

ROSE: A Retinal OCT-Angiography Vessel Segmentation Dataset and New Model

Yuhui Ma, Huaying Hao, *Member, IEEE*, Jianyang Xie^{ID}, Huazhu Fu^{ID}, *Senior Member, IEEE*, Jiong Zhang^{ID}, Jianlong Yang^{ID}, Zhen Wang, Jiang Liu^{ID}, *Senior Member, IEEE*, Yalin Zheng^{ID}, *Member, IEEE*, and Yitian Zhao^{ID}

Abstract—Optical Coherence Tomography Angiography (OCTA) is a non-invasive imaging technique that has been increasingly used to image the retinal vasculature at capillary level resolution. However, automated segmentation of retinal vessels in OCTA has been under-studied due to various challenges such as low capillary visibility and high vessel complexity, despite its significance in understanding many vision-related diseases. In addition, there is no publicly available OCTA dataset with manually graded vessels for training and validation of segmentation algorithms. To address these issues, for the first time in the field of retinal image analysis we construct a dedicated Retinal OCTA SEgmentation dataset (ROSE), which consists of 229 OCTA images with vessel annotations at either centerline-level or pixel level. This dataset with the source code has been released for public access to assist researchers in the community in undertaking research in related topics. Secondly, we introduce a novel split-based coarse-to-fine vessel segmentation network for OCTA images (OCTA-Net), with the ability to detect thick and thin vessels separately. In the OCTA-Net, a split-based coarse segmentation module is first utilized to produce a preliminary confidence map of vessels, and a split-based refined segmentation module is then used to optimize the shape/contour of the retinal microvasculature. We perform a thorough evaluation of the state-of-the-art vessel seg-

mentation models and our OCTA-Net on the constructed ROSE dataset. The experimental results demonstrate that our OCTA-Net yields better vessel segmentation performance in OCTA than both traditional and other deep learning methods. In addition, we provide a fractal dimension analysis on the segmented microvasculature, and the statistical analysis demonstrates significant differences between the healthy control and Alzheimer's Disease group. This consolidates that the analysis of retinal microvasculature may offer a new scheme to study various neurodegenerative diseases.

Index Terms—Optical coherence tomography angiography, vessel segmentation, deep network, benchmark.

I. INTRODUCTION

THE vasculature is an essential structure in the retina, and its morphological changes can be used not only to identify and classify the severity of systemic, metabolic, and hematologic diseases [1], but also to facilitate a better understanding of disease progression, and assessment of therapeutic effects [2]. Color fundus is the most commonly used retinal imaging technique: however, it is difficult with this method to capture microvasculatures (thin vessels and capillaries), which are surrounded in the fovea and parafovea regions, as shown in the green rectangle area of Fig. 1 (A). Fluorescein angiography and indocyanine green angiography can resolve the retinal vasculature including capillaries, but they are invasive techniques and may cause severe side effects and even death [3].

In contrast, Optical Coherence Tomography Angiography (OCTA) is a newly emerging non-invasive imaging technique, with the ability to produce high-resolution 3D images of the retinal vasculature, and has been increasingly accepted as an invaluable imaging tool to observe retinal vessels [4], [5]. By means of OCTA imaging technology, such as the RTVue XR Avanti SD-OCT system (Optovue, Inc, Fremont, California, USA), equipped with AngioVue software (version 2015.1.0.90), *en face* images of retinal vascular plexus at different depths can be generated by using the maximum projection of OCTA flow signals within specific slabs. Fig. 1 (B-D) show superficial vascular complexes (SVC), deep vascular complexes (DVC), and the inner retinal vascular plexus that includes both SVC and DVC (SVC+DVC). In brief, the SVC extended from 3 μm below the internal limiting membrane (ILM) to 15 μm below the inner plexiform layer (IPL); the DVC extended from 15 to 70 μm below the IPL; and the inner retina extended from 3 μm beneath the ILM

Manuscript received November 16, 2020; accepted December 2, 2020. Date of publication December 7, 2020; date of current version March 2, 2021. This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LZ19F010001, Grant LQ20F030002, and Grant LQ19H180001; in part by the Key Research and Development Program of Zhejiang Province under Grant 2020C03036; and in part by the Ningbo 2025 S&T Megaprojects under Grant 2019B10033 and Grant 2019B10061. Yuhui Ma and Huaying Hao contributed equally to this work. (Corresponding authors: Yalin Zheng; Yitian Zhao.)

Yuhui Ma is with the Cixi Institute of Biomedical Engineering, Ningbo Institute of Materials Technology, Chinese Academy of Sciences, Ningbo 315201, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China.

Huaying Hao, Jianyang Xie, Jianlong Yang, and Yitian Zhao are with the Cixi Institute of Biomedical Engineering, Ningbo Institute of Materials Technology, Chinese Academy of Sciences, Ningbo 315201, China (e-mail: yitian.zhao@nimte.ac.cn).

Huazhu Fu is with the Inception Institute of Artificial Intelligence, Abu Dhabi, United Arab Emirates.

Jiong Zhang is with the Keck School of Medicine, University of Southern California, Los Angeles, CA 90007 USA.

Zhen Wang is with the Department of Neurology, The First Affiliated Hospital, Wenzhou Medical University, Wenzhou 325035, China.

Jiang Liu is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China.

Yalin Zheng is with the Department of Eye and Vision Science, University of Liverpool, Liverpool L69 3BX, U.K. (e-mail: yzheng@liverpool.ac.uk).

Digital Object Identifier 10.1109/TMI.2020.3042802

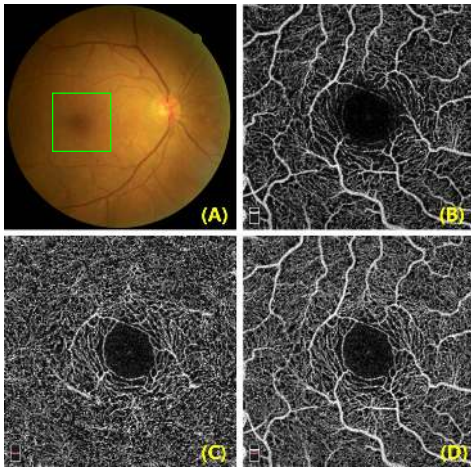


Fig. 1. Illustration of the macula in color fundus image and fovea-centred (green rectangle area) $3 \times 3 \text{ mm}^2$ OCTA images of the same eye: (A) color fundus, (B) superficial vascular complexes (SVC), (C) deep vascular complexes (DVC), and (D) the inner retina vascular plexus including both SVC and DVC (SVC+DVC).

to the outer boundary of the outer plexiform layer [6]. These plexuses were distinguished and separated automatically, using the proprietary tool supplied with the device. OCTA enables observation of microvascular details down to capillary level, permitting quantitative assessment of the microvascular and morphological vessel structures in the retina.

By extracting microvascular structures from different OCTA depth layers, one can obtain their corresponding *en face* projections to analyze their respective variations. In particular, the microvasculature distributed within the parafovea is of great interest, as any abnormality there often indicates the presence of some diseases such as early stage glaucomatous optic neuropathy, diabetic retinopathy, and age-related macula degeneration [7]–[11]. More recently, several studies have shown that the morphological changes of microvasculature revealed by OCTA are associated with Alzheimer’s Disease and Mild Cognitive Impairment [12], [13]. A new avenue is thereby opened up to study the relation between the appearance of retinal vessels and various neurodegenerative diseases. Thus, automatic vessel detection and quantitative analysis of OCTA images are of great value for the early diagnosis of vascular-related diseases affecting retinal circulation, and the assessment of disease progression. However, automated vessel segmentation in OCTA images has been explored only rarely, and remains a challenging task, despite the fact that many medical segmentation approaches - particularly deep learning based techniques [14], [15] - have achieved great success in segmenting blood vessels.

There is no publicly available OCTA dataset with manual vessel annotations, which hinders the further validation of OCTA segmentation techniques. To our best knowledge, only a few automated methods have been developed to segment the retinal vessels from OCTA images based on fully automatic thresholding schemes, such as the methods proposed in [16]–[20]. Recently, several deep learning-based methods were developed for vessel segmentation in OCTA images, and each using its own private OCTA dataset. For instances, Eladawi *et al.* [8] proposed an automatic method on 47 OCTA

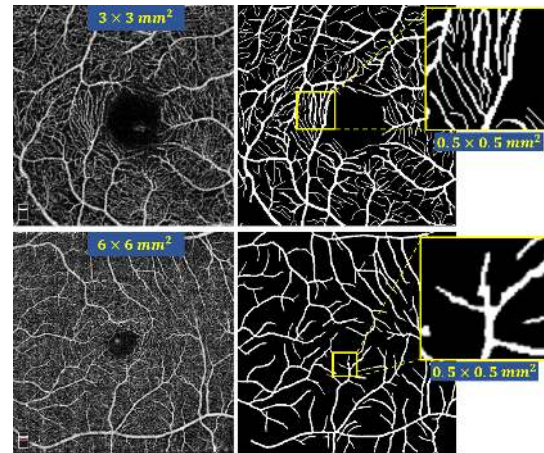


Fig. 2. Comparison of two OCTA images with different scan modes of $3 \times 3 \text{ mm}^2$ (in top row) and $6 \times 6 \text{ mm}^2$ (in bottom row), respectively. Columns 1 to 3 respectively show the original OCTA images, manually annotated vessel networks by experts, and zoomed image patches of the same size ($0.5 \times 0.5 \text{ mm}^2$) in both scans.

images. Zhang *et al.* [11] set up the first 3D OCTA microvascular segmentation approach to directly extract 3D capillary networks from OCTA volume data for subsequent shape modeling and analysis. However, they mainly evaluated the *test re-test reliability* of their 3D framework on 360 OCTA volume images, as there were no manually annotated 3D vessel networks available. Mou *et al.* trained a deep network [1] with 30 OCTA images, but the small dataset used suggests that this method may not be universally applicable across different pathological scenarios. Li *et al.* [21] more recently introduced an image projection network that can achieve 3D-to-2D vessel segmentation: they evaluated their method on 316 OCTA images. Although these methods achieve usable segmentations for OCTA vessel analysis, the privacy of their datasets makes a unified evaluation benchmark impossible.

In addition, Eladawi [8] and Li *et al.* [21] set up their segmentation frameworks based on the OCTA data within a $6 \times 6 \text{ mm}^2$ fovea-centered field of view (FOV). It is worth noting that the scan density for the $3 \times 3 \text{ mm}^2$ is higher than that for the $6 \times 6 \text{ mm}^2$ scans [22], $6 \times 6 \text{ mm}^2$ scans has relatively wider area of scan coverage, and is more like to detect the presence of pathological features, e.g., microaneurysms and non-perfusion. By contrast, the $3 \times 3 \text{ mm}^2$ protocol has a higher scan resolution, and thus is able to delineate the foveal avascular zone (FAZ) and capillaries more clearly than $6 \times 6 \text{ mm}^2$ scans. Previous findings by Zhang *et al.* [11] have shown that small capillaries play a much more important role in distinguishing different disease groups compared with relatively large vessels. It is therefore necessary to established a dedicated OCTA dataset focusing on more detailed capillary networks within a $3 \times 3 \text{ mm}^2$ FOV, and is the main motivation we construct our ROSE dataset. Fig. 2 demonstrates a comparison between $3 \times 3 \text{ mm}^2$ and $6 \times 6 \text{ mm}^2$ FOVs. We may clearly observe the richer capillary information graded by experts from the $3 \times 3 \text{ mm}^2$ scans, while $6 \times 6 \text{ mm}^2$ scan has wider FOV but poorer delineation of microvascular. Therefore, the proposed OCTA-Net will be fully evaluated using the well-established $3 \times 3 \text{ mm}^2$ OCTA dataset for more detailed microvascular study.

The OCTA imaging process typically produces images with a low signal-to-noise ratio (SNR) [5]. Additionally, varying vessel appearances at different depth layers, projection [23], [24], motion and shadow artifacts [25]–[29], and the potential existence of pathologies [11] increase the challenge for achieving accurate segmentations, particularly for densely connected capillaries, which can easily result in discontinuous segmentations. Most deep learning-based methods are region-based [30], a technique which is prone to produce imprecise and discontinuous vessel segmentation results, and existing methods do not perform well when required to detect subtle differences in microvascular networks with different vessel thicknesses and imaging depths.

A. Contributions

In order to mitigate the issues of lacking a public retinal OCTA dataset and effective microvascular segmentation methods, we make the following contributions in this work.

- For the first time in the retinal image analysis field, we establish a publicly available retinal OCTA dataset, with precise manual annotations of retinal microvascular networks, in order to promote relevant research in the community.
- We introduce a novel split-based coarse-to-fine vessel segmentation network for blood vessel segmentation in OCTA, aimed at detecting thick and thin vessels separately. In our method, a split-based coarse segmentation (SCS) module is first utilized to produce a preliminary confidence map of vessels, and a split-based refined segmentation (SRS) module is then used to optimize towards the finer vessels, with a view to obtaining more accurate overall segmentation results.
- We give a full evaluation/benchmarking of OCTA microvascular segmentation, both quantitatively and qualitatively. Comparative analysis shows that the proposed OCTA-Net works robustly on different types of retinal images and yields accurate vessel segmentations.

To further promote developments in this field, the code, baseline models, and evaluation tools, are publicly available at <https://imed.nimte.ac.cn/dataofrose.html>.

II. RELATED WORKS

In the past two decades, we have witnessed a rapid development of retinal blood vessel segmentation methods for color fundus images (as evidenced by extensive reviews [2], [31]). As blood vessels are curvilinear structures distributed across different orientations and scales, the conventional vessel segmentation methods are mainly based on various filters, including Hessian matrix-based filters [32], matched filters [33], multi-oriented filters [34], symmetry filters [2], and tensor-based filters [35]. These filter-based methods aim to suppress non-vascular or non-fiber structures and image noise, thus enhance the curvilinear structures, thereby simplifying the subsequent segmentation problem.

By contrast, automated vessel segmentation from OCTA images is relatively unexplored tasks, and most of the existing methods are based on thresholding schemes. Yousefi *et al.* [16] developed a hybrid Hessian/intensity based method to segment and quantify shape and diameter of vessels. Gao *et al.* [17]

binarized *en face* retinal OCTA images according to mean reflectance projection and maximum decorrelation projection. Camino *et al.* [18] set an optimized reflectance-adjusted threshold to segment the vascular. Sarabi *et al.* [19] developed a three-step algorithm including adaptive thresholding processes to construct the vessel mask. Wu *et al.* [20] proposed an optimized approach based on an improved vascular connectivity analysis (VCA) algorithm to extract the vascular network. However, these thresholding-based approaches are sensitive to noise distributed in *en face* OCTA images and hard to perform well on regions with no significant intensity difference.

In recent years, deep learning-based methods have made significant progress in the fields of medical image segmentation. In particular, many deep neural networks have been modified and applied for blood vessel segmentation [1], [14], [36], [37] and have yielded promising results. However, the extraction of vessels from OCTA images has been relatively unexplored. We will review and discuss the most relevant vessel segmentation works in this section.

A method based on Convolutional Neural Network (CNN) [36] was proposed to enhance training samples for better retinal vessel detection: subsequently, a Conditional Random Field (CRF) was incorporated into the CNN by Fu *et al.* for retinal vessel detection [14]. Wang *et al.* [38] applied the U-Net [39] for retinal vessel segmentation in fundus images of pathological conditions. Xiao *et al.* [40] modified ResU-Net [41] by introducing a weighted attention mechanism for high-quality retinal vessel segmentation. Gu *et al.* [15] proposed a context encoder network (CE-Net), which consists of dense atrous convolution and residual multi-kernel pooling modules for retinal vessel image segmentation. Jin *et al.* [42] integrated deformable convolution into the DUNet, which is designed to extract context information and enable precise localization by combining low-level features with high-level ones. Yan *et al.* [43] proposed a three-stage deep model to segment thick and thin vessels separately, which achieves accurate vessel segmentation for both types of vessels. However, there have been very few deep learning methods for vessel segmentation in OCTA images. Mou *et al.* [1] proposed a channel and spatial attention network (CS-Net) for curvilinear structures (including vessels in some example OCTA images) where they applied spatial and channel attention to further integrate local features with their global dependencies adaptively. Li *et al.* [21] presented an image projection network, which is a novel end-to-end architecture that can achieve 3D-to-2D image segmentation in OCTA images.

III. DATASET

Our constructed **Retinal OCT-Angiography vessel SEgmentation (ROSE)** dataset comprises of two subsets, named as ROSE-1 and ROSE-2, which were acquired by two different devices. All the data described in this paper are acquired from studies that have appropriate approvals from the institutional ethics committees, and written informed consent was obtained from each participant in accordance with the Declaration of Helsinki. The diagnosis result of each subject is provided in the dataset.

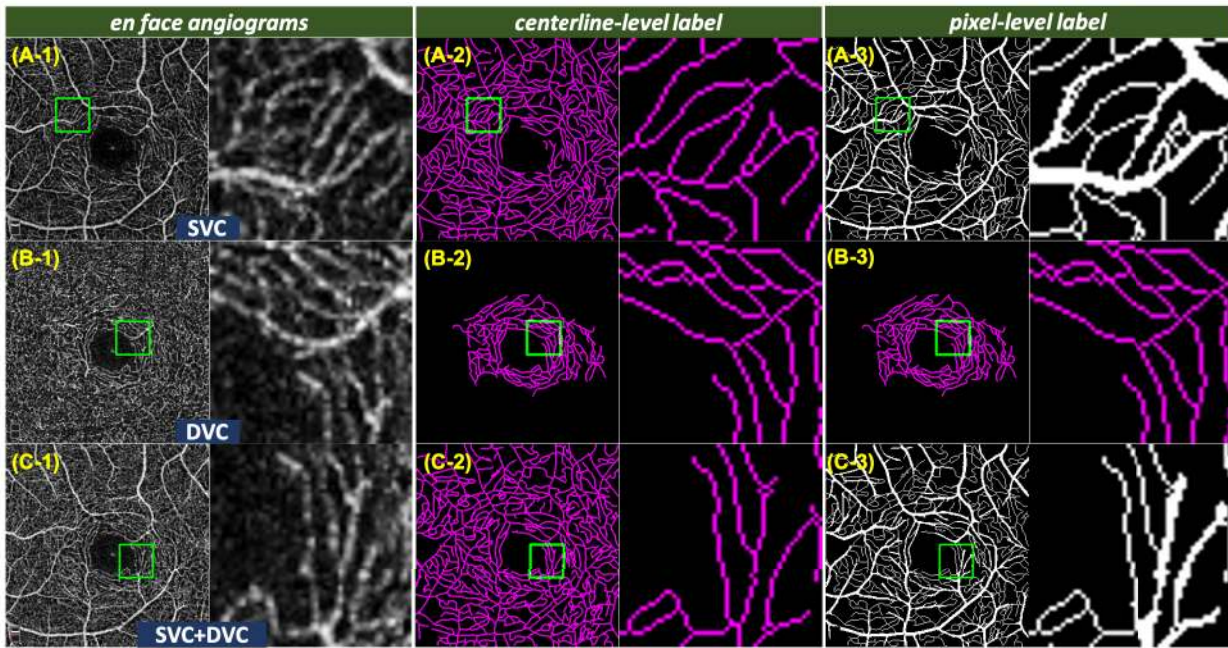


Fig. 3. Illustration of OCTA images from ROSE-1 and their manual annotations. From top to bottom: *en face* of the SVC, DVC and SVC+DVC, respectively. From left to right: *en face*, centerline-level labels, and pixel-level labels, respectively.

A. ROSE-1

The ROSE-1 set consists of a total of 117 OCTA images from 39 subjects (including 26 with Alzheimer’s disease (AD) and 13 healthy controls). The mean age was 68.4 ± 7.4 years of the AD group and 63.0 ± 10.2 years of the control group. Participants with known eye disease, such as glaucoma, age-related macular degeneration, high myopia, etc, and with known systematic disease, e.g., diabetes, were excluded from this study. The diagnosis of AD were based on the NINCDS-ADRDA criteria [44] and participants did not undergo PET imaging or lumbar puncture for assessment of biomarker status. All the OCTA scans were captured by the RTVue XR Avanti SD-OCT system (Optovue, USA) equipped with AngioVue software, with an image resolution of 304×304 pixels. The scan area was $3 \times 3 \text{ mm}^2$ centered on the fovea, within an annular zone of 0.6 mm - 2.5 mm diameter around the foveal center. The SVC, DVC and SVC+DVC angiograms of each participant were obtained.

Two different types of vessel annotations were made by image experts and clinicians for the ROSE-1 dataset, and the consensus of them was then used as the ground truth:

(1) **Centerline-level annotation.** The centerlines of vessels were manually traced using ImageJ software [45] by our experts on the SVC, DVC, and SVC+DVC images individually, as shown in Fig. 3 A-2, B-2, and C-2;

(2) **Pixel-level annotation.** We first invited an image expert to grade the complete microvascular segments with varying diameters in the SVC and SVC+DVC images at pixel level. Since it is difficult for a human expert to perceive the diameters of small capillaries located around the macula region, we asked the expert to grade the small capillaries at centerline level. The combination of these different labels is defined as the final pixel-level annotation, as shown in Fig. 3 A-3, and C-3.

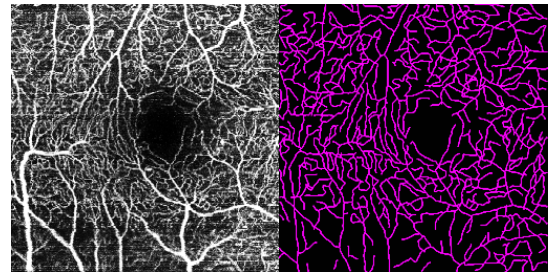


Fig. 4. Example of an OCTA image and its centerline-level manual annotations in ROSE-2.

(Note that, Fig. 3 B-3 is also the centerline-level label of the DVC, as it is difficult to obtain pixel-level grading in this layer.) ROSE-1 dataset were further analysed in Section Discussion to demonstrate the significance of retinal vasculature in the management of neurodegenerative diseases.

B. ROSE-2

The ROSE-2 subset contains a total of 112 OCTA images taken from 112 eyes, acquired by a Heidelberg OCT2 system with Spectralis software (Heidelberg Engineering, Heidelberg, Germany). These images are from eyes with various macula diseases. All the images in this dataset are *en face* angiograms of the SVC within a $3 \times 3 \text{ mm}^2$ area centred at the fovea. These OCTA images were reconstructed from 512×512 repeated A-scans, with the Heidelberg automated real time (ART) and TruTrack system employed to reduce artefacts and noise. Each image was resized into a grayscale image with 840×840 pixels. All the visible retinal vessels were manually traced using an in-house program written in Matlab (Mathworks R2018, Natwick) by an experienced ophthalmologist. An example OCTA image and its corresponding centerline-level annotation are shown in Fig. 4. It should be noted that only the centerlines

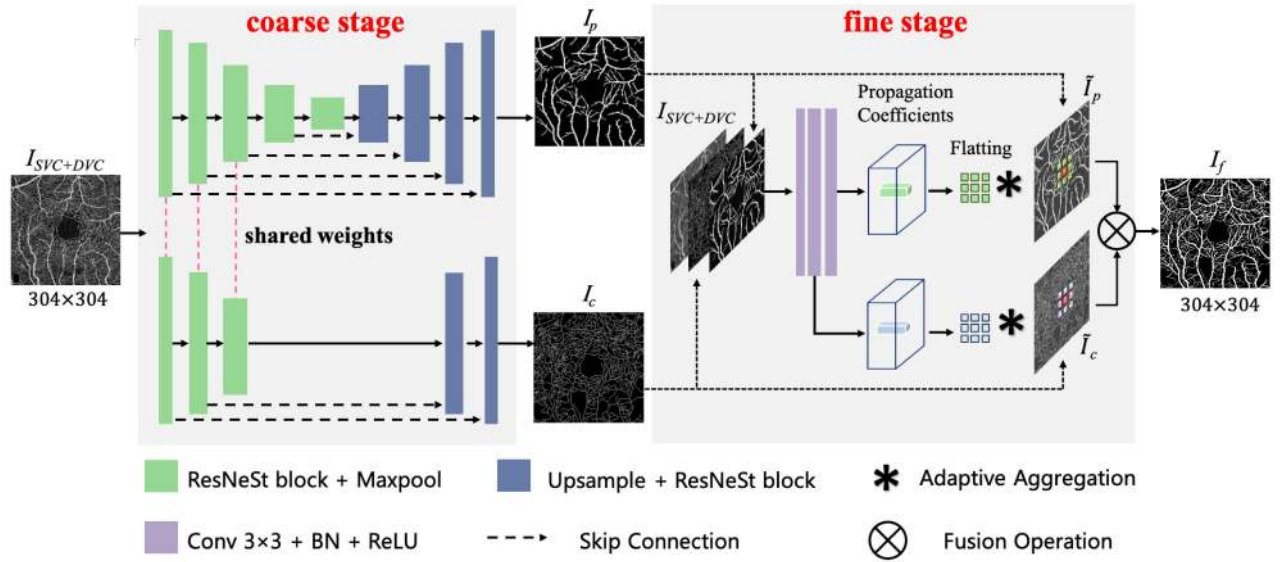


Fig. 5. Architecture of the proposed OCTA-Net network (with an example of *en face* of the angioram of SVC+DVC in the ROSE-1 set). The SCS module (coarse stage) is designed to produce two preliminary confidence maps that segment pixel-level and centerline-level vessels, respectively. The SRS module (fine stage) is then introduced as a fusion network to obtain the final refined segmentation results.

are annotated at single pixel level for ROSE-2. For better visualization, all the centerlines are widened to 7-pixel wide in the illustration figures.

IV. PROPOSED METHOD

In this section, we introduce a novel split-based coarse-to-fine network, named as OCTA-Net, for retinal vessel segmentation in OCTA images. The pipeline of OCTA-Net has two indispensable stages - coarse stage and fine stage, as illustrated in Fig. 5. In the coarse stage, a split-based coarse segmentation (SCS) module is designed to produce preliminary confidence maps. In the fine stage, a split-based refined segmentation (SRS) module is used to fuse these vessel confidence maps to produce the final optimized results.

A. Coarse Stage: Split-Based Coarse Segmentation Module

Since the ROSE-1 set has both pixel-level and centerline-level vessel annotations for each *en face* OCTA image, we design a split-based coarse segmentation (SCS) module with a partial shared encoder and two decoder branches (for pixel-level and centerline-level vessel segmentation, respectively), to balance the importance of both pixel-level and centerline-level vessel information, as illustrated in the coarse stage of Fig. 5. It should be noted that for the ROSE-2 set and *en face* of the DVC layer in the ROSE-1 set on, the designed SCS module only consists of one encoder and one decoder (same architecture as pixel-level vessel segmentation) due to only centerline-level annotations for them.

1) *Pixel-Level Vessel Segmentation*: Pixel-level vessel segmentation is a U-shape network including five encoder layers and the symmetric decoder layers. A ResNet-style structure with split attention module, ResNeSt block [46], is used as the backbone of each encoder-decoder layer. The primary point of ResNeSt block is to regard a series of representations as the combination of different feature groups, then apply

channel-wise attention to these groups. The detailed structure of ResNeSt block is illustrated as Fig. 6 (A). The input $X \in \mathbb{R}^{H \times W \times C}$ of the block is equally split into two cardinal groups which are then fed into two cardinal blocks with same structure respectively. In each cardinal block, the cardinal group is further equally divided and input to two parallel branches. Each branch consists of 1×1 and 3×3 convolutional layers followed by batch normalization (BN) and ReLU layers, and outputs feature maps with size of $H \times W \times C/4$.

In addition, a split attention module is applied in each cardinal block to integrate these feature maps from the two branches, as illustrated in Fig. 6 (B). The module first fuses vessel feature maps from the two branches (denoted as U_1 and U_2) via an element-wise summation, then adopts global pooling (gp) to generate channel-wise statistics s :

$$s = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W [U_1(i, j) + U_2(i, j)]. \quad (1)$$

Two fully connected (FC) layers followed by a softmax layer are then applied to s to obtain a_1 and a_2 , the channel-wise soft attention weights of U_1 and U_2 respectively, as illustrated in Fig. 6 (B).

With the weight vectors a_1 and a_2 , the weighted results of U_1 and U_2 and output of the cardinal block V are: $V_1 = a_1 \cdot U_1$, $V_2 = a_2 \cdot U_2$, $V = V_1 + V_2$, where \cdot represents channel-wise product. Next, outputs of both cardinal blocks (denoted as V^1 and V^2) are fused by concatenation along the channel dimension and one 1×1 convolution $\mathcal{F}_{1 \times 1}$: $Z = \mathcal{F}_{1 \times 1}([V^1, V^2])$. Therefore, the final output Y of the ResNeSt block is produced using a shortcut connection: $Y = Z + T(X)$, where T represents an appropriate transformation, such as stride convolution, combined convolution with pooling or even identity mapping.

2) *Centerline-Level Vessel Segmentation*: Compared with pixel-level annotation, vessel annotation at centerline level aims to grade the vessels in regions with poor contrast,

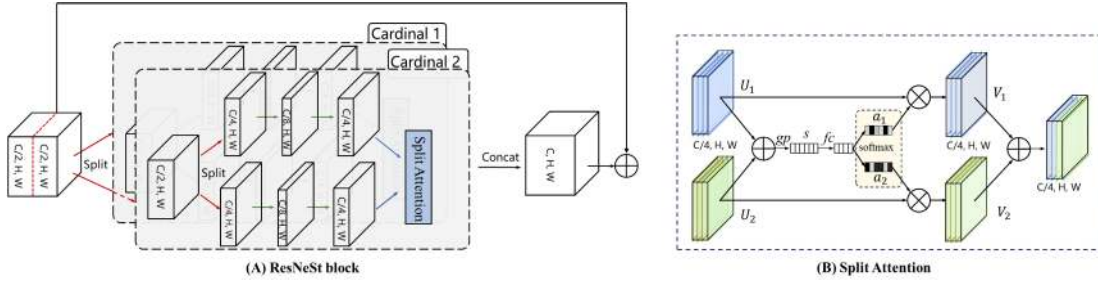


Fig. 6. ResNeSt [46] block with split attention module. (A) ResNeSt block. (B) Split attention module.

more complex topological structures, and relatively smaller diameters. On one hand, considering the differences between centerline-level and pixel-level vessels, the features used for pixel-level vessel segmentation might not be suitable for centerline-level vessel segmentation. The deeper architecture might be detrimental to closer attention to low-level features, which are of great significance for centerline-level vessel segmentation. On the other hand, pixel- and centerline-level vessel segmentation may reveal shared features after feature extraction, due to spatial dependencies between the two types of vessel annotations. Based on the above considerations, we append several ResNeSt blocks followed by an up-sampling layer in the third encoder layer of the backbone, as the decoder of the centerline-level vessel segmentation network. Finally, the outputs of the decoder are processed by one 1×1 convolutional layer with a Sigmoid function to achieve the centerline-level segmentation map.

B. Fine Stage: Split-Based Refined Segmentation Module

In order to further recover continuous details of small vessels, we introduce the fine stage to adaptively refine the vessel prediction results from the coarse stage. Inspired by [47], a split-based refined segmentation (SRS) module is proposed as the fine stage. The structure of our SRS module is illustrated in the fine stage of Fig. 5. In order to fully integrate pixel-level and centerline-level vessel information from the SCS module, the predicted pixel-level and centerline-level vessel maps and the original (single channel) OCTA image are first concatenated as input (total 3 channels) to the SRS module. In addition, the SRS module will produce adaptive propagation coefficients to refine the pixel-level and centerline-level maps respectively. In the SRS module, a mini network including three convolutional layers with 3×3 kernels is designed to refine the pixel-level map from the coarse stage. Besides, one additional 3×3 convolutional layer is appended to the second layer of the mini network to refine the centerline-level map from the coarse stage. BN and ReLU layers are adopted after each convolutional layer. Finally, the refined pixel- and centerline-level maps are then merged into a complete vessel segmentation map, by choosing the larger value from the two maps at each pixel. As in the case of the coarse stage, for the ROSE-2 set and *en face* of the DVC in the ROSE-1 set, no additional convolutional layer is appended to the mini network and only the centerline-level map from the coarse stage is refined.

The detailed channel configuration of the SRS module is shown in Fig. 5. For vessel refinements, the module produces normalized $m \times m$ local propagation coefficient maps for all the positions, formulated as:

$$w_i^p = \frac{\exp(h_i^p)}{\sum_{i=1}^{m \times m} \exp(h_i^p)}, \quad p \in 1, 2, \dots, m \times m, \quad (2)$$

where h_i^p is the confidence value at position i for its neighbor p , and $m \times m$ is the size of propagation neighbors. Finally, the local propagation coefficient vector w_i^p at position i will be multiplied by the confidence map of thick or thin vessels from the front model and aggregate to the center point to generate the refinement result, denoted as:

$$\mathbf{g}_i = \sum_{p=1}^{m \times m} w_i^p \cdot \mathbf{f}_i^p, \quad (3)$$

where \mathbf{f}_i^p is the confidence vector of the neighbor p at position i from the SCS module, and \mathbf{g}_i is the final predicted vector at position i . Note that the propagation coefficient maps can learn the spatial relationship between position i and its neighbors to refine the structure information of vessels. In addition, the final vessel map must be similar to that before refinement. To achieve this goal, the coefficient of position i should be far larger than that of its neighbors, and we adopt a reasonable method for initialization of network parameters following [47]:

$$\begin{cases} k_l(a, c) = \varepsilon, \varepsilon \sim \mathcal{N}(0, \sigma^2) \text{ and } \sigma \ll 1 \\ b_l(c) = \begin{cases} 1 & l = L, c = (m \times m + 1)/2 \\ 0 & \text{others} \end{cases} \\ (l = 1, 2, \dots, L) \end{cases} \quad (4)$$

where L is the number of convolutional layers, k_l and b_l represent convolutional kernels and the bias of layer l respectively, c is the channel of a layer and a is the position in kernels.

V. EXPERIMENTS

A. Experimental Setting

The proposed method was implemented with PyTorch. Both the coarse and the fine stage were trained with 200 epochs and with the following settings: Adam optimization with the initial learning rate of 0.0005, batch size of 2 and weight decay of 0.0001. For more stable training, we adopted poly learning rate policy with a poly power of 0.9.

For the coarse stage, we set $r = 16$ as the reduction ratio of the FC layers in the split attention modules, and selected mean square error (MSE) as the loss function:

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N (p_i - g_i)^2 \quad (5)$$

where N is the number of all pixels, p_i and g_i represents the i -th pixel of the prediction map and the ground truth respectively. For the fine stage, we set $m = 3$ as the size of aggregation neighbors. Considering that there is a large imbalance between vessel regions and the background in *en face* OCTA images, we replaced MSE loss with Dice coefficient loss to further optimize the vessel segmentation in the fine stage:

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N p_i g_i + \epsilon}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2 + \epsilon} \quad (6)$$

where the parameter ϵ is a small positive constant used to avoid numerical problems and accelerate the convergence of the training process. For training and inference of the proposed method, the ROSE-1 subset was split into 90 images for training and 27 images for testing, and the ROSE-2 subset was split into 90 images for training and 22 images for testing. Data augmentation was conducted by randomly rotation of an angle from -10° to 10° during all training stages.

We train our OCTA-Net separately for ROSE-1 and ROSE-2, so the annotation differences will not affect the reliability of the evaluation results. In addition, for ROSE-1 dataset, we have both pixel- and centerline-level annotations, so the proposed method learns features from both types of manual annotations.

B. Evaluation Metrics

To achieve comprehensive and objective assessment of the segmentation performance of the proposed method, the following metrics are calculated and compared:

- Area Under the ROC Curve (AUC);
- Sensitivity (SEN) = $TP / (TP + FN)$;
- Specificity (Specificity) = $TN / (TN + FP)$;
- Accuracy (ACC) = $(TP + TN) / (TP + TN + FP + FN)$;
- *Kappa* score = $(Accuracy - p_e) / (1 - p_e)$;
- False Discovery Rate (FDR) = $FP / (FP + TP)$;
- *G-mean* score [48] = $\sqrt{\text{Sensitivity} \times \text{Specificity}}$;
- Dice coefficient (Dice) = $2 \times TP / (FP + FN + 2 \times TP)$;

where TP is true positive, FP is false positive, TN is true negative, and FN is false negative. p_e in *Kappa* score represents opportunity consistency between the ground truth and prediction, and it is denoted as:

$$p_e = \frac{((TP + FN)(TP + FP) + (TN + FP)(TN + FN))}{(TP + TN + FP + FN)^2} \quad (7)$$

The use of sensitivity and specificity is not adequate for the evaluation of this segmentation task, since over-segmentation still leads to high sensitivity, and the vast majority of pixels do not belong to vessels. Specifically, for centerline-level vessel detection in the DVC images from the ROSE-1 and all images from the ROSE-2, a three-pixel tolerance region around the manually traced centerlines is considered a true positive, which

follows the evaluating methods for extracting one pixel-wide curves in [49].

C. Performance Comparison and Analysis

We have thoroughly evaluated the proposed method over our ROSE dataset, and compared it to existing state-of-the-art segmentation methods to demonstrate the superiority of our OCTA-Net in the segmentation of OCTA microvasculature.

Comparison methods. In order to verify the superiority of our method, we compared our method with other state-of-the-art segmentation methods on both ROSE-1 and ROSE-2, including three conventional methods: infinite perimeter active contour (IPAC) [50], trainable COSFIRE filters [32], and curvelet denoising based optimally oriented flux enhancement (COOF) [11] for their effectiveness in detecting vessels with irregular and oscillatory boundaries; and six deep learning approaches: U-Net [39], ResU-Net [41], CE-Net [15], DUNet [42], CS-Net [1] and three-stage networks [43]. For [50] and [32], the parameters were tuned to achieve segmentation results of all vessels in both ROSE-1 and ROSE-2 subsets. For deep learning approaches, all hyper-parameters were manually adjusted to yield the best performances.

• **Subjective comparisons.** Fig. 7 presents the respective vessel segmentation results produced by the proposed method, backbone, and the other three selected state-of-the-art segmentation networks. We can see that U-Net achieves relatively low performance, due to its over-segmentation at regions with high density. CE-Net and CS-Net achieve better performance than U-Net. However, they are not able to preserve fine capillaries well in terms of producing weak vessel responses. In contrast, the proposed method yields more visually informative results. The benefit of the proposed method for segmentation can be observed from the representative regions (green patches). It is clear from visual inspection that our method has identified more complete and thinner vessels particularly in ROSE-1 subset (shown in purple). It achieves relatively uniform responses in both thick and thin vessels, and provides more sensitive and accurate segmentation on capillaries, as demonstrated in the segmentation results of the DVC and SVC+DVC angiograms.

In contrast, all the methods yield very similar segmentation performance in ROSE-2. Therefore, to better evaluate the performance of the proposed method, we provide quantitative results in the following subsections.

• **Performance on the SVC layer in ROSE-1.** We will first evaluate the vessel segmentation performance on each of the plexus layers of the ROSE-1 subset. Table I quantifies the segmentation performance in SVC images of the different approaches. Overall, our method achieves the best performance in terms of almost all the metrics, with the single exception that its FDR score is 0.0038 lower than that of the three-stage network [43]. Nevertheless, the proposed method is able to correctly identify the majority of vessels using our two-stage architecture. Besides, the last column of Table I also gives paired t-test results in terms of AUC. All $p < 0.05$ demonstrate that the proposed method performs significantly better than the other compared methods.

• **Performance on the DVC layer in ROSE-1.** For the DVC images in the ROSE-1 subset, we first adopted the

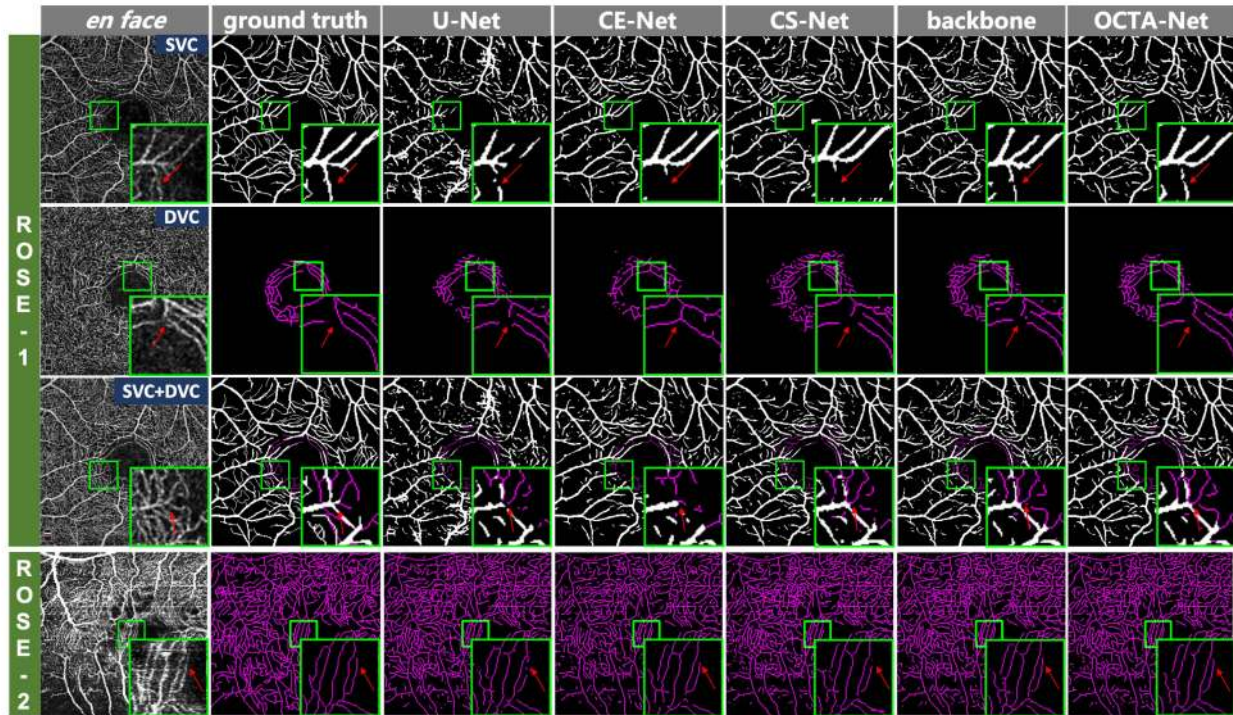


Fig. 7. Vessel segmentation results of different methods on different layers of ROSE-1 and ROSE-2. From left to right: *en face* angiograms (original images), manual annotations, vessel segmentation results obtained by U-Net, CE-Net, CS-Net, backbone (ResNeSt) and the proposed method (OCTA-Net), respectively. Note, the color purple indicates the segmentation results at centerline level.

TABLE I

SEGMENTATION RESULTS OBTAINED USING DIFFERENT METHODS ON THE SVC LAYER FROM ROSE-1

Methods	ROSE-1 (SVC)							
	AUC	ACC	G-mean	Kappa	Dice	FDR	Time (s)	p-value
IPAC [50]	0.8420	0.8245	0.7517	0.4664	0.5751	0.4816	-	<0.001
COSFIRE [32]	0.9286	0.9227	0.7883	0.7089	0.7517	0.0471	-	<0.01
COOF [11]	0.8689	0.8530	0.8161	0.5684	0.6606	0.4121	-	<0.001
U-Net [39]	0.9218	0.8955	0.8068	0.6476	0.7116	0.2627	0.018	<0.001
ResU-Net [41]	0.9252	0.9098	0.8229	0.6911	0.7461	0.2107	0.021	<0.01
CE-Net [15]	0.9292	0.9121	0.8256	0.6978	0.7511	0.1995	0.025	<0.001
DUNet [42]	0.9334	0.9118	0.8249	0.6970	0.7505	0.2006	0.266	<0.001
CS-Net [1]	0.9392	0.9152	0.8304	0.7093	0.7608	0.1883	0.045	<0.001
three-stage [43]	0.9341	0.9179	0.8318	0.7168	0.7663	0.1787	33.255	<0.01
OCTA-Net	0.9453	0.9182	0.8361	0.7201	0.7697	0.1775	0.059	-

TABLE II

SEGMENTATION RESULTS OBTAINED USING DIFFERENT METHODS ON THE DVC IMAGES FROM ROSE-1

Methods	ROSE-1 (DVC)							
	AUC	ACC	G-mean	Kappa	Dice	FDR	Time (s)	p-value
IPAC [50]	0.7563	0.7522	0.7684	0.0636	0.0911	0.9510	-	<0.001
COSFIRE [32]	0.8520	0.9130	0.8523	0.2199	0.2405	0.8516	-	<0.001
COOF [11]	0.8162	0.6678	0.7847	0.0649	0.1003	0.9465	-	<0.001
U-Net [39]	0.9186	0.9895	0.8442	0.6553	0.6605	0.3776	0.018	<0.01
ResU-Net [41]	0.9425	0.9885	0.8553	0.6510	0.6567	0.4002	0.021	<0.05
CE-Net [15]	0.9505	0.9843	0.8503	0.5707	0.5783	0.5147	0.025	<0.05
DUNet [42]	0.9631	0.9864	0.8369	0.6555	0.5850	0.4930	0.266	0.105
CS-Net [1]	0.9671	0.9882	0.8155	0.5825	0.5884	0.4710	0.045	0.991
three-stage [43]	0.9576	0.9894	0.8399	0.6569	0.6622	0.3780	33.255	<0.05
OCTA-Net	0.9673	0.9909	0.8811	0.7028	0.7074	0.3492	0.059	-

TABLE III

SEGMENTATION RESULTS OBTAINED USING DIFFERENT METHODS ON THE SVC+DVC ANGIOGRAMS FROM ROSE-1

Methods	ROSE-1 (SVC+DVC)							
	AUC	ACC	G-mean	Kappa	Dice	FDR	Time (s)	p-value
IPAC [50]	0.7941	0.8007	0.7054	0.3982	0.5223	0.5211	-	<0.001
COSFIRE [32]	0.8800	0.8981	0.7256	0.6125	0.6671	0.0988	-	<0.001
COOF [11]	0.8217	0.7762	0.7742	0.4306	0.5685	0.5465	-	<0.001
U-Net [39]	0.9039	0.8865	0.8050	0.6308	0.7012	0.2892	0.018	<0.001
ResU-Net [41]	0.9108	0.8997	0.8188	0.6689	0.7309	0.2248	0.021	<0.001
CE-Net [15]	0.9155	0.8990	0.8203	0.6678	0.7300	0.2479	0.025	<0.001
DUNet [42]	0.9250	0.9046	0.8213	0.6819	0.7403	0.2223	0.266	<0.001
CS-Net [1]	0.9311	0.9073	0.8263	0.6919	0.7488	0.2137	0.045	<0.001
three-stage [43]	0.9248	0.9090	0.8275	0.6967	0.7524	0.2049	33.255	<0.001
OCTA-Net	0.9375	0.9099	0.8338	0.7022	0.7576	0.2087	0.059	-

ResNeSt backbone to obtain preliminary vessel segmentation results at the coarse stage. Afterwards, the initial segmentations are combined with their original images as inputs of the fine stage for producing final vessel segmentations. Table II demonstrates the segmentation results achieved by our method and the state-of-the-art methods. Although the improvement on AUC is not significant when compared with CS-Net and DUNet, the proposed network outperformed all the other compared approaches. In particular, it significantly outperforms other methods by a large margin, with an increase of about 12.0% and 11.9% in kappa and Dice, respectively, and a reduction of about 13.2% in FDR when compared with CS-Net. These performance improvements are consistent with the segmentation results shown in the middle row of Fig. 7, where the proposed method successfully extracts small capillaries from macula regions with promising continuity and integrity, while other methods produces relatively lower capillary responses.

• **Performance on the SVC+DVC angiograms in ROSE-1.** Table III shows the results of using different segmentation approaches on the SVC+DVC images. Again, the proposed method achieves overall the best performance, with a single exception at the FDR score, where a performance score of 0.0988 is obtained using the method by

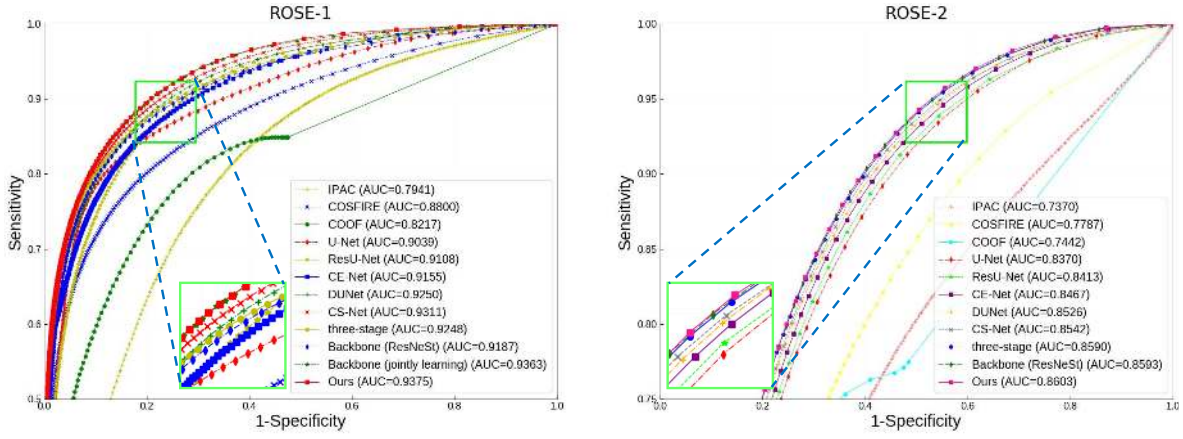


Fig. 8. ROC curves of different methods in terms of SVC+DVC angiograms segmentation performance on ROSE-1 and that of SVC of ROSE-2 datasets.

TABLE IV
SEGMENTATION RESULTS OBTAINED USING
DIFFERENT METHODS ON ROSE-2

Methods	ROSE-2							
	AUC	ACC	G-mean	Kappa	Dice	FDR	Time (s)	p-value
IPAC [50]	0.7370	0.8592	0.8207	0.4758	0.5515	0.5590	-	<0.001
COSFIRE [32]	0.7787	0.9212	0.7742	0.5699	0.6142	0.3891	-	<0.001
COOF [11]	0.7442	0.8945	0.8117	0.5498	0.6112	0.4620	-	<0.001
U-Net [39]	0.8370	0.9319	0.8001	0.6177	0.6564	0.3542	0.018	<0.001
ResU-Net [41]	0.8413	0.9339	0.8074	0.6348	0.6725	0.3337	0.021	<0.001
CE-Net [15]	0.8467	0.9377	0.8248	0.6708	0.7066	0.2930	0.025	<0.001
DUNet [42]	0.8526	0.9372	0.8213	0.6576	0.6935	0.3132	0.266	<0.001
CS-Net [1]	0.8542	0.9385	0.8235	0.6658	0.7010	0.3025	0.045	<0.001
three-stage [43]	0.8590	0.9384	0.8280	0.6671	0.7024	0.3071	33.255	0.244
OCTA-Net	0.8603	0.9386	0.8315	0.6724	0.7077	0.3019	0.059	-

Azzopardi *et al.* [32]. However, detection rate of conventional methods of of Azzopardi *et al.* [32] and Zhao *et al.* [50] are significantly lower than those of all the deep learning-based methods. This shows that the conventional methods have yet to solve the problems as posed by the high degree of anatomical variations across the population, and the varying scales of vessels within an image. Moreover, motion artefacts, noise, poor contrast and low resolution in OCTA exacerbate these problems. By contrast, deep learning-based methods extract a set of higher-level discriminative representations, which are derived from both local and global appearance features and thus can achieve better performance. All $p < 0.001$ demonstrate the significant performance of our method.

• **Performance on ROSE-2.** ROSE-2 only provides centerline-level manual annotation, and includes *en face* OCTA images of the SVC. Therefore, as with the DVC images in ROSE-1 subset, for ROSE-2, our method adopts a ResNeSt backbone at the coarse stage to obtain the preliminary vessel segmentation results. Then, the final results are obtained at the fine stage, using the original image and the preliminary results from the coarse stage as input. Table IV presents the performance of the different segmentation methods. It shows that our method achieves the best AUC, ACC, Kappa and Dice respectively. The statistical analysis reveals that our OCTA-Net outperforms significantly other methods: all $p < 0.001$ with a single exception that when compared with three-stage model.

Moreover, Table I - IV also report the inference time cost of deep learning methods. We don't list the time cost of three conventional methods here because it is unfair to compare these methods run on CPU or even with MATLAB codes that

are not optimized in efficiency with the deep learning methods run on GPU. For fair comparison of inference time, we test all these trained deep learning models with PyTorch. We observed that although consisting of a coarse stage and a fine stage, the proposed two-stage framework is capable to achieve better segmentation performance within a relatively shorter time on both ROSE-1 and ROSE-2.

In order to illustrate the vessel segmentation performance of different methods in a more intuitive manner, we have also provided ROC curves for both the ROSE-1 and ROSE-2 subsets, as shown in Fig. 8. Due to limited space, here we show only the segmentation performances on the SVC+DVC images in ROSE-1. Compared with the conventional methods such as the algorithms proposed by Zhao *et al.* [50] and Azzopardi *et al.* [32], deep learning based methods demonstrate their superiority in segmenting OCTA images. This is because the introduction of excellent modules such as ResNet and attention blocks, are usually helpful in improving the AUC score of the encoder-decoder architecture. In addition, there are two reasons that our two-stage architecture achieves the best ROC curve (shown in red in both the subfigures of Fig. 8). Firstly, using ResNeSt as the backbone of the encoder-decoder architecture further strengthens performance for feature extraction, which improves the extraction of vessel information at different complexities. Secondly, the fine stage can adjust local details on the basis of preliminary results from the coarse stage, which additionally refines the segmentation accuracy of the coarse stage.

VI. DISCUSSION AND CONCLUSION

A. Ablation Studies

In this paper, the proposed vessel segmentation method consists of a ResNeSt backbone, joint learning from pixel-level and centerline-level vessel segmentation, and two-stage training. To validate the effectiveness of these components, we carried out the following ablation studies. U-Net [39] is treated as the baseline encoder-decoder method. Then, we gradually evaluated how each of these components affect the results.

• **Ablation for ResNeSt backbone.** To discuss the performance of the ResNeSt backbone, we compare segmentation performance of the original U-Net, ResU-Net [41] (the modified U-Net with residual blocks in the encoder) and

TABLE V
ABLATION STUDIES OF OUR TWO-STAGE METHOD ON BOTH THE ROSE-1 AND ROSE-2 SUBSETS

Methods	ROSE-1 (SVC+DVC)						ROSE-2					
	AUC	ACC	G-mean	Kappa	Dice	FDR	AUC	ACC	G-mean	Kappa	Dice	FDR
U-Net [39]	0.9039	0.8865	0.8050	0.6308	0.7012	0.2892	0.8370	0.9319	0.8001	0.6177	0.6564	0.3542
ResU-Net [41]	0.9108	0.8997	0.8188	0.6689	0.7309	0.2248	0.8413	0.9339	0.8074	0.6348	0.6725	0.3337
Backbone (ResNeSt)	0.9187	0.9064	0.8228	0.6887	0.7461	0.2112	0.8593	0.9382	0.8293	0.6692	0.7047	0.3052
Backbone (joint learning)	0.9363	0.9091	0.8341	0.7003	0.7564	0.2135	-	-	-	-	-	-
Ours	0.9375	0.9099	0.8338	0.7022	0.7576	0.2087	0.8603	0.9386	0.8315	0.6724	0.7077	0.3019

our proposed encoder-decoder architecture (with ResNeSt as the backbone), as shown in Table V. For both the ROSE-1 and ROSE-2 subsets, our encoder-decoder architecture with ResNeSt as the backbone achieves the best performance on AUC, ACC, Kappa, Dice and FDR in comparison with the original U-Net and ResU-Net. This indicates that the ResNeSt backbone is superior in feature extraction, which reveals more information about vessels with different characteristics.

- **Ablation for joint learning at pixel-level and centerline-level vessel segmentation.** In addition, we compared joint learning from pixel-level and centerline-level vessel segmentation with only one segmentation branch (with ResNeSt as the backbone) of all vessels for the SVC+DVC images in the ROSE-1 subset, so as to demonstrate the advantages of joint learning from both pixel-level and centerline-level vessel segmentation. Comparisons of both performance are illustrated in Table V. We can observe that joint learning achieves higher scores in terms of AUC, ACC, Kappa and Dice than learning from only one segmentation branch. This suggests that joint learning could help to improve both pixel-level and centerline-level vessel segmentation performance by highlighting the relevant topological distinctions between pixel-level and centerline-level vessels.

- **Ablation for two-stage training.** Furthermore, we analysed the impact of the fine stage on the coarse stage in our two-stage procedure. At the coarse stage, for the ROSE-1 subset, pixel-level and/or centerline-level vessel segmentation results are treated as the preliminary segmentation results, while for the ROSE-2 subset, vessel segmentation results produced by the ResNeSt backbone are treated as the preliminary segmentation results. At the fine stage, final vessel segmentation results are derived from both the original images and preliminary results from the coarse stage. Accordingly, we made a comparison between the preliminary segmentation results of the coarse stage and final vessel segmentation results of the fine stage. As illustrated in Table V, final vessel segmentation performance at the fine stage for the most parts shows improvement when compared with results from the coarse stage. The last column of Fig. 7 also indicates that some details of microvasculature are optimized in the fine stage.

B. Clinical Evaluation

It has been suggested that the retina may serve as a window for monitoring and assessing cerebral microcirculation [51] and neurodegeneration conditions [13]. OCT images have been utilized to observe neurodegenerative changes occurring in the ganglion cell-inner plexiform layer (GC-IPL) thickness and retinal nerve fiber layer (RNFL) thickness of AD and MCI

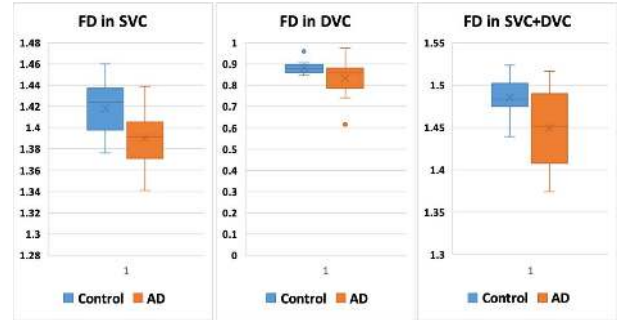


Fig. 9. Descriptive results of the FD of AD patients and controls in the SVC, DVC and SVC+DVC, respectively.

patients [52]. Recently, contributions of vascular biomarkers such as length, density and tortuosity, to the diagnosis of MCI and AD are increasingly recognized [52], [53]. OCTA, as an extension of OCT, that can provide in vivo, noninvasive visualization of the retinal vessels in different layers. With the simultaneous occurrence of both neurodegeneration and microvascular changes in the brain, many studies [51], [54], [55] have suggested that the macula microvasculature may provide vital information on the changes in the cerebral microcirculation during the subclinical phase. Changes in the retinal capillary network may indicate the onset and progression of retinopathy, and fractal dimension (FD) is a well-known measure for characterizing the geometric complexity of retinal vasculature that will be a promising biomarker for vascular diseases [56]–[58] as well as neurodegenerative [59] and cerebrovascular diseases [60]. In particular, Cheung *et al.* [56] indicated that the conventional structural measures (e.g., branching angle and vascular tortuosity) represent only one of the many aspects of the retinal vascular geometry and are lack of single global measure that can summarize the branching pattern of the vasculature as a whole. On the other hand, FD reflects the overall complexity [58], and has been used as a global measure of the geometric pattern of the retinal vasculature potentially representing the complex branching pattern of the microvasculature [59].

In this work, we performed an FD analysis by applying a box-counting method named Fraclab [61] on the segmentation results of the SVC, DVC, and SVC+DVC images in ROSE-1 using the proposed method. The box-plots in Fig. 9 show the statistical analysis results on the ROSE-1 dataset, including 39 images of normal and 78 images of AD subjects, and each subject has three OCTA angiograms: SVC, DVC, and SVC+DVC. We have done three statistical tests with one for each of SVC, DVC and SVC+DVC in order to avoid the potential correlation issues amongst them. It can be observed

that the AD group has reduced FD in the SVC, DVC and SVC+DVC when compared with the control group. Student's t-test was employed to assess the differences between the AD and control groups and results confirmed that the differences between the AD and control participants are significant in the SVC ($p=0.004<0.05$), DVC ($p=0.028<0.05$) and SVC+DVC ($p=0.007<0.05$), respectively. All the three tests have demonstrated significant differences between AD and healthy groups. These results are consistent with the previous findings that retinal microvascular changes may reflect neurodegenerative changes [13], [52].

C. Projection Artefacts

Projection artefacts are one of the main limitations of OCTA images [62]. In this work we used the exported OCTA images as it is and didn't applied any further artifacts removal method. However, Optovue has incorporated their own proprietary three-dimensional projection artifact removal technique into their AngioVue software, and so the ROSE-1 dataset we used in fact have undergone some artifacts removal. In real scenario, projection artefacts may still exist in OCTA images and require further technological improvements, which however is not the focus of the paper. In our experience if a segmentation works on noisy images it will often works well (if not better) in more clean images. We expect that the proposed approach will continue to work on images free of projection artefacts.

VII. CONCLUSIONS

In this paper, we have presented a novel Retinal OCTA SEgmentation dataset (ROSE) dataset, which is a large, carefully designed and systematically manually-annotated dataset for vessel segmentation in OCTA images. To our best knowledge, this is the first OCTA dataset released to the research community for the vessel segmentation task. It contains two sub-sets, where the images were acquired by two different devices. All the vessels were manually annotated by human experts at either centerline level and/or pixel level. All of the images contained in the dataset were eventually used for clinical diagnostic purposes. To ensure the utmost protection of patient privacy, the identities of all patients have been removed and cannot be reconstructed. We plan to keep growing the dataset with more challenging situations and various types of eye and neurodegenerative diseases, such as diabetic retinopathy and Parkinson's disease.

In addition to the new dataset, we further proposed a novel two-stage framework for vessel segmentation in OCTA images. In the coarse stage, a split-based coarse segmentation (SCS) module has been designed to achieve the preliminary segmentation results: ResNeSt block is used as the backbone of the framework. In the fine stage, a split-based refined segmentation (SRS) module has been adopted to improve the vessel segmentation results by utilizing both the original images and the preliminary results from the coarse stage.

The experimental results on the ROSE dataset show that our vessel segmentation approach outperforms other state-of-the-art methods. Small capillaries are major components of the retinal microvasculature and many of them are very

tiny segments with only 2 – 4 pixel width. Due to the high noise ratio and low capillary visibility in OCTA images, large vessels can be easily extracted while the segmentation of small capillaries becomes challenging. Thus, small increases of metric values may already indicate significant improvement of segmentation with more extracted capillaries. We have conducted paired t-tests on the segmentation performance of our method and other methods. The statistical analysis in Section V.C demonstrates competitive performance of the proposed method. In particular, the improvements of small capillaries in quantitative evaluation metrics are quite significant also due to the extreme imbalance between foreground and background regions. The sub-analysis on AD shows the great potential of exploring retinal microvascular-based analysis for the diagnosis of various neurodegenerative diseases.

REFERENCES

- [1] L. Mou *et al.*, "CS-Net: Channel and spatial attention network for curvilinear structure segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, 2019, pp. 721–730.
- [2] Y. Zhao *et al.*, "Automatic 2-D/3-D vessel enhancement in multiple modality images using a weighted symmetry filter," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 438–450, Feb. 2018.
- [3] M. T. Witmer, G. Parlitsis, S. Patel, and S. Kiss, "Comparison of ultra-widefield fluorescein angiography with the Heidelberg spectralis noncontact ultra-widefield module versus the optos optomap," *Clin. Ophthalmol.*, vol. 7, pp. 389–394, Feb. 2013.
- [4] R. Leitgeb, "En face optical coherence tomography: A technology review," *Biomed. Opt. Exp.*, vol. 10, no. 5, pp. 2177–2201, 2019.
- [5] Y. Jia *et al.*, "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography," *Opt. Exp.*, vol. 20, no. 4, pp. 4710–4725, 2012.
- [6] J. P. Campbell *et al.*, "Detailed vascular anatomy of the human retina by projection-resolved optical coherence tomography angiography," *Sci. Rep.*, vol. 7, no. 1, p. 42201, Mar. 2017.
- [7] Y. Zhao *et al.*, "Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 51–63, Jan. 2017.
- [8] N. Eladawi *et al.*, "Automatic blood vessels segmentation based on different retinal maps from OCTA scans," *Comput. Biol. Med.*, vol. 89, pp. 150–161, Oct. 2017.
- [9] M. Alam, D. Toslak, J. I. Lim, and X. Yao, "Color fundus image guided artery-vein differentiation in optical coherence tomography angiography," *Investigative Ophthalmol. Visual Sci.*, vol. 59, no. 12, pp. 4953–4962, 2018.
- [10] J. Zhang, A. H. Kashani, and Y. Shi, "3D surface-based geometric and topological quantification of retinal microvasculature in OCT-angiography via Reeb analysis," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*. Shenzhen, China: Springer, 2019, pp. 57–65.
- [11] J. Zhang *et al.*, "3D shape modeling and analysis of retinal microvasculature in OCT-angiography images," *IEEE Trans. Med. Imag.*, vol. 39, no. 5, pp. 1335–1346, May 2020.
- [12] D. C. DeBuc, G. M. Somfai, and A. Koller, "Retinal microvascular network alterations: Potential biomarkers of cerebrovascular and neural diseases," *Amer. J. Physiol. Heart Circulatory Physiol.*, vol. 312, no. 2, pp. H201–H212, Feb. 2017.
- [13] S. P. Yoon *et al.*, "Retinal microvascular and neurodegenerative changes in Alzheimer's disease and mild cognitive impairment compared with control participants," *Ophthalmol. Retina*, vol. 3, no. 6, pp. 489–499, Jun. 2019.
- [14] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*. Athens, Greece: Springer, 2016, pp. 132–139.
- [15] Z. Gu *et al.*, "CE-net: Context encoder network for 2D medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2281–2292, Oct. 2019.
- [16] S. Yousefi, T. Liu, and R. K. Wang, "Segmentation and quantification of blood vessels for OCT-based micro-angiograms using hybrid shape/intensity compounding," *Microvascular Res.*, vol. 97, pp. 37–46, Jan. 2015.

- [17] S. S. Gao et al., "Compensation for reflectance variation in vessel density quantification by optical coherence tomography angiography," *Investigat. Ophthalmol. Vis. Sci.*, vol. 57, no. 10, pp. 4485–4492, Aug. 2016.
- [18] A. Camino, M. Zhang, L. Liu, J. Wang, Y. Jia, and D. Huang, "Enhanced quantification of retinal perfusion by improved discrimination of blood flow from bulk motion signal in OCTA," *Transl. Vis. Sci. Technol.*, vol. 7, no. 6, p. 20, Dec. 2018.
- [19] M. S. Sarabi, J. K. Gahm, M. M. Khansari, J. Zhang, A. H. Kashani, and Y. Shi, "An automated 3D analysis framework for optical coherence tomography angiography," *BioRxiv*, Art. no. 655175, May 2019, doi: 10.1101/655175.
- [20] S. Wu et al., "An optimized segmentation and quantification approach in microvascular imaging for OCTA-based fibrovascular regression monitoring," to be published, doi: 10.21203/rs.3.rs-51265/v1.
- [21] M. Li et al., "Image projection network: 3D to 2D image segmentation in OCTA images," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3343–3354, Nov. 2020.
- [22] N. A. Iafe, N. Phasukkijwatana, X. Chen, and D. Sarraf, "Retinal capillary density and foveal avascular zone area are age-dependent: Quantitative analysis using optical coherence tomography angiography," *Investigative Ophthalmol. Vis. Sci.*, vol. 57, no. 13, pp. 5780–5787, 2016.
- [23] T. T. Hormel, J. Wang, S. T. Bailey, T. S. Hwang, D. Huang, and Y. Jia, "Maximum value projection produces better en face OCT angiograms than mean value projection," *Biomed. Opt. Exp.*, vol. 9, no. 12, pp. 6412–6424, 2018.
- [24] A. Zhang, Q. Zhang, and R. K. Wang, "Minimizing projection artifacts for accurate presentation of choroidal neovascularization in OCT micro-angiography," *Biomed. Opt. Exp.*, vol. 6, no. 10, pp. 4130–4143, 2015.
- [25] R. F. Spaide, J. G. Fujimoto, and N. K. Waheed, "Image artifacts in optical coherence angiography," *Retina*, vol. 35, no. 11, p. 2163, 2015.
- [26] F. K. Chen, R. D. Viljoen, and D. M. Bukowska, "Classification of image artefacts in optical coherence tomography angiography of the choroid in macular diseases," *Clin. Experim. Ophthalmol.*, vol. 44, no. 5, pp. 388–399, Jul. 2016.
- [27] A. Camino, Y. Jia, J. Yu, J. Wang, L. Liu, and D. Huang, "Automated detection of shadow artifacts in optical coherence tomography angiography," *Biomed. Opt. Exp.*, vol. 10, no. 3, pp. 1514–1531, 2019.
- [28] K. G. Falavarjani, M. Al-Sheikh, H. Akil, and S. R. Sada, "Image artefacts in swept-source optical coherence tomography angiography," *Brit. J. Ophthalmol.*, vol. 101, no. 5, pp. 564–568, May 2017.
- [29] I. C. Holmen et al., "Prevalence and severity of artifacts in optical coherence tomographic angiograms," *JAMA Ophthalmol.*, vol. 138, no. 2, pp. 119–126, 2020.
- [30] Z. Zhang, H. Fu, H. Dai, J. Shen, Y. Pang, and L. Shao, "Et-net: A generic Edge-attention guidance network for medical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, 2019, pp. 442–450.
- [31] M. M. Fraz et al., "Blood vessel segmentation methodologies in retinal images—A survey," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 407–433, Oct. 2012.
- [32] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable COS-FIRE filters for vessel delineation with application to retinal images," *Med. Image Anal.*, vol. 19, no. 1, pp. 46–57, Jan. 2015.
- [33] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. W. Pluijm, R. Duits, and B. M. T. H. Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE Trans. Med. Imag.*, vol. 35, no. 12, pp. 2631–2644, Dec. 2016.
- [34] J. Zhang, Y. Chen, E. Bekkers, M. Wang, B. Dashtbozorg, and B. M. T. H. Romeny, "Retinal vessel delineation using a brain-inspired wavelet transform and random forest," *Pattern Recognit.*, vol. 69, pp. 107–123, Sep. 2017.
- [35] S. Cetin and G. Unal, "A higher-order tensor vessel tractography for segmentation of vascular structures," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 2172–2185, Oct. 2015.
- [36] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
- [37] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation," 2018, *arXiv:1802.06955*. [Online]. Available: <http://arxiv.org/abs/1802.06955>
- [38] X. Wang et al., "Retina blood vessel segmentation using a U-Net based convolutional neural network," in *Proc. Comput. Sci., Int. Conf. Data Sci.*, 2018, pp. 8–9.
- [39] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, 2015, pp. 234–241.
- [40] X. Xiao, S. Lian, Z. Luo, and S. Li, "Weighted res-UNet for high-quality retina vessel segmentation," in *Proc. 9th Int. Conf. Inf. Technol. Med. Edu. (ITME)*, Oct. 2018, pp. 327–331.
- [41] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018.
- [42] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "DUNet: A deformable network for retinal vessel segmentation," *Knowl. Based Syst.*, vol. 178, pp. 149–162, Aug. 2019.
- [43] Z. Yan, X. Yang, and K.-T. Cheng, "A three-stage deep learning model for accurate retinal vessel segmentation," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 4, pp. 1427–1436, Jul. 2019.
- [44] B. Dubois et al., "Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria," *Lancet Neurol.*, vol. 6, no. 8, pp. 734–746, Aug. 2007.
- [45] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, Jul. 2012.
- [46] H. Zhang et al., "ResNeSt: Split-attention networks," 2020, *arXiv:2004.08955*. [Online]. Available: <http://arxiv.org/abs/2004.08955>
- [47] R. Zhang, S. Tang, M. Lin, J. Li, and S. Yan, "Global-residual and local-boundary refinement networks for rectifying scene parsing predictions," in *Proc. 26th Int. Joint Conf. Artif. Intell. (IJCAI)*, Aug. 2017, pp. 3427–3433.
- [48] J.-H. Ri, G. Tian, Y. Liu, W.-H. Xu, and J.-G. Lou, "Extreme learning machine with hybrid cost function of G-mean and probability for imbalance learning," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 9, pp. 2007–2020, Sep. 2020.
- [49] P. Guimaraes, J. Wigdahl, and A. Ruggeri, "A fast and efficient technique for the automatic tracing of corneal nerves in confocal microscopy," *Transl. Vis. Sci. Technol.*, vol. 5, no. 5, p. 7, Sep. 2016.
- [50] Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images," *IEEE Trans. Med. Imag.*, vol. 34, no. 9, pp. 1797–1807, Sep. 2015.
- [51] C. S. Lee et al., "Associations between recent and established ophthalmic conditions and risk of Alzheimer's disease," *Alzheimer's Dementia*, vol. 15, no. 1, pp. 34–41, Jan. 2019.
- [52] J. A. van de Kreeke et al., "Optical coherence tomography angiography in preclinical Alzheimer's disease," *Brit. J. Ophthalmol.*, vol. 104, no. 2, pp. 157–161, 2020.
- [53] D. S. Grewal, B. W. Polascik, G. C. Hoffmeyer, and S. Fekrat, "Assessment of differences in retinal microvasculature using OCT angiography in Alzheimer's disease: A twin discordance report," *Ophthalmic Surg., Lasers Imag. Retina*, vol. 49, no. 6, pp. 440–444, 2018.
- [54] A. London, I. Benhar, and M. Schwartz, "The retina as a window to the brain—From eye research to CNS disorders," *Nature Rev. Neurol.*, vol. 9, no. 1, pp. 44–53, 2013.
- [55] M. Bulut et al., "Evaluation of optical coherence tomography angiographic findings in Alzheimer's type dementia," *Brit. J. Ophthalmol.*, vol. 102, no. 2, pp. 233–237, Feb. 2018.
- [56] N. Cheung et al., "Quantitative assessment of early diabetic retinopathy using fractal analysis," *Diabetes Care*, vol. 32, no. 1, pp. 106–110, Jan. 2009.
- [57] G. N. Thomas et al., "Measurement of macular fractal dimension using a computer-assisted program," *Invest. Ophthalmol. Vis. Sci.*, vol. 55, no. 4, pp. 2237–2243, Apr. 2014.
- [58] T. J. MacGillivray, N. Patton, F. N. Doubal, C. Graham, and J. M. Wardlaw, "Fractal analysis of the retinal vascular network in fundus images," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2007, pp. 6455–6458.
- [59] D. C. DeBuc, G. M. Somfai, E. Arthur, M. Kostic, S. Oropesa, and C. M. Santiesteban, "Investigating multimodal diagnostic eye biomarkers of cognitive impairment by measuring vascular and neurogenic changes in the retina," *Frontiers Physiol.*, vol. 9, p. 1721, Dec. 2018.
- [60] S. Lemmens, A. Devulder, K. Van Keer, J. Bierkens, P. De Boever, and I. Stalmans, "Systematic review on fractal dimension of the retinal vasculature in neurodegeneration and stroke: Assessment of a potential biomarker," *Frontiers Neurosci.*, vol. 14, p. 16, Jan. 2020.
- [61] K. J. Falconer, "Fractal geometry: Mathematical foundations and applications," *Biometrics*, vol. 46, no. 3, p. 886, 1990.
- [62] A. Shahlaee, W. A. Samara, J. Sridhar, S. K. Kasi, J. Hsu, and A. C. Ho, "Accentuation of optical coherence tomography angiography projection artefacts on hyper-reflective retinal layers," *Acta Ophthalmol.*, vol. 96, no. 7, pp. e883–e884, Nov. 2018.