



OPEN ACCESS

EDITED BY

Yanda Meng,
University of Exeter, United Kingdom

REVIEWED BY

Karthik Karunakaran,
Prediscan Medtech, India
Hao Wang,
Clemson University, United States

*CORRESPONDENCE

Mingguang He
✉ mingguang.he@polyu.edu.hk
Danli Shi
✉ dani.shi@polyu.edu.hk

RECEIVED 19 December 2025

REVISED 09 March 2026

ACCEPTED 16 March 2026

PUBLISHED 07 April 2026

CITATION

Zhao Z, Huang S, Song F, Zhang W, Lu Y, Shang X, Zheng Y, Shi D and He M (2026) Automated interpretation of fundus fluorescein angiography with multi-retinal vascular lesion segmentation. *Front. Med.* 13:1762735. doi: 10.3389/fmed.2026.1762735

COPYRIGHT

© 2026 Zhao, Huang, Song, Zhang, Lu, Shang, Zheng, Shi and He. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Automated interpretation of fundus fluorescein angiography with multi-retinal vascular lesion segmentation

Ziwei Zhao^{1,2}, Shoujin Huang³, Fan Song^{1,2}, Weiyi Zhang^{1,2}, Yan Lu⁴, Xianwen Shang¹, Yingfeng Zheng⁵, Danli Shi^{1,2,*} and Mingguang He^{1,2,6*}

¹School of Optometry, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR, China, ²Research Center for SHARP Vision, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR, China, ³Shenzhen Technology University, Shenzhen, China, ⁴Foshan Second People's Hospital, Foshan, China, ⁵State Key Laboratory of Ophthalmology, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Zhongshan Ophthalmic Center, Guangdong Provincial Clinical Research Center for Ocular Diseases, Sun Yat-sen University, Guangzhou, China, ⁶Center for Eye and Vision Research, Hong Kong, Hong Kong SAR, China

Purpose: Fundus fluorescein angiography (FFA) is essential for diagnosing and managing retinal vascular diseases, yet its evaluation is time-consuming and subject to inter-observer variability. We aim to develop a deep-learning-based framework for multi-lesion segmentation in FFA images and to evaluate its primary performance on standard 55° images.

Methods: A dataset of 428 standard 55° FFA images and 53 ultra-wide-field (UWF) FFA images was annotated for non-perfusion areas (NPA), microaneurysms (MA), neovascularization (NV) and laser spots. A residual U-Net framework was trained using a data-pooling strategy. The primary analysis was performed on the 55° test set, whereas the UWF subset was analyzed only as a preliminary exploratory observation because of its limited sample size. Performance was assessed using Dice score, Intersection over Union (IoU) and recall.

Results: On 55° FFA images, the model achieved Dice scores of 0.65 ± 0.24 for NPA, 0.70 ± 0.13 for MA, 0.73 ± 0.23 for NV, and 0.70 ± 0.17 for laser spots. In an exploratory analysis of UWF images, overlap-based performance was lower for NPA (0.48 ± 0.21 , $p = 0.02$) and MA (0.58 ± 0.19 , $p = 0.01$), whereas laser spot segmentation was similar. NV segmentation in CNV achieved a Dice score of 0.90 ± 0.09 . NPA segmentation was better in RVO than in DR, whereas MA segmentation was better in DR than in RVO. Moreover, NV segmentation was significantly stronger in the venous phase (0.77 ± 0.17) and late phase (0.75 ± 0.24) than in the arteriovenous phase (0.50 ± 0.32 , both $p < 0.05$). Exploratory UWF analysis showed lower overlap-based performance, highlighting the likely impact of domain shift and the need for larger UWF-specific datasets.

Conclusion: This study has established a highly consistent and reproducible framework for multi-lesion segmentation in FFA images. The model may help standardize lesion quantification, reduce manual grading burden, and support future development of larger multi-center FFA datasets.

KEYWORDS

convolutional neural networks, deep learning, lesion segmentation, retinal fluorescein angiography, retinal vascular disease

Introduction

Retinal vascular diseases, including diabetic retinopathy (DR), retinal vein occlusion (RVO) with its subtype branch retinal vein occlusion (BRVO), and neovascular age-related macular degeneration (nAMD) coupled with choroidal neovascularization (CNV), pose a substantial global health challenge. As leading causes of vision loss and blindness, these diseases affect millions and impose considerable personal and socioeconomic costs (1). Specifically, DR prevalence in diabetics is 27.0%, contributing to 0.4 million cases of blindness (2), and the projected incidence of AMD by 2040 is 288 million, with about 10% expected to develop neovascular complications (3).

Fundus fluorescein angiography (FFA) is a critical diagnostic tool to detect and characterize retinal vascular abnormalities such as microaneurysms (MA), neovascularization (NV), and capillary non-perfusion areas (NPA). Segmenting and quantifying these lesions are critical for accurate diagnosis, disease staging, and monitoring. For instance, in DR, the extent of NPA indicates the severity of ischemia and predicts the risk of disease progression, thereby determining the necessity for laser photocoagulation (4). Similarly, measuring MAs provides valuable prognostic information (5). In DR, MA turnover rates serve as key indicators for disease progression (6–8). While historically overlooked in RVO, recent studies link MA to persistent macular edema (ME) in BRVO (9, 10). Notably, MAs in collateral vessels independently predict retinal edema (11). Thus, precise identification of these biomarkers helps identify patients at high risk of vision-threatening complications (12).

However, accurate interpretation of FFA images is inherently complex and subject to inter-observer variability (13). While precise quantification of lesions (e.g., measuring the exact area of non-perfusion) holds critical prognostic value for risk stratification, it is rarely performed in daily practice because manual delineation is prohibitively time-consuming (14). An automated segmentation system addresses these challenges by enabling highly consistent and reproducible quantification without adding to the clinical burden, thereby streamlining the workflow for busy clinicians (15). Furthermore, it may assist less experienced ophthalmologists by highlighting subtle or scattered lesions and reducing the risk of missed findings.

Recent work has explored deep learning (DL) for DR grading and lesion detection (16). Pan et al.'s study achieved multi-label detection in DR using FFA, with sensitivities of 79.7% for NPA, 98.0% for MA, 84.0% for leakage, and 80.3% for laser scars (17). However, their focus has predominantly been on image-level lesion classification without specific localization or delineation. Deep learning approaches, including convolutional neural networks (CNNs), have shown promise in segmenting FFA lesions, as summarized in Table 1. These studies DLs potential or the potential of deep learning for accurate lesion identification in FFA images of specific disease datasets (15, 18–25). Nevertheless, limitations exist: research has been focused only on 30°, 55° or ultra-widefield (UWF) FFA images, targeting specific lesions within particular retinal diseases such as NPA in DR (18, 19, 24), NPA in BRVO (20, 22), and CNV in nAMD (25). This narrows their application and prevents the identification of multiple lesions across a range

of diseases. Meanwhile, as DL models advance in accuracy, there is an accompanying surge in the volume of data required to train, validate, and test these models.

To address these gaps, we assembled a pixel-level annotated FFA dataset containing NPA, MA, NV, and post-photocoagulation laser spots across multiple retinal vascular conditions. Rather than claiming architectural novelty, the main novelty of this study lies in its clinical task formulation: a unified framework for concurrently segmenting multiple morphologically distinct lesions across DR, RVO, and CNV. In addition, to our knowledge, this is the first study to explicitly incorporate laser-spot segmentation as a dedicated class, which is clinically relevant for estimating the true ischemic burden in previously treated eyes. Overall, we propose an automated, highly consistent and reproducible approach for concurrent lesion segmentation in FFA images.

Materials and methods

The flow chart of the study is shown in Figure 1.

Dataset

The dataset was collected from multiple tertiary hospitals and comprised 428 55° and 53 UWF images from 481 eyes of 481 patients. Standard 55° images were captured using Zeiss FF450 Plus (Carl Zeiss Meditec, North America) and Heidelberg Spectralis (Heidelberg Engineering, Heidelberg, Germany) cameras at a resolution of 768 × 768 pixels. UWF images were captured by the Optos California (Optos plc, Dunfermline, Scotland, UK) at a higher resolution of 1,536 × 1,024 pixels. Because the UWF subset was small, it was retained only as a preliminary exploratory cohort and was not used to support formal claims of cross-modality generalizability. The dataset included DR, CRVO, BRVO, nAMD, CNV, normal eye, and other diseases, and covered arteriovenous, venous, and late phases. Images with moderate obscuration, including regional masking by hemorrhage, were retained to better reflect real-world clinical practice. Only images of extremely poor quality—where retinal structures were nearly indiscernible and lacked clinical reference value—were excluded. All patient data were anonymized and de-identified following the Declaration of Helsinki. Individual consent was waived due to the retrospective nature and the thorough anonymization process. The study was approved by the Institutional Review Board of the Hong Kong Polytechnic University (HSEARS20240301004).

Labeling and image preparation

Two experienced ophthalmologists (Z. Z. and F. S.) annotated the dataset using Labelme (Ver 3.16.2) independently. For each image, graders used the polygon tool to outline the optic disc, macular region, non-perfusion areas (NPA), neovascularization (NV), and laser spots (Laser), while the point tool was applied to identify microaneurysms (MA). They also determined the most

TABLE 1 Summary of previous studies on lesion segmentation performance in FFA images.

Year	Author	Modality	Target	Disease	Test model	Best F1 score
2020	Joan et al.	UWF	NPA	DR	U-net	0.661
2020	Jin et al.	30 °	NPA	DME	U-net	N/A
2021	Tang et al.	55 °	NPA	RVO	CE-net, U-Net, DeepLabv3+	0.883
2021	Chen et al.	55 °	Leakage	CSC	attention gated network	0.811
2022	Kanato et al.	55 °	NPA	RVO	U-Net, PSPNet, DeepLab v3+	0.736
2022	Xiang et al.	55 °	NPA	Not Specified	ResNet50+VGG16+MLFB+SAT	0.772
2022	Lee et al.	UWF	NPA NV	DR	DeepLab v3+, SegNet, U-net	0.67 0.87
2023	David et al.	55 °	NV Leakage	nAMD	U-net based	0.73 0.65
2023	Zhao et al.	55 °	NPA	DR & BRVO	ResNet-152, Unet-VGG16	0.90

FFA, fundus fluorescein angiography; UWF, ultra-widefield imaging; DR, diabetic retinopathy; RVO, retinal vein occlusion; BRVO, branch retinal vein occlusion; CSC, central serous chorioretinopathy; nAMD, neovascular age-related macular degeneration; NPA, non-perfusion areas; NV, neovascularization.

likely diagnosis and the phase of the image. NPAs were defined as regions exceeding the physiological intercapillary distance yet devoid of fluorescein filling. For NV, the annotation target included the hyperfluorescent neovascular complex together with its associated leakage cloud when present, because progressive leakage is a defining angiographic feature of active neovascularization. The laser-spot class predominantly comprised old pan-retinal photocoagulation scars, which were grouped as a single category because they shared the same clinical meaning of previously treated, non-functioning retinal tissue.

Following Joan M. Nunez do Rio et al., non-informative peripheral areas in UWF images were excluded by defining a ‘gradable area’—a zone on UWF-FFA images where lesions and anatomical structures are discernible (18). The initial inter-grader Dice score was 0.69. Subsequently, the ground truth (GT) was established through consensus between the 2 graders and confirmed by a retinal specialist.

Before feeding each input image into the network, we performed min–max normalization to scale pixel intensities to the range of 0 to 1. UWF images were zero-padded to a square canvas before resizing to the network input size, rather than being directly warped, whereas standard 55 ° images were resized after normalization. Data augmentation methods included Gaussian modifications, brightness and contrast adjustments, resolution reduction, and gamma shifts, with assigned probabilities of 10%, 15%, 25%, and 10%, respectively. Random horizontal and vertical flips were also applied to both the images and labels. These augmentation steps were intended to improve robustness while preserving clinically relevant lesion morphology.

Network architecture

Our network architecture was built on the U-net framework (26), incorporating residual units (27), in both the encoding and decoding stages as shown in Figure 1. In the encoder, input features were down sampled through 3 x 3 convolutions with a stride of 2 and then passed through repeated residual blocks, with shortcut summation between the block input and output. In the

decoder, features from the skip connections were concatenated with incoming features before repeated residual processing, and the encoded representation was progressively restored.

Instance normalization and ReLU activation were applied after all convolutional operations. The channel numbers were 32, 64, 128, 256, and 320. Three residual units were used at each up sampling/down sampling stage. The network was configured with 1 input channel and 10 output channels. For the segmentation task, the five target classes were modeled as five binary segmentation tasks.

Training and testing details

During training, the 55 ° and UWF images were combined using a data-pooling strategy rather than a dedicated domain-adaptation module. The primary objective of this design was to learn a unified representation from the larger 55 ° dataset while allowing limited exposure to UWF images. Given the small size of the UWF cohort, its role was exploratory and its results were interpreted cautiously. We employed the SGD optimizer with the Poly learning rate decay algorithm. The learning rate was set to 0.01, with a weight decay of 3e-5, batch size of 1, and momentum of 0.99. The U-net model was trained for 500 epochs, and the final model was selected according to the best mean Dice score on the validation set. The loss function combined Dice loss and cross-entropy loss. Loss was computed separately for each category and then summed for backpropagation. Training was performed on a Tesla V100 with an Intel(R) Xeon(R) Gold 6248R CPU (Intel Corporation, Santa Clara, CA, USA).

Performance metrics

For image inference, we implemented a testing augmentation strategy, involving random horizontal and vertical flips applied to the input image into the network to obtain diverse outcomes, respectively. Subsequently, the predicted probability values were averaged to derive the final segmentation result.

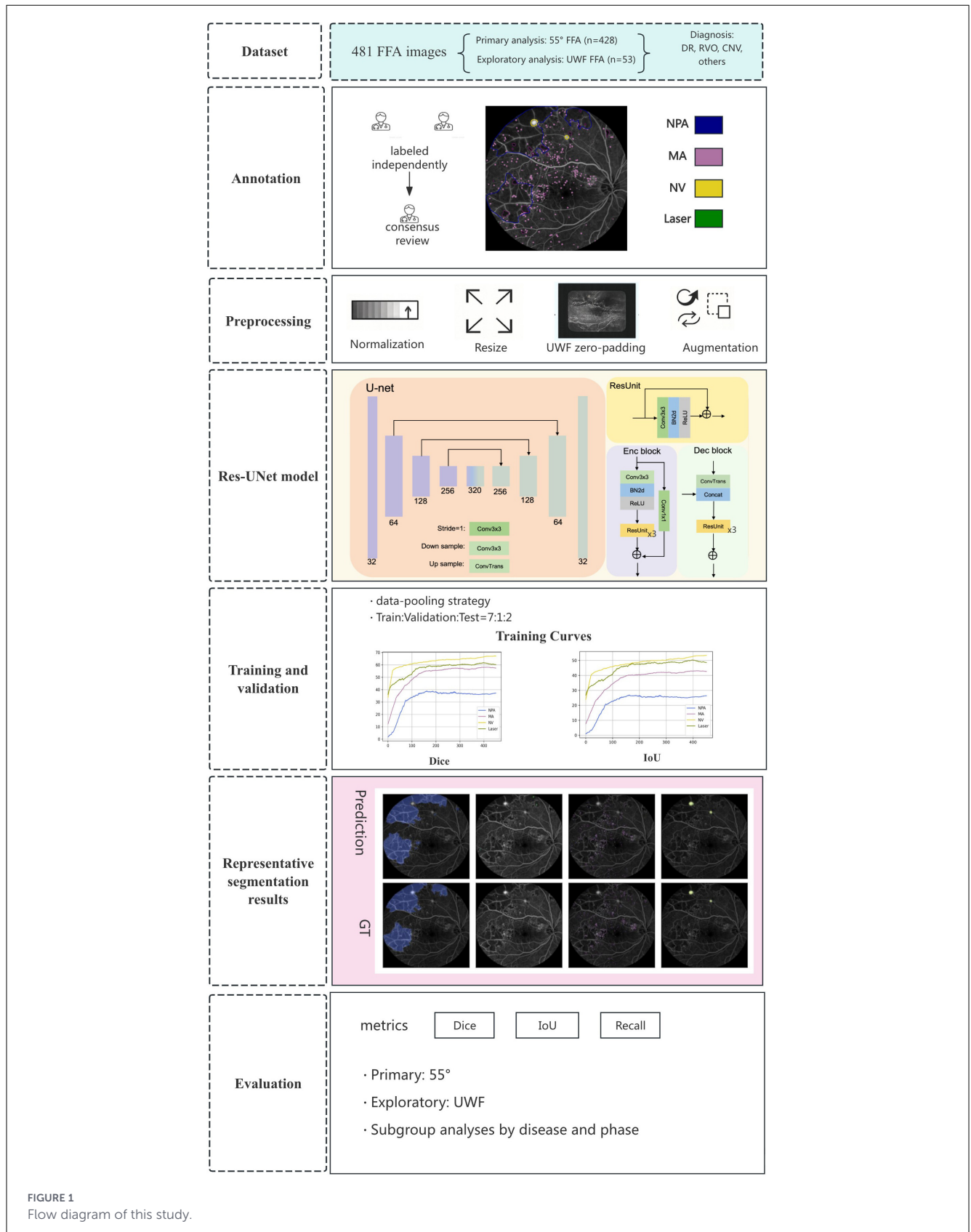


FIGURE 1
 Flow diagram of this study.

The performance of the model was quantitatively evaluated using the Dice score and the Intersection over Union (IoU). Pixel-level sensitivity (recall) was additionally considered as a

complementary detection-oriented metric, because overlap-based metrics alone may underrepresent clinical usefulness when a lesion is detected but its boundary is imperfectly delineated.

The Dice score, also known as the F1 score in binary segmentation, measures the similarity between the predicted ground-truth masks and is defined as:

$$DSC = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (1)$$

Here, X and Y represent the predicted and ground truth pixel sets, respectively. The Dice coefficient ranges from 0 (no overlap) to 1 (perfect agreement).

The IoU metric, also known as the Jaccard index, evaluates the overlap between the predicted segmentation and the ground truth masks and is defined as:

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} \quad (2)$$

Similarly, X and Y denote the predicted and ground truth pixel sets. IoU values range from 0 to 1, with 1 indicating perfect overlap.

Statistical analysis

Statistical analyses were performed using R software. Data are presented as mean \pm standard deviation (SD). Group comparisons were conducted using two-tailed independent t -tests, with statistical significance set at $p < 0.05$.

Results

Dataset compilation

Our dataset comprised 428 55° FFA images and 53 UWF-FFA images, totaling 481 FFA images. For a focused study, we prioritized the 55° images due to their larger quantity, while also including UWF images for preliminary model performance evaluation on a smaller sample size. For the 55° FFA images, 305 were allocated for training, 38 for validation, and 85 for testing. Correspondingly, the UWF FFA images were divided into 40 for training, 3 for validation, and 10 for testing.

Distribution of diseases

The dataset comprised diverse retinal pathologies, including 296 images of DR, 29 of CRVO, 57 of BRVO, 58 of nAMD or CNV, 16 of normal fundus, and 25 of other conditions requiring further examination.

Label frequency

The comprehensive distribution of label frequencies and pixel counts is detailed in [Table 2](#). Regarding lesion density, the mean number of labels per image was 3.2 for NPA, 120.5 for MA, 0.95 for NV, and 26.5 for Laser.

Primary performance on 55° FFA images

Model performance metrics, including the Dice, IoU and recall scores for 55° , UWF FFA, and the overall dataset are presented in [Table 3](#). The detection of NPA, MA, and NV on 55° FFA images yielded Dice scores of 0.65 ± 0.24 , 0.70 ± 0.13 , and 0.73 ± 0.23 , respectively. For laser spots, a target not previously emphasized in FFA segmentation studies, the Dice score was 0.70 ± 0.17 .

Exploratory results on UWF-FFA images

Given the very limited UWF test set ($n = 10$), these results should be interpreted as preliminary exploratory observations rather than robust evidence of cross-modality generalizability. Overlap-based performance for NPA, MA, and NV was lower on UWF than on 55° images ([Table 3](#), [Supplementary Figure S6](#)), with Dice scores of 0.48 ± 0.21 ($p = 0.02$), 0.58 ± 0.19 ($p = 0.01$), and 0.50 ± 0.34 ($p = 0.14$), respectively. This pattern likely reflects domain shift related to peripheral distortion, image geometry, and limited UWF training exposure. Laser performance remained numerically similar, but this finding should also be interpreted cautiously because of the small sample size.

Results by disease type

In DR, the Dice scores for NPA, MA, NV, and Laser were 0.59 ± 0.22 , 0.73 ± 0.09 , 0.66 ± 0.25 , and 0.70 ± 0.11 , respectively. In RVO, the corresponding Dice scores were 0.77 ± 0.25 , 0.53 ± 0.20 , 0.71 ± 0.29 , and 0.70 ± 0.31 . Additionally, in CNV, the Dice score for NV was notably high at 0.90 ± 0.09 . [Table 4](#) details performance for each label by disease. [Figure 2](#) highlights the differences in Dice scores, demonstrating a significantly better segmentation performance for NPA in RVO compared to DR, and a superior performance for MA segmentation in DR over RVO ($p < 0.01$). Regarding NV segmentation, performance in CNV was superior to DR and numerically higher than in RVO; only the comparison with DR reached statistical significance ($p = 0.02$).

Results by phase

As illustrated in [Figure 3](#), segmentation of NPA, MA, and Laser showed no significant differences across arteriovenous, venous, and late phases. In contrast, NV segmentation was significantly stronger in the venous phase (Dice = 0.77 ± 0.17 , $p = 0.01$) and late phase (Dice = 0.75 ± 0.24 , $p = 0.02$) than in the arteriovenous phase (Dice = 0.50 ± 0.32). Detailed model performance metrics by phase are provided in [Table 5](#).

Qualitative evaluation

Representative examples of model predictions and ground-truth masks for different labels in 55° FFA images are shown in [Supplementary Figures S1–S5](#). Overall, the segmentation

TABLE 2 Distribution of label frequency and pixel counts in 55° and UWF FFA images.

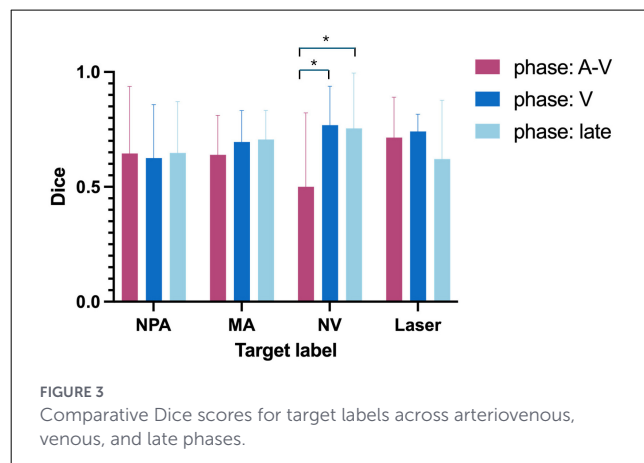
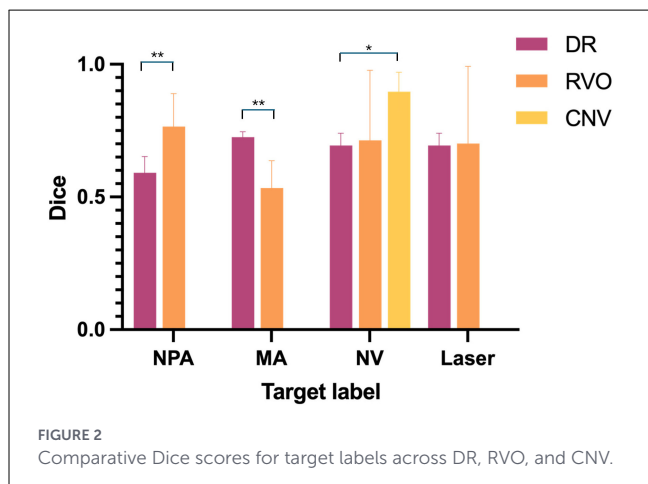
Label name	Frequency counts			Pixel counts		
	55°	UWF	All	55°	UWF	All
NPA	1,024	503	1,527	20,946,340	8,731,017	29,677,357
MA	41,741	16,234	57,975	1,519,689	486,512	2,006,201
NV	359	100	459	2,305,044	388,791	2,693,835
Laser	9,509	3,253	12,762	5,659,824	1,257,354	6,917,178

TABLE 3 Model performance metrics on 55° FFA, UWF FFA, and the overall dataset.

Label name	55°			UWF			All		
	Dice	IoU	Recall	Dice	IoU	Recall	Dice	IoU	Recall
NPA	0.65 ± 0.24	0.53 ± 0.24	0.74 ± 0.23	0.48 ± 0.21	0.34 ± 0.17	0.47 ± 0.21	0.64 ± 0.24	0.51 ± 0.24	0.71 ± 0.25
MA	0.70 ± 0.13	0.55 ± 0.14	0.79 ± 0.12	0.58 ± 0.19	0.43 ± 0.17	0.70 ± 0.11	0.69 ± 0.14	0.54 ± 0.14	0.78 ± 0.12
NV	0.73 ± 0.23	0.62 ± 0.26	0.88 ± 0.18	0.50 ± 0.34	0.38 ± 0.30	0.79 ± 0.16	0.72 ± 0.25	0.61 ± 0.26	0.87 ± 0.18
Laser	0.70 ± 0.17	0.56 ± 0.17	0.72 ± 0.20	0.74 ± 0.03	0.59 ± 0.04	0.80 ± 0.09	0.70 ± 0.17	0.56 ± 0.16	0.73 ± 0.20

TABLE 4 Model performance metrics by disease category.

Label name	DR			RVO			CNV		
	Dice	IoU	Recall	Dice	IoU	Recall	Dice	IoU	Recall
NPA	0.59 ± 0.22	0.45 ± 0.20	0.68 ± 0.23	0.77 ± 0.25	0.67 ± 0.28	0.81 ± 0.27	---	---	---
MA	0.73 ± 0.09	0.55 ± 0.10	0.79 ± 0.09	0.53 ± 0.20	0.43 ± 0.17	0.74 ± 0.18	---	---	---
NV	0.66 ± 0.25	0.54 ± 0.26	0.82 ± 0.20	0.71 ± 0.29	0.61 ± 0.29	0.97 ± 0.02	0.90 ± 0.09	0.82 ± 0.14	0.97 ± 0.08
Laser	0.70 ± 0.11	0.54 ± 0.11	0.71 ± 0.15	0.70 ± 0.31	0.60 ± 0.28	0.74 ± 0.33	---	---	---

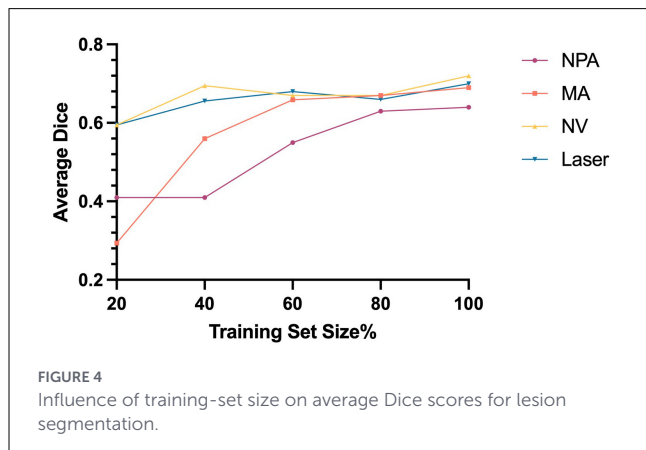


performance was satisfactory across lesion types. NPA recognition is particularly strong, especially in BRVO cases. MA predictions generally aligned with ground truth, although false positives occurred in some cases, such as mistaking parts of laser spots with window defects for MA or interpreting punctate hyperfluorescence from BRVO depigmentation as MA. The model accurately identified typical NV, but performance decreased in early-stage

lesions or when leakage was subtle, occasionally confusing prominent vascular leakage with NV. Laser spot segmentation was largely satisfactory; however, lower Dice scores were mainly observed in atypical cases, and some predictions correctly identified laser spots that had been missed in the original ground truth. In UWF examples (Supplementary Figure S5), NPA detection was less accurate in the peripheral retina, and MA/NV performance was

TABLE 5 Model performance metrics by phase.

Label	A-V			V			late		
name	Dice	IoU	Recall	Dice	IoU	Recall	Dice	IoU	Recall
NPA	0.65 ± 0.29	0.53 ± 0.29	0.77 ± 0.22	0.63 ± 0.23	0.49 ± 0.23	0.68 ± 0.25	0.65 ± 0.22	0.51 ± 0.23	0.72 ± 0.26
MA	0.64 ± 0.17	0.49 ± 0.16	0.72 ± 0.16	0.70 ± 0.14	0.55 ± 0.14	0.78 ± 0.11	0.71 ± 0.13	0.56 ± 0.14	0.82 ± 0.10
NV	0.50 ± 0.32	0.39 ± 0.32	0.78 ± 0.26	0.77 ± 0.17	0.65 ± 0.21	0.86 ± 0.18	0.75 ± 0.24	0.65 ± 0.26	0.92 ± 0.11
Laser	0.71 ± 0.18	0.58 ± 0.19	0.75 ± 0.16	0.74 ± 0.08	0.59 ± 0.09	0.76 ± 0.15	0.62 ± 0.26	0.49 ± 0.22	0.64 ± 0.28



inferior to that on 55 ° images, further supporting the exploratory nature of the UWF analysis.

Ablation study: impact of training set size on model performance

To evaluate the impact of training-set size, we further divided the original training set into subsets, containing 20%, 40%, 60%, and 80% of the full dataset. As depicted in Figure 4, average model performance for the labels NPA, MA, NV, and Laser consistently improved as the amount of the training data increased.

Discussion

In this work, we developed a unified deep learning framework for simultaneous segmentation of NPA, MA, NV, and laser spots across multiple retinal vascular diseases. Rather than proposing a novel backbone architecture, the key novelty of this study lies in the clinical task formulation: a single framework that addresses multiple morphologically distinct lesions across DR, RVO, and CNV. To our knowledge, this is also the first study to explicitly segment post-photocoagulation laser spots as a dedicated class, which is clinically important for reducing false positives and estimating the true residual ischemic burden in treated eyes. On standard 55 ° images, the model achieved overall satisfactory performance and approached the range of human inter-grader agreement.

The UWF findings should be interpreted conservatively. Because the exploratory UWF cohort included only 53 images, with 10 test images, it is statistically insufficient to support strong claims of cross-modality generalizability. We therefore repositioned UWF as a preliminary exploratory analysis rather than as a major validation arm. The lower UWF Dice scores, particularly for NPA, most likely reflect a combination of peripheral distortion, resolution and aspect-ratio differences, and limited modality-specific training data. These findings are better understood as evidence of domain shift and as justification for future UWF-specific or multi-center expansion, rather than as proof of deployable UWF performance.

When comparing our results with highly specialized prior studies, our overall 55 ° NPA Dice score of 0.65 is lower than the best disease-specific binary models. This difference is likely explained by three factors. First, Tang et al. focused exclusively on RVO, in which NPA often has clearer and more continuous boundaries than the patchy, diffuse NPA frequently seen in DR (18–20, 22, 24). Second, our dataset was intentionally not over-curated and retained clinically realistic obscuration such as hemorrhagic masking. Third, our framework jointly optimizes multiple lesion classes, which introduces an inherent performance trade-off compared with single-task models. It is worth noting that we grouped retinal and choroidal neovascularization under a single ‘NV’ category. Despite biological differences, they share key angiographic features—hyperfluorescence and leakage—allowing the model to successfully extract shared visual representations. Regarding phase-dependent performance, the superior NV performance in venous and late phases is consistent with our annotation protocol, because the NV label included associated fluorescein leakage rather than only the vascular scaffold itself. Regarding NPA quantification, we acknowledge that the arteriovenous phase remains the clinical gold standard due to optimal contrast, and late-phase measurements were included primarily to test model stability against leakage artifacts. Our ablation study confirms that larger training sets correlate with higher performance, a trend supported by Zhao et al., who achieved superior NPA segmentation with the largest dataset of 1,295 images (15).

Establishing a definitive ground truth for FFA segmentation is inherently challenging due to subjective interpretation. Inter-grader variability is a well-documented issue; for instance, Joan et al. reported a Dice score of 0.65 for NPA between experts, and Lee et al. noted similarly variable IoUs (18, 24). Consistent with these findings, our internal assessment revealed an inter-grader Dice score of 0.69. This creates a ceiling effect for supervised learning:

when experts do not completely agree on the precise boundary of diffuse or microscopic lesions, the model cannot be expected to consistently exceed the consensus it is trained on. In this context, the model's performance suggests that it is approximating expert consensus rather than learning an absolute biological truth. Therefore, rather than creating a new ground truth, the model offers value by providing a standardized, reproducible baseline. This capability is crucial for mitigating the inconsistency inherent in manual grading and reducing the labor burden.

Clinically, a consistent segmentation system could reduce manual grading burden, support less experienced readers, and facilitate more standardized quantification of ischemic and neovascular lesions. Quantification of NPA is particularly relevant for risk stratification and treatment planning. In CRVO, eyes with NPAs exceeding 10-disc areas on 55° FFA images are at a higher risk of developing anterior segment NV (14), while an initial NPA over 75-disc areas on UWF FFA images greatly increases the chances of progressing to ischemic CRVO within the 1 year (28, 29). By enabling more reproducible lesion quantification, automated segmentation may help identify eyes that warrant closer surveillance or earlier intervention.

Our study has several limitations. First, the strongest evidence in this manuscript comes from the 55° dataset; the UWF cohort was too small for robust validation and is reported only as an exploratory observation. Second, the lack of external validation may limit generalizability. Third, incorporating paired color fundus images could better differentiate hemorrhage from NPA. Future work should expand the dataset across centers and camera platforms and may benefit from semi-supervised learning, active learning, and federated learning to reduce annotation burden while improving robustness.

In conclusion, our study demonstrates the feasibility of a deep learning-based framework for multi-lesion segmentation in FFA images across a variety of retinal vascular diseases, with the most robust evidence derived from standard 55° images. The model exhibits overall satisfactory performance for NPA, MA, NV, and laser spots and provides a highly consistent and reproducible approach to lesion quantification. With broader external validation and larger multi-center datasets, this framework may support more standardized FFA interpretation and improve risk stratification in retinal vascular disease.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of the Hong Kong Polytechnic University (HSEARS20240301004). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal

guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

ZZ: Writing – original draft, Writing – review & editing. SH: Methodology, Software, Writing – review & editing. FS: Methodology, Validation, Writing – review & editing. WZ: Methodology, Software, Writing – review & editing. YL: Resources, Writing – review & editing. XS: Writing – review & editing, Methodology. YZ: Resources, Writing – review & editing. DS: Funding acquisition, Project administration, Supervision, Writing – review & editing. MH: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the PolyU-HuiKang Joint Laboratory for Biotech Innovation (P0052690), the Global STEM Professorship Scheme (P0046113), the Start-up Fund for RAPs under the Strategic Hiring Scheme (P0048623) from HKSAR, and the Henry G. Leong Endowed Professorship in Elderly Vision Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We thank the InnoHK HKSAR Government for providing valuable support. Danli Shi and Mingguang He disclose support for the research and publication of this work from the PolyU-HuiKang Joint Laboratory for Biotech Innovation (P0052690), the Start-up Fund for RAPs under the Strategic Hiring Scheme from HKSAR (P0048623), the Global STEM Professorship Scheme (P0046113), and Henry G. Leong Endowed Professorship in Elderly Vision Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of

artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2026.1762735/full#supplementary-material>

References

- Eckert KA, Carter MJ, Lansingh VC, Wilson DA, Furtado JM, Frick KD, et al. A simple method for estimating the economic cost of productivity loss due to blindness and moderate to severe visual impairment. *Ophthalmic Epidemiol.* (2015) 22:349–55. doi: 10.3109/09286586.2015.1066394
- Thomas RL, Halim S, Gurudas S, Sivaprasad S, Owens DR. IDF diabetes atlas: a review of studies utilising retinal photography on the global prevalence of diabetes related retinopathy between 2015 and 2018. *Diabetes Res Clin Pract.* (2019) 157:107840. doi: 10.1016/j.diabres.2019.107840
- Wong WL, Su X, Li X, Cheung CM, Klein R, Cheng CY, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health.* (2014) 2:e106–16. doi: 10.1016/S2214-109X(13)70145-1
- Nicholson L, Ramu J, Chan EW, Bainbridge JW, Hykin PG, Talks SJ, et al. Retinal nonperfusion characteristics on ultra-widefield angiography in eyes with severe nonproliferative diabetic retinopathy and proliferative diabetic retinopathy. *JAMA Ophthalmol.* (2019) 137:626–31. doi: 10.1001/jamaophthalmol.2019.0440
- Dubow M, Pinhas A, Shah N, Cooper RF, Gan A, Gentile RC, et al. Classification of human retinal microaneurysms using adaptive optics scanning light ophthalmoscope fluorescein angiography. *Invest Ophthalmol Vis Sci.* (2014) 55:1299–309. doi: 10.1167/iovs.13-13122
- Akram MU, Khalid S, Khan S. A Identification and classification of microaneurysms for early detection of diabetic retinopathy. *Pattern Recognit.* (2013) 46:107–16. doi: 10.1016/j.patcog.2012.07.002
- Pappuru RK, Ribeiro L, Lobo C, Alves D, Cunha-Vaz J. Microaneurysm turnover is a predictor of diabetic retinopathy progression. *Br J Ophthalmol.* (2019) 103:222–6. doi: 10.1136/bjophthalmol-2018-311877
- Xu J, Zhang X, Chen H, Li J, Zhang J, Shao L, et al. Automatic analysis of microaneurysms turnover to diagnose the progression of diabetic retinopathy. *IEEE Access.* (2018) 6:9632–42. doi: 10.1109/ACCESS.2018.2808160
- Hasegawa T, Kawano T, Maruko I, Koizumi H, Iida T. Clinical findings of eyes with macular edema associated with branch retinal vein occlusion refractory to ranibizumab. *Retina.* (2018) 38:1347–53. doi: 10.1097/IAE.0000000000001703
- lesato Y, Imai A, Hirano T, Toriyama Y, Murata T. Effect of leaking capillaries and microaneurysms in the perifoveal capillary network on resolution of macular edema by anti-vascular endothelial growth factor treatment. *Jpn J Ophthalmol.* (2016) 60:86–94. doi: 10.1007/s10384-016-0425-5
- Esaki Y, Hirano Y, Kurobe R, Yasuda Y, Tomiyasu T, Suzuki N, et al. Morphologic classifications and locations of microaneurysms and clinical relevance in branch retinal vein occlusion. *Clin Ophthalmol.* (2020) 14:1909–19. doi: 10.2147/OPHT.S258958
- Fernández-Carneado J, Almazán-Moga A, Ramírez-Lamelas DT, Cuscó C, Alonso de la Fuente JJ, Pastor JC, et al. Quantification of microvascular lesions in the central retinal field: could it predict the severity of diabetic retinopathy? *J Clin Med.* (2023) 12:3948. doi: 10.3390/jcm12123948
- Resnikoff S, Felch W, Gauthier TM, Spivey B. The number of ophthalmologists in practice and training worldwide: a growing gap despite more than 200 000 practitioners. *Br J Ophthalmol.* (2012) 96:783–7. doi: 10.1136/bjophthalmol-2011-301378
- A randomized clinical trial of early panretinal photocoagulation for ischemic central vein occlusion. The central vein occlusion study group N report. *Ophthalmol.* (1995) 102:1434–44. doi: 10.1016/S0161-6420(95)30848-2
- Zhao X, Lin Z, Yu S, Xiao J, Xie L, Xu Y, et al. An artificial intelligence system for the whole process from diagnosis to treatment suggestion of ischemic retinal diseases *Cell Rep Med.* (2023) 4:101197. doi: 10.1016/j.xcrm.2023.101197
- Gao Z, Jin K, Yan Y, Liu X, Shi Y, Ge Y, et al. End-to-end diabetic retinopathy grading based on fundus fluorescein angiography images using deep learning. *Graefes Arch Clin Exp Ophthalmol.* (2022) 260:1663–73. doi: 10.1007/s00417-021-05503-7
- Pan X, Jin K, Cao J, Liu Z, Wu J, You K, et al. Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning. *Graefes Arch Clin Exp Ophthalmol.* (2020) 258:779–85. doi: 10.1007/s00417-019-04575-w
- Nunez do Rio JM, Sen P, Rasheed R, Bagchi A, Nicholson L, Dubis AM, et al. Deep learning-based segmentation and quantification of retinal capillary non-perfusion on ultra-wide-field retinal fluorescein angiography. *J Clin Med.* (2020) 9:2537. doi: 10.3390/jcm9082537
- Jin K, Pan X, You K, Wu J, Liu Z, Cao J, et al. Automatic detection of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography for decision making using deep learning. *Sci Rep.* (2020) 10:15138. doi: 10.1038/s41598-020-71622-6
- Tang Z, Zhang X, Yang G, Zhang G, Gong Y, Zhao K, et al. Automated segmentation of retinal nonperfusion area in fluorescein angiography in retinal vein occlusion using convolutional neural networks. *Med Phys.* (2021) 48:648–58. doi: 10.1002/mp.14640
- Chen M, Jin K, You K, Xu Y, Wang Y, Yip CC, et al. Automatic detection of leakage point in central serous chorioretinopathy of fundus fluorescein angiography based on time sequence deep learning. *Graefes Arch Clin Exp Ophthalmol.* (2021) 259:2401–11. doi: 10.1007/s00417-021-05151-x
- Masayoshi K, Katada Y, Ozawa N, Ibuki M, Negishi K, Kurihara T, et al. Automatic segmentation of non-perfusion area from fluorescein angiography using deep learning with uncertainty estimation. *Inform Med.* (2022) 32:101060. doi: 10.1016/j.imu.2022.101060
- Xiang D, Yan S, Guan Y, Cai M, Li Z, Liu H, et al. Semi-supervised dual stream segmentation network for fundus lesion segmentation. *IEEE Trans Med Imaging.* (2022) 42:713–25. doi: 10.1109/TMI.2022.3215580
- Lee PK, Ra H, Baek J. Automated segmentation of ultra-widefield fluorescein angiography of diabetic retinopathy using deep learning. *Br J Ophthalmol.* (2023) 107:1859–63. doi: 10.1136/bjo-2022-321063
- Holomic D, Seeböck P, Gerendas BS, Mylonas G, Najeeb BH, Schmidt-Erfurth U, et al. Segmentation of macular neovascularization and leakage in fluorescein angiography images in neovascular age-related macular degeneration using deep learning. *Eye.* (2023) 37:1439–44. doi: 10.1038/s41433-022-02156-6
- Ronneberger O, Fischer P, Brox T. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, proceedings, part III.* Cham: Springer (2015). p. 234–41. doi: 10.1007/978-3-319-24574-4_28
- He K, Zhang X, Ren S, Sun J. In: *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14.* Cham: Springer (2016). p. 630–45. doi: 10.1007/978-3-319-46493-0_38
- Thomas AS, Thomas MK, Finn AP, Fekrat S. USE OF THE ISCHEMIC INDEX ON WIDEFIELD FLUORESCIN ANGIOGRAPHY TO CHARACTERIZE A CENTRAL RETINAL VEIN OCCLUSION AS ISCHEMIC OR NONISCHEMIC. *Retina.* (2019) 39:1033–8. doi: 10.1097/IAE.0000000000002126
- Nicholson L, Vazquez-Alfageme C, Patrao NV, Triantafyllopoulou I, Bainbridge JW, Hykin PG, et al. Retinal nonperfusion in the posterior pole is associated with increased risk of neovascularization in central retinal vein occlusion. *Am J Ophthalmol.* (2017) 182:118–25. doi: 10.1016/j.ajo.2017.07.015