

Image Projection Network: 3D to 2D Image Segmentation in OCTA Images

Mingchao Li, Yerui Chen, Zexuan Ji, Keren Xie, Songtao Yuan, Qiang Chen, and Shuo Li

Abstract—We present an image projection network (IPN), which is a novel end-to-end architecture and can achieve 3D-to-2D image segmentation in optical coherence tomography angiography (OCTA) images. Our key insight is to build a projection learning module (PLM) which uses a unidirectional pooling layer to conduct effective features selection and dimension reduction concurrently. By combining multiple PLMs, the proposed network can input 3D OCTA data, and output 2D segmentation results such as retinal vessel segmentation. It provides a new idea for the quantification of retinal indicators: without retinal layer segmentation and without projection maps. We tested the performance of our network for two crucial retinal image segmentation issues: retinal vessel (RV) segmentation and foveal avascular zone (FAZ) segmentation. The experimental results on 316 OCTA volumes demonstrate that the IPN is an effective implementation of 3D-to-2D segmentation networks, and the uses of multi-modality information and volumetric information make IPN perform better than the baseline methods.

Index Terms—End to end, 3D to 2D, image projection network, biomedical volumetric image segmentation, optical coherence tomography angiography.

I. INTRODUCTION

The quantification of retinal indicators plays a vital role in the clinical decision and quantitative study of ophthalmopathy [1]–[5]. For example, Early Treatment Diabetic Retinopathy Study (ETDRS) grids assess early macular thickening in diabetic patients by establishing the retinal thickness zoning [6], and retinal vessel density is used to diagnose the health of retinal vascular system [7]. These indicators provide a basis for clinical diagnosis of retinal diseases and significantly improve the decision-making efficiency of doctors. These conveniences are inseparable from the advancement of observation tools and the improvement of algorithm performance.

Optical coherence tomography (OCT) is a noninvasive optical imaging modality that uses coherent light to capture 3D structural data of retina with micrometer-resolution [8], [9], as shown in Fig. 1(b). Comparing to color fundus imaging

technology, OCT can acquire more detailed information about retinal structures and thus becomes a leading modality in the clinic observation of retinopathy.

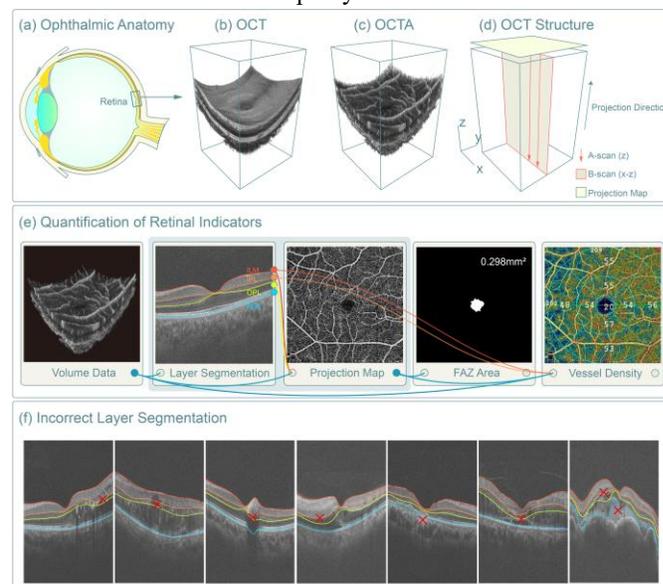


Fig. 1. (a-d) The structures of OCT and OCTA. (e) Quantification of retinal parameters which depend on layer segmentation and projection maps. (f) Layer segmentation error caused by the destruction of the retinal structure.

Building on OCT platform, OCT angiography (OCTA) provides blood flow information of retina and choroid [10]–[12], as shown in Fig. 1(c). This promising tool can clearly display the blood vessel and microvascular system, which makes it possible to quantify the blood vessel and the foveal avascular zone (FAZ) [7]. It overcomes the shortcomings that traditional OCT cannot provide blood flow information finely and has gradually become a common observation tool in clinical routine.

Both OCT and OCTA can provide 3D data, but most retinal indicators, such as the vessel density and the FAZ area, are quantified on the projection maps rather than 3D space. This is because the flat structure of the retina makes the projection maps more informative and convenient for doctors to observe. However, the generation of these quantitative indicators

This work was supported in part by National Natural Science Foundation of China (61671242), Key R&D Program of Jiangsu Science and Technology Department (BE2018131), and “111” Program B13022.

M. C. Li, Y. R. Chen, Z. X. Ji and *Q. Chen are with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (correspondence e-mail: chen2qiang@njust.edu.cn).

K. R. Xie, *S. T. Yuan are with Department of Ophthalmology, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, China (correspondence e-mail: yuansongtao@vip.sina.com).

S. Li is with the University of Western Ontario, London ON, Canada. (slishuo@gmail.com)

Copyright (c) 2019 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

depends on the retinal layer segmentation and projection images obtained by the layer segmentation (Fig. 1(e)). For example, in order to obtain a clear OCTA projection image for measuring the area of FAZ, it is often necessary to eliminate the interference of choroidal blood flow signals, so the projection maps are usually generated by the maximum value from the internal limiting membrane (ILM) layer to the outer plexiform layer (OPL). Besides, the statistics of retinal thickness, vessel density and other indicators also rely on layer segmentation to define the statistical interval.

The retinal layer segmentation plays an important role in quantifying retinal indicators. Although a large number of layer segmentation algorithms have been developed [13]–[18], and have achieved high accuracy in the normal population, they will often fail when retinal diseases destroy the retinal layer structure. In clinical practice, most of the subjects of OCT and OCTA data are from patients with retinal diseases, and the phenomenon of layer segmentation algorithm failure is widespread, as shown in Fig. 1(f). The failure of layer segmentation leads to the difficulty of quantifying indicators and has become a bottleneck in the field of retinal disease analysis.

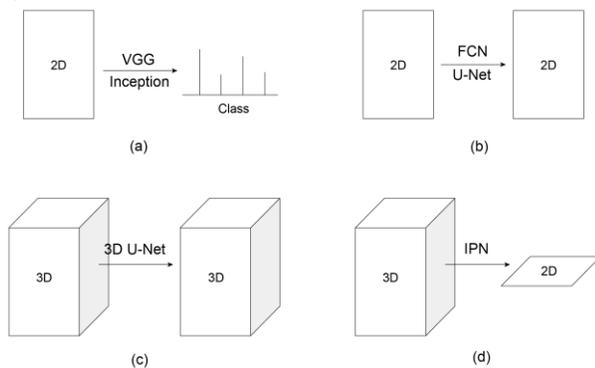


Fig. 2. Task patterns in mainstream end-to-end deep learning frameworks (a-c) and our proposed IPN (d).

Is layer segmentation necessary for quantifying retinal indicators? In this paper, we explore how to quantify the retinal indicators using deep learning without layer segmentation. Deep learning has made significant achievements in image classification and semantic segmentation. It can fulfill the task of classification or segmentation end to end. Fig. 2 shows several mainstream end-to-end networks, including: (1) The classification network represented by VGG [19] and Inception [20], [21] that can achieve the task from 2D to category. (2) The two-dimensional segmentation network represented by fully convolutional networks (FCN) [22] and U-Net [23] that can achieve 2D to 2D semantic segmentation. (3) The three-dimensional segmentation network represented by 3D U-Net [24] that is characterized by the segmentation results from 3D volume data to 3D labels. However, these networks cannot obtain the 2D retina indicators directly from the 3D OCTA data. Alternatively, they need 3D pixel-to-pixel labels, which are labor-intensive and difficult to be obtained. Therefore, we propose a novel end-to-end architecture named image projection network (IPN), as shown in Fig. 3. IPN can summarize the effective features in 3D data along the projection

direction and output the segmentation results on a 2D plane, to realize the semantic segmentation from 3D to 2D.

We tested the performance of our network on two crucial issues in the quantification of retinal indicators: (1) Retinal vessel (RV) segmentation, (2) Foveal avascular zone (FAZ) segmentation. The experiments on 316 OCTA volumes from different patients demonstrate that our network is effective and achieves the highest segmentation accuracy in these issues.

The main contributions of this paper can be highlighted as: (1) We propose IPN, which is a novel 3D-to-2D segmentation network that can be applied to the multiple segmentation challenges in OCTA images. (2) We introduce a projection learning module that uses a unidirectional pooling layer to conduct effective features selection and dimension reduction concurrently. (3) The segmentation process of our network does not need retinal layer segmentation and projection maps. (4) The uses of multi-modality information and volumetric information make the segmentation results more accurate and robust.

II. RELATED WORK

OCTA is a relatively new, non-invasive imaging technique that generates volumetric angiography images in a matter of seconds [25]. The blood flow signals provided by OCTA can be used to quantify blood vessels and foveal avascular zone. The automated segmentation of the blood vessels and the foveal avascular zone has become two crucial issues in the quantification of retinal indicators [7].

1) Retinal vessel segmentation

The retinal vessel reflects the health of retina and is one of the determining factors in an ophthalmic examination [26]. The retinal vessel segmentation plays a crucial role in diagnosing relevant diseases [27]–[32]. At present, most of the RV segmentation tasks are carried out on the color fundus images [26], [33]–[38]. Due to the limited ability of OCT images to present blood information, only a few approaches [39]–[43] have been reported. The clear 3D display of the vascular system in OCTA has reinvigorated interest in the quantification of retinal vessels.

The widely used quantitative indicators of retinal vessels, such as vessel density [7] and vessel tortuosity [44], require RV segmentation to obtain binary blood vessel image. Several RV segmentation methods in OCTA images have been reported. A simple method is based on adaptive thresholding binarization to obtain RV with a higher intensity than that of the background [7]. However, the threshold-based method will generate a lot of background noise. Eladawi et al. [45] used a joint Markov-Gibbs random field model to segment the retinal vessels in different OCTA projection maps. Li et al. [46] developed an algorithm based on top-hat filter and optimally oriented flux to detect the capillary plexus. These methods [45] [46] can remove background noise well in the projection images and obtain the binary vessel images. However, they still have two limitations: (1) Lack of the precise distinction between large vessels and capillary plexus. Sometimes, they also generate large vessel masks by thresholding filter images. However, these large

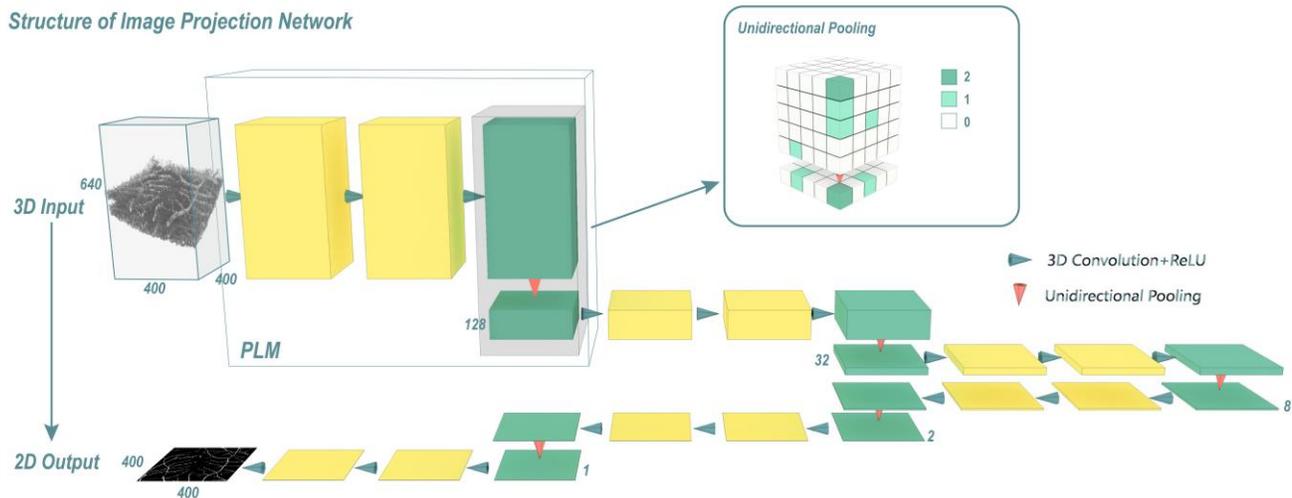


Fig. 3. The structure of IPN for 3D-to-2D end-to-end image segmentation.

vessel masks with noise and rough boundaries cannot be considered as the precise segmentation of large vessels. (2) Rely on projection maps. The projection maps in different retinal plexus have high specificity and are lack of tolerance to incorrect layer segmentation, which will affect the performance of quantitative indicators. In this paper, we try to use IPN to segment the large vessels to achieve the accurate segmentation of the large blood vessels, which can avoid the impact of layer segmentation errors and does not depend on projection maps.

2) Foveal avascular zone segmentation

Before the appearance of OCTA, due to the limited ability of imaging equipment to display the retinal vascular system, it was difficult to describe the strict boundary of FAZ, and only a few works on FAZ segmentation were reported [47]–[49]. OCTA makes it possible to quantify FAZ, and then a large number of studies have focused on the differences of FAZ in different populations [29], [50]–[56]. With the increasing attention to FAZ, the automatic FAZ segmentation in OCTA has become necessary research. Lu et al. [5] designed an algorithm based on an active contour model to detect FAZ. Díaz et al. [57] developed an automatic FAZ segmentation algorithm based on morphological operation in four different projection maps. Guo et al. [58] used a fully convolutional network to segment avascular area including FAZ and non-perfusion areas. All of these works are based on OCTA projection images, which need layer segmentation to eliminate the interference of choroidal blood flow signals. Different from these methods, our network can learn the spatial structure information of FAZ and the end-to-end generation mode also avoids the retinal layer segmentation and can achieve higher segmentation accuracy.

3) End-to-end Network for Semantic Segmentation

End-to-end networks have achieved unprecedented success in semantic segmentation. One of the simplest and most popular frameworks is FCN [22], which achieves excellent results on natural images. Based on FCN, many new structures have been proposed to improve segmentation accuracy, such as across-layer connections, which can solve the loss of resolution. A successful FCN named U-net [23], added a contracting path to capture context and a symmetric expanding path to identify the

location of objects, and achieves outstanding performance on biomedical segmentation. DeepLab [59] used dilated convolution kernels to expand the receptive field of the convolutional neural network. In order to realize three-dimensional semantic segmentation in medical images, the U-net was extended to 3D U-Net [24].

The above networks implement end-to-end semantic segmentation from pixels to pixels. However, they cannot summarize the effective features along a certain dimension to generate a specific projection map, which is the 3D-to-2D learning mode we need. Ji’s method [60] is one of the representative tasks to achieve 2D segmentation using volumetric information. However, instead of using a 3D-to-2D segmentation network, they use a classification network to classify each column in the 3D volume as a feature vector and then obtains the 2D segmentation results by splicing and voting. The limited use of volumetric information and the occupation of a large number of computing resources are still their limitations.

To address the above limitations, we proposed a novel end-to-end framework, called image projection network (IPN) which achieves 3D to 2D semantic segmentation like a projection process. In the following section, we will describe and discuss the implementation and development of our IPN architecture.

III. METHODS

Before designing the IPN from 3D to 2D, we first design a projection network from 2D to 1D as a pre-experiment to verify whether the neural network can summarize the effective features along the projection direction.

A. 2D-to-1D IPN

We use the framework of the classical VGG model for reference, remove all the full connection layers, and change the original pooling layer to the unidirectional pooling layer. The purpose of those changes is to select effective features along the projection direction and reduce the data dimension to condense the data onto the two-dimensional projection plane.

1) *Unidirectional pooling*

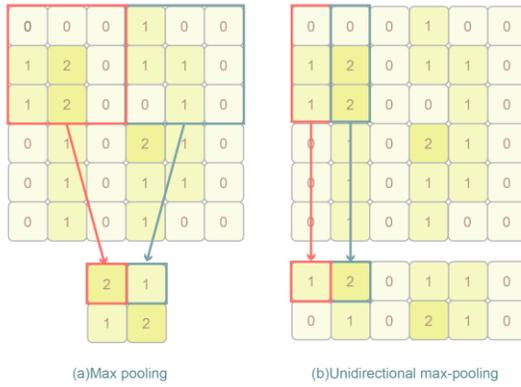


Fig. 4. The structures of 3×3 max-pooling and unidirectional max-pooling.

The pooling layer can reduce network parameters and control over-fitting, especially max pooling, which plays a role in feature selection to a certain extent. Thus, our network uses the max-pooling layer to reduce the image dimension. Fig. 4(a) shows a 3×3 max-pooling layer. After pooling, the image size in all directions decreases. However, we expect to reduce the image size only in the projection direction, while the image size in other directions remains unchanged. We change the size of pooling kernels and propose a unidirectional pooling, as shown in Fig. 4(b). By adjusting the size of pooling, the unidirectional pooling keeps the dimensions unchanged except the projection direction. After several unidirectional pooling, the size of 2D data in the projection direction will be gradually compressed and eventually become a 1D projection vector.

2) *The structure and limitation*

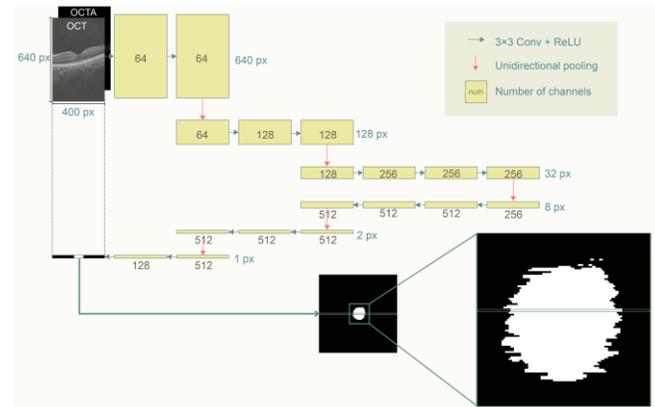


Fig. 5. The structure of 2D-to-1D IPN and an example of FAZ segmentation results.

After introducing the unidirectional pooling, the final structure of 2D-to-1D IPN is designed as shown in Fig. 5. We use this network to train each B-scan of OCT and OCTA images for FAZ segmentation. The input of the network is 2D B-scan image and the output is a 1D vector. The final segmentation result is obtained by splicing, which has poor spatial continuity because it contains a lot of jagged edges. Although segmentation results show the limitations of 2D-to-1D IPN, they imply the effectiveness of this projection network, which paves the way for the design of 3D to 2D networks.

TABLE I

THE ARCHITECTURE PARAMETERS OF IMAGE PROJECTION NETWORK

| | Channel number | PLM parameter | Output size |
|---------|----------------|---------------|-----------------------------|
| Input | 2 | - | $640 \times 100 \times 100$ |
| PLM1 | 32 | 5 | $128 \times 100 \times 100$ |
| PLM2 | 64 | 4 | $32 \times 100 \times 100$ |
| PLM3 | 128 | 4 | $8 \times 100 \times 100$ |
| PLM4 | 256 | 4 | $2 \times 100 \times 100$ |
| PLM5 | 512 | 2 | $1 \times 100 \times 100$ |
| Conv6 | 256 | - | $1 \times 100 \times 100$ |
| Conv7 | 128 | - | $1 \times 100 \times 100$ |
| Conv8 | 2 | - | $1 \times 100 \times 100$ |
| Softmax | 2 | - | $1 \times 100 \times 100$ |

B. *3D-to-2D IPN*

To avoid the poor spatial continuity in the 2D-to-1D network and obtain better segmentation results, we further proposed the 3D-to-2D image projection network by making full use of 3D volumetric information.

The structure of the 3D-to-2D IPN is shown in Fig. 3. Unlike 2D-to-1D IPN, 3D-to-2D IPN uses 3D convolution instead of 2D convolution and the unidirectional pooling extends from 2D to 3D but still occurs only in the projection direction. With this change, IPN can input three-dimensional images and output two-dimensional labels. Besides, we modularize the network structure and construct the projection learning module (PLM), as shown in Fig. 6(a).

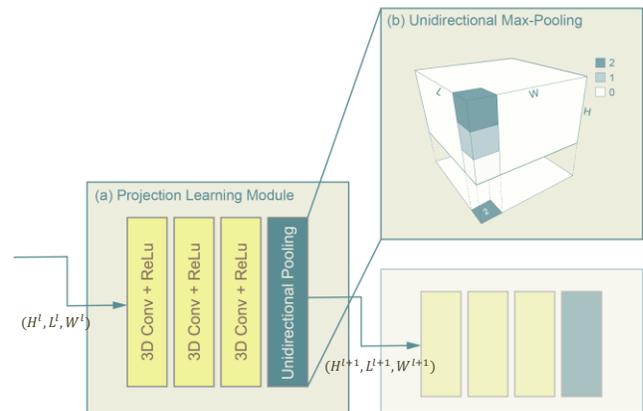


Fig. 6. The structure of the PLM and the unidirectional pooling in 3D-to-2D IPN.

1) *Projection learning module*

PLM consists of three 3D convolution layers and one unidirectional pooling layer. The convolution layers are used to extract image features and the unidirectional pooling layer is used to select effective features along the projection direction. In this paper, the kernel size of the 3D convolution is $3 \times 3 \times 3$, and the activation function is ReLU function. We use the unidirectional max-pooling shown in Fig. 6(b). Assume that the input size of the l th PLM is (H^l, L^l, W^l) , the output size is $(H^{l+1}, L^{l+1}, W^{l+1})$, the kernel size is k^l , and the stride size is d^l , then the relationship between output size and input size is as follows:

$$\begin{cases} H^{l+1} = \frac{H^l - k^l}{d^l} + 1 \\ L^{l+1} = L^l \\ W^{l+1} = W^l \end{cases} \quad (1).$$

In this paper, we make d^l equal to k^l . Then Eq. (1) is simplified to Eq. (2). After several PLM modules, the three-dimensional data is finally compressed into a plane, which is the projection plane we need.

$$\begin{cases} H^{l+1} = \frac{H^l}{k^l} \\ L^{l+1} = L^l \\ W^{l+1} = W^l \end{cases} \quad (2).$$

PLM is the primary component module of the 3D-to-2D IPN, which plays a vital role in the process of projection learning. To explore its influence on the network, we will discuss the relationship between the number of PLM modules and the segmentation accuracy and speed in Sec. IV.

At the end of the network, we use convolution layers to reduce the number of channels to aggregate the 2D plane information obtained by the PLM module. The parameter settings of the 3D-to-2D IPN are shown in Table I. For different segmentation tasks, there are some various implementation details.

2) Details for RV segmentation

OCTA can display the strength of blood flow signals, and the blood flow in retinal vessels has a higher intensity, as shown in Fig. 7. More attention has been paid to the distribution of blood vessels, so the task of RV segmentation is to detect the location of these vessels in the direction of the projection map and to obtain the segmentation results, as shown in Fig. 7(a). The retinal vessels are mainly distributed in the surface layer of the retina, although artifacts can be produced in the deep layer of the retina, as shown in Fig. 7 (b), the position of the artifacts is consistent with that of the real vessels in the projection direction, so they do not affect the display of the retinal vessels in the projection maps.

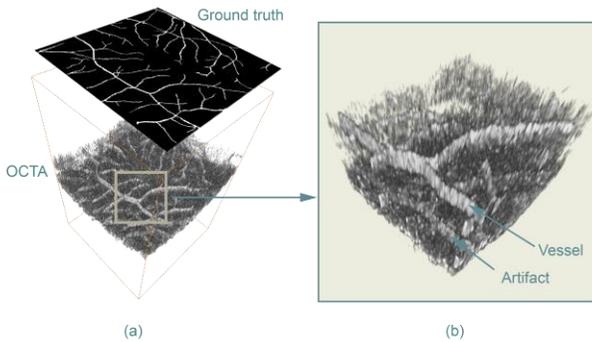


Fig. 7. Blood flow intensity information in OCTA and ground truth of RV segmentation.

The input data is 3D OCT and OCTA images, which occupy two channels when entering the network. The size of OCT and OCTA data we use is $640 \text{ px} \times 400 \text{ px} \times 400 \text{ px}$. Due to the limitation of computing resources, we need to split them into blocks. The size of blocks is $640 \text{ px} \times 100 \text{ px} \times 100 \text{ px}$, and each OCT and OCTA data is divided into 16 blocks during the test stage. Because of the uniform distribution of blood vessels in the projection map, the random sampling method is used in the training stage.

3) Details for FAZ segmentation

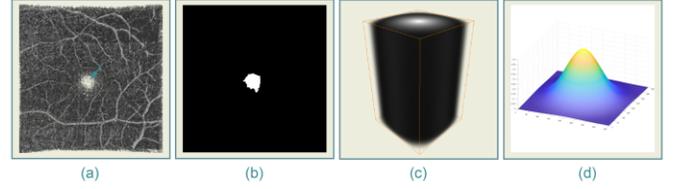


Fig. 8. (a) Foveal avascular zone in OCTA. (b) Ground truth. (c) Distance map. (d) Sampling weight map in the training process.

FAZ is a region devoid of retinal vessels within the fovea. Figs. 8(a) and (b) show a FAZ in three-dimensional OCTA data and the ground truth, respectively. Since the OCTA images are often taken from the center of the fovea, FAZ is located in the center of the images. Considering the importance of the location information for FAZ segmentation, based on the two-input channel of the RV segmentation network, we add a distance map channel (Fig. 8(c)) as the third channel. The distance map is generated as follows:

$$\begin{cases} \alpha(x, y, z) = \sqrt{(x - x_c)^2 + (y - y_c)^2} \\ \text{Dismap}(x, y, z) = \left(1 - \frac{\alpha(x, y, z)}{\max \alpha(x, y, z)}\right)^\gamma \end{cases} \quad (3),$$

where (x_c, y_c) is the center of projection images; γ is a scaling factor. In this work, $x_c = y_c = 200$, $\gamma = 2$.

An essential challenge in FAZ segmentation is the imbalance of positive and negative samples. FAZ occupies only a small part of the projection map, while the background accounts for a large part. Therefore, instead of the random sampling in RV segmentation, we adopt normal distribution sampling centered at the projection center, as shown in Fig. 8(d). In this way, the center position has a higher probability of being selected as training data, thus increasing the proportion of positive samples.

Other settings are consistent with the RV segmentation.

IV. EXPERIMENTS AND RESULTS

A. Data

The proposed IPN was validated on a database that includes 316 OCTA volumes and the corresponding OCT volumes from 293 subjects. These OCT volumes and OCTA volumes, providing structural and fluid information of the retina, respectively, are from the same commercial 70 kHz spectral domain OCT system with a center wavelength of 840 nm (RTVue-XR, Optovue, CA). In this system, the OCTA volumes are obtained from multiple OCT volumes through the split-spectrum amplitude-decorrelation (SSADA) algorithm [12]. Each OCT volume has a size of $640 \text{ px} \times 400 \text{ px} \times 400 \text{ px}$ corresponding to a $2 \text{ mm} \times 6 \text{ mm} \times 6 \text{ mm}$ volume centered at the retinal macular region. The OCTA volume size is $160 \text{ px} \times 400 \text{ px} \times 400 \text{ px}$. The vertical resolution of the OCTA image is 1/4 of that of the OCT image, so we used bilinear interpolation to stretch the size of OCTA volume into $640 \text{ px} \times 400 \text{ px} \times 400 \text{ px}$ to match the OCT images.

The 316 eyes were included and imaged from Jiangsu Province Hospital between March 2018 to September 2018. The average age of the subjects was 49.07 ± 17.56 years old. To ensure the diversity of the data, the images were collected not only from normal retina but also from various retinopathies that

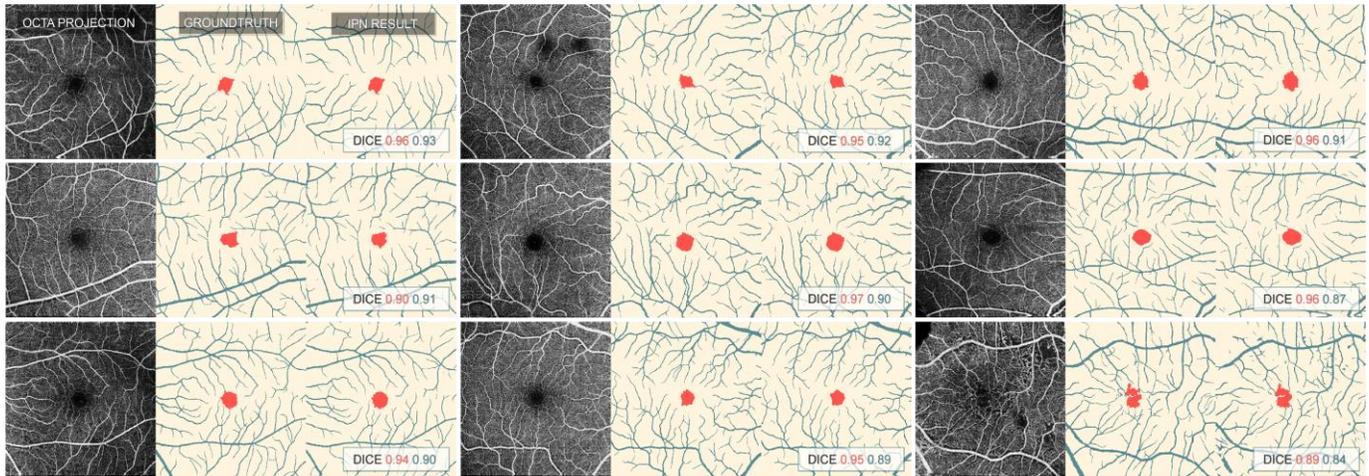


Fig. 9. The results of FAZ segmentation and RV segmentation using IPN. Blue areas represent retinal vessels and the blue number represent its Dice score. Red areas represent FAZ and the red number represent its Dice score.

are encountered commonly in clinical practice, such as age-related macular degeneration (AMD), retinal vein occlusion (RVO), central serous chorioretinopathy (CSC), choroidal neovascularization (CNV), diabetic retinopathy (DR) and so on. From the dataset, we randomly selected 200 eyes as the training set for training network, 16 eyes as the validation set for model selection and threshold selection, and 100 eyes as the testing set for performance evaluation. Table II lists the detailed

TABLE II

GENDER AND PATHOLOGY DISTRIBUTION IN THE TRAINING SET, TESTING SET AND VALIDATION SET

| | | Train | Test | Val | Total |
|-----------|--------|-------|------|-----|-------|
| Gender | Male | 126 | 69 | 11 | 195 |
| | Female | 74 | 47 | 5 | 121 |
| Pathology | AMD | 32 | 12 | 4 | 48 |
| | DR | 18 | 12 | 4 | 34 |
| | CSC | 9 | 5 | 0 | 14 |
| | CNV | 7 | 6 | 0 | 13 |
| | RVO | 8 | 4 | 0 | 12 |
| | Other | 59 | 26 | 1 | 86 |
| | Normal | 67 | 35 | 7 | 109 |
| Total | | 200 | 100 | 16 | 316 |

distribution of retinal diseases and gender in training, testing and validation sets. Other conditions not listed include retinal detachment, optic atrophy, epiretinal membrane, retinitis pigmentosa, retinoschisis, etc.

The ground truth of RV and FAZ is drawn on the OCTA maximum projection maps between the ILM layer and OPL. Because of the exclusion of artifacts and choroidal blood flow signals, the maximum projection maps display inner retinal vessels more clearly than the average projection maps. The ILM layer and OPL were generated by software (OCTExplorer 3.8). Ten students and three experts participated in the manufacture and revision of the ground truth. The ground truth drawing of the RV is as follows: (1) The large blood vessels in the inner retina are the segmentation targets in this study, which have a relatively obvious vascular topology and high signal intensity. (2) Capillary plexus, lesion signals and background noise are

excluded. The capillary plexus in the inner retina with $6 \text{ mm} \times 6 \text{ mm}$ field is different from the large vessel target, because it has no clear vascular topology and relatively low signal intensity.

B. Implementation Details

We use Adam stochastic optimization of the TensorFlow framework. To enable the training of 3D networks, we use the memory efficient cuDNN3 convolution layer implementation. The models are run on 1 NVIDIA GeForce GTX 1080Ti GPU. We use cross-entropy loss function with batch size 3, and the maximal number of iterations 20k and initial learning 10^{-4} . The standard normal initialization method is used to initialize the network with a variance of 0.02.

In order to ensure that the network is not over-fitting, we calculate the Dice coefficient of the validation set during the training process and save the best model when the Dice coefficient of the validation set arrives at the maximum value. In the test stage, we output the target probability image of the validation set and the test set. We determine the final threshold by maximizing the Dice coefficient on the validation set, and this threshold is used to obtain the final binary results of the test set for evaluation. The training process of IPN takes ~ 10 hours, and the testing speed is ~ 3 s/volume. The test speed of the network is related to the number of PLM modules, and we will discuss their relationship in Sec. IV-D.

C. Evaluation metrics

To assess network performance, we adopt five metrics to evaluate the segmentation results quantitatively:

- (1) Dice coefficient (DICE): $2TP/(2TP + FP + FN)$
- (2) Jaccard index (JAC): $TP/(TP + FP + FN)$
- (3) Balance-accuracy (BACC): $(TPR + TNR)/2$
- (4) Precision (PRE): $TP/(TP + FP)$
- (5) Recall (REC): $TP/(TP + FN)$

where TP is true positive, FP is false positive, TN is true negative, and FN is false negative. Especially in our segmentation task, since the area of RV and FAZ is smaller than that of background, using accuracy to evaluate the results may

TABLE III
QUANTITATIVE COMPARISON OF DIFFERENT METHODS ON THE TEST SET (MEAN ± SD)

| No. | Issue | Method | DICE (%) | JAC (%) | BACC (%) | PRE (%) | REC (%) |
|-----|------------------|-------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| 1 | FAZ | 2D-to-1D | 81.74±15.67 | 71.63±19.38 | 90.59±9.82 | 86.21±15.38 | 81.33±19.62 |
| | | 3D-to-2D | 88.61±11.61 | 81.23±16.35 | 94.71±6.52 | 89.92±13.54 | 89.56±13.04 |
| 2 | RV | OCT | 83.02±4.20 | 71.17±5.68 | 88.70±3.16 | 88.37±3.29 | 78.57±6.41 |
| | | OCTA | 87.78±2.73 | 78.33±4.25 | 91.52±2.34 | 92.30±2.28 | 83.85±4.77 |
| | | OCT+OCTA | 88.15±2.77 | 78.92±4.34 | 91.71±2.28 | 92.68±2.39 | 84.20±4.62 |
| | | OCT+OCTA+D* | 88.61±11.61 | 81.23±16.35 | 94.71±6.52 | 89.92±13.54 | 89.56±13.04 |
| 3 | FAZ | OCT | 72.25±21.52 | 60.28±22.60 | 88.00±12.70 | 76.20±23.04 | 76.36±25.45 |
| | | OCTA | 83.07±19.25 | 74.78±23.22 | 90.99±10.54 | 89.26±18.65 | 82.16±21.06 |
| | | OCT+OCTA | 83.92±16.62 | 75.10±20.08 | 91.71±9.15 | 88.80±15.45 | 83.58±18.33 |
| | | OCT+OCTA+D* | 88.61±11.61 | 81.23±16.35 | 94.71±6.52 | 89.92±13.54 | 89.56±13.04 |
| | RV | PRO ¹ +FCN | 76.07±4.27 | 61.56±5.28 | 85.24±3.31 | 80.54±3.63 | 72.43±6.75 |
| | | PRO ² +FCN | 81.70±2.26 | 69.12±3.18 | 88.36±1.92 | 85.74±2.66 | 78.18±3.91 |
| | | PRO ¹ +U-Net | 82.27±4.21 | 70.09±5.73 | 88.34±2.97 | 87.45±4.22 | 77.93±5.93 |
| | | PRO ² +U-Net | 86.92±2.34 | 76.94±3.60 | 91.20±1.97 | 91.01±2.75 | 83.34±4.04 |
| FAZ | IPN | 88.15±2.77 | 78.92±4.34 | 91.71±2.28 | 92.68±2.39 | 84.20±4.62 | |
| | Lu et al. [5] | 71.73±25.30 | 60.87±25.82 | 88.60±14.10 | 71.84±27.25 | 77.48±28.09 | |
| | Díaz et al. [57] | 79.70±21.25 | 69.88±21.15 | 94.64±11.39 | 73.53±21.67 | 89.62±22.58 | |
| | IPN | 88.61±11.61 | 81.23±16.35 | 94.71±6.52 | 89.92±13.54 | 89.56±13.04 | |
| 4 | RV | IPN-U | 78.84±2.27 | 65.13±3.05 | 87.96±1.88 | 79.63±2.14 | 78.18±3.77 |
| | | IPN-UC | 87.45±2.93 | 77.81±4.49 | 91.39±2.52 | 91.84±2.32 | 83.65±5.15 |
| | | IPN | 88.15±2.77 | 78.92±4.34 | 91.71±2.28 | 92.68±2.39 | 84.20±4.62 |
| | FAZ | IPN-U | 84.79±11.66 | 75.13±15.40 | 93.71±6.43 | 84.51±14.32 | 87.61±12.87 |
| | | IPN-UC | 86.51±13.96 | 78.42±18.25 | 93.39±7.88 | 89.35±14.64 | 86.92±15.75 |
| | | IPN | 88.61±11.61 | 81.23±16.35 | 94.71±6.52 | 89.92±13.54 | 89.56±13.04 |

D* is distance map. PRO¹ is OCTA full-projection map. PRO² is OCTA maximum-projection map between ILM layer and OPL.

lead to overestimation and loss of significance. In order to evaluate accuracy when the positive and negative samples are unbalanced, we use the balance-accuracy instead of the general accuracy to evaluate the results. In this metric, TPR is true positive rate, and TNR is true negative rate.

D. Performance

Fig. 9 shows several typical results of FAZ segmentation and RV segmentation using the proposed IPN. The last case in Fig. 9 has the lowest Dice score in the RV segmentation due to the ambiguity of some vessels under the influence of DR disease. The average Dice scores for RV segmentation and FAZ segmentation in the test set are 0.88 and 0.89, respectively. The Quantitative and qualitative results demonstrate that the proposed IPN is effective.

We consider that the effectiveness of IPN is mainly attributed to the learning ability of PLM. The number of PLMs determines the size of the network, which further affects the segmentation accuracy and speed. In this paper, we explore the relationship between the number of PLMs and the network performance.

We change the number of PLMs to explore its impact on the segmentation results. In these experiments, we reduce the channel number of all convolution layers in the PLM module to 64. The size of unidirectional pooling kernel is obtained by factorization. We evaluate the test speed of the network and the Dice score of the segmentation results on the test set, as shown in Fig. 10. It shows that with the increase of the PLM number,

the testing speed of the network slows down gradually, and the Dice score of the test set gains when the number of PLMs is small. Taking into account the accuracy and speed, we set the number of the PLM modules as 5.

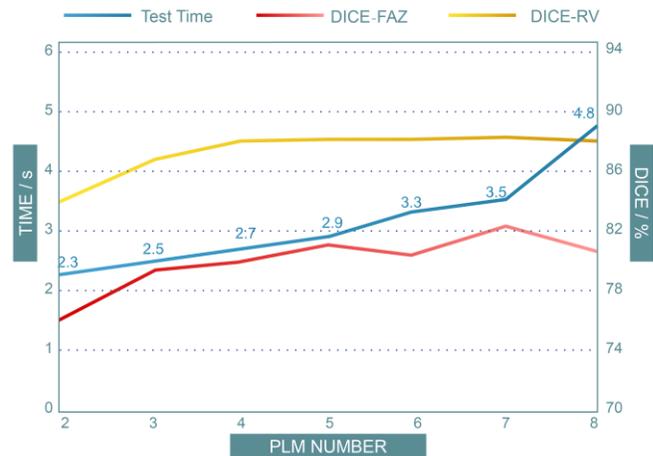


Fig. 10. The relationship between the PLM number and the network performance.

In order to explore the strengths and weaknesses of our network in detail, we design the following comparative experiments: (1) We compare the 3D-to-2D IPN with 2D-to-1D IPN in Sec. IV-E1). (2) We discuss the effects of single modality input and multiple modality input in Sec. IV-E2). (3) We compare our IPN with the reported methods and the mainstream deep learning methods in Sec. IV-E3). It shows that

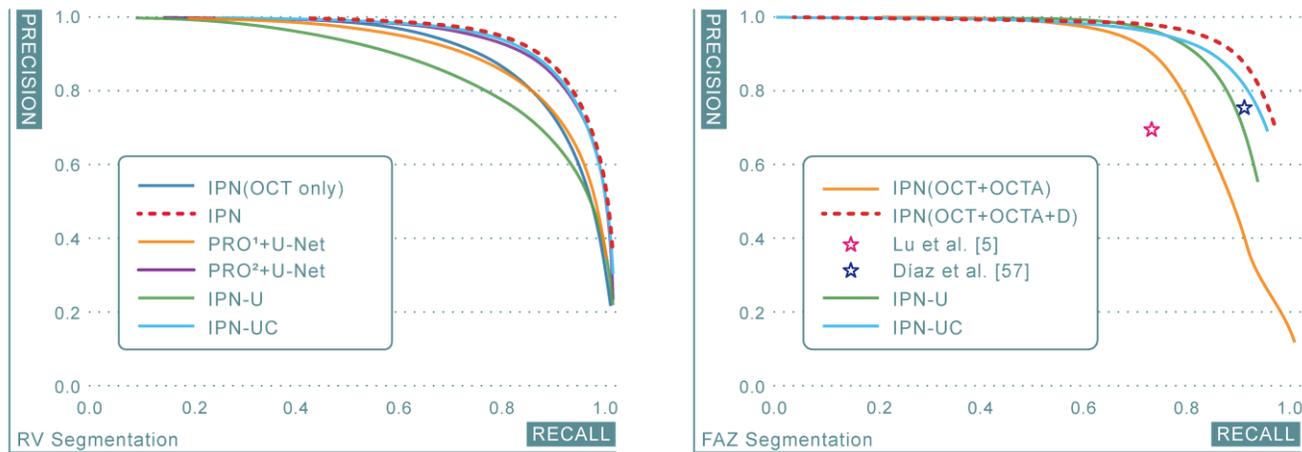


Fig. 11. Precision-recall curves of different methods on the RV segmentation (left) and FAZ segmentation (right).

the proposed IPN achieves state-of-the-art performance in the two segmentation tasks. (4) In Sec. IV-E4), we further explore the network structure. We replace the unidirectional pooling layer with the standard max-pooling layer and add decoding path in the network, and discuss the impact of these changes. Besides, we give the precision-recall curves of segmentation results for the different networks, as shown in Fig. 11. Table III lists the quantitative results in all comparative experiments.

E. Comparative Experiments

1) 3D-to-2D IPN vs. 2D-to-1D IPN

The quantitative results in Table III-1 show that 3D-to-2D IPN performs better than 2D-to-1D IPN. Fig. 12 shows an example of the FAZ segmentation results, which indicates that 3D-to-2D IPN has better spatial continuity than 2D-to-1D IPN because the segmentation result of 2D-to-1D IPN has jagged edges.

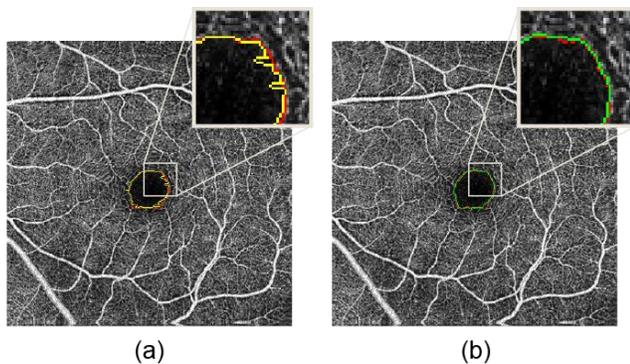


Fig. 12. FAZ segmentation results using 2D-to-1D IPN (a) and 3D-to-2D IPN (b). The red line represents the ground truth; the yellow and green lines represent the results of 2D-to-1D IPN and 3D-to-2D IPN, respectively.

2) Multi-channel vs. Single-channel

We give different inputs in our IPN for each segmentation problem. Individually, we input OCT, OCTA and the distance map for FAZ segmentation, while input OCT and OCTA for RV segmentation. Our purpose is to make full use of multimodality information and sufficient prior knowledge to improve network performance. In this experiment, we compared the effects of different inputs. The quantitative results of different inputs are shown in Table III-2. It shows that the

segmentation performance of the multi-channel input is better than that of the single-channel input. Fig. 13 and Fig. 14 show the results of different inputs on RV segmentation and FAZ segmentation, respectively.

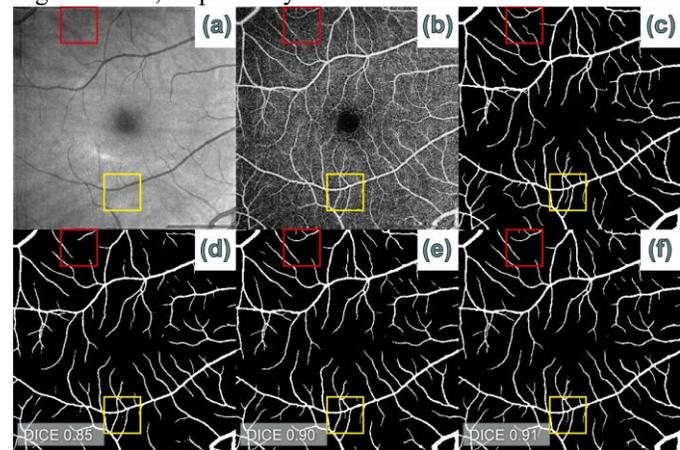


Fig. 13. An example of RV segmentation results. (a) OCT projection map. (b) OCTA projection map between ILM and OPL. (c) The ground truth. (d) The result with the input of OCT only. (e) The result with the input of OCTA only. (f) The result with the input of OCT and OCTA.

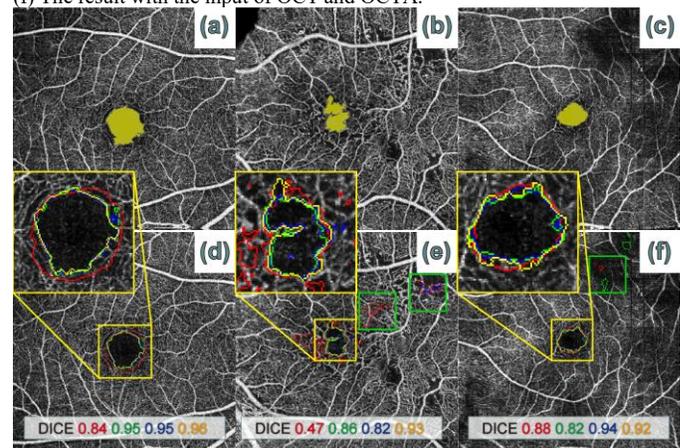


Fig. 14. Three examples of FAZ segmentation results. Yellow areas in (a)-(c) represent the ground truth. Colored lines in (d-f) represent the results of IPN with different input: OCT (red line), OCTA (green line), OCT+OCTA (blue line), OCT+OCTA+Distance map (yellow line).

In Fig. 13, we can see that the result of two-channel input is similar to that of the only OCTA input and much better than

that of the only OCT input, which indicates that OCTA plays a better role than OCT in RV segmentation. Besides, we can find that even if only input OCT, our IPN also identifies some inconspicuous vessels, which are not evident in the OCT projection image, as shown in the red box and the yellow box in Fig. 13. It demonstrates that making full use of volumetric information rather than relying on the information-limited projection maps is one of the advantages of our network.

In FAZ segmentation, it can be seen from Fig. 14(d) that only the OCT images cannot obtain the accurate contour of FAZ, because OCT cannot clearly display the complete microvascular system. In addition, from Figs. 14 (e)(f), it can be seen that the distance map plays an important role, which is related to the location specificity of FAZ. Without the distance map, the network mistakenly assumes that the areas with weak blood flow signals belong to FAZ, such as the weakening of local signals due to turbid refractive media and non-perfusion zone, as shown in the green box in Fig. 14(e)(f).

3) IPN vs. others

We compared the proposed IPN with other methods. First, we use two popular 2D segmentation frameworks (FCN and U-Net) for RV segmentation in two types of projection maps. Then, two reported methods [5], [57] are used to compare our IPN for FAZ segmentation. The quantitative results are shown in Table III-3.

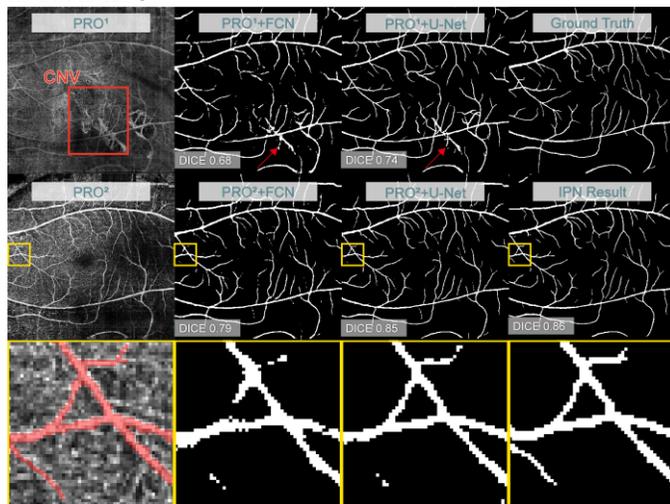


Fig. 15. An example of RV segmentation results in the case of CNV. PRO¹ is OCTA full-projection map. PRO² is OCTA maximum-projection map between ILM and OPL.

In RV segmentation, we conduct experiments in two types of projection maps. One is the full-projection map (Fig.15, PRO¹), which directly averages the three-dimensional data along the projection direction. Another is the maximum-projection map of the inner retina (Fig.15, PRO²), that is, to get the maximum along the projection direction from the ILM layer to OPL. This projection method eliminates the blood flow signal of choroid and removes the artifacts, which can display the retinal vessels more clearly than the full-projection map. To generate the maximum-projection maps, we used a public layer segmentation software (OCTExplorer 3.8) to segment ILM layer and OPL. It takes about 5 minutes for each eye.

We successively train FCN and U-Net using the above projection maps, and the performance on the test set is shown

in Table III-3. It indicates that the performance of the maximum-projection map is better than that of the full-projection map, the performance of U-Net is better than that of FCN, and the performance of IPN is the best. Fig. 15 shows an example of one eye with CNV. Because the full-projection map contains the blood flow signal of the choroid, a part of CNV is segmented as the blood vessel of the inner retina. The maximum-projection map does not include these signals and thus does not misclassify the choroidal blood flow information. IPN also has no such misclassification because it implicitly learns the location information of the retinal layers. Besides, the use of volumetric information can help IPN better understand the shape of retinal vessels, thereby better maintaining the integrity of the retinal vessel topology. As shown in the yellow box in Fig. 15, the methods based on projection maps lack tolerance to background noise, manifested as loss of a portion of the blood vessel, or the blood vessel is discontinuous. In contrast, the vessel structure in IPN segmentation results is more complete and its edges are smooth.

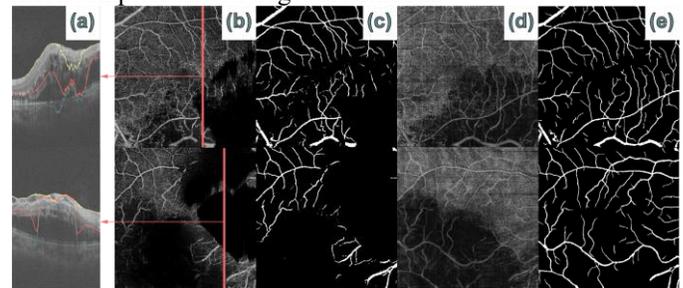


Fig. 16. Two examples of RV segmentation with incorrect layer segmentation. (a) B-scan. (b) OCTA maximum-projection map between ILM layer and OPL. (c) The U-Net result using the maximum-projection map. (d) OCTA full-projection map. (e) IPN result.

Another limitation for RV segmentation depending on projection maps is that the generation of the projection maps depends on the retinal layer segmentation. When the layer segmentation algorithm is no longer effective due to retinopathy, the RV segmentation results using projection maps will become terrible, as shown in Fig. 16(c). IPN does not need layer segmentation and is more robust in these cases (Fig. 16(e)).

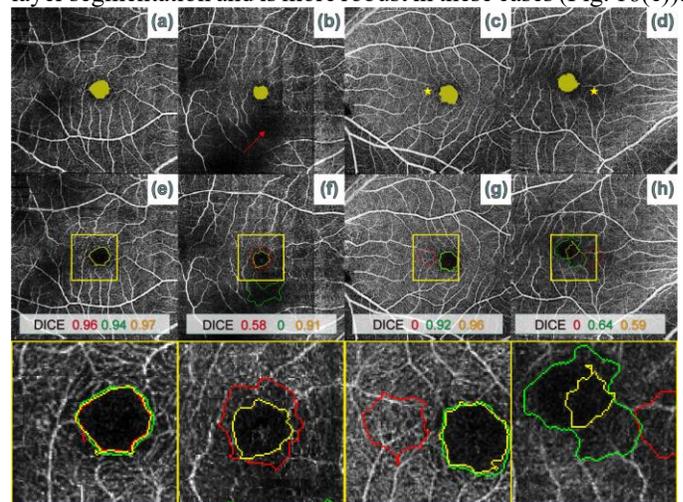


Fig. 17. Four examples of FAZ segmentation using the proposed IPN and other methods. Yellow areas in (a)-(d) represent the ground truth. Colored lines in (e)-(h) represent the results of different methods: Lu et al. [5] (red line), Díaz et al. [57] (green line), IPN (yellow line). The red arrow in (b) indicates that the

blood flow signal is weakened due to the turbid medium. The yellow star in (c)-(d) represents the center of the image.

For FAZ segmentation, we compared our network with the two reported algorithms, [5] and [57]. The first method [5] is based on an active shape model, generalized gradient vector flow (GGVF). It works well in the projection image with clear capillary plexus, as shown in Fig.17(e). However, it is difficult to achieve high segmentation accuracy in the whole dataset. On the one hand, when the foveal center deviates from the central position of the image, the seed area will be misjudged, as shown in Figs. 17(c)(d). On the other hand, the weak signal region will make the initial contour of the active contour model larger, which will result in the wrong segmentation, as shown in Fig. 17(f).

The other method [57] is based on edge detection and morphological processing. Its advantage is that it can be applied to several types of projection maps. However, it takes the area factor as the basis for judging FAZ and does not consider the overall distribution of blood vessels, which reduces the distinction between small FAZ area and the weak blood flow signal area caused by turbid refractive media, as shown in Fig. 17(f).

Compared with these methods, IPN utilizes the intensity distribution information and the location information of three-dimensional data, and achieves better performance in most cases. The introduction of the distance map helps IPN to eliminate some misconceptions, but also bring some limitations. For instance, when the foveal center deviates from the center of the image, the distance map will also produce wrong guidance. The case with the lowest Dice scores in FAZ segmentation, as shown in Fig. 17(h), is caused by the severe deviation of the foveal center and the wrong guidance of the distance map.

4) More Exploration

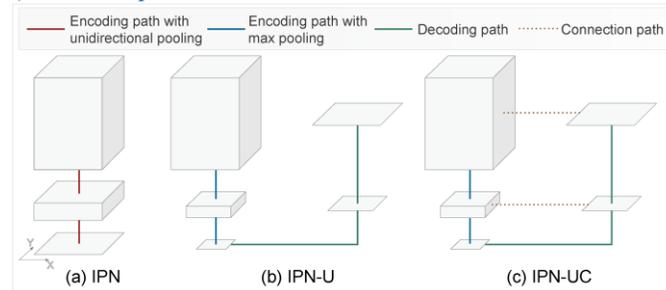


Fig. 18. Three network architectures for 3D to 2D segmentation. Their encoding paths use the same number and the same scale of the convolutional layers. The pooling layer uses the unidirectional pooling in IPN (a) and the standard max-pooling in IPN-U (b) and IPN-UC (c). Their decoding paths use the same number and the same scale of the deconvolutional layers with the convolutional layers.

Up to now, our IPN has achieved a better performance than the baseline methods in RV segmentation and FAZ segmentation. This is mainly due to the idea of the 3D-to-2D learning mode that can make full use of the OCTA volume data. We introduced the unidirectional pooling in IPN, which can easily handle scale transformation and supports the high-resolution representation in the XY direction. In this section, we discuss other implementation for 3D-to-2D segmentation without using unidirectional pooling.

First, we replace the unidirectional pooling by a standard max-pooling in the original IPN, so the final output of the 2D

plane is smaller in the XY direction than the original projection map. We add the decoding path, like U-Net’s upsampling, to restore the size of the 2D plane. We call this network as IPN-U (Fig. 18(b)). The performance of IPN-U is worse than that of IPN, as shown in Table III-4. The yellow arrows in Fig. 19(g) and Fig. 20(g) show that the segmentation results of IPN-U are misaligned during splicing, which implies that it is difficult to recover the whole information in the XY direction using only the decoding path.

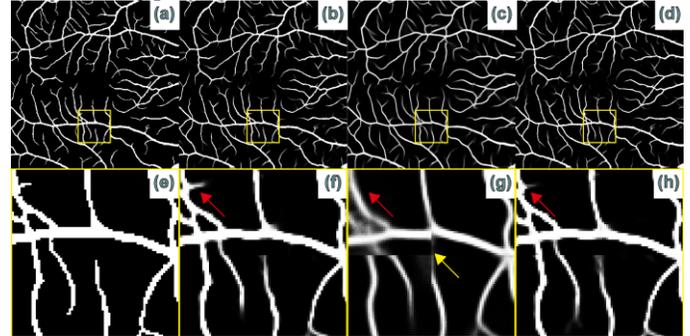


Fig. 19. An example of RV segmentation. (a) Ground truth. (b)-(d) Probability result of IPN, IPN-U and IPN-UC, respectively.

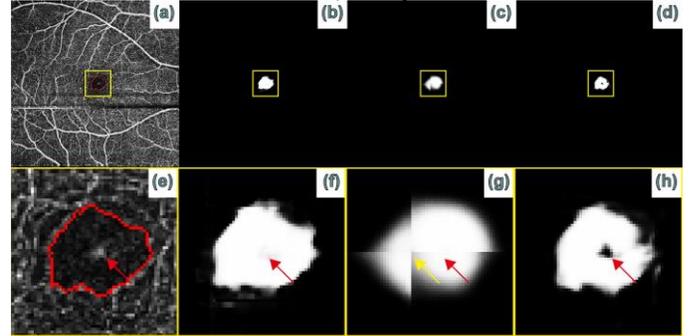


Fig. 20. An example of FAZ segmentation. (a) OCTA maximum-projection map and the ground truth (red line). (b)-(d) Probability result of IPN, IPN-U and IPN-UC, respectively.

Then, to help the network restore the information in the XY direction, we add the connection structure between the encoding path and the decoding path in IPN-U. The connection structure we used is the convolution with the size of $h \times 1 \times 1$, where h is the height of the convolution block. We call this network as IPN-UC (Fig. 18(c)). The results in Table III-4 indicate that the performance of IPN-UC has not surpassed that of IPN. The example of RV segmentation in Fig. 19(h) shows that the connection structure can recover most of the lost information in the XY direction, but some small structures are still missing. Another example (Fig. 20) shows that due to the small receptive field of the connected structure, IPN-UC is susceptible to local noise, so a hole appears in the segmentation result, marked by the red arrow in Fig. 20(h).

Compared with IPN-UC, IPN uses the unidirectional pooling to maintain the high-resolution representation in the XY direction for better performance. Table III-4 and Fig. 11 both illustrate that IPN has the best performance in the network structure we discussed. Therefore, we use IPN as our proposed 3D-to-2D segmentation framework in this paper.

F. Limitations

In RV segmentation, we use IPN to segment the large vessels

in the inner retina. However, we have not implemented the segmentation of capillary plexus and the quantification of vessel density, which are also important for the analysis of the retinal vessel. The reason is that it is difficult for the pixel-level annotation of capillary plexus in our dataset, especially in the disease condition and the low image quality, where the capillary plexus often presents an incomplete topology. After the ground truth is clear, we will have more discussion on the segmentation of capillary plexus and the quantification of vessel density.

In FAZ segmentation, we introduced the distance map as one input of IPN to weaken the negative effects of the non-perfused area and the weak signals away from the image center. However, when the fovea severely occasionally deviates from the center of the image, the distance map plays the opposite role, as shown in Fig. 17(d). We believe that it is related to the distribution of FAZ locations in the training set. The inclusion of a broader set of examples in the training set can help to improve this situation.

Our network needs to occupy a large amount of computing resources in the training stage. Because of the limited computing resources, we have to divide the OCTA volume into small blocks, which leads to the need for splicing the predicted results in order to achieve complete results. Fortunately, the testing speed is acceptable and faster than the speed of layer segmentation algorithms. With the increase of input size and the optimization of network structure, the segmentation accuracy and the time efficiency of our network are expected to be further improved.

V. CONCLUSION

We propose a novel 3D to 2D segmentation network, IPN, which can be applied to RV segmentation and FAZ segmentation in OCTA images. The key insight of our network is the construction of the projection learning module, in which we introduced a unidirectional pooling to summarize the effective information of 3D data into 2D projection plane. Our experimental results indicate that IPN is an effective implementation of 3D to 2D segmentation networks. It does not need retinal layer segmentation or projection maps in the segmentation process. The uses of multi-modality information and volumetric information make the proposed method more accurate and robust than the baseline methods.

In the future, we will try to explore the IPN structure with better performance to solve other quantification problems in OCTA images, such as the segmentation of capillary plexus and the segmentation of lesions. Moreover, our framework also has the potential for other application in different fields that needs more practice. Finally, our code is publicly available at https://github.com/chaosallen/IPN_tensorflow.

REFERENCES

[1] S. A. Agemy, N. K. Sripsema, C. M. Shah, R. C. Gentile, Y. Hsiao, and Q. Zhou, "Retinal Vascular Perfusion Density Mapping Using Optical Coherence Tomography Angiography in Normal and Diabetic Retinopathy Patients," *Retina*, vol. 35, no. 11, pp. 2353–2363, 2015.

[2] Y. Jia, S. T. Bailey, D. J. Wilson, O. Tan, M. L. Klein, C. J. Flaxel, et al., "Quantitative Optical Coherence Tomography Angiography of

Choroidal Neovascularization in Age-Related Macular Degeneration," *Ophthalmology*, vol. 121, no. 7, pp. 1435–1444, 2014.

[3] P. K. Yu, Z. Mammo, C. Balaratnasingam, and D. Yu, "Quantitative Study of the Macular Microvasculature in Human Donor Eyes," *Invest Ophthalmol Vis Sci*, vol. 59, no. 1, pp. 108–116, 2018.

[4] T. D. Nafia and A. Handayani, "Quantification of Retinal Vascular Tortuosity: Evaluation on Different Numbers of Sampling Points," *2018 2nd Int. Conf. Biomed. Eng.*, pp. 39–43, 2018.

[5] Y. Lu, J. M. Simonett, J. Wang, M. Zhang, T. Hwang, M. Ahmed, et al., "Evaluation of Automatically Quantified Foveal Avascular Zone Metrics for Diagnosis of Diabetic Retinopathy Using Optical Coherence Tomography Angiography," *Investig. Ophthalmology Vis. Sci.*, vol. 59, no. 6, pp. 2212–, 2018.

[6] P. Massin, A. Erginay, B. Haouchine, A. B. Mehidi, M. Paques, and A. Gaudric, "Retinal thickness in healthy and diabetic subjects measured using optical coherence tomography mapping software," *Eur. J. Ophthalmol.*, vol. 12, no. 2, pp. 102–108, 2002.

[7] C. Lavia, S. Bonnin, M. Maule, A. Erginay, R. Tadayoni, and A. Gaudric, "Vessel Density of Superficial, Intermediate, and Deep Capillary Plexuses Using Optical Coherence Tomography Angiography," *Retina*, vol. 39, no. 2, p. 1, 2018.

[8] M. Adhi and J. S. Duker, "Optical coherence tomography-current and future applications," *Curr. Opin. Ophthalmol.*, vol. 24, no. 3, pp. 213–221, 2013.

[9] W. Geitzenauer, C. K. Hitzenberger, and U. M. Schmidt-Erfurth, "Retinal optical coherence tomography: Past, present and future perspectives," *Br. J. Ophthalmol.*, vol. 95, no. 2, pp. 171–177, 2011.

[10] R. A. LEITGEB, "En face optical coherence tomography: a technology review [Invited]," *Biomed. Opt. Express*, vol. 10, no. 5, pp. 2177–2201, 2019.

[11] R. F. Spaide, J. G. Fujimoto, N. K. Waheed, S. R. Sadda, and G. Staurenghi, "Optical coherence tomography angiography," *Prog. Retin. Eye Res.*, vol. 64, pp. 1–55, 2018.

[12] Y. Jia, O. Tan, J. Tokayer, B. Potsaid, Y. Wang, J. Jonathan, et al., "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography," *Opt. Express*, vol. 20, no. 4, p. 4710, 2012.

[13] S. J. Chiu, M. J. Allingham, P. S. Mettu, S. W. Cousins, J. A. Izatt, and S. Farsiu, "Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema," *Biomed. Opt. Express*, vol. 6, no. 4, p. 1172, 2015.

[14] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," *Opt. Express*, vol. 18, no. 18, p. 19413, 2010.

[15] L. Fang, D. Cunefare, C. Wang, H. G. Robyn, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express*, vol. 8, no. 5, pp. 2732–2744, 2017.

[16] K. Gao, W. Kong, S. Niu, D. Li, and Y. Chen, "Automatic retinal layer segmentation in SD-OCT images with CSC guided by spatial characteristics," *Multimed. Tools Appl.*, pp. 1–12, 2019.

[17] A. Lang, A. Carass, M. Hauser, E. S. Sotirchos, P. A. Calabresi, H. S. Ying, et al., "Retinal layer segmentation of macular OCT images using boundary classification," *Biomed. Opt. Express*, vol. 4, no. 7, p. 1133, 2013.

[18] A. Yazdanpanah, G. Hamarneh, B. R. Smith, and M. V Sarunic, "Segmentation of Intra-Retinal Layers from Optical Coherence Tomography Images using an Active Contour Approach," *IEEE Trans. Med. Imaging*, vol. 30, no. 2, pp. 484–496, 2010.

[19] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Comput. Sci.*, 2014.

[20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," 2016.

[21] C. Szegedy, V. Vanhoucke, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2014.

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2014.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Int. Conf. Med. Image Comput. Comput. Interv.*, 2015.

- [24] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," *MICCAI*, 2016.
- [25] T. E. De Carlo, A. Romano, N. K. Waheed, and J. S. Duker, "A review of optical coherence tomography angiography (OCTA)," *Int. J. Retin. Vitreol.*, pp. 1–5, 2015.
- [26] B. Sheng, P. Li, S. Mo, H. Li, X. Hou, Q. Wu, *et al.*, "Retinal Vessel Segmentation Using Minimum Spanning Superpixel Tree Detector," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2707–2719, 2019.
- [27] M. D. Abramoff, S. Member, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 169–208, 2010.
- [28] C. Heneghan, J. Flynn, M. O. Keefe, and M. Cahill, "Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis," *Med. Image Anal.*, vol. 6, no. 4, pp. 407–429, 2002.
- [29] J. S. A. Romo, R. E. Linderman, A. Pinhas, J. Carroll, R. B. Rosen, and T. Y. P. Chui, "Novel Development of Parafoveal Capillary Density Deviation Mapping using an Age-Group and Eccentricity Matched Normative OCT Angiography Database," *Transl. Vis. Sci. Technol.*, vol. 8, no. 3, 2019.
- [30] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and R. L. Kennedy, "Measurement of Retinal Vessel Widths From Fundus Images Based on 2-D Modeling," *IEEE Trans. Med. Imaging*, vol. 23, no. 10, pp. 1196–1204, 2004.
- [31] Q. Mirsharif, F. Tajeripour, and H. Pourreza, "Automated characterization of blood vessels as arteries and veins in retinal images," *Comput. Med. Imaging Graph.*, vol. 37, no. 7–8, pp. 607–617, 2013.
- [32] H. M. Pakter, S. C. Fuchs, M. K. Maestri, L. B. Moreira, L. M. D. Ricardi, V. F. Pamplona, *et al.*, "Computer-Assisted Methods to Evaluate Retinal Vascular Caliber: What Are They Measuring?," *Invest Ophthalmol Vis Sci*, vol. 52, no. 2, pp. 810–815, 2011.
- [33] Q. Jin, Q. Chen, Z. Meng, B. Wang, and R. Su, "Construction of Retinal Vessel Segmentation Models Based on Convolutional Neural Network," *Neural Process. Lett.*, 2019.
- [34] Q. P. Lau, M. L. Lee, W. Hsu, and T. Y. Wong, "Simultaneously Identifying All True Vessels From Segmented Retinal Images," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1851–1858, 2013.
- [35] J. Odstreilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, *et al.*, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Process.*, vol. 7, no. 4, pp. 373–383, 2013.
- [36] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Iterative Vessel Segmentation of Fundus Images," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1738–1749, 2015.
- [37] B. Al-diri, A. Hunter, and D. Steel, "An Active Contour Model for Segmenting and Measuring Retinal Vessels," *IEEE Trans. Med. Imaging*, vol. 28, no. 9, pp. 1488–1497, 2009.
- [38] C. Ding, Y. Xia, and Y. Li, "Supervised Segmentation of Vasculature in Retinal Images Using Neural Networks," *IEEE Int. Conf. Orange Technol.*, pp. 49–52, 2014.
- [39] M. Niemeijer, M. K. Garvin, B. Van Ginneken, M. Sonka, and M. D. Abr, "Vessel Segmentation in 3D Spectral OCT Scans of the Retina," *Proc. SPIE - Int. Soc. Opt. Eng.*, vol. 6914, pp. 1–8, 2008.
- [40] R. Kafieh, S. Member, H. Rabbani, and F. Hajizadeh, "An Accurate Multimodal 3-D Vessel Segmentation Method Based on Brightness Variations on OCT Layers and Curvelet Domain Fundus Image Analysis," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2815–2823, 2013.
- [41] Z. Hu, M. Niemeijer, M. D. Abramoff, K. Lee, and M. K. Garvin, "Automated Segmentation of 3-D Spectral OCT Retinal Blood Vessels by Neural Canal Opening False Positive Suppression," *Int. Conf. Med. Image Comput. Comput. Interv.*, pp. 33–40, 2010.
- [42] Z. Hu, M. Niemeijer, M. D. Abramoff, and M. K. Garvin, "Multimodal Retinal Vessel Segmentation From Spectral-Domain Optical Coherence Tomography and Fundus Photography," *IEEE Trans. Med. Imaging*, vol. 31, no. 10, pp. 1900–1911, 2012.
- [43] S. Youse, T. Liu, and R. K. Wang, "Segmentation and quantification of blood vessels for OCT-based micro-angiograms using hybrid shape/intensity compounding," *Microvasc. Res.*, vol. 97, pp. 37–46, 2015.
- [44] H. Lee, M. Lee, H. Chung, and H. C. Kim, "Quantification of Retinal Vessel Tortuosity in Diabetic Retinopathy using Optical Coherence Tomography Angiography," *Retina*, p. 1, 2017.
- [45] N. Eladawi, M. Elmogy, O. Helmy, A. Aboelfetouh, A. Riad, H. Sandhu, *et al.*, "Automatic blood vessels segmentation based on different retinal maps from OCTA scans," *Comput. Biol. Med.*, vol. 89, no. August, pp. 150–161, 2017.
- [46] S. Brook, "Automated segmentation and quantification of OCT angiography for tracking angiogenesis progression," *Biomed. Opt. Express*, vol. 8, no. 12, pp. 5604–5616, 2017.
- [47] M. Eladawy, M. S. Sherif, M. Elbably, and N. M. A. Salem, "Automatic Detection and Measurement of Foveal Avascular Zone," *Radio Sci. Conf.*, 2003.
- [48] S. H. M. Alipour, H. Rabbani, and M. Akhlaghi, "A new combined method based on curvelet transform and morphological operators for automatic detection of foveal avascular zone," *Signal, Image Video Process.*, vol. 8, no. 2, pp. 205–222, 2013.
- [49] Y. Zheng, J. S. Gandhi, A. N. Stangos, C. Campa, D. M. Broadbent, and S. P. Harding, "Automated Segmentation of Foveal Avascular Zone in," *Investig. Ophthalmology Vis. Sci.*, vol. 51, no. 7, pp. 1–7, 2010.
- [50] M. Al-sheikh, H. Akil, M. Pfau, and S. R. Satta, "Swept-Source OCT Angiography Imaging of the Foveal Avascular Zone and Macular Capillary Network Density in Diabetic Retinopathy," *Investig. Ophthalmology Vis. Sci.*, vol. 57, no. 8, pp. 3907–, 2016.
- [51] R. Mastropasqua, L. Toto, P. A. Mattei, M. Di Nicola, I. A. L. Zecca, P. Carpineto, *et al.*, "Reproducibility and repeatability of foveal avascular zone area measurements using swept-source optical coherence tomography angiography in healthy subjects Rodolfo," *Eur J Ophthalmol*, vol. 27, no. 3, pp. 336–341, 2016.
- [52] I. Y. Febo, O. B. Ocak, B. S. Yilmaz, A. Inal, B. Gokyigit, and M. Taskapili, "Comparison of quantitative measurement of foveal avascular zone and macular vessel density in eyes of children with amblyopia and healthy controls: an optical coherence tomography angiography study," *J. AAPOS*, vol. 21, no. 3, pp. 224–228, 2017.
- [53] J. J. Park, B. T. Soetikno, and A. A. Fawzi, "Characterization of the Middle Capillary Plexus Using Optical Coherence Tomography Angiography in Healthy and Diabetic Eyes," *Retina*, vol. 36, no. 11, pp. 2039–2050, 2017.
- [54] L. Giselle, J. S. A. Romo, R. E. Linderman, B. D. Krawitz, S. Mo, A. Zaki, *et al.*, "Within-subject assessment of foveal avascular zone enlargement in different stages of diabetic retinopathy using en face OCT reflectance and OCT angiography," *Biomed. Opt. Express*, vol. 9, no. 12, pp. 24–28, 2018.
- [55] K. G. Falavarjani, H. Shenazandi, D. Naseri, P. Anvari, K. Pegah, A. Farzaneh, *et al.*, "Original Article Foveal Avascular Zone and Vessel Density in Healthy Subjects: An Optical Coherence Tomography Angiography Study," *J Ophthalmic Vis Res.*, vol. 13, no. 3, pp. 260–265, 2018.
- [56] W. Ang, D. A. H. Uang, T. H. S. H. Wang, and Y. A. L. I. J. Ia, "Three-dimensional structural and angiographic evaluation of foveal ischemia in diabetic retinopathy: method and validation," *Biomed. Opt. Express*, vol. 10, no. 7, pp. 3522–3532, 2019.
- [57] M. Díaz, J. Novo, P. Cutrin, F. Gómez-Ulla, M. G. Penedo, and M. Ortega, "Automatic segmentation of the Foveal Avascular Zone in ophthalmological OCT-A images," *PLoS One*, 2018.
- [58] Y. Guo, A. Camino, J. Wang, D. Huang, T. S. Hwang, and Y. Jia, "MEDnet, a neural network for automated detection of avascular area in OCT angiography," *Biomed. Opt. Express*, vol. 9, no. 11, pp. 5147–5158, 2018.
- [59] L. Chen, G. Papandreou, S. Member, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, 2016.
- [60] Z. Ji, Q. Chen, S. Niu, T. Leng, and D. L. Rubin, "Beyond Retinal Layers: A Deep Voting Model for Automated Geographic Atrophy Segmentation in SD-OCT Images," *Transl. Vis. Sci. Technol.*, vol. 7, no. 1, p. 1, 2018.