VasFAZ-Net: A Hybrid Loss-Optimized EfficientNet-scSE Network for Retinal Vasculature and FAZ Segmentation in OCTA

Nisan Pranavah Raja, Varun P. Gopi,
Bibin Francis, Chelli Devi N

Abstract

Accurate segmentation of fine retinal vessels and the Foveal Avascular Zone (FAZ) in OCTA images remains challenging due to low contrast, noise, and class imbalance. To address this, a lightweight deep learning framework called VasFAZ-Net is proposed, which uses an EfficientNet-B0 encoder, a parallel spatial and channel squeeze-and-excitation (scSE) decoder, and a hybrid Dice–Focal loss function, trained on the OCTA-500 dataset with both 3 mm and 6 mm scans. The proposed model achieved Dice scores of up to 88.48% for vessel segmentation and 97.98% for FAZ segmentation with only 8.8 million parameters and an inference time of 18.2 ms per image. These results demonstrate that the proposed method provides accurate, computationally efficient, and clinically deployable multi-target OCTA segmentation for retinal disease analysis.

Keywords: Optical coherence tomography angiography (OCTA), automated retinal segmentation, deep learning, spatial–channel attention, foveal avascular zone (FAZ), retinal microvasculature.

MSC 2020: 68T45, 68U10, 92C55, 62H35.

1 Introduction

Optical coherence tomography angiography (OCTA) is a non-invasive imaging technique that enables high-resolution visualization of the reti-

doi:10.56415/csjm.v34.01

nal microvasculature while also providing volumetric structural information. By capturing depth-resolved scans, OCTA distinguishes the different retinal layers and their vascular networks [1]. Compared to conventional techniques such as fluorescein angiography and indocyanine green angiography, OCTA offers a safer alternative, as the latter require intravenous dye administration that is time-consuming, invasive, and may trigger adverse reactions ranging from mild nausea to severe anaphylaxis. In contrast, OCTA is rapid, risk-free, and capable of producing three-dimensional angiographic images, making it highly relevant for early clinical screening. This approach is particularly useful for identifying and tracking eye diseases, including age-related macular degeneration (AMD), Diabetic Retinopathy (DR), retinal vein occlusion (RVO), glaucoma, and Choroidal Neovascularization (CNV).

Precise identification of retinal blood vessels and the foveal avascular zone (FAZ) is essential for reliable OCTA analysis and subsequent disease detection. Once segmented, quantitative biomarkers such as vessel length, density, diameter index, and FAZ area and perimeter can be extracted to provide objective measures of retinal health [2]. These metrics are particularly useful in tracking disease progression. For instance, an enlargement of the FAZ area and perifoveal intercapillary space has been correlated with the severity of Diabetic Retinopathy (DR), from normal eyes through diabetes without retinopathy, non-proliferative DR, and proliferative DR. Despite its importance, manual annotation of vessels and the FAZ is labor-intensive and impractical due to the intricate and fine-scale structure of retinal capillaries. Consequently, automated OCTA segmentation has emerged as a rapidly evolving research direction with significant promise for enhancing clinical decision-making.

Automated analysis of OCTA images faces several key challenges, like the fact that the OCTA images are low-contrast, noisy, and in many cases affected by motion and shadowing artifacts [3]. All of these factors hinder the accurate delineation of vessel margins. These are particularly problematic in small capillary beds, where the dense and irregular vascular architecture makes background separation difficult. The separation of arteries, veins, and capillaries adds another level of complexity to the problem [2]. Public datasets such as ROSE [4], OCTA-500 [5],

and OCTAGON [6] exist but are limited in size and diversity.

Contributions

- A simple U-Net-style encoder-decoder with an EfficientNet-B0 [7] encoder and a parallel spatial+channel attention decoder with embedded SE (scSE).
- A binary-per-target training setup (1-channel sigmoid head) with a joint Dice+Focal objective; the models are trained separately for artery, vein, capillary, and FAZ.
- The proposed framework demonstrates strong performance in vessel-type segmentation (capillaries, arteries, and veins) and FAZ delineation on the OCTA-500 dataset, evaluated across both 6 mm × 6 mm, and 3 mm × 3 mm fields of view (FOV), while preserving an efficient trade-off between computational speed and memory usage (∼8.8M parameters; ∼18.2 ms per 400 × 400 image).

2 Related Works

Several methods have been proposed for the automated segmentation of retinal vessels and FAZs in OCTA images. Early works relied on classical image processing pipelines. For example, Diaz et al. [6] used edge detection and morphological operations for FAZ delineation, while Xu et al. [2] applied line detection and clustering for vessel and capillary segmentation in diabetic patients. Other morphological approaches, such as those by Liu et al. [8], combined watershed and region-growing techniques; however, their performance degraded on poor-quality scans. With the rise of deep learning, more accurate solutions have been introduced. Le et al. [9] proposed an ensemble of U-Net and LSTM for AO-OCT vasculature segmentation, and Lin et al. [10] designed a joint learning network for FAZ delineation and disease diagnosis across multiple datasets. Giarratano et al. [3] compared handcrafted filters with CNN- and U-Net–based models, while Liang et al. [11] leveraged unsupervised vessel segmentation as prior knowledge for FAZ detection. Ryu et al. [12] combined segmentation with feature extraction and demographic data for DR severity assessment. Public datasets have ac-

celerated progress. Hao et al. [13] released the ROSE dataset and proposed VAFF-Net, later extended in works such as Huang et al. [14] with residual U-Nets. Li et al. [5] introduced OCTA-500, enabling vessel subtype classification (arteries, veins, and capillaries). Subsequent studies [15],[16] validated U-Net variants, CNNs, and hybrid approaches on OCTA-500, reporting Dice scores up to 0.92. More recently, Chinkamol et al. [1] reduced manual labeling effort via weakly supervised learning, showing competitive results across multiple OCTA projections.

Recent studies have further explored automated segmentation of the foveal avascular zone (FAZ) and retinal vasculature in OCTA images using deep learning techniques. Saeidian et al. [17] proposed a framework combining FAZ segmentation with diabetic retinopathy (DR) classification, demonstrating the clinical relevance of accurate FAZ delineation. Earlier work by Mirshahi et al. [18] focused specifically on FAZ segmentation using convolutional neural networks, highlighting the challenges posed by image noise and low contrast. More recent approaches have investigated multi-domain feature fusion and wide-field OCTA analysis for DR assessment [19], reporting improved segmentation performance at the cost of increased model complexity. In addition to supervised learning, unsupervised strategies have also been explored for OCTA vessel segmentation, although such methods often struggle with fine capillary structures. Despite these advances, accurately segmenting arteries, veins, capillaries, and the FAZ within a lightweight and computationally efficient framework remains an open challenge.

OCTA research has increasingly moved toward more integrated analysis frameworks that combine segmentation with additional diagnostic objectives and advanced feature modeling. Task-specific network designs have been introduced to jointly address DR grading and lesion segmentation in ultra-wide OCTA images, illustrating the benefit of shared representations across related clinical tasks. At the same time, deep learning models with enhanced feature aggregation mechanisms have been developed for the segmentation and quantitative assessment of retinal microstructures, aiming to improve structural continuity and boundary definition in OCTA images. Cross-domain feature fusion strategies have also been explored to improve robustness across heterogeneous data distributions, particularly for lesion segmentation in

DR, although such approaches are primarily optimized for lesion-level analysis rather than detailed vascular subtype or FAZ segmentation. These developments reflect ongoing efforts to improve OCTA segmentation performance while highlighting remaining challenges related to model efficiency and task specificity.

Overall, while deep learning dominates current methods, challenges remain in handling artifacts, vessel complexity, and dataset limitations. Our VasFAZ-Net addresses these gaps through a lightweight EfficientNet-B0 encoder and scSE-based decoder tailored for capillary and FAZ boundaries.

2.1 Research Gaps

Although recent OCTA segmentation methods have made various improvements, several fundamental limitations remain unaddressed. Many existing approaches address either FAZ delineation or vessel segmentation separately and do not provide explicit support for simultaneously differentiating arteries, veins, and capillaries, which is often required for comprehensive vascular analysis. Furthermore, recent accuracy gains often come at the expense of complex multi-branch or transformer-based architectures, which significantly increase computational costs and limit real-time applicability in resource-constrained clinical settings [20]. Unsourced or weakly supervised learning methods alleviate the annotation burden, yet most of them lack preservation of fine capillary structure and accurate boundary delineation in the presence of noise and various imaging artifacts [21]. Severe class imbalance inherent to OCTA data is also typically addressed implicitly, rather than directly integrated into the model optimization strategy. These unmet issues reveal a notable literature gap for a segmentation framework that jointly addresses multi-structure delineation, computational efficiency, and robustness.

2.2 Motivation

The motivation for this work is to develop a compact, computationally efficient OCTA segmentation framework that delivers accurate delineation of fine vascular structures and the FAZ without resorting to overly complex model designs. Such a balance between complexity and

accuracy will be imperative in enabling reliable extraction of clinically relevant biomarkers while satisfying the practical constraints of routine clinical use. Given the need to place equal emphasis on both segmentation accuracy and efficiency, the proposed approach aims to enable real-time OCTA analysis and support its wider clinical adoption.

3 Materials and Methods

This section describes the dataset, preprocessing pipeline, proposed VasFAZ-Net architecture, training procedure, and evaluation strategy. To ensure clarity and reproducibility, the methodology is presented in algorithmic form.

3.1 Materials

The proposed framework was evaluated using the publicly available OCTA-500 dataset [5], which comprises 500 scans acquired with a 70 kHz spectral-domain OCT system operating at a central wavelength of 840 nm. To minimize inter-eye correlation and ensure subject-level independence, only one eye per participant was included in the analysis. The dataset comprises both healthy controls and pathological cases, encompassing a range of retinal disorders, including DR, AMD, retinal vein occlusion (RVO), and CNV.
The database is divided into two subsets based on the field of view (FOV):

1. **OCTA_6mm:** 300 scans (6 mm × 6 mm FOV; IDs 10001–10300)
2. **OCTA_3mm:** 200 scans (3 mm × 3 mm FOV; IDs 10301–10500)

Each scan is accompanied by manual annotations for arteries, veins, capillaries, and the foveal avascular zone (FAZ). In this study, the B5 projection map was utilized, which corresponds to the maximum projection between the internal limiting membrane (ILM) and the outer plexiform layer (OPL). This projection highlights the superficial capillary plexus (SCP) and offers a comprehensive representation of the retinal vasculature, making it particularly suitable for evaluating vascular integrity and detecting pathological alterations, such as capillary dropout, vascular occlusion, and neovascularization, in DR and related diseases.

3.2 Image Augmentation Techniques

Image augmentation was applied to artificially increase the dataset's diversity and enhance the deep learning model's robustness for retinal vessel segmentation. During training, random horizontal and vertical flips were performed with probability $p = 0.5$ to account for variability in vessel orientation. Brightness and contrast adjustment with low probability $p = 0.1$ and a limit of 0.1 were performed together with random gamma correction ($p = 0.5$) in order to increase robustness to illumination changes. Geometric variability was enhanced by random rotations up to $\pm 45^\circ$ with a probability of $p = 0.5$, ensuring invariance to vessel orientation and spatial alignment. Mild elastic transformations were also applied, with $p = 0.5$, to simulate natural anatomical variability. Coarse dropout ($p = 0.5$) was utilized to randomly occlude small regions in an image, thereby enhancing tolerance to localized noise, artifacts, and signal loss. All these augmentations collectively enhanced the model's ability to generalize across diverse imaging conditions while preserving the anatomical fidelity of the retinal vasculature.

3.2.1 Challenges in Augmentation

Excessive or inappropriate transformations can distort vessel morphology and reduce accuracy. Hence, augmentation parameters were carefully tuned to balance diversity and data fidelity, ensuring that the model learned robust yet clinically meaningful features.

3.3 Methods

This work introduces a lightweight yet effective architecture specifically designed for segmenting OCTA images. The proposed architecture is presented in Fig. 1a. The design builds on the widely adopted U-Net [22], which has been a standard backbone in retinal image analysis [3],[9],[11],[12],[14]. To improve accuracy and efficiency, we replace the encoder with a pre-trained EfficientNet-B0 [7], chosen for its favorable balance between accuracy and computational cost. EfficientNet-B0 (5.3M parameters, 0.39B FLOPs) is significantly lighter than larger variants such as EfficientNet-B7 (66M parameters, 37B FLOPs), while outperforming conventional encoders like ResNet-50 [23] and DenseNet-

169 [24]. This makes it well-suited for medical imaging tasks with limited annotated data.

3.3.1 EfficientNet-B0 in the Encoder

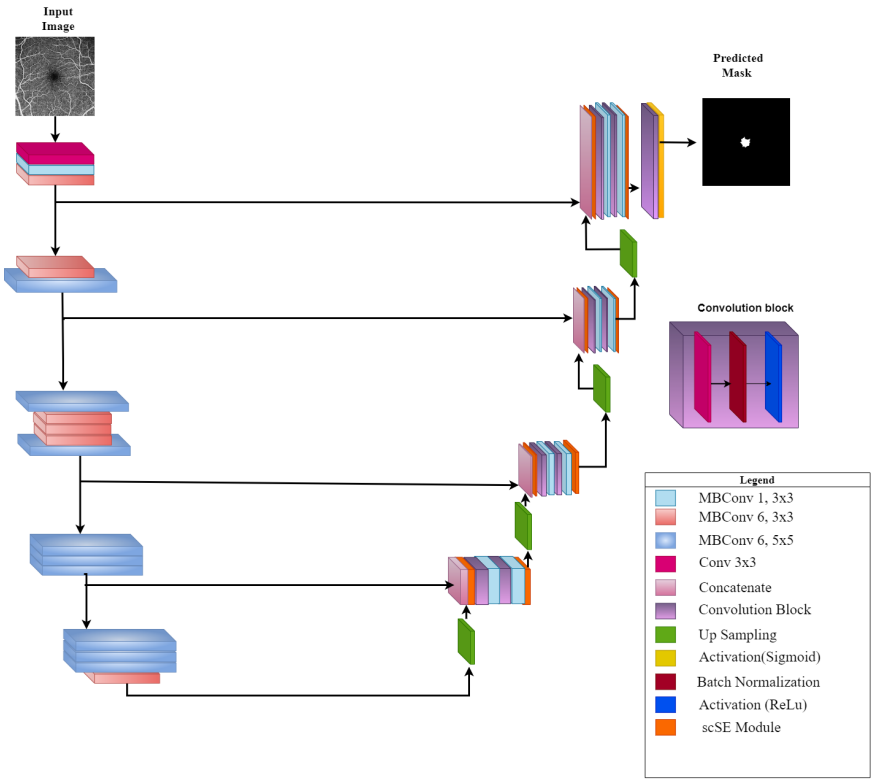
The EfficientNet-B0 encoder leverages Mobile Inverted Bottleneck Convolutions (MBConv), which expand feature maps before depthwise convolutions, reducing parameters while preserving discriminative detail. Pre-training on ImageNet provides transferable features, mitigating the challenges of small medical datasets. This ensures robust multi-scale feature extraction with reduced memory and computational demand.

3.3.2 Modified Decoder with Parallel Spatial and Channel Attention

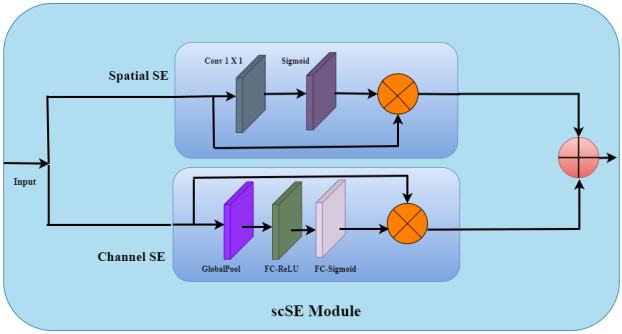
To complement the encoder, we propose a decoder enhanced with a spatial-channel-squeeze-excitation (scSE) block (Fig. 1b). Traditional U-Net decoders concatenate features via skip connections; however, our decoder incorporates attention mechanisms to selectively emphasize informative features. Unlike conventional sequential spatial–channel attention designs [25], we apply spatial and channel (sc) attention in parallel, each coupled with a squeeze-and-excitation (SE) block. This independent processing improves robustness, reduces overfitting, and captures fine structural details of capillaries and FAZ. Unlike prior SE-based methods such as Hu et al. [26], which perform channel excitation sequentially, our scSE design applies spatial and channel recalibration concurrently, enabling simultaneous emphasis of global context and localized vascular features while reducing redundancy.

3.3.3 Squeeze-and-Excitation (SE) Operation

The SE operation [26] enhances feature recalibration by first "squeezing" global context via average pooling, then "exciting" channel responses through fully connected layers and sigmoid activation. Embedding SE within both attention branches ensures stronger feature discrimination, particularly for complex, mesh-like vasculature.



(a) Proposed network architecture



(b) Proposed scSE block

Figure 1. (a) Proposed network architecture, (b) Proposed scSE block

3.4 Model Optimization

Training was conducted using the Adam optimizer in combination with a hybrid loss function comprising Dice and Focal terms. All models were trained for 300 epochs with a mini-batch size of 16, and the optimal checkpoint was selected based on the highest validation Dice score. The Adam optimizer was configured with a learning rate of 0.001, $\epsilon_{\text{opt}} = 10^{-8}$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The proposed architecture contains approximately 8.79 million trainable parameters.

The Dice loss [27] is defined as:

$$L_{\text{Dice}} = 1 - \frac{2\sum_i y_i \hat{y}_i + \epsilon_{\text{dice}}}{\sum_i y_i + \sum_i \hat{y}_i + \epsilon_{\text{dice}}},$$

and the Focal loss [28] as:

$$L_{\text{Focal}} = -\alpha_t (1 - p_t)^{\gamma} \log(p_t).$$

The combined loss is:

$$L_{\text{total}} = 0.5\, L_{\text{Dice}} + 0.5\, L_{\text{Focal}}.$$

We trained four independent binary models (capillary, artery, vein, FAZ), each with a sigmoid output head. Predictions were thresholded at 0.5 per class. Focal loss hyperparameters were $\gamma = 2.0$, $\alpha = 0.25$, with `reduction=mean` and `eps` $= 10^{-6}$. The Dice loss applied smoothing $\varepsilon = 10^{-7}$. Results consistently showed gains across Dice, IoU, and accuracy, confirming the robustness of the objective.

4 Results

4.1 Experimental setup and performance measures

All experiments were implemented in PyTorch using the torchvision API. Training and inference were performed on an NVIDIA RTX 3090 Ti GPU (24 GB) with 128 GB of system RAM. Both training and evaluation were conducted on the publicly available OCTA-500 dataset [5]. For the OCTA_6mm subset, 240 scans were used for training, 10 for

validation, and 50 for testing. For the OCTA_3mm subset, the split included 140 training, 10 validation, and 50 testing images. All splits were performed at the patient level to ensure subject independence, and only one eye per volunteer was included. The dataset provides six projection maps (B1–B6) [5]. In this study, the B5 projection map was utilized, corresponding to the maximum projection between the ILM and OPL layers. This projection highlights the superficial capillary plexus, providing a more comprehensive visualization of the retinal vasculature compared to other modalities. The B5 projection was chosen because it emphasizes the superficial capillary plexus (SCP) with higher vessel contrast and fewer projection artifacts. This is crucial for proper segmentation of the retinal vasculature and the FAZ. Besides, the analysis performed on 6 mm × 6 mm and 3 mm × 3 mm fields of view (FOV) will permit an evaluation of the proposed method at different clinically relevant scales, preserving the original native image resolution to avoid resampling bias. The OCTA_6mm subset covers a 6 mm × 6 mm region centered on the fovea (400×400 pixels), whereas the OCTA_3mm subset captures a smaller 3 mm × 3 mm region (304×304 pixels). Unless otherwise specified, 6 mm × 6 mm scans were processed at 400×400 resolution and 3 mm × 3 mm scans at 304×304 without additional resampling. The performance of the proposed framework was assessed using five widely adopted evaluation metrics: Dice coefficient, intersection over union (IoU), sensitivity, specificity, and accuracy. The Dice coefficient quantifies the overlap between predicted segmentations and ground truth annotations, whereas IoU measures the spatial agreement between them. Sensitivity and specificity, derived from the confusion matrix, evaluate the model's ability to correctly identify the presence or absence of the region of interest (ROI). Accuracy provides an overall measure of classification performance across all categories.

These metrics are formally defined as follows:

$$\text{Dice} = \frac{2 \cdot |S_O \cap G_T|}{|S_O| + |G_T|}, \tag{1}$$

$$IoU = \frac{|S_O \cap G_T|}{|S_O \cup G_T|}, \tag{2}$$

$$Accuracy = \frac{t_p + t_n}{t_p + f_p + t_n + f_t}, \tag{3}$$

$$Sensitivity = \frac{t_p}{t_p + f_t}, \tag{4}$$

$$Specificity = \frac{t_n}{t_n + f_p}, \tag{5}$$

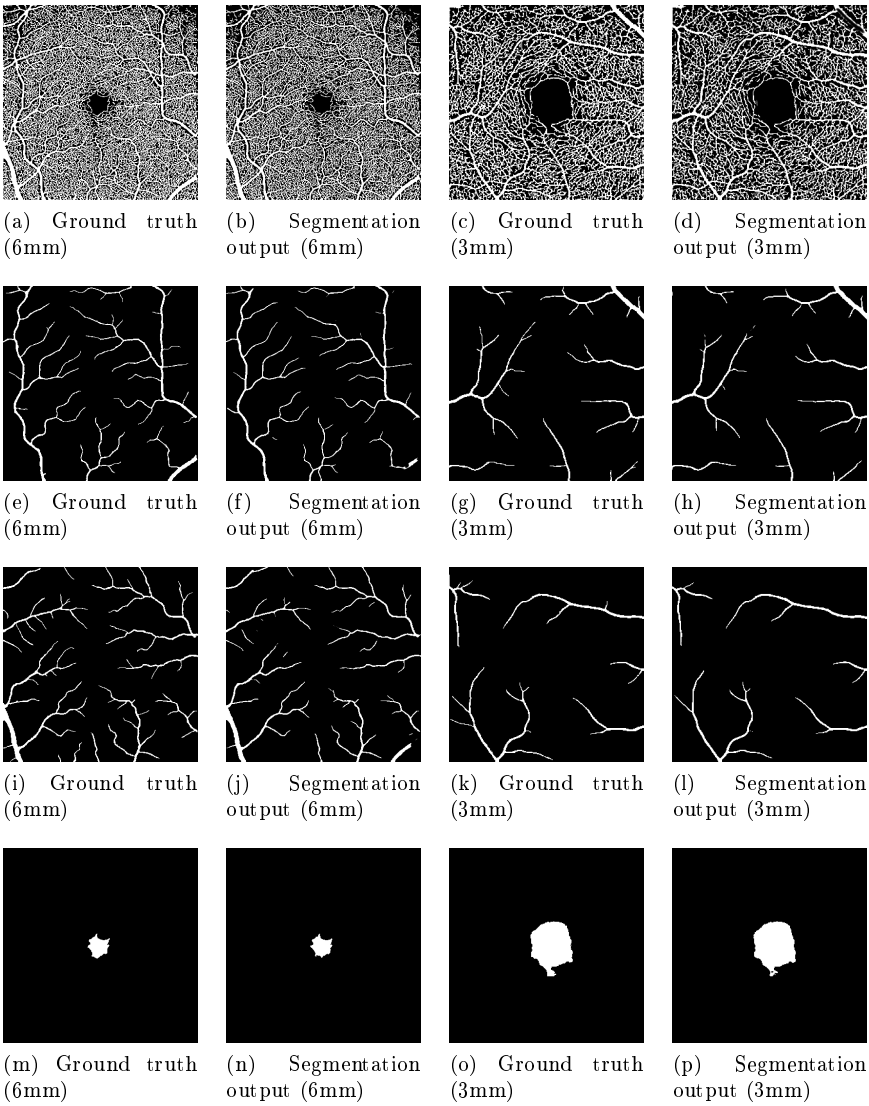
where S_O denotes the segmented output, G_T the ground truth, and t_p, f_p, t_n, and f_t correspond to true positive, false positive, true negative, and false negative, respectively.

4.2 Qualitative Segmentation Results

Figure 2 (a–p) illustrates the segmentation outputs for capillaries, arteries, veins, and the foveal avascular zone (FAZ) using the OCTA-500 dataset. The visual comparison includes both the reference annotations and the predictions generated from the OCTA_6mm and OCTA_3mm subsets. As observed, the proposed framework closely reproduces the ground-truth structural patterns, demonstrating reliable delineation of vascular structures and the FAZ across different fields of view (FOV).

4.3 Quantitative Results

Table 1 reports the segmentation performance for capillary, artery, vein, and FAZ across both OCTA-500 subsets (OCTA_3mm and OCTA_6mm). Dice and IoU are reported as percentages for consistency with prior work. On the OCTA_6mm test set, Dice values ranged from 88.38% to 94.55%, while the OCTA_3mm subset yielded slightly higher values (89.59%–97.98%). This difference is expected since OCTA_3mm scans focus on a smaller FOV with greater detail, whereas the larger OCTA_6mm area makes segmentation more sensitive to small discrepancies [1], [5], [6]. Across all experiments, FAZ segmentation consistently achieved higher accuracy than vessel segmentation. This is attributed to the relatively simpler structure of the FAZ compared to the dense, mesh-like vasculature of retinal arteries, veins, and capillaries, which remain more challenging to delineate.



(a) Ground truth (6mm) (b) Segmentation output (6mm) (c) Ground truth (3mm) (d) Segmentation output (3mm)

(e) Ground truth (6mm) (f) Segmentation output (6mm) (g) Ground truth (3mm) (h) Segmentation output (3mm)

(i) Ground truth (6mm) (j) Segmentation output (6mm) (k) Ground truth (3mm) (l) Segmentation output (3mm)

(m) Ground truth (6mm) (n) Segmentation output (6mm) (o) Ground truth (3mm) (p) Segmentation output (3mm)

Figure 2. Segmentation results: capillary (a–d), artery (e–h), vein (i–l), and FAZ (m–p)

Table 1. Comparison results of vasculature, and FAZ segmentation with state-of-the-art methods

	OCTA_6mm test set					OCTA_3mm test set				
Model	**Dice**	**IoU**	**Accuracy**	**SE**	**SP**	**Dice**	**IoU**	**Accuracy**	**SE**	**SP**
Results of Capillary Segmentation										
Attention U-Net [25]	87.47	77.59	90.29	89.10	90.97	89.65	81.24	91.44	91.52	91.41
UNet 3+ [29]	87.11	77.16	89.96	88.41	90.87	87.55	77.85	89.61	90.14	89.32
IPN-V2 [5]	87.64	77.65	87.82	87.12	88.01	90.06	81.52	90.32	89.83	89.90
U-Net++ [30]	87.34	77.52	90.08	89.17	90.61	88.45	79.29	90.49	89.84	90.97
UNETR [20]	84.63	73.35	88.11	84.87	90.08	87.22	77.33	88.46	90.75	88.89
MS-Net [31]	85.95	75.36	89.48	89.33	89.57	88.20	78.89	90.13	89.38	90.66
AV-Net [32]	87.44	77.68	90.32	89.04	90.86	89.53	81.05	91.34	91.50	91.17
TransUNet [33]	83.31	71.39	86.98	83.58	89.88	86.51	76.22	88.01	89.92	87.78
CS-Net [34]	87.16	76.90	90.63	89.27	90.14	90.04	81.57	91.67	90.99	90.74
U-Net [22]	84.61	73.33	88.08	84.89	90.05	86.98	76.96	88.35	90.64	86.79
Proposed work	**88.38**	**79.62**	**90.96**	**90.28**	**91.65**	**90.48**	**81.98**	**91.95**	**92.68**	**91.63**
Results of Artery Segmentation										
Attention U-Net [25]	86.49	76.20	98.83	83.84	**99.54**	88.57	79.50	99.15	88.50	99.56
IPN-V2 [5]	**88.52**	78.96	98.69	86.32	99.47	87.34	78.12	98.92	86.95	99.53
UNet 3+ [29]	83.88	72.24	98.58	83.07	99.31	86.75	76.60	99.03	85.71	99.54
MS-Net [31]	85.84	75.19	98.82	83.61	99.53	88.51	79.40	99.19	89.44	99.55
AV-Net [32]	82.72	70.53	98.49	81.31	99.30	85.44	74.58	98.95	83.18	99.56
TransUNet [33]	84.67	73.41	98.51	83.27	99.27	85.52	74.70	98.89	83.21	99.51
UNETR [20]	85.78	75.10	98.57	83.25	99.41	86.71	76.53	98.98	85.64	99.49
CS-Net [34]	79.89	67.22	98.33	78.54	99.24	82.12	69.98	98.73	79.76	99.46
U-Net [22]	85.64	74.89	98.74	85.10	99.38	87.07	77.11	99.04	86.96	99.50
Proposed work	88.40	**79.23**	**98.99**	**87.47**	99.52	**90.70**	**83.01**	**99.31**	**90.70**	**99.64**
Results of Vein Segmentation										
MS-Net [31]	84.99	73.92	98.67	86.05	99.28	87.48	77.75	99.26	**90.74**	99.52
IPN-V2 [5]	88.22	**80.63**	**99.15**	85.92	**99.60**	87.25	80.48	99.19	84.32	99.66
UNETR [20]	84.21	72.72	98.53	84.02	99.21	85.21	74.23	99.17	84.01	99.56
U-Net++ [30]	84.24	73.79	98.57	84.14	99.29	86.35	75.98	99.21	84.22	99.66
AV-Net [32]	83.57	71.78	98.47	81.36	99.33	70.62	54.58	98.24	71.88	99.04
TransUNet [33]	83.78	72.08	98.45	83.44	99.16	84.27	72.81	99.12	83.97	99.53
U-Net [22]	85.10	74.07	98.54	86.75	99.14	84.82	73.65	99.12	83.52	99.59
CS-Net [34]	80.48	67.78	98.03	82.44	98.84	80.91	68.21	98.93	78.31	99.56
Attention U-Net [25]	86.43	76.11	98.72	85.08	99.41	85.68	74.95	99.16	84.77	99.59
Proposed work	**88.48**	79.38	98.87	**87.53**	99.46	**89.59**	**81.15**	**99.39**	88.40	**99.72**
Results of FAZ Segmentation										
UNet 3+ [29]	87.57	77.89	99.73	83.05	99.93	96.33	92.93	99.69	96.17	99.84
IPN-V2 [5]	93.25	88.93	99.78	93.17	99.91	97.56	95.02	99.80	96.72	99.92
CS-Net [34]	86.16	78.03	99.66	85.83	99.87	95.48	91.74	99.65	95.33	99.84
TransUNet [33]	88.69	79.67	99.71	87.63	99.87	95.56	91.49	99.61	95.37	99.81
U-Net++ [30]	81.31	68.51	99.60	77.79	99.85	96.31	92.88	99.63	96.37	99.77
MS-Net [31]	90.25	82.24	99.81	86.79	**99.96**	96.93	94.05	99.77	95.72	**99.95**
UNETR [20]	89.71	81.34	99.78	88.74	99.89	96.44	93.12	99.65	96.36	99.84
AV-Net [32]	86.13	75.64	99.68	87.83	99.82	95.75	91.85	99.60	95.48	99.79
U-Net [22]	88.57	79.49	99.74	87.58	99.88	96.14	92.57	99.67	96.07	99.82
Proposed work	**94.55**	**89.89**	**99.87**	**94.13**	99.94	**97.98**	**96.05**	**99.83**	**97.27**	99.94

5 Discussion

5.1 Ablation Study

An ablation study was conducted to evaluate the contribution of each architectural component. The baseline is a standard U-Net, to which the EfficientNet-B0 encoder, channel attention, and spatial attention were added incrementally. Results for capillary, vein, FAZ, and artery segmentation are summarized in Table 2. The plain U-Net achieved only moderate performance. Adding spatial or channel attention individually led to marginal improvements, while combining both gave a slight boost in accuracy. Incorporating EfficientNet-B0 as the encoder provided a clear performance gain across all metrics. The best results were obtained when both attention modules were integrated with U-Net + EfficientNet-B0, i.e., the full proposed architecture. This configuration consistently achieved the highest Dice and IoU scores on both the OCTA_6mm and OCTA_3mm datasets, with only minor exceptions, in which U-Net + EfficientNet-B0 alone reported slightly higher sensitivity or specificity. Overall, the ablation confirms that EfficientNet-B0 is the primary contributor to the accuracy improvements, while the attention modules further enhance fine-grained delineation of vessels and FAZs.

5.2 Comparison of computational complexity

Evaluating segmentation models requires not only assessing their accuracy but also examining their computational efficiency. Key factors, such as memory usage, the number of trainable parameters, and inference speed, directly impact their practicality for clinical deployment. In this study, the proposed architecture is benchmarked against seven representative networks: U-Net [22], U-Net 3+ [29], U-Net++ [30], CS-Net [34], Attention U-Net [25], AV-Net [32], and IPN-V2 [5]. All methods were benchmarked on the same GPU and image resolution, using the identical timing protocol described above (PyTorch 1.10.1, AMP, batch size = 1, 400×400 inputs). As summarized in Table 3, our model yields a peak GPU memory footprint of < 1 GB, a moderate parameter count of 8.79 M, and a model-only throughput of 55 images/s (i.e., 18.2 ms per image). Under identical conditions, competing baselines

Table 2. Ablation study of vasculature and FAZ segmentation

Methods	6mm					3mm				
	Dice	IoU	Accuracy	SE	SP	Dice	IoU	Accuracy	SE	SP
Results of capillary segmentation										
U-Net	84.61	73.33	88.08	84.89	90.05	86.98	76.96	88.35	90.64	86.79
U-Net + SA	84.41	73.03	87.56	87.25	87.71	86.83	76.73	89.00	89.04	89.03
U-Net + CA	84.28	72.83	87.61	85.97	88.61	87.02	77.03	89.23	88.61	89.73
U-Net + SA + CA	84.47	73.12	87.94	84.96	89.85	86.92	76.87	88.76	89.88	88.06
U-Net + EfficientNet-B0	85.98	75.41	89.37	87.00	**91.72**	89.21	80.52	90.72	91.26	90.33
U-Net + EfficientNet-B0 + SA	86.17	75.70	89.50	87.33	90.73	88.98	80.15	90.55	90.84	90.34
U-Net + EfficientNet-B0 + CA	86.00	75.44	89.40	86.95	90.82	88.77	79.81	90.42	90.17	90.58
Proposed work	**88.38**	**79.62**	**90.96**	**90.28**	91.65	**90.48**	**81.98**	**91.95**	**92.68**	**91.63**
Results of artery segmentation										
U-Net	85.64	74.89	98.74	85.10	99.38	87.07	77.11	99.04	86.96	99.50
U-Net + SA	84.64	73.37	98.68	82.77	99.41	86.78	76.65	99.01	87.24	99.46
U-Net + CA	85.87	75.24	98.78	84.51	99.44	87.25	77.39	99.04	88.54	99.44
U-Net + SA + CA	85.96	75.38	98.76	86.25	99.34	88.10	78.37	99.11	88.28	99.53
U-Net + EfficientNet-B0	87.12	77.19	98.88	85.82	99.48	87.95	78.49	99.07	**91.05**	99.38
U-Net + EfficientNet-B0 + SA	87.80	78.25	98.92	**88.35**	99.40	89.97	81.77	99.26	88.67	**99.67**
U-Net + EfficientNet-B0 + CA	86.84	76.74	98.85	86.32	99.42	88.88	79.99	99.16	89.67	99.53
Proposed work	**88.40**	**79.23**	**98.99**	87.47	**99.52**	**90.70**	**83.01**	**99.31**	90.70	99.64
Results of vein segmentation										
U-Net	85.10	74.07	98.54	86.75	99.14	84.82	73.65	99.12	83.52	99.59
U-Net + SA	84.85	73.69	98.57	83.52	99.33	85.75	75.05	99.15	86.20	99.55
U-Net + CA	85.45	74.59	98.61	85.13	99.29	85.72	75.02	99.13	88.27	99.46
U-Net + SA + CA	85.62	74.87	98.62	86.01	99.25	84.70	73.46	99.09	85.58	99.50
U-Net + EfficientNet-B0	87.61	77.96	98.73	**89.74**	99.20	89.38	80.79	99.20	**89.62**	99.57
U-Net + EfficientNet-B0 + SA	87.75	77.78	**98.90**	87.51	99.42	89.13	80.39	99.36	87.95	99.71
U-Net + EfficientNet-B0 + CA	87.37	77.58	98.74	87.26	99.34	88.34	79.12	99.31	87.98	99.65
Proposed work	**88.48**	**79.38**	98.87	87.53	**99.46**	**89.59**	**81.15**	**99.39**	88.40	**99.72**
Results of FAZ segmentation										
U-Net	88.57	79.49	99.74	87.58	99.88	96.14	92.57	99.67	96.07	99.82
U-Net + SA	90.72	83.03	99.79	89.90	99.91	96.86	93.91	99.73	95.77	99.90
U-Net + CA	89.29	80.66	99.75	89.53	99.87	96.25	92.77	99.68	95.17	99.88
U-Net + SA + CA	89.08	80.31	99.75	87.10	99.91	96.20	92.69	99.67	94.74	99.90
U-Net + EfficientNet-B0	93.08	87.05	99.84	91.76	99.94	97.00	94.18	99.76	97.15	99.87
U-Net + EfficientNet-B0 + SA	93.55	87.89	99.84	93.42	99.91	97.72	95.55	99.82	96.60	**99.95**
U-Net + EfficientNet-B0 + CA	93.49	87.78	99.85	91.41	**99.95**	97.54	95.20	99.81	96.26	**99.95**
Proposed work	**94.55**	**89.89**	**99.87**	**94.13**	99.94	**97.98**	**96.05**	**99.83**	**97.27**	99.94

SA: Spatial Attention, CA: Channel Attention

achieve approximately 2–7 images/s. These results highlight a favorable speed-memory trade-off for the proposed architecture without compromising segmentation accuracy. Baselines were re-timed under the same protocol (same GPU, input size, and batch size).

5.3 Comparison with State-of-the-Art Methods

The performance of the proposed model was benchmarked against recent state-of-the-art methods. We adopted a two-fold comparison strategy. First, for studies with publicly available code [22], [25], [29], [30],

Table 3. Comparison of Computational complexity

Method	GPU used	Image resolution	Memory (GB)	Parameters (million)	Speed (ms)
CS-Net [34]	RTX 3090 Ti	400×400	1.0	8.4	144
U-Net [22]	RTX 3090 Ti	400×400	2.0	**5.5**	238
AV-Net [32]	RTX 3090 Ti	400×400	2.4	6.5	158
Attention U-Net [25]	RTX 3090 Ti	400×400	2.8	5.6	197
U-Net 3+ [29]	RTX 3090 Ti	400×400	4.9	16.1	243
U-Net++ [30]	RTX 3090 Ti	400×400	5.3	12.1	255
IPN-V2 [5]	RTX 3090 Ti	400×400	7.3	7.6	427
Proposed work	RTX 3090 Ti	400×400	**0.8**	8.79	**18.2**

[32], [34], the networks were re-implemented under comparable parameter settings and trained/tested on the 3mm and 6mm datasets. Table 4 summarizes the results across capillary, artery, vein, and FAZ segmentation using Dice, IoU, accuracy, sensitivity, and specificity. In capillary segmentation, our method consistently achieved the best or second-best performance across all metrics on both subsets, closely followed by CS-Net [34], which is tailored for curvilinear structures. For arteries, our network achieved competitive scores across all metrics, with a marginal drop in specificity (0.02%) on OCTA_6mm compared to Attention U-Net [25]. For veins and FAZ, the proposed architecture outperformed all baselines across both field-of-view subsets. Overall, these results highlight the robustness of our method, particularly in handling the fine, mesh-like structure of capillaries, which remain the most challenging to segment accurately. To further validate the robustness and generalizability of the proposed model, VasFAZ-Net was evaluated on the ROSE-O dataset [13]. As this dataset provides annotations for retinal vessels and the FAZ, we specifically assessed the model's performance on FAZ segmentation. Despite being trained solely on the OCTA-500 dataset, VasFAZ-Net achieved strong cross-dataset performance, attaining a Dice score of 92.17%, an IoU of 86.49%, an accuracy of 99.14%, a sensitivity of 92.22%, and a specificity of 99.44% on the 39 OCTA 3 mm test images. These results underscore the model's robustness and adaptability across varying imaging datasets and acquisition protocols.

The second comparison involved published results on the OCTA-500 dataset, where direct code implementation was not possible. To ensure fairness, only studies reporting separate metrics for OCTA_6mm and

Table 4. Comparison of segmentation performance on OCTA_6mm and OCTA_3mm test sets

		6mm					3mm				
Method	**ROI**	**Dice**	**IoU**	**ACC**	**SE**	**SP**	**Dice**	**IoU**	**ACC**	**SE**	**SP**
Ma et al., 2022 [35]	Retinal vessels	88.69	-	92.53	-	-	91.68	-	93.45	-	-
Zhong et al., 2022 [36]	Retinal vessels	89.20	80.61	-	-	-	92.66	86.35	-	-	-
Ning et al., 2023 [37]	Retinal vessels	89.09	-	98.25	-	-	91.47	-	98.99	-	-
Liu et al., 2022 [38]	Retinal vessels	-	-	-	-	-	91.20	-	98.80	-	-
	FAZ	-	-	-	-	-	95.40	91.20	-	-	-
Hu et al., 2022 [39]	Retinal vessels	89.72	81.17	-	89.08	99.08	91.13	83.78	-	89.86	99.49
	FAZ	90.51	84.24	-	92.09	99.89	98.43	96.93	-	98.07	99.96
Xiao et al., 2022 [40]	Retinal vessels + FAZ	-	-	-	-	-	98.76	97.55	-	-	-
	FAZ	-	-	-	-	-	97.94	96.00	-	-	-
Khan et al., 2023 [16]	Retinal vessels	91.60	85.20	-	-	-	-	-	-	-	-
	Retinal vessels + FAZ	86.64	79.04	-	-	-	-	-	-	-	-
Proposed work	Capillary	88.38	79.62	90.96	90.28	91.65	90.48	81.98	91.95	92.68	91.63
	Artery	88.40	79.23	98.99	87.47	99.52	90.70	83.01	99.31	90.70	99.64
	Vein	88.48	79.38	98.87	87.53	99.46	89.59	81.15	99.39	88.40	99.72
	FAZ	94.55	89.89	99.87	94.13	99.94	97.98	96.05	99.83	97.27	99.94

OCTA_3mm subsets were considered. However, significant variability exists across works in terms of target regions and evaluation metrics. For example, several approaches focus exclusively on retinal vasculature segmentation [35]–[37], whereas others extend their scope to include delineation of the foveal avascular zone (FAZ) [16], [38]–[40]. Likewise, some studies report results only on one subset (e.g., OCTA_6mm [16] or OCTA_3mm [38],[40]). Metrics also differ, with Dice and IoU being the most common [16], [36], [40], while accuracy, sensitivity, and specificity are reported less frequently [35], [37], [39]. Given this variability, a one-to-one comparison remains difficult. Nevertheless, Table 4 provides a consolidated overview of segmentation performance, showing that our method consistently improves performance across both subsets and all vascular structures compared to the majority of recent literature.

5.4 Limitations and Future Directions

Despite its promising results, the proposed method has certain limitations.

Although VasFAZ-Net demonstrates high segmentation accuracy, occasional errors occur in the presence of motion artifacts, projection shadows, and low-contrast areas. Such artifacts can cause broken vessel segments or minor over-smoothing of fine capillaries. Future work will

focus on incorporating artifact-aware learning and domain adaptation strategies to improve segmentation reliability under degraded imaging conditions.

The approach was validated only on the OCTA-500 dataset, restricting population diversity. Broader multi-institutional studies with scans from different devices are required to ensure robustness. While arteries and veins are delineated with high accuracy, the fine, mesh-like capillary network remains more challenging and requires further refinement for reliable clinical deployment. Future work should extend beyond segmentation to include quantitative biomarkers such as vessel length, density, diameter, and FAZ area/perimeter, which are critical for disease diagnosis and progression analysis. Additionally, lightweight architectures with lower computational demands should be explored to facilitate real-time screening and integration into clinical workflows.

6 Conclusion

This study presents an automated method for retinal vessel and FAZ segmentation using a modified U-Net with EfficientNet-B0 encoding and an attention-based scSE decoder. The model achieves competitive performance on the OCTA-500 dataset, with average Dice scores of 88.38 (capillary), 88.48 (vein), 88.40 (artery), and 94.55 (FAZ) on the OCTA_6mm subset, and 90.48, 90.70, 89.59, and 97.98, respectively, on the OCTA_3mm subset. Compared to recent approaches, the network demonstrates superior segmentation efficiency for arteries, veins, and FAZ while maintaining low memory usage (0.8 GB) and fast inference (18.2 ms). Overall, the proposed framework presents a straightforward yet effective solution for OCTA segmentation, with considerable potential for future adoption in clinical applications.

References

[1] A. Chinkamol, V. Kanjaras, P. Sawangjai, Y. Zhao, T. Sudhawiyangkul, C. Chantrapornchai, C. Guan, and T. Wilaiprasitporn, "OCTAve: 2D en face optical coherence tomography angiog-

raphy vessel segmentation in weakly-supervised learning with locality augmentation," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 6, pp. 1931–1942, 2023.

[2] B. Xu, J. Chen, S. Zhang, S. Shen, X. Lan, Z. Chen, Z. Yan, and B. Xu, "Association between the severity of diabetic retinopathy and optical coherence tomography angiography metrics," *Frontiers in Endocrinology*, vol. 12, p. 777552, 2021.

[3] Y. Giarratano, E. Bianchi, C. Gray, A. Morris, T. MacGillivray, B. Dhillon, and M. O. Bernabeu, "Automated segmentation of optical coherence tomography angiography images: Benchmark data and clinically relevant metrics," *Translational Vision Science & Technology*, vol. 9, no. 13, p. 5, 2020.

[4] Y. Ma, H. Hao, J. Xie, H. Fu, J. Zhang, J. Yang, Z. Wang, J. Liu, Y. Zheng, and Y. Zhao, "ROSE: A retinal OCT-Angiography vessel segmentation dataset and new model," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 928–939, 2021.

[5] M. Li, K. Huang, Q. Xu, J. Yang, Y. Zhang, Z. Ji, K. Xie, S. Yuan, Q. Liu, and Q. Chen, "OCTA-500: A retinal dataset for optical coherence tomography angiography study," *Medical Image Analysis*, vol. 93, p. 103092, 2024.

[6] M. Díaz, J. Novo, P. Cutrín, F. Gómez-Ulla, M. G. Penedo, and M. Ortega, "Automatic segmentation of the foveal avascular zone in ophthalmological OCT-A images," *PLOS ONE*, vol. 14, no. 2, pp. 1–22, 2019.

[7] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the International Conference on Machine Learning*. PMLR, 2019, pp. 6105–6114.

[8] J. Liu, S. Yan, N. Lu, D. Yang, C. Fan, H. Lv, S. Wang, X. Zhu, Y. Zhao, Y. Wang, Z. Ma, and Y. Yu, "Automatic segmentation of foveal avascular zone based on adaptive watershed algorithm in retinal optical coherence tomography angiography images," *Journal of Innovative Optical Health Sciences*, vol. 15, no. 1, p. 2242001, 2022.

[9] C. T. Le, D. Wang, R. Villanueva, Z. Liu, D. X. Hammer, Y. Tao, and O. J. Saeedi, "Novel application of Long Short-Term Memory Network for 3D to 2D retinal vessel segmentation in adaptive Optics—Optical coherence tomography volumes," *Applied Sciences*, vol. 11, no. 20, p. 9475, 2021.

[10] L. Lin, Z. Wang, J. Wu, Y. Huang, J. Lyu, P. Cheng, J. Wu, and X. Tang, "BSDA-Net: A boundary shape and distance aware joint learning framework for segmenting and classifying OCTA images," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Springer International Publishing, 2021, pp. 65–75.

[11] Z. Liang, J. Zhang, and C. An, "Foveal avascular zone segmentation of OCTA images using deep learning approach with unsupervised vessel segmentation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1200–1204.

[12] G. Ryu, K. Lee, D. Park, I. Kim, S. H. Park, and M. Sagong, "A deep learning algorithm for classifying diabetic retinopathy using optical coherence tomography angiography," *Translational Vision Science & Technology*, vol. 11, no. 2, p. 39, 2022.

[13] J. Hao, T. Shen, X. Zhu, Y. Liu, A. Behera, D. Zhang, B. Chen, J. Liu, J. Zhang, and Y. Zhao, "Retinal structure detection in OCTA image via voting-based multitask learning," *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3969–3980, 2022.

[14] K.-W. Huang, Y.-R. Yang, Z.-H. Huang, Y.-Y. Liu, and S.-H. Lee, "Retinal vascular image segmentation using improved UNet based on residual module," *Bioengineering*, vol. 10, no. 6, p. 722, 2023.

[15] Q. Li, X.-r. Zhu, G. Sun, L. Zhang, M. Zhu, T. Tian, C. Guo, S. Mazhar, J.-K. Yang, and Y. Li, "Diagnosing diabetic retinopathy in OCTA images based on multilevel information fusion using a deep learning framework," *Computational and Mathematical Methods in Medicine*, vol. 2022, no. 1, p. 4316507, 2022.

[16] A. Khan, J. Hao, Z. Dong, and J. Li, "Adaptive deep clustering

network for retinal blood vessel and foveal avascular zone segmentation," *Applied Sciences*, vol. 13, no. 20, p. 11259, 2023.

[17] J. Saeidian, H. Riazi-Esfahani, H. Azimi, H. Farrokhpour, A. Momeni, M. Jamalitootakani, A. Mirshahi, H. Faghihi, R. Sadeghi, and E. Khalili Pour, "Automated FAZ segmentation and diabetic retinopathy classification using OCTA images," *BMC ophthalmology*, vol. 25, no. 1, p. 602, 2025.

[18] R. Mirshahi, P. Anvari, H. Riazi-Esfahani, M. Sardarinia, M. Naseripour, and K. G. Falavarjani, "Foveal avascular zone segmentation in optical coherence tomography angiography images using a deep learning approach," *Scientific reports*, vol. 11, no. 1, p. 1031, 2021.

[19] G. Jia, F. Ma, S. Li, Z. Zhang, H. Liu, Y. Guo, and J. Meng, "Multi-domain fusion network: A novel approach to OCTA image segmentation in diabetic retinopathy," *Biomedical Signal Processing and Control*, vol. 109, p. 107945, 2025.

[20] A. Hatamizadeh, Y. Tang, V. Nath, D. Yang, A. Myronenko, B. Landman, H. R. Roth, and D. Xu, "UNETR: Transformers for 3D medical image segmentation," in *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2022, pp. 1748–1758.

[21] A. Alksas, A. Sharafeldeen, H. M. Balaha, M. Z. Haq, A. Mahmoud, M. Ghazal, N. S. Alghamdi, M. Alhalabi, J. Yousaf, H. Sandhu, and A. El-Baz, "Advanced OCTA imaging segmentation: Unsupervised, non-linear retinal vessel detection using modified self-organizing maps and joint MGRF modeling," *Computer Methods and Programs in Biomedicine*, vol. 254, p. 108309, 2024.

[22] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, 2015, pp. 234–241.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision*

and Pattern Recognition (CVPR). IEEE, 2015, pp. 770 – 778.

[24] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE, 2017, pp. 2261 – 2269.

[25] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[26] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.* IEEE, 2017, pp. 7132 – 7141.

[27] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.* Springer International Publishing, 2017, pp. 240 – 248.

[28] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *2017 IEEE International Conference on Computer Vision (ICCV).* IEEE, 2017, pp. 2999 – 3007.

[29] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: A full-scale connected UNet for medical image segmentation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2020, pp. 1055–1059.

[30] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.* Springer International Publishing, 2018, pp. 3 – 11.

[31] B. Zhang, Y. Wang, C. Ding, Z. Deng, L. Li, Z. Qin, Z. Ding, L. Bian, and C. Yang, "Multi-scale feature pyramid fusion network

for medical image segmentation," *International Journal of Computer Assisted Radiology and Surgery*, vol. 18, no. 2, pp. 353 – 365, 2023.

[32] M. Alam, D. Le, T. Son, J. I. Lim, and X. Yao, "AV-Net: Deep learning for fully automated artery-vein classification in optical coherence tomography angiography," *Biomedical Optics Express*, vol. 11, no. 9, pp. 5249–5257, 2020.

[33] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation," *arXiv preprint arXiv:2102.04306*, 2021.

[34] L. Mou, Y. Zhao, L. Chen, J. Cheng, Z. Gu, H. Hao, H. Qi, Y. Zheng, A. Frangi, and J. Liu, "CS-Net: Channel and spatial attention network for curvilinear structure segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2019, Proceedings*. Springer International Publishing, 2019, pp. 721–730.

[35] Z. Ma, D. Feng, J. Wang, and H. Ma, "Retinal OCTA image segmentation based on global contrastive learning," *Sensors*, vol. 22, no. 24, p. 9847, 2022.

[36] Y. Zhong, M. Xu, and M. Wu, "Dive into Plane: Lightweight & Modular Linear Projection Cross-dimensional Network for Retinal Vessel Segmentation in OCTA Images," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2022, pp. 897 – 901.

[37] H. Ning, C. Wang, X. Chen, and S. Li, "An accurate and efficient neural network for OCTA vessel segmentation and a new dataset," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 1966 – 1970.

[38] Y. Liu, A. Carass, L. Zuo, Y. He, S. Han, L. Gregori, S. Murray, R. Mishra, J. Lei, P. A. Calabresi, S. Saidha, and J. L. Prince, "Disentangled representation learning for OCTA vessel segmenta-

tion with limited training data," *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3686–3698, 2022.

[39] K. Hu, S. Jiang, Y. Zhang, X. Li, and X. Gao, "Joint-Seg: Treat foveal avascular zone and retinal vessel segmentation in OCTA images as a joint task," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–13, 2022.

[40] P. Xiao, X. Hu, K. Ma, G. Wang, Z. Feng, Y. Huang, and J. Yuan, "OMSN and FAROS: OCTA microstructure segmentation network and fully annotated retinal OCTA segmentation dataset," *arXiv preprint arXiv:2212.13059*, 2022.

Nisan Pranavah Raja, Varun P. Gopi,
Bibin Francis, Chelli Devi N

Received August 29, 2025
Revised January 5, 2026
Accepted January 12, 2026

Nisan Pranavah Raja
ORCID: https://orcid.org/0009-0006-4518-6787
National Institute of Technology Tiruchirappalli
Department of Electronics and Communication Engineering, National Institute of Technology Tiruchirappalli, Tamil Nadu-620015, India.
E-mail: 408122004@nitt.edu

Varun P. Gopi
ORCID: https://orcid.org/0000-0001-5593-3949
National Institute of Technology Tiruchirappalli
Department of Electronics and Communication Engineering, National Institute of Technology Tiruchirappalli, Tamil Nadu-620015, India.
E-mail: varun@nitt.edu

Bibin Francis
ORCID: https://orcid.org/0000-0002-9828-6443
National Institute of Technology Tiruchirappalli
Department of Electronics and Communication Engineering, National Institute of Technology Tiruchirappalli, Tamil Nadu-620015, India.
E-mail: bibin@nitt.edu

Chelli Devi N
ORCID: https://orcid.org/0000-0002-3430-7428
Department of Biomedical Engineering, Kalasalingam Academy of Research and Education
Kalasalingam Academy of Research and Education, Krishnankoil, 620015, Tamil Nadu, India
E-mail: n.chellidevi@klu.ac.in