function: $\varphi(t) = 0.5 \cdot (1 + \cos(\frac{t}{10000\pi}))$. During the training, confidence in estimating weights for challenging samples improves gradually, allowing the $\mathcal{L}_{PO}$ to assign larger weights to difficult samples.

Image-level consideration loss: Perceptual optimization loss, focused on pixel-level classification, may under-perform in class-imbalanced tasks. In contrast, the Dice coefficient loss proves more robust in such scenarios. Dice coefficient measures the overlap between model predictions and targets, with values closer to 1 indicating better segmentation. The Dice coefficient loss function [4] is expressed as:

$$\mathcal{L}_{Dice} = -\frac{2}{K} \sum_{k=1}^{K} \left[\sum_{i=1}^{I} \frac{u_i^k v_i^k}{\sum_{j=1}^{J} (u_j^k + v_j^k)} \right], \tag{5}$$

where $u \in R^{I \times K}$ is the output of softmax activation, $v \in R^{I \times K}$ is the one-hot encoding of the gold standard, I is the number of pixels in the training batch, and K is the number of classes.

Finally, combining $\mathcal{L}_{Dice}$ and $\mathcal{L}_{PO}$ in a hybrid manner, $\mathcal{L}_{Total} = \mathcal{L}_{Dice} + \mathcal{L}_{PO}$, not only accelerates convergence but also enhances the overall stability of network training.

3 Experiments and results

3.1 Datasets

The segmentation performance of MUSE-Net is evaluated over two in-house OCT image datasets. **HRF-1** consists of 140 OCT volumes acquired using the Heidelberg OCT system (Heidelberg Engineering, Heidelberg), capturing images within $4.5 \times 4.5\ mm^2$ and $6 \times 6\ mm^2$ areas centered at the fovea. **HRF-2** consists of 140 OCT volumes obtained using the SVision OCT system (SVision Imaging), capturing images within a $6 \times 6\ mm^2$ area centered at the fovea. Considering the majority of HRFs spanning 2-4 B-Scans [15], this study randomly selects 8 consecutive B-Scans from each OCT volume for manual annotation. Each dataset yields 1,120 images with HRF segmentation ground truth. A 3-fold cross-validation approach is adopted to evaluate the performance and ensure that OCT images from the same individual volume do not simultaneously appear in the test or training set. The HRF areas were initially annotated by one senior ophthalmologist and later reviewed and refined by another senior ophthalmologist. All images were acquired with regulatory approvals and patient consent as appropriate, following the Declaration of Helsinki.

3.2 Implementation details

Our approach is implemented using the PyTorch framework with NVIDIA GeForce GTX 3090. The model is trained with the SGD optimizer, and the Nesterov momentum term is set to 0.99. The initial learning rate is 0.01, and the model is

Table 1. HRF segmentation results with standard deviations of different methods.

Methods	HRF-1				HRF-2			
	Dice(%)	IOU(%)	Sen(%)	Pre(%)	Dice(%)	IOU(%)	Sen(%)	Pre(%)
U-Net [11]	65.04.4	54.84.3	70.04.7	63.73.8	66.51.0	53.91.1	69.91.5	69.62.5
U-Net++ [19]	65.44.0	54.83.8	72.13.4	63.04.1	66.141.1	53.31.1	70.31.8	68.60.4
Res U-Net [12]	64.94.8	55.14.5	63.55.1	70.14.0	64.51.8	51.81.9	62.93.2	73.11.1
SW-3DUNet [15]	55.34.0	40.94.2	69.02.5	49.34.8	46.54.6	31.73.6	51.27.1	49.98.7
SA-Net [16]	65.13.6	55.23.3	63.84.8	70.31.9	65.50.5	52.80.4	67.91.3	70.01.3
nnUNet [7]	67.82.5	56.71.9	73.26.0	74.08.7	68.61.4	56.11.6	71.70.8	73.92.5
MUSE-Net	**72.41.6**	**60.62.2**	**76.73.5**	**75.11.0**	**73.21.0**	**59.61.3**	**72.71.3**	**77.41.2**

trained for 300 epochs with a batch size of 2. A polynomial learning rate adjustment strategy is employed, with the momentum set to 0.9. The nnUNet [7] is utilized for data augmentation during training.

3.3 Experimental results

To benchmark our model's performance, we compared it with several state-of-the-art methods in the medical image segmentation field, including U-Net [11], U-Net++ [19], Res U-Net [12], SW-3DUNet [15], SA-Net [16], and nn-UNet [7]. We used Dice coefficient (Dice), intersection over union (IoU), precision (Pre), and Sensitivity (Sen) to evaluate the segmentation performance.

Comparison with state-of-the-arts: The quantitative results are reported in Table 1. It can be seen that the proposed method surpasses the state-of-the-art ones in HRF segmentation with large margins. Specifically, our method achieves outstanding performance on the test set for HRF-1, with a Dice of 72.4%, an IOU of 60.6%, a Sen of 76.7% and a Pre of 75.1%, and for HRF-2, with a Dice of 73.2%, an IOU of 59.6%, a Sen of 72.7% and a Pre of 77.4%, respectively.

The proposed method performs well in identifying complete regions of HRF and avoiding false segmentations, as verified in Fig. 4. In contrast, the state-of-the-art methods, such as U-Net [11] and SA-Net [16], over-segment the background inside and outside the retina into HRF. Especially in areas with complex shapes where HRF converges, it is easy to fail to identify them correctly. Both the quantitative and qualitative results demonstrate that the proposed method has achieved state-of-the-art performance in accurately segmenting HRF.

Clinical application: To better demonstrate the clinical value of the proposed method, we show an example of a visual comparison of HRF in a patient with DME undergoing treatment, as shown in Fig. 5. Our method enables the extraction and precise computation of pathological quantification information, including the distribution and quantity characteristics of HRF, such as quantity (Q) and volume (V). This offers crucial insights into the spatio-temporal dynamic changes of lesions before and after treatment.

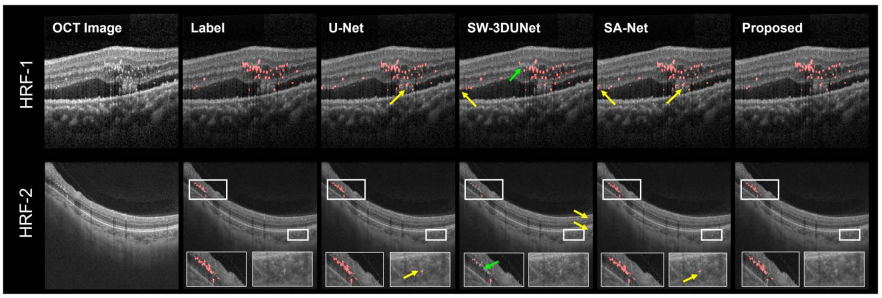

Fig. 4. Illustrative HRF segmentation results. The green arrow indicates under-segmentation, and the yellow arrows are over-segmentation.

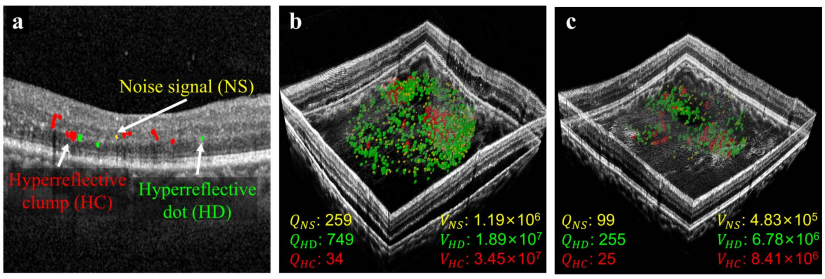

Fig. 5. An example of a visual comparison of HRF and its relevant parameter measurements in a patient with DME before and after the treatment. (a) OCT B-scan with segmentation results. (b-c) 3D visualization of HRF in an OCT volume, at the patient's initial visit and six months later (after treatment). For this case, the number parameters (Q_{NS}, Q_{HD}, Q_{HC}), and volume parameters (V_{NS}, V_{HD}, and V_{HC}) are also calculated based on the segmentation.

Ablation study: To validate the effectiveness of multi-dimensional semantic information enhancement in improving HRF segmentation, we systematically integrate specific modules into the training framework. Sequentially, we introduce M1: joint loss, M2: single-slice semantic enhancement, and M3: multi-dimensional semantic enhancement into the backbone. Table 2 summarizes the experimental results. They show a progressive enhancement in the network performance for HRF segmentation by incorporating different modules and components. The complementary local information, global contextual information and class imbalance handling of these modules further boosts the segmentation performance of the network. Despite minor Sen and Pre coefficients variations on the HRF-1 dataset, the optimal Dice score (72.4%) was achieved when considering both Sen and Pre coefficients.

Table 2. Ablation study of our MUSE-Net for HRF segmentation.

Methods	HRF-1				HRF-2			
	Dice(%)	IOU(%)	Sen(%)	Pre(%)	Dice(%)	IOU(%)	Sen(%)	Pre(%)
Backbone	67.82.5	56.71.9	73.26.0	74.08.7	68.61.4	56.11.6	71.70.8	73.92.5
Backbone + M1	68.91.7	58.42.3	77.42.1	72.11.2	68.90.4	56.40.4	72.10.3	74.20.8
Backbone + M2	69.13.1	58.93.4	**78.52.1**	71.84.1	69.10.7	56.61.1	71.81.7	75.01.5
Backbone + M3	71.52.2	60.12.8	74.83.1	**75.72.3**	71.60.8	58.01.0	72.01.3	75.00.5
MUSE-Net	**72.41.6**	**60.62.2**	76.73.5	75.11.0	**73.21.0**	**59.61.3**	**72.71.3**	**77.41.2**

4 Conclusions

This study introduces a novel approach to tackle the challenges of HRF segmentation in OCT images. It incorporates attention-based multi-dimensional semantic information enhancement modules and class-imbalance-insensitive joint losses for accurate HRF segmentation. The findings indicate that both single-slice and multi-dimensional semantic enhancement modules, along with joint losses, improve HRF segmentations. Compared to the state-of-the-arts, the proposed approach exhibits superior performance. The application of the proposed method to a clinical image also shows significant changes in the relevant parameters that may indicate the treatment efficacy. Nonetheless, further research is needed to validate the effectiveness of the proposed method in large-scale clinical studies.

Acknowledgment This work was supported in part by the National Natural Science Foundation of China (62272444, 62371442, 62103398), Zhejiang Provincial Natural Science Foundation of China (LR22F020008, LZ23F010002, LR24F010002), China Postdoctoral Science Foundation (2023M743629).

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bolz, M., Schmidt-Erfurth, U., Deak, G., Mylonas, G., Kriechbaum, K., Scholda, C., Vienna, D.R.R.G., et al.: Optical coherence tomographic hyperreflective foci: a morphologic sign of lipid extravasation in diabetic macular edema. Ophthalmology **116**(5), 914–920 (2009)
2. Borrelli, E., Grosso, D., Barresi, C., Lari, G., Sacconi, R., Senni, C., Querques, L., Bandello, F., Querques, G.: Long-term visual outcomes and morphologic biomarkers of vision loss in eyes with diabetic macular edema treated with anti-vegf therapy. American Journal of Ophthalmology **235**, 80–89 (2022)
3. Cao, D., Leong, B., Messinger, J.D., Kar, D., Ach, T., Yannuzzi, L.A., Freund, K.B., Curcio, C.A.: Hyperreflective foci, optical coherence tomography progression indicators in age-related macular degeneration, include transdifferentiated retinal pigment epithelium. Investigative Ophthalmology & Visual Science **62**(10), 34–34 (2021)

4. Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C.: The importance of skip connections in biomedical image segmentation. In: International Workshop on Deep Learning in Medical Image Analysis, International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis. pp. 179–187. Springer (2016)
5. Fragiotta, S., Abdolrahimzadeh, S., Dolz-Marco, R., Sakurada, Y., Gal-Or, O., Scuderi, G., et al.: Significance of hyperreflective foci as an optical coherence tomography biomarker in retinal diseases: characterization and clinical implications. Journal of Ophthalmology **2021** (2021)
6. Goel, S., Sethi, A., Pfau, M., Munro, M., Chan, R.V.P., Lim, J.I., Hallak, J., Alam, M.: Automated region of interest selection improves deep learning-based segmentation of hyper-reflective foci in optical coherence tomography images. Journal of Clinical Medicine **11**(24), 7404 (2022)
7. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
8. Klein, R., Lee, K.E., Danforth, L., Tsai, M.Y., Gangnon, R.E., Meuer, S.E., Wong, T.Y., Cheung, C.Y., Klein, B.E.: The relationship of retinal vessel geometric characteristics to the incidence and progression of diabetic retinopathy. Ophthalmology **125**(11), 1784–1792 (2018)
9. Moraes, G., Fu, D.J., Wilson, M., Khalid, H., Wagner, S.K., Korot, E., Ferraz, D., Faes, L., Kelly, C.J., Spitz, T., et al.: Quantitative analysis of oct for neovascular age-related macular degeneration using deep learning. Ophthalmology **128**(5), 693–705 (2021)
10. Okuwobi, I.P., Ji, Z., Fan, W., Yuan, S., Bekalo, L., Chen, Q.: Automated quantification of hyperreflective foci in sd-oct with diabetic retinopathy. IEEE Journal of Biomedical and Health Informatics **24**(4), 1125–1136 (2019)
11. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)
12. Schlegl, T., Bogunovic, H., Klimscha, S., Seeböck, P., Sadeghipour, A., Gerendas, B., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U.: Fully automated segmentation of hyperreflective foci in optical coherence tomography images. arXiv preprint arXiv:1805.03278 (2018)
13. Wang, X., Han, S., Chen, Y., Gao, D., Vasconcelos, N.: Volumetric attention for 3d medical image segmentation and detection. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22. pp. 175–184. Springer (2019)
14. Wei, J., Yu, S., Du, Y., Liu, K., Xu, Y., Xu, X.: Automatic segmentation of hyperreflective foci in oct images based on lightweight dbr network. Journal of Digital Imaging pp. 1–10 (2023)
15. Xie, S., Okuwobi, I.P., Li, M., Zhang, Y., Yuan, S., Chen, Q.: Fast and automated hyperreflective foci segmentation based on image enhancement and improved 3d u-net in sd-oct volumes with diabetic retinopathy. Translational Vision Science & Technology **9**(2), 21–21 (2020)
16. Yao, C., Zhu, W., Wang, M., Zhu, L., Huang, H., Chen, H., Chen, X.: Sanet: a self-adaptive network for hyperreflective foci segmentation in retinal oct images. In: Medical Imaging 2021: Image Processing. vol. 11596, pp. 809–815. SPIE (2021)

17. Zhang, J., Zhang, J., Zhang, C., Zhang, J., Gu, L., Luo, D., Qiu, Q.: Diabetic macular edema: current understanding, molecular mechanisms and therapeutic implications. Cells **11**(21), 3362 (2022)
18. Zheng, Z., Zhong, Y., Wang, J., Ma, A., Zhang, L.: Farseg++: Foreground-aware relation network for geospatial object segmentation in high spatial resolution remote sensing imagery. IEEE Transactions on Pattern Analysis and Machine Intelligence (2023)
19. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. pp. 3–11. Springer (2018)