Dense Dilated Network with Probability Regularized Walk for Vessel Detection

Lei Mou, Li Chen, Jun Cheng, Zaiwang Gu, Yitian Zhao and Jiang Liu

Abstract—The detection of retinal vessel is of great importance in the diagnosis and treatment of many ocular diseases. Many methods have been proposed for vessel detection. However, most of the algorithms neglect the connectivity of the vessels, which plays an important role in the diagnosis. In this paper, we propose a novel method for retinal vessel detection. The proposed method includes a dense dilated network to get an initial detection of the vessels and a probability regularized walk algorithm to address the fracture issue in the initial detection. The dense dilated network integrates newly proposed dense dilated feature extraction blocks into an encoder-decoder structure to extract and accumulate features at different scales. A multi-scale Dice loss function is adopted to train the network. To improve the connectivity of the segmented vessels, we also introduce a probability regularized walk algorithm to connect the broken vessels. The proposed method has been applied on three public data sets: DRIVE, STARE and CHASE_DB1. The results show that the proposed method outperforms the state-of-the-art methods in accuracy, sensitivity, specificity and also area under receiver operating characteristic curve.

Index Terms—Vessel segmentation, encoder-decoder, deep learning, regularized walk, vessel reconnection

I. INTRODUCTION

Retinal vessels are commonly analyzed in the diagnosis and treatment of various ocular diseases such as age-related

L. Mou is with School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430081, China, Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan 430081, China and also Cixi Institute of Biomedical Engineering, Chinese Academy of Sciences, Zhejiang 315201, China. E-mail: moulei@nimte.ac.cn.

L. Chen is with School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430081, China and Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan 430081, China. E-mail: chenli@wust.edu.cn.

J. Cheng is with Cixi Institute of Biomedical Engineering, Chinese Academy of Sciences, Zhejiang 315201, China and also with UBTech Research, UBTech Robotics Corp Ltd, Guangdong 518055, China. Email: sam.j.cheng@gmail.com.

Z. Gu is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Guangdong 518055, China Email: guzaiwang@gmail.com.

Y. Zhao is with the Cixi Institute of Biomedical Engineering, Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, Zhejiang 315201, China. E-mail: yitian.zhao@nimte.ac.cn.

J. Liu is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Guangdong 518055, China and also with the Cixi Institute of Biomedical Engineering, Chinese Academy of Sciences, Zhejiang 315201, China. Email: liuj@sustech.edu.cn.

macular degeneration, diabetic retinopathy, glaucoma, hypertension, arteriosclerosis and multiple sclerosis [1] [2] [3]. Clinicians have found that the structure of retinal blood vessels is associated with diabetes [4] [5] which affects the lining of the blood vessels in eyes. Poor circulation in the retinal vessels often compounds these problems by causing the generation of fragile new vessels. Further, vessel segmentation is also the basis for subsequent retinal artery/venous vessel classification task, which is of direct assistance to the ophthalmologist in terms of diagnosis and treatment of eye disease [6], [7]. Therefore, it is important for the clinicians to segment and analyze the retinal vessels, especially the tiny ones. However, manual demarcation of retinal vessels is a costly and time-consuming task. In recent years, extensive work has been done to segment the vessel automatically. Li *et al.* used the Hessian matrix and random walks to segment retinal vessel [8]. Nguyen *et al.* utilized a multi-scale line detection scheme for vessel segmentation [9]. Bankhead *et al.* developed a fast retinal vessel detection method by using wavelets [10] and edge location refinement [11]. Soares *et al.* used supervised classification to apply wavelets to 2-D retinal blood vessel segmentation [12]. Läthén *et al.* proposed a quadrature filtering method with multi-scale analysis for retinal vessel segmentation [13]. Zhao *et al.* [14] further proposed a vessel segmentation model based on saliency guided infinite perimeter active contour. A more detailed review of traditional vessel detection methods can be found in [15] [16]. A common limitation of these traditional vessel detection algorithms is that they often require some manually determined parameters tuned based on local data set, which makes the algorithms less scalable.

With the rapid development of deep learning, many algorithms based on deep learning have been proposed for retinal blood vessel detection. Fu *et al.* proposed a loss function based on multi-scale output combined with conditional random field (CRF) to improve the detection of vessels in retinal images with lesions [17]. Liskowski *et al.* utilized global contrast normalization and zero-phase component analyzing techniques to enhance training samples for better blood vessel detection [18]. Recently, Gu *et al.* proposed context encoder which utilizes dense atrous convolution and residual multi-kernel pooling for biomedical image segmentation [19]. Zhang *et al.* [20] proposed ET-Net which focuses on the guidance of edge information to segment optic discs, retinal vessels and lung organs. An efficient AG-Net [21] has also been proposed to take attention mechanism into CNN to detect vessels in fundus. Compared with the traditional methods, deep learning-based algorithms have better scalability and achieve better accuracy. The emergence of the encoder-decoder structure enables the
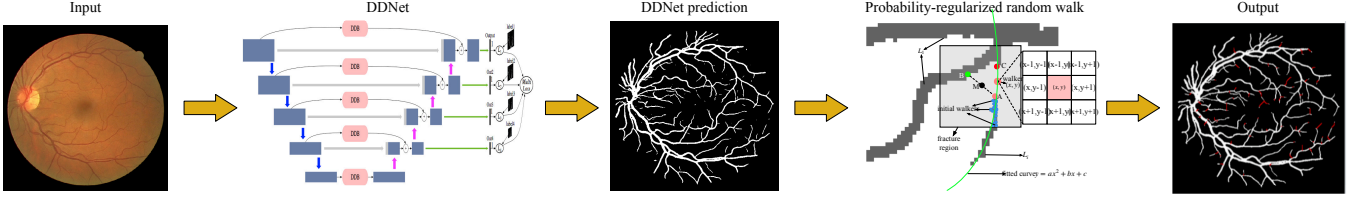
Fig. 1. The overall flow chart of the proposed method. It includes a dense dilated convolutional network (DDNet) and a regularized walk algorithm. The DDNet is based on encoder-decoder structure, which is integrated with newly proposed dense dilated blocks (DDB) between the encoder and decoder to compute an initial map of the vessels. Then a probability regularized walk is used to improve the results from the initial map.

network to extract deeper features while extracting the context information of the vessels [22] [23], thereby obtaining more accurate segmentation. In order to achieve higher segmentation accuracy, recent semantic segmentation algorithms [24], [25], [26] adopted the state-of-the-art classification structures such as ResNet [27], ResNeXt [28] and DenseNet [29]. In addition, multi-scale fusion strategy has been used to train the network for feature extraction at different resolutions [30] [31] [32]. In convolutional networks, pooling operations are used to obtain a larger receptive field, which leads to some loss of low-level information. To avoid this issue, Yu and Koltun [33] designed a scheme called dilated (or atrous) convolution which is able to increase the receptive field without pooling. Prior works show that atrous convolution can be combined with batch normalization to achieve more accurate results [34] [35]. Recently, generative adversarial network [36] has also been proposed to improve vessel detetion.

Although these deep learning-based methods show better performance in vessel detection [17] [12] [18], they always fail to detect small or tiny vessels, which lead to the fracture of the vessels. As a result, these methods often have poor performance on the detection of tiny blood vessels. Clinically, tiny vessels provide an indispensable reference for the diagnosis of diseases like neovascular diseases. The retinal vessels provide a unique opportunity for cerebral small vessel disease study. For example, subjects with smaller retinal arteriolar-to-venular ratio tended to have more white matter lesions [37]. Vessel connectivity has a major impact on the screening of vascular lesions. Therefore, improving the detection of tiny vessels in fundus images and effectively connecting fractured segmentations are crucial for ocular disease detection.

In this paper, we propose a dense dilated network (DDNet) and a probability regularized walk (PRW) algorithm for vessel detection. Fig. 1 shows the overall architecture of the proposed method, which includes the DDNet to obtain an initial map of retinal vessels and the PRW algorithm to connect the broken vessels in the initial map. In the vessel reconnection, we propose to integrate the probability output of the deep learning and the local vessel directions into the regularized walk algorithm for fractured vessel reconnection. Our main contributions are as follows:

1) We propose a dense dilate network for retinal vessel detection.
2) We propose a probability regularized walk algorithm to connect broken vessels from the initial vessel detection. It integrates the output of the dense dilated network and the local vessel directions into the regularized walk.
3) Experimental results show that the proposed method outperforms other methods for retinal vessel detection.

The rest of the paper is organized as follows. Section II reviews the related work. Section III introduces our method including DDNet and PRW. In Section V, we introduce the experimental results. The last section concludes the paper.

II. RELATED WORKS

A. Encoder-Decoder

The encoder-decoder structure is one of the most popular frameworks for semantic segmentation. A notable characteristic of the encoder-decoder framework is that it is an end-to-end learning algorithm. Many segmentation methods are based on the encoder-decoder structure, such as SegNet [23] and U-Net [22]. The encoder extracts the shallow and deep features of the images, which can be implemented in a variety of ways. Many different modules such as VGG16 [38], GoogLeNet [39] and ResNet [27] can be used in the encoder. In order to preserve the spatial information, Yu et al. [40] proposed a dilated residual network to extract features in the encoder. Zhao et al. [30] proposed the image cascade network (ICNet) by utilizing the advantages of image pyramid in PSPNet [41] to achieve reasonable performance in computation and segmentation of high-resolution images. In a recent work, Zhou et al. [42] proposed a hybrid block of dilated convolutions with different rates between the encoder and the decoder to extract better features. Pohlen et al. proposed the full-resolution residual network [43] model to enhance the final classification performance by combining different stages of information. The internal structure of the decoder is also important to improve the segmentation performance of the model. Recent work shows that unpooling and deconvolution operations in the decoder are beneficial for certain segmentation tasks [44].

B. Vessel Detection

Retinal vessels are widely used in the diagnosis of various ocular diseases. In recent years, a variety of methods have been proposed to detect retinal vessels from fundus images, including unsupervised approaches [45] and supervised approaches [46]. The earlier traditional method [47] used support vector machines. Zhao et al. [15] [16] proposed infinite perimeter active contour model with hybrid region information and a weighted symmetry filter to detect vessels. Some other methods used a variety of filters to enhance the blood vessels
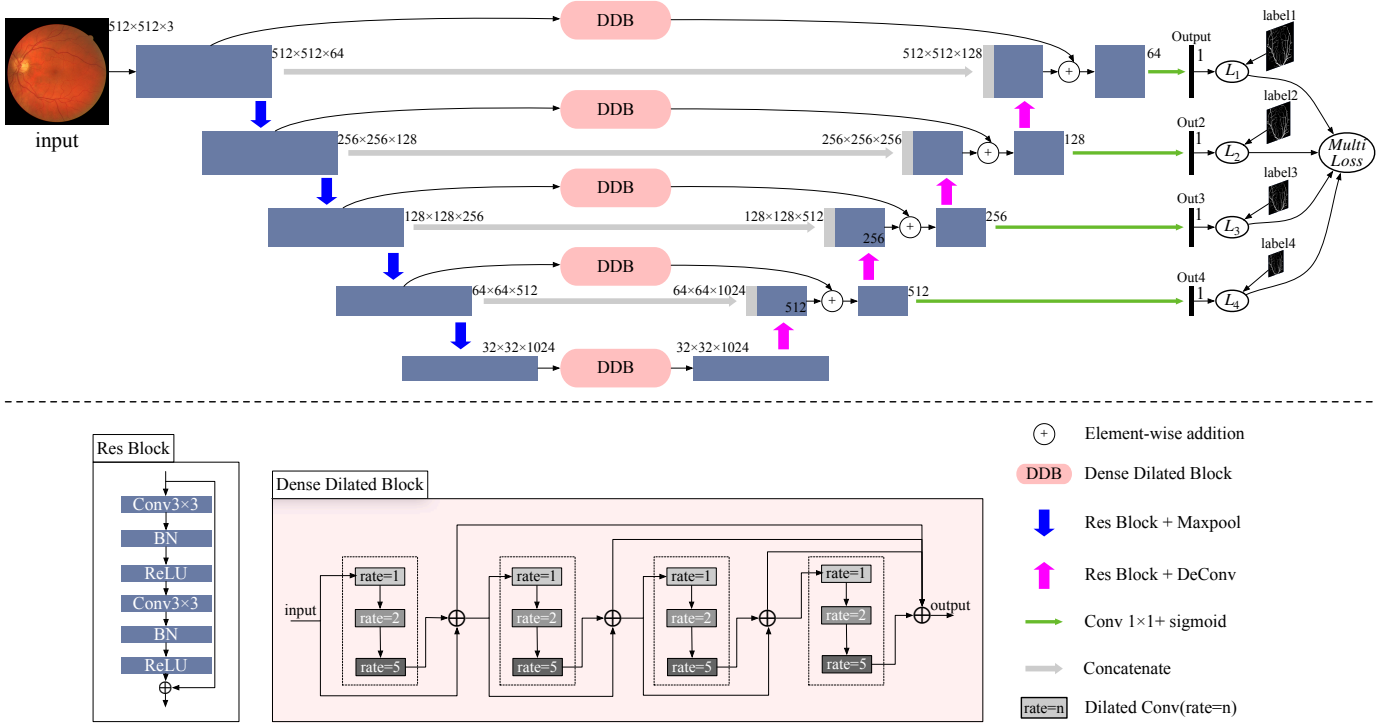
Fig. 2. Diagram of the DDNet structure. In DDNet, there are three modules in total. The first module is the encoder-decoder infrastructure in which residual blocks are used to extract image features for vessel segmentation. The second module is dense dilated block. In this module, we compute features in different scales and combine them for segmenting blood vessels. The last module is the multi-scale dice loss.

[48] [13] [49]. Marin *et al.* [45] employed a gray-level vector and moment invariant features using a neural network. Staal *et al.* [50] detected vessels by detecting and grouping the ridge features. Soares *et al.* [12] used a multi-scale Gabor transform to extract features. Zhou *et al.* [51] proposed to learn more discriminative features by improving dense CRF. Zhang *et al.* [52] proposed orientation scores for robust vessel detection. In the field of deep learning, encoder-decoder structures [22] have shown to be promising for vessel detection. Maninis *et al.* [53] proposed a multi-task structure for both vessel detection and optic disc segmentation. Fu *et al.* [17] proposed DeepVessel model to improve the sensitivity of segmentation by introducing side-output layers and CRF. Recently, Zhang *et al.* [54] introduced an edge-aware mechanism to convert the task into a multi-class task. Oliveira *et al.* [55] and Wu *et al.* [56] proposed to segment retinal vessels based on conventional FCN and multi-scale architecture, respectively. Compared with [54][56], we have taken the vessel connectivity into consideration using probability regularized walk, however [54]-[56] and other methods have not taken this into account. In [54], the method has considered the importance of thin vessels by distinguishing the thin vessels from thick vessels to force the networks to treat background pixels. However, this is different from our method to improve connectivity.

C. Vessel Reconnection

Although many methods have been proposed for vessel detection, few of them deal with the vessel connection in the detection. Favali *et al.* [57] matched vessel segments by utilizing the vessel connectivity based on spectral clustering. Joshi

et al. [58] presented a method for automatic identification and reconnection of broken vessels in the segmentation. Zhang *et al.* [59] connected the broken curvilinear structures using cortically inspired completion in ocular images.

Random walk is a technology that has applications in many fields. Some methods employed random walk for object segmentation [60] [61] [62]. In recent years, random walk has been widely used in vessel segmentation. Li *et al.* achieved considerable results using the hessian-based filter and random walk for vessel segmentation in [8]. M'hiri *et al.* [63] applied the similar method to coronary arteries. Zhu *et al.* [64] segmented the vessel by using random walk with adaptive cylinder flux.

III. DENSE DILATED NETWORK

Our proposed DDNet is an end-to-end deep neural network that mainly contains three parts. As shown in Fig. 2, the first part is the backbone encoder-decoder structure for feature extraction and up-sampling. The second part is the dense dilated block (DDB), which has been integrated with the backbone. The design details of DDB are given in III-B. The last part is the computation of the loss.

A. Backbone Encoder-Decoder

Since the encoder-decoder structure has shown promising results for semantic segmentation, we use the encoder-decoder structure as the backbone of our DDNet. Following the work of Zhang *et al.* [65], we replace the original blocks in the U-Net with ResNet [27] blocks to extract feature in our method.
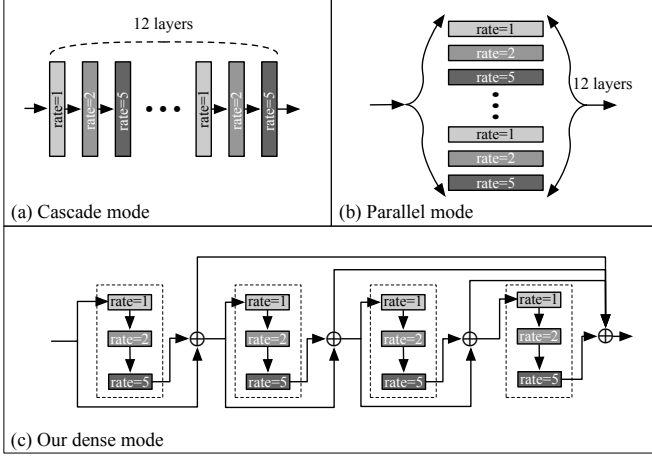
Fig. 3. Diagram of different modes of dilated convolution. (a) cascade mode, and (b) parallel mode. (c) dense mode.

Our results show that the replacement with ResNet blocks improves the semantic segmentation as it enables feature reuse and mitigates the gradient vanishing problem.

B. Dense Dilated Block

Recent studies in deep learning show that increasing the receptive field of the convolutional filter is important to salient feature extraction with more spatial information [66]. Earlier methods often use pooling layers and increase kernel size to enlarge the receptive field. However, pooling operation will potentially lead to the loss of some spatial information while the increased kernel size requires more computation. To overcome this limitation, dilated convolutions are adopted to replace the original convolutions [67]. Dilated convolution is a useful approach to adjust receptive field of feature points without pooling or increasing the kernel size. Meanwhile, convolutional layers in a dense mode are found to support feature reuse and make the network more efficient [29].

Motivated from the above observation, we propose dense dilated block (DDB) to organize atrous convolutions in a dense form. As shown in Fig. 3(c), the proposed DDB contains four major blocks and each major block has three dilated convolutions. Wang *et al.* [35] have successfully demonstrated the benefits of the dilated convolution for the segmentation. In our paper, we follow the work to use three dilated convolutions with rate 1, 2, and 5. Our method is different from existing cascade mode [33] and parallel mode [68]. Fig. 3 compares the three different modes. The benefits of this organization are twofold: first, it reuses the features of the early layers and makes the network more efficient; second, we adopt dilated convolution in each layer of DDB to extract deeper features that avoiding the reduction of the feature map resolution when increasing the receptive field. The depth-concatenate in our dense mode establishes the connection between different layers and makes full use of the features in different scales.

C. Multi-Scale Dice Loss

The loss function is an important factor in deep learning. Recent work [67] demonstrates that the multi-scale loss effec-

tively assists the training of the models. In the classification task, Li *et al.* [69] combined the loss of local and global features to improve the model. In the semantic segmentation task, Fu *et al.*[31] up-sampled the output of each layer in the decoder to the same size of the original image, and then used the multi-scale method for auxiliary training. In this paper, we adopt a multi-scale loss for the training as well. We resize the labels to the same size of each layer of the decoder to formulate the objective function. The ground-truth is then resized to the same size of the corresponding layer. Let $\mathcal{W}$ be the parameters of the convolutional layers, N as the number of decoder layers and $w = (w_1, w_2, \ldots, w_N)$ as the corresponding weights. The multi-scale loss is computed as follows.

$$\mathcal{L}\left(\mathcal{W}, w\right) = \sum_{n=1}^{N} \lambda_n L^{(n)}\left(\mathcal{W}, w_n\right), \tag{1}$$

where λ_n is the weight of nth layer in decoder, $L^{(n)}$ represents the loss of the nth layer. In our paper, we use equal weights as that in [31] and set $\lambda_n = 1/N$.

The use of the loss function is flexible and can be a mean square error (MSE) loss, a cross entropy (CE) loss or a Dice loss. It is also possible to apply different loss functions to different layers. We observed that there is a data imbalance between the foreground and the background in the vessel segmentation task. Recent work [70] and [71] have shown that the Dice coefficient can effectively reduce the negative effect of data imbalance. Thus, we choose the Dice coefficient for the final loss computation. In our paper, the multi-scale Dice coefficient loss $L_{dice}^{(n)}$ is computed as:

$$L_{dice}^{(n)} = 1 - \frac{2\sum_{i=1}^{K} p_i g_i + \epsilon}{\sum_{i=1}^{K} p_i^2 + \sum_{i=1}^{K} g_i^2 + \epsilon}, \tag{2}$$

where K denotes the number of pixels, $p_i \in [0,1]$ and $g_i \in \{0,1\}$ denote the predicted probability and ground truth value of ith pixel, respectively. The parameter ϵ ($\epsilon = 1.0$ in this paper) is a Laplace smoothing factor used to avoid numerical problems and accelerate the convergence of the training process. We set $N = 4$ in this work, which means that the encoder and decoder have four down-sampling layers and up-sampling layers, respectively. The gradient of $L_{dice}^{(n)}$ with respect to p_i is expressed as:

$$\begin{aligned}
\frac{\partial L_{dice}^{(n)}}{\partial p_i} = &- \frac{2g_i}{\left(\sum_{i=1}^{K} p_i^2 + \sum_{i=1}^{K} g_i^2 + \epsilon\right)} \\
&+ \frac{2p_i\left(2\sum_{i=1}^{K} p_i g_i + \epsilon\right)}{\left(\sum_{i=1}^{K} p_i^2 + \sum_{i=1}^{K} g_i^2 + \epsilon\right)^2}.
\end{aligned} \tag{3}$$

Then we use the standard stochastic gradient descent for back-propagation and the optimization.

IV. PROBABILITY REGULARIZED WALK

Although many deep learning based methods have been proposed for retinal vessel segmentation, most of the algorithms did not consider the structural connectivity. Therefore,

fractured vessel segments often appear in the results [59]. The broken vessel segmentation often post challenge to the subsequent analysis of the vessels such as the topology, the tortuosity, vessel lengths, etc. To obtain a better geometric description of the vessels, we propose a PRW algorithm to connect the broken vessel segments.

Random walk is one of the most widely known and used method in graph theory [72]. Let $G = (V, E)$ denote an undirected graph with a set of vertices V and a set of edges E. A random walk process in such graph can be characterized by a series of transition probabilities among its vertices. Many early work applied random walk to semantic segmentation [61] [8] [73]. These methods are often based on the similarity of pixels in the neighbour. In deep learning, we often compute the probabilities of pixels belonging to foreground or background, which are binarized with threshold to get the vessel map. Clearly, the binarization may discard some information such as the difference between the probabilities and the threshold values. Inspired from that, PRW integrates the probability outputs of deep learning into the regularized walk. In the vessel reconnection tasks, we treat each pixel in the image as a separate node in the graph, where the similarity between the two nodes can be evaluated based on the DDNet prediction probability. Since PRW aims to connect broken vessels, we expect the broken vessels to lie along the direction of the remaining vessels. Therefore, we integrate the local vessel directions as well. It should be pointed out that we only borrowed and simulated the walking process of the walker in the PRW instead of the random walk algorithm in the strict sense.

The PRW takes the probability map of DDNet and the binary vessel map as two inputs. Firstly, we label all the connected components in the binary map using morphological operations (bwlabel function in Matlab). For simplicity, we denote the largest connected component as L_1 and the rest of components as L_i $(i = 2, \ldots, N_c)$, where N_c is the number of components. Next, we connect each L_i to L_1 iteratively to obtain the final vessel map.

To connect L_i to L_1, we first detect the fractured vessels between L_i and L_1. As illustrated in Fig. 4, two points A and B with minimum distance in between are found from L_i and L_1. Mathematically,

$$d_{AB} = \underset{\substack{(x_A, y_A) \in L_i \\ (x_B, y_B) \in L_1}}{\arg \min} \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}. \quad (4)$$

where d_{AB} denotes the distance between A and B. Note that we calculate the d_{AB} only using the centerline pixels of L_i and L_1. Let L_i and L_1 contain m and n pixels $(m \ll n)$, respectively. Thus, the time complexity of the operation is: $O(m \times n)$.

In the second step, we determine a region of interest (ROI) where possible missing vessels are detected. Our initial method of detecting ROI is to find all fractured and isolated vessels by morphological operation. Denote the center point between A and B as M and a $l \times l$ square centered at M is determined as the ROI. In this paper, we set $l = 100$ empirically. With this threshold, we do not connect L_i too far away from L_1, i.e., $d_{AB} > l$.
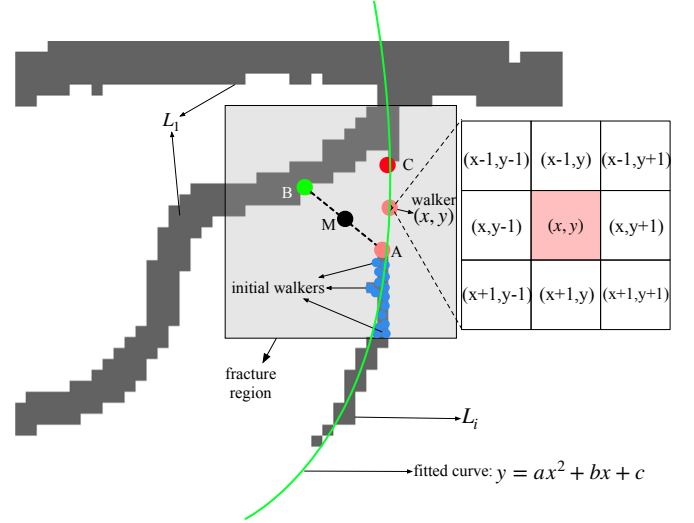

Fig. 4. Illustration of regularized walk. Grey area represents fracture region. In the fracture region, the pink point represents the walker with its position (x, y) during the migration. A is the ending point of L_i, and B is the closest point to A in L_1, and M is the middle point of A and B. The blue dots at the intersection of L_i and the fracture region indicate the initial location of the walkers. The green line represents the curve that is fitted by points in L_i. The square on the right side of the figure represents the 8-connected neighbors of the walker (shown in pink), and the numerical values represent the coordinates.

In the third step, we calculate the main walking direction of the walker. Based on the observation that the missing vessels are usually on the same line of the remaining vessels L_i, we fit the vessel using polynomial curves. In this paper, a second order polynomial $y = ax^2 + bx + c$ is adopted. All the points in L_i are used to fit the curve and determine the polynomial function. Then the intersecting point between the polynomial curve and L_1 is determined as C with coordinates of (x_c, y_c). Fig. 4 illustrates the process.

In the fourth step, we use all points that belong to both L_i and ROI as the initial walkers (seed pixels) and let these points walk iteratively until they arrive L_1. We optimize this walking process by the predicted probability of the neural network (P_{nn}) and the direction probability (P_d) of the current walker. The detailed description of P_{nn} and P_d is as follows.

An important issue in the fourth step is the walking direction in each movement. Intuitively, we hope the walker to move toward point C. However, as discussed, the binarization discards some information and a straight line or a fitted curve may discard the local changes. Therefore, we propose to regularize the walking direction using the probability map from the DDNet. The probability of a walker at (x, y) walking toward (x', y') is computed as $P(x', y')$:

$$P(x', y') = \alpha P_d(x', y') + (1 - \alpha) P_{nn}(x', y'), \quad (5)$$

where the first item $P_d(x', y')$ is a probability which denotes the association between the main walking direction and the direction from (x, y) to (x', y'). The second item $P_{nn}(x', y')$ denotes the probability output of DDNet. It indicates the probability of the pixel at (x', y') being a vessel pixel. α controls the trade-off between the two items.

In order to let the walker walks toward C, we design P_d such that a pixel with smaller distance to C has larger probability. For a pixel (x', y') where $x' \neq x_c$ or $y' \neq y_c$, we compute P_d as

$$P_d(x', y') = \frac{1}{\sqrt{(x' - x_c)^2 + (y' - y_c)^2}}, \qquad (6)$$

We set $\alpha = 0.2$ empirically in (5) such that neither item will dominate the results. In each step, we let the walker walk toward its 8-connected neighbor with largest probability $P(x', y')$:

$$(x', y') = \underset{(x', y') \in \Omega}{\arg\max} \{P(x', y')\}, \qquad (7)$$

where Ω is consisted of the 8-connected neighbors of (x, y), i.e., $\Omega = \{(x - 1, y - 1), (x - 1, y), (x - 1, y + 1), (x, y - 1), (x, y + 1), (x + 1, y - 1), (x + 1, y), (x + 1, y + 1)\}$. Additionally, if there is always $P_{nn} < \epsilon_{nn}$ (ϵ_{nn} is set to be 0.1 in this paper) around the walker in the walking process, we stop the movement of the walker and determine that the current ROI is not a fractured vessel or the terminal point A we selected is in the wrong direction. When point A is not the right point for connecting, the corresponding probability at the point will be low. In this case, the algorithm will terminate the connect operation to L_1 to avoid mis-connection. Meanwhile, it may still have the chance to be connected to L_1 as the walker can walk from point C in L_1. In order to obtain the newly created vessels without holes, we use a 3×3 kernel to simulate walkers during walking process. When all the walker activities are completed, there are many paths established in the ROI, or there is no path (indicating that the two regions are not expected to be connected). For time complexity, assume that the DDNet prediction binary map contains k fracture regions (ROI). Within each ROI, the PRW requires a maximum of $l \times l$ loops to traverse each pixel to reconstruct the blood vessel, where l is the length of the ROI. Therefore, the time complexity of the PRW is $O(k \times l^2)$.

V. EXPERIMENTS

A. Dataset

We use the following three public datasets in our experiments: DRIVE [50], STARE [74], and CHASE_DB1[75].

The DRIVE[1] dataset contains 40 color fundus images, which were originally divided into 20 images for training and 20 images for testing. The images were acquired using a Canon CR5 non-mydriatic 3-CCD camera with a $45°$ field of view (FOV). Each image in this dataset has a dimension of 565×584. We follow the same partition of the images in our training and testing.

The STARE[2] dataset comprises of 20 color fundus images. The images were captured by a Topcon TRV-50 fundus camera at $35°$ FOV. Half of images contain pathological indications and the other half is from healthy subjects. Each image has a dimension of 700×605. However, unlike the DRIVE dataset, there is no fixed partition of training and testing sets for

STARE. In this paper, we adopt the k-fold ($k = 4$) cross-validation [76] for the training and testing phases, similar to that in [77]. Therefore, we have 15 images for training and the remaining 5 images for testing in each fold.

The CHASE_DB1[3] dataset contains 28 images from both the left and right eyes, and the images were captured with a $30°$ FOV. Each image has a dimension of 999×960. For convenience, we still use the cross-validation method as STARE. In each fold, 21 images are used for training and the rest of 7 images are used for testing.

Considering that there are two sets of manual ground truth provided for the datasets, we follow [17] and [78] to use the manual annotation from the first observer as the ground truth for all the images.

B. Evaluation Metrics

In order to quantitatively evaluate our method, we compute Specificity (Spe), Sensitivity (Sen) and Accuracy (Acc) as our evaluation metrics:

$$Spe = \frac{TN}{TN + FP}, \qquad (8)$$

$$Sen = \frac{TP}{TP + FN}, \qquad (9)$$

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}, \qquad (10)$$

where TP, FN, TN, and FP represent true positive, false negative, true negative and false positive, respectively. In addition, we also introduce the area under receiver operating characteristic (ROC) curve (AUC) metric.

To evaluate the performance of the PRW algorithm in the fracture region, we define a new metric Err to explain how the PRW reconnects the vessel successfully:

$$Err = \frac{1}{N_r} \sum_{i=1}^{N_r} \left(\frac{FP}{TP + FP} \right)^{(i)}, \qquad (11)$$

where $\left(\frac{FP}{TP+FP} \right)^{(i)}$ denotes the pixel error of the ith ROI detected by PRW, and N_r is the number of ROIs. The metric Err explains the error rate when the regularized walk algorithms are used to recover broken vessels. A smaller Err indicates a better performance of the regularized walk algorithm.

C. Training and Testing Phases

We implemented DDNet on PyTorch framework with a single NVIDIA GPU (GeForce GTX 1080). During the training, we employ adaptive moment estimation (Adam) for optimizing the deep model. The initial learning rate is set to 0.001 and a weight decay of 0.0005. We follow the same poly learning rate policy in [41] and the maximum epoch is 1000. Due to GPU memory limitations, we set the batch size to 6 throughout the training. All training images are resized to 512×512. The output of the DDNet is a 1-channel map. We use the OTSU thresholding method [79] to binarize the output after

---

[1]http://www.isi.uu.nl/Research/Databases/DRIVE/

[2]http://www.ces.clemson.edu/$\sim$ahoover/stare/probing/index.html

[3]http://blogs.kingston.ac.uk/retinal/chasedb1/

TABLE I
Acc, Sen AND TIME CONSUMPTION OF PRW AT DIFFERENT ROI THRESHOLDS.

Threshold	0	20	40	60	80	**100**	120	140	160
Acc	0.9594	0.9594	0.9595	0.9599	0.9603	**0.9605**	0.9605	0.9605	0.9605
Sen	0.8126	0.8126	0.8127	0.8128	0.8130	**0.8131**	0.8131	0.8131	0.8131
Time(s)	-	-	3.786	20.695	54.425	**81.308**	117.128	164.363	252.148

TABLE II
ERROR RATES (Err) OF PRW AT DIFFERENT α VALUES.

α	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
Err	0.0	0.1145	0.1023	0.0197	**0.0173**	0.0188	0.1043	0.1169	0.2031	0.2165	0.2327

the testing. PRW algorithm is implemented on the Matlab platform. The PRW algorithm takes the output of the DDNet as input and delivers the binary map of reconnected vessels.

We perform data augmentation on the training and validation images with a random rotation of an angle from $-45°$ to $45°$ and a random contrast enhancement in the range $[-2, 2]$. The labels of each dataset are resized to 256×256, 128×128, 64×64 during the training. In the testing phase, the input image is resized to 512×512, without any enhancements. It is worth pointing out that the proposed DDNet reached an average inference time of 145.61ms with a total of 56.03M parameters and a total of 476.08G FLOPs in the testing phase.

D. Hyperparameters Setting

We introduce two hyperparameters l and α in the PRW algorithm, which control the detection range of the ROI and the movement behavior of the walkers. For ROI detection, a small threshold causes the ROI to fail to cover the fracture region, and an excessive threshold increases the time complexity. During the movement of walkers, an excessive α will cause the force of P_d to be much larger than P_{nn}, thus causing the walker to break away from the real trajectory. Therefore, the selection of appropriate hyperparameters will promote the performance and efficiency of vascular reconnection.

Firstly, we set a series of thresholds l from 0 to 160 with a step of 20 and implement the PRW algorithm at different values. Then, we compute the Acc and Sen scores and time consumption (Time) for each set of experiments. The comparison results are shown in Fig. 5 and Table I. As shown in the figure, further increasing the threshold l does lead to higher Acc and Sen, though the time consumption is increased. Therefore, we set $l = 100$. Secondly, we set a series of α from 0.05 to 0.5 with a step of 0.05 and implement PRW in each setting. We compute the reconnection error rate (Err) for different α values. The lower the Err, the better the performance of the algorithm. Table II shows the Err of the PRW at different α values. From Table II, we can conclude that when $\alpha = 0.2$, the error rate is the lowest. It should be explained that when $\alpha = 0$, $Err = 0$. The reason for this is that when $\alpha = 0$, $P_d = 0$, so that $P(x', y') = P_{nn}$. That is to say, the probability of the walker around the initial position is the largest, so that the walker will not be driven to escape from the initial position to reach L_1. Similarly, we reached the same conclusion on STARE and CHASE_DB1.
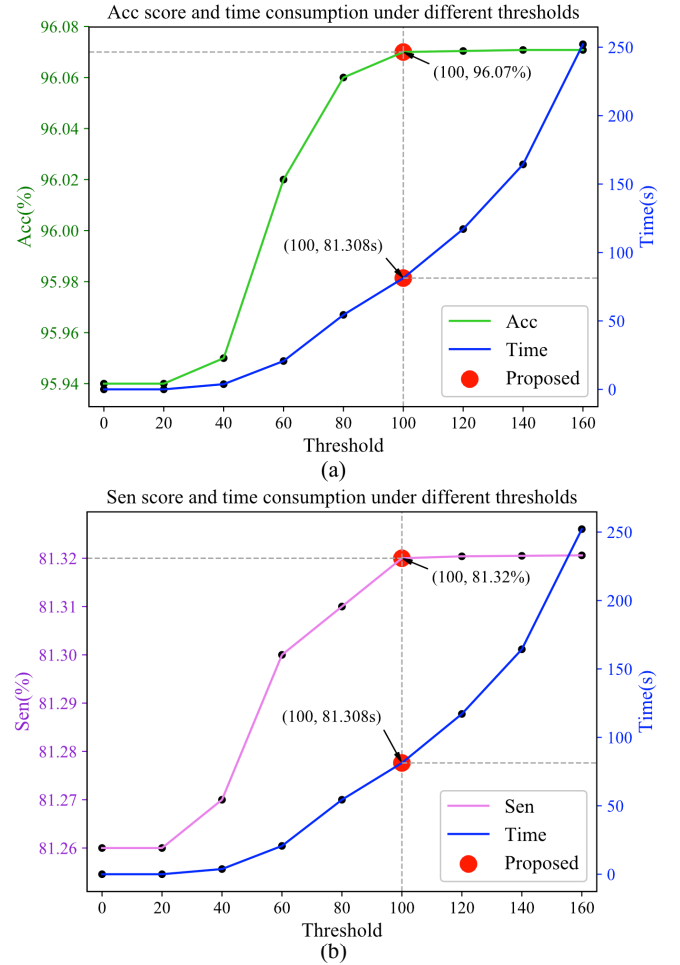

Fig. 5. Acc and Sen scores and time consumption results at different thresholds. (a) The green line with a left vertical axis indicates the Acc scores at different thresholds, and the blue line with a right vertical axis shows the time consumption of PRW at different thresholds. (b) The pink line with a left vertical axis indicates the Sen scores at different thresholds and the blue line with a right vertical axis shows the time consumption of PRW at different thresholds. The red point indicates the Acc and Sen score and time consumption under the threshold in this paper.

TABLE III
ABLATION STUDY FOR REPLACING ORIGINAL BLOCK WITH RESNET BLOCK

Method	Acc	Sen	Spe	AUC
U-Net	0.9485	0.7356	0.9602	0.9641
ResUNet	**0.9501**	**0.7887**	**0.9632**	**0.9665**

TABLE IV
ABLATION STUDY FOR THE DENSE DILATED BLOCK.

Method	Acc	Sen	Spe	AUC
ResUNet	0.9501	0.7887	0.9632	0.9665
ResUNet + Cascade Mode	0.9578	0.8109	0.9780	0.9781
ResUNet + Parallel Mode	0.9564	0.8064	0.9783	0.9776
ResUNet + Dense Mode	**0.9594**	**0.8126**	**0.9788**	**0.9796**

TABLE V
ABLATION STUDY FOR MULTI-SCALE DICE LOSS.

Method	Acc	Sen	Spe	AUC
DDNet + CE	0.9561	0.8089	0.9781	0.9780
DDNet + MSE	0.9573	0.8100	0.9775	0.9784
DDNet + M-net Loss	0.9590	0.8121	0.9779	0.9789
DDNet + MSD	**0.9594**	**0.8126**	**0.9788**	**0.9796**

TABLE VI
ABLATION STUDY FOR PROBABILITY REGULARIZED WALK.

Method	Acc	Sen	Spe
DDNet	0.9594	0.8126	**0.9788**
DDNet + CRW	0.9601	0.8128	0.9782
DDNet + PRW	**0.9607**	**0.8132**	0.9783

It can be observed from the first experimental result that a larger threshold l enables the ROI to cover fracture region. However, the number of fracture region is limited in the prediction map, and when $l > 100$, the AUC score hardly increases. From the analysis of the second experimental result, we conclude that the vessel reconnection error rate is the lowest when $\alpha = 0.2$, and the inference time required for PRW increases exponentially when $\alpha > 0.2$. Based on the above two experimental results, we can conclude that the PRW algorithm performs best in the vessel reconnection task when $l = 100$ and $\alpha = 0.2$.

E. Ablation Studies

In the proposed method, we have introduced the ResNet blocks, the DDB, the multi-scale dice loss, and the PRW algorithm to improve the baseline encoder-decoder method. To justify the effectiveness of these modules, we conduct the following ablation studies. For simplicity, the DRIVE dataset is used. We start from the baseline U-Net [22] approach and evaluate how these modules affect the results.

1) Ablation study for replacing original blocks with ResNet blocks: To justify the use of the ResNet blocks, we replace the original blocks in the U-Net with ResNet blocks. We denote the modied U-Net as ResUNet. Table III shows the segmentation results. Our results show that we achieve 0.9501, 0.7887, 0.9632, and 0.9665 on Acc, Sen, Spe and AUC in comparison to 0.9485, 0.7356, 0.9602, and 0.9641 in the original U-Net.

2) Ablation study for dense dilated block: To justify the use of the proposed DDB module, we compare the dense mode with cascade mode, parallel mode as well as the ResUNet without additional blocks. The ResUNet is taken as baseline and the multi-scale dice loss is used in all experiments. For fair comparison, we set up 12 dilated convolutional layers in cascade mode and the parallel mode. The dilation rates are set to 1, 2 and 5, as shown in Fig. 3. The comparison results are shown in Table IV. The results show that our DDB module is effective to improve the results and it performs better than the other two modes as well.

3) Ablation study for multi-scale dice loss: To justify the use of multi-scale dice loss (MSD), we compare it with cross entropy loss (CE) and mean square error loss (MSE). Additionally, we set up another comparative experiment comparing

the multi-scale loss in M-Net [31] (named as M-net Loss) with the proposed multi-scale label loss. In this ablation setting, the backbone is our proposed DDNet. We show the comparison results in Table V. The results show that the multi-scale dice loss achieves highest AUC score of 0.9891. Therefore, the use of dice loss in the vessel segmentation task contributes to the performance improvement.

4) Ablation study for probability regularized walk: To justify the benefits of the proposed PRW, we compare the PRW with conventional walk (CRW). In addition, the results without any random walk are also given. All results are shown in Table VI. The results show that PRW leads to highest Acc and Sen at 0.9607 and 0.8132, respectively. Since the PRW algorithm may misclassify some true negative (TN) pixels, Spe has a 0.08% decrease. Table VII compares the Err by CRW and PRW in the DRIVE, STARE, and CHASE_DB1 datasets. We can observe a large drop of the error Err from CRW to PRW. We can conclude that the PRW effectively reconnects the interrupted region.

Fig. 6 shows an example of detected vessels before and after the PRW as well as that after the CRW. In the figure, the red indicates the broken vessels recovered by both algorithms. The green arrow indicates a scenario where the proposed PRW method works well while the CRW recovers the vessels incorrectly.

F. Comparison With State-of-the-art Methods

To justify the performance improvement compared with the state-of-the-art algorithms, we finally compare our method with other methods proposed by Azzopardi [49], Zhao [15], Roychowdhury [80], Zhao [16], Xie [67], Fu [17], Zhou [51] and Ronneberger [22]. For the DRIVE dataset, the results are from the original papers. For STARE and CHASE_DB1, the results are obtained by ourselves, where the original implementations of [67] and [17] are used to obtain the results for Xie's method and Fu's method. For [22], we replace the CE loss with the Dice loss from the original U-Net code as Dice loss works better. For [51], we implement the code and carefully fine-tune the parameters to get the results. As we mentioned in Section V-A, we use four-fold cross-validation for the STARE and CHASE_DB1 datasets.

Table VIII shows the Acc, Sen, Spe and AUC of these methods, the DDNet and the DDNet with PRW on the three

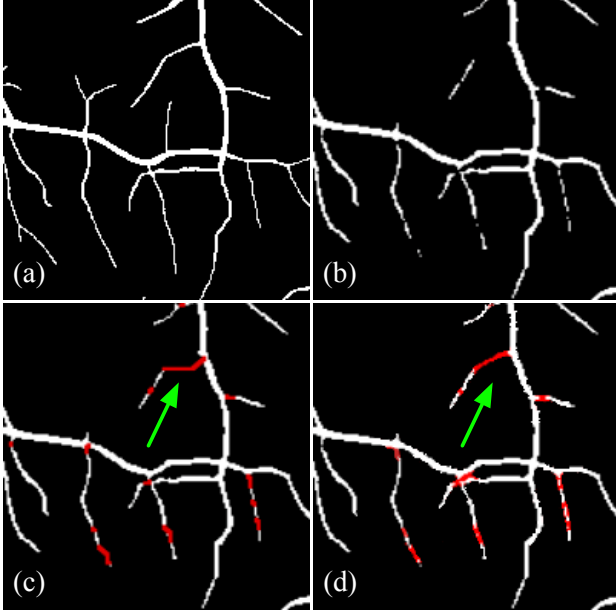| Methods	DRIVE	STARE	CHASE_DB1
	Err	Err	Err
DDNet+CRW	0.1390	0.1286	0.1054
DDNet+PRW	**0.0173**	**0.0195**	**0.0160**

Fig. 6. Illustration of vessel reconnection. (a) manual ground truth. (b) result of DDNet. (c) DDNet+CRW. (d) DDNet+PRW.

datasets. From Table VIII, the proposed DDNet and PRW improve the accuracy of vessel segmentation consistently in all three datasets. Fig. 7 shows results from three examples. From the figure, the proposed method using DDNet and PRW segments vessels more accurately.

VI. CONCLUSIONS

In this paper, we propose a dense dilated convolution network and a probability regularized walk algorithm to segment vessels and reconnect fractured vessels, respectively. The proposed DDNet integrates the context information of different layers and utilizes multi-scale dice loss to improve the segmentation of vessels. Dense Dilated Block (DDB) effectively improves the feature representation of deeper convolutional layers. And the efficient reuse of deep features facilitates efficiency of parameter usage. The probability regularized walk algorithm effectively reconnects the fractured blood vessels, which makes the segmentation results better. The experiments show that the proposed method is able to achieve better vessel detection results than other methods. A limitation of our approach is that the PRW is proposed as a post-processing of DDNet. In future, we will explore to integrate the random walk within the deep learning.

REFERENCES

[1] S. T. Venkataraman, J. G. Flanagan, and C. Hudson, "Vascular reactivity of optic nerve head and retinal blood vessels in glaucoma—a review," *Microcirculation*, vol. 17, no. 7, pp. 568–581, 2010.

[2] J. Flammer, S. Orgül, V. P. Costa, N. Orzalesi, G. K. Krieglstein, L. Metzner Serra, J. Renard, and E. Stefánsson, "The impact of ocular blood flow in glaucoma," *Progress in Retinal and Eye Research*, vol. 21, no. 4, pp. 359–393, 2002.

[3] K. A. Vermeer, F. M. Vos, H. G. Lemij, and A. M. Vossepoel, "A model based method for retinal blood vessel detection," *Computers in Biology and Medicine*, vol. 34, no. 3, pp. 209–219, 2004.

[4] H. Jelinek and M. J. Cree, *Automated image detection of retinal pathology*, Crc Press, 2009.

[5] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images," *British Journal of Ophthalmology*, vol. 83, no. 8, pp. 902–910, 1999.

[6] Y. Zhao, J. Xie, P. Su, Y. Zheng, Y. Liu, J. Cheng, and J. Liu, "Retinal artery and vein classification via dominant sets clustering-based vascular topology estimation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2018, pp. 56–64, Springer, Cham.

[7] Y. Zhao, J. Xie, H. Zhang, Y. Zheng, Y. Zhao, H. Qi, Y. Zhao, P. Su, J. Liu, and Y. Liu, "Retinal vascular network topology reconstruction and artery/vein classification via dominant set clustering," *IEEE Transactions on Medical Imaging*, 2019.

[8] Y. Li, H. Gong, W. Wu, G. Liu, and G. Chen, "An automated method using hessian matrix and random walks for retinal blood vessel segmentation," in *2015 8th International Congress on Image and Signal Processing (CISP)*. IEEE, 2015, pp. 423–427.

[9] U. T. V. Nguyen, A. Bhuiyan, L. A. F. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern recognition*, vol. 46, no. 3, pp. 703–715, 2013.

[10] I. Daubechies, *Ten lectures on wavelets*, vol. 61, Siam, 1992.

[11] P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, "Fast retinal vessel detection and measurement using wavelets and edge location refinement," *PloS ONE*, vol. 7, no. 3, pp. e32435, 2012.

[12] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.

[13] G. Läthén, J. Jonasson, and M. Borga, "Blood vessel segmentation using multi-scale quadrature filtering," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 762–767, 2010.

[14] Y. Zhao, J. Zhao, J. Yang, Y. Liu, Y. Zhao, Y. Zheng, L. Xia, and Y. Wang, "Saliency driven vasculature segmentation with infinite perimeter active contour model," *Neurocomputing*, vol. 259, pp. 201–209, 2017.

[15] 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 Transactions on Medical Imaging*, vol. 34, no. 9, pp. 1797–1807, 2015.

[16] Y. Zhao, Y. Zheng, Y. Liu, Y. Zhao, L. Luo, S. Yang, T. Na, Y. Wang, and J. Liu, "Automatic 2-d/3-d vessel enhancement in multiple modality images using a weighted symmetry filter," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 438–450, 2018.

[17] 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 *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2016, pp. 132–139, Springer, Cham.

[18] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2369–2380, 2016.

[19] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "Ce-net: Context encoder network for 2d medical image segmentation," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2019.

[20] Z. Zhijie, Fu. Huazhu, D. Hang, S. Jianbing, P. Yanwei, and S. Ling, "Et-net: A generic edge-attention guidance network for medical image segmentation," in *Medical Image Computing and Computer Assisted Intervention*, 2019.

[21] S. Zhang, H. Fu, Y. Yan, Y. Zhang, Q. Wu, M. Yang, M. Tan, and Y. Xu, "Attention guided network for retinal image segmentation," in *Medical Image Computing and Computer Assisted Intervention*, 2019.

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

[23] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
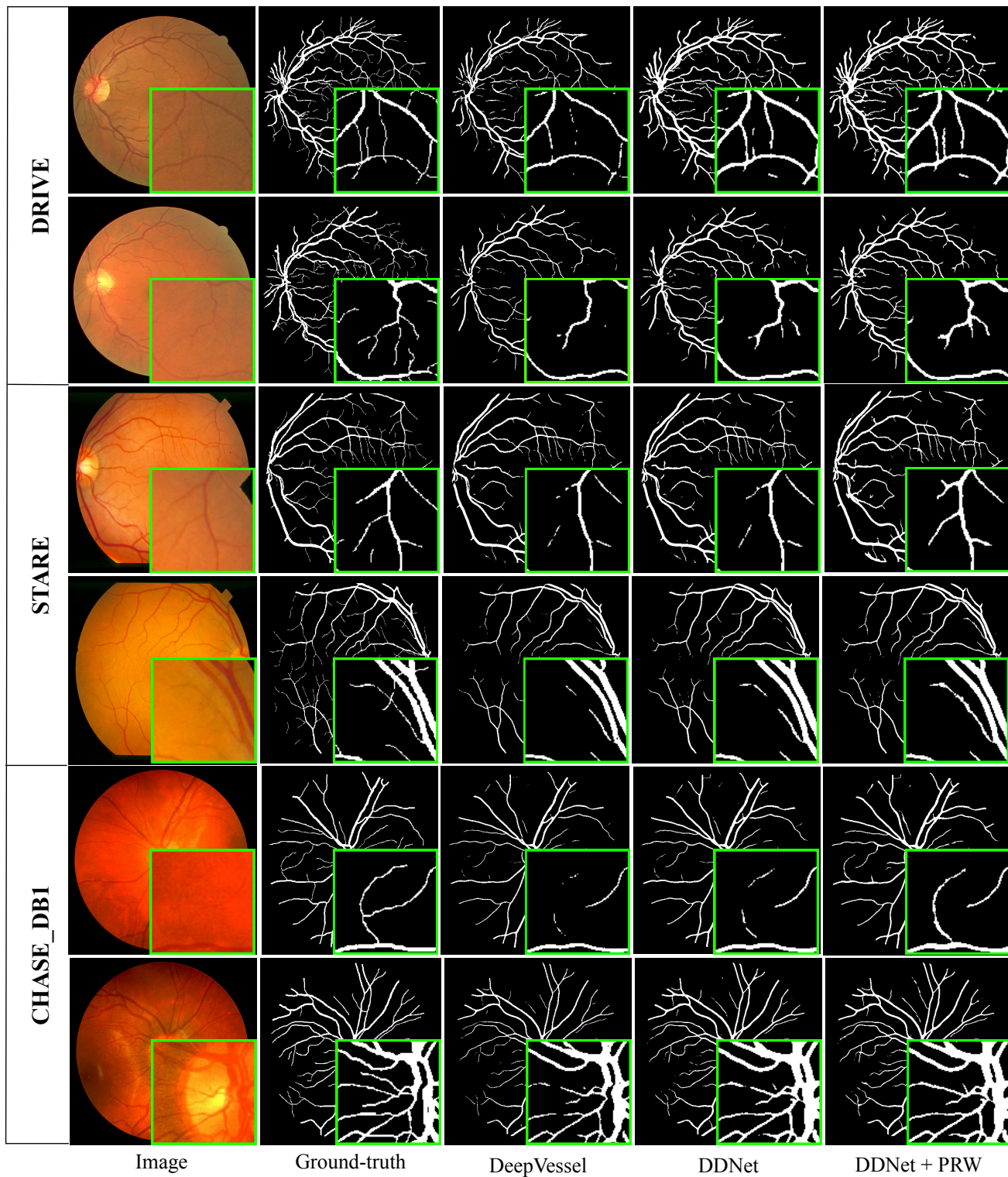
Fig. 7. Visual illustration of segmentation results on DRIVE, STARE, CHASE_DB1. From left to right: original image, ground-truth, results by DeepVessel [17], DDNet and DDNet + PRW. We show a local enlargement with a green box.

TABLE VIII
PERFORMANCE COMPARISON ON THE DRIVE, STARE AND CHASE_DB1 DATASETS.

Methods	DRIVE				STARE				CHASE_DB1			
	Acc	Sen	Spe	AUC	Acc	Sen	Spe	AUC	Acc	Sen	Spe	AUC
2nd Human Observer	0.9472	0.7760	0.9724	-	0.9349	0.8952	0.9384	-	0.9545	0.8105	0.9711	-
Azzopardi [49]	0.9442	0.7655	0.9704	0.9614	0.9497	0.7716	0.9701	0.9563	0.9387	0.7585	0.9587	0.9487
Ronneberger [22]	0.9501	0.7356	0.9602	0.9641	0.9517	0.7101	0.9682	0.9615	0.9499	0.7094	0.9767	0.9613
Zhao [15]	0.9540	0.7420	**0.9820**	0.8620	0.9560	0.7800	0.9780	0.8740	-	-	-	-
Roychowdhury [80]	0.9494	0.7395	0.9782	0.9672	0.9560	0.7317	**0.9842**	0.9673	0.9467	0.7615	0.9575	0.9623
Zhao [16]	0.9580	0.7740	0.9790	0.9750	0.9570	0.7880	0.9760	0.9590	-	-	-	-
Xie [67]	0.9435	0.7364	0.9730	0.9774	0.9402	0.7116	0.9724	0.9801	0.9380	0.7151	0.9679	0.9798
Fu [17]	0.9533	0.7603	0.9776	0.9789	0.9609	0.7412	0.9701	0.9790	0.9581	0.7130	**0.9812**	0.9806
Zhou [51]	0.9469	0.8078	0.9674	-	0.9585	0.8065	0.9761	-	0.9520	0.7553	0.9751	-
Our DDNet	**0.9594**	0.8126	0.9788	**0.9796**	**0.9685**	0.8391	0.9769	**0.9858**	**0.9637**	0.8268	0.9773	**0.9812**
Our DDNet+PRW	0.9607	0.8132	0.9783	-	0.9698	0.8398	0.9761	-	0.9648	0.8275	0.9768	-

[24] G. Lin, A. Milan, C. Shen, and I. D. Reid, "Refinenet: Multi-path refinement networks for high-resolution semantic segmentation.," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5186–5177.

[25] P. Bilinski and V. Prisacariu, "Dense decoder shortcut connections for single-pass semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6596–6605.

[26] M. Yang, K. Yu, C. Zhang, Z. Li, and K. Yang, "Denseaspp for semantic segmentation in street scenes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3684–3692.

[27] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[28] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5987–5995.

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

[30] H. Zhao, X. Qi, X. Shen, J. Shi, and J. Jia, "Icnet for real-time semantic segmentation on high-resolution images," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 405–420.

[31] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018.

[32] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[33] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," *arXiv preprint arXiv:1511.07122*, 2015.

[34] L. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.

[35] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, and G. Cottrell, "Understanding convolution for semantic segmentation," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018, pp. 1451–1460.

[36] T. Zhang, H. Fu, Y. Zhao, J. Cheng, M. Guo, Z. Gu, B. Yang, Y. Xiao, S. Gao, and J. Liu, "Skrgan: Sketching-rendering unconditional generative adversarial networks for medical image synthesis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2019, pp. 777–785, Springer, Cham.

[37] M. K. Ikram, F. J. De Jong, E. J. Van Dijk, N. D. Prins, A. Hofman, M. M. B. Breteler, and P. T. V. M. De Jong, "Retinal vessel diameters and cerebral small vessel disease: the rotterdam scan study," *Brain*, vol. 129, no. 1, pp. 182–188, 2005.

[38] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[39] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[40] F. Yu, V. Koltun, and T. A. Funkhouser, "Dilated residual networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 472–480.

[41] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6230–6239.

[42] L. Zhou, C. Zhang, and M. Wu, "D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 182–186.

[43] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe, "Full-resolution residual networks for semantic segmentation in street scenes," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3309–3318, 2017.

[44] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1520–1528.

[45] D. Marín, A. Aquino, M. E. Gegúndez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on Medical Imaging*, vol. 30, no. 1, pp. 146–158, 2010.

[46] J. I. Orlando and M. Blaschko, "Learning fully-connected crfs for blood vessel segmentation in retinal images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2014, pp. 634–641, Springer, Cham.

[47] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Transactions on Medical Imaging*, vol. 26, no. 10, pp. 1357–1365, 2007.

[48] T. Jerman, F. Pernuš, B. Likar, and Ž. Špiclin, "Enhancement of vascular structures in 3d and 2d angiographic images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 9, pp. 2107–2118, 2016.

[49] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable cosfire filters for vessel delineation with application to retinal images," *Medical image analysis*, vol. 19, no. 1, pp. 46–57, 2015.

[50] J. Staal, M. D Abràmoff, M. Niemeijer, M. A Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[51] L. Zhou, Q. Yu, X. Xu, Y. Gu, and J. Yang, "Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement," *Computer Methods and Programs in Biomedicine*, vol. 148, pp. 13–25, 2017.

[52] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. W. Pluim, R. Duits, and B. M. ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE Transactions on Medical Imaging*, vol. 35, no. 12, pp. 2631–2644, 2016.

[53] K.K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2016, vol. 9901, pp. 140–148, Springer, Cham.

[54] Y. Zhang and A. Chung, "Deep supervision with additional labels for retinal vessel segmentation task," in *Medical Image Computing and Computer Assisted Intervention*, 2018, pp. 83–91.

[55] A. Oliveira, S. Pereira, and C. Silva, "Retinal vessel segmentation

based on fully convolutional neural networks," *Expert Systems with Applications*, vol. 112, pp. 229–242, 2018.

[56] Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, "Multiscale network followed network model for retinal vessel segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2018, pp. 119–126, Springer, Cham.

[57] M. Favali, S. Abbasi-Sureshjani, B. ter Haar Romeny, and A. Sarti, "Analysis of vessel connectivities in retinal images by cortically inspired spectral clustering," *Journal of Mathematical Imaging and Vision*, vol. 56, no. 1, pp. 158–172, 2016.

[58] V. S. Joshi, M. K. Garvin, J. M. Reinhardt, and M. D. Abramoff, "Identification and reconnection of interrupted vessels in retinal vessel segmentation," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2011, pp. 1416–1420.

[59] J. Zhang, E. Bekkers, D. Chen, T. T. M. Berendschot, J. Schouten, J. P. W. Pluim, Y. Shi, B. Dashtbozorg, and B. M. t. H. Romeny, "Reconnection of interrupted curvilinear structures via cortically inspired completion for ophthalmologic images," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 5, pp. 1151–1165, 2018.

[60] M. Meila and J. Shi, "Learning segmentation by random walks," in *Advances in Neural Information Processing Systems*, 2001, pp. 873–879.

[61] L. Grady, "Random walks for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.

[62] M. Cheng and G. Zhang, "Connectedness of random walk segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 200–202, 2011.

[63] F. M'hiri, L. Duong, C. Desrosiers, and M. Cheriet, "Vesselwalker: Coronary arteries segmentation using random walks and hessian-based vesselness filter," in *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. IEEE, 2013, pp. 918–921.

[64] N. Zhu and A. C. S. Chung, "Random walks with adaptive cylinder flux based connectivity for vessel segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2013, pp. 550–558, Springer, Cham.

[65] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, 2018.

[66] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning.," in *AAAI*, 2017, vol. 4, p. 12.

[67] S. Xie and Z. Tu, "Holistically-nested edge detection," in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1395–1403.

[68] L. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 801–818.

[69] W. Li, X. Zhu, and S. Gong, "Person re-identification by deep joint learning of multi-loss classification," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 2017, pp. 2194–2200.

[70] F. Milletari, N. Navab, and S. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.

[71] C. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *In: Cardoso M. et al. (eds) Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2017, ML-CDS 2017. LNCS*, vol. 10553, pp. 240–248. Springer, Cham, 2017.

[72] László L. et al., "Random walks on graphs: A survey," *Combinatorics, Paul erdos is eighty*, vol. 2, no. 1, 1993.

[73] X. Dong, J. Shen, L. Shao, and L. Van Gool, "Sub-markov random walk for image segmentation," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 516–527, 2016.

[74] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, pp. 203–210, 2000.

[75] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.

[76] T. Fushiki, "Estimation of prediction error by using k-fold cross-validation," *Statistics and Computing*, vol. 21, no. 2, pp. 137–146, 2011.

[77] J. Mo and L. Zhang, "Multi-level deep supervised networks for retinal vessel segmentation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12(12), pp. 2181–2193, 2017.

[78] Y. Lin, H. Zhang, and G. Hu, "Automatic retinal vessel segmentation via deeply supervised and smoothly regularized network," *IEEE Access*, vol. 7, pp. 57717–57724, 2018.

[79] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[80] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Iterative vessel segmentation of fundus images," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1738–1749, 2015.