A Global and Local Enhanced Residual U-Net for Accurate Retinal Vessel Segmentation

Sheng Lian, Lei Li, Guiren Lian, Xiao Xiao, Zhiming Luo, and Shaozi Li

Abstract—Retinal vessel segmentation is a critical procedure towards the accurate visualization, diagnosis, early treatment, and surgery planning of ocular diseases. Recent deep learning-based approaches have achieved impressive performance in retinal vessel segmentation. However, they usually apply global image pre-processing and take the whole retinal images as input during network training, which have two drawbacks for accurate retinal vessel segmentation. First, these methods lack the utilization of the local patch information. Second, they overlook the geometric constraint that retina only occurs in a specific area within the whole image or the extracted patch. As a consequence, these global-based methods suffer in handling details, such as recognizing the small thin vessels, discriminating the optic disk, etc. To address these drawbacks, this study proposes a **G**lobal and **L**ocal enhanced residual **U**-n**E**t (*GLUE*) for accurate retinal vessel segmentation, which benefits from both the globally and locally enhanced information inside the retinal region. Experimental results on two benchmark datasets demonstrate the effectiveness of the proposed method, which consistently improves the segmentation accuracy over a conventional U-Net and achieves competitive performance compared to the state-of-the-art.

Index Terms—Retinal vessel segmentation, deep learning, weighted Res-UNet, global and local enhance

1 Introduction

According to reports in [1], [2], [3], Diabetic Retinopathy (DR), Glaucoma and Age-related Macular Degeneration (AMD) are the leading cause of blindness in the aging population. These statistical results promote researchers to develop the automatic diagnosis systems for retinal pathology. The retina is the only human body part which can observe the microcirculation through a noninvasive fundus examination. With the complex fundoscope system, digital retinal images can provide a magnified view of retina area, including retinal vessel branches, optical disk and macula. The subtle changes and abnormalities in retinal vessel structures can be deemed as an important signal for diagnosing multiple diseases, including DR [4], AMD [2], cardiovascular disease [5], hypertension [6] and many chronic eye diseases [7]. The task of retinal vessel segmentation is to classify pixels that belong to the vessel region in given retinal images, which plays an important role in ophthalmologists' diagnosing procedure.

Manual retinal vessel segmentation is a time consuming and tedious task even for a well-trained ophthalmologist. However, the diagnosis of retinal diseases, especially some acute diseases, require an urgent feedback. As such, computer-aided automatic retinal vessel segmentation can help for reducing medical costs, avoiding delayed treatment and improving efficiency. Although the quality of retinal imaging has improved significantly with the improvement of new imaging technology [8], automatic retinal vessel segmentation is still a challenging task. The main challenges can be briefly summarized as follows:

- *Indistinct small vessels*. Small blood vessels located at the end of branches have extremely low contrast which are even indistinguishable for professional ophthalmologists.
- *Low contrast at optic disk area*. Optic disk area is usually brighter and with lower contrast, which makes it hard to recognize the retinal vessels in this area.
- *Complex bifurcated structure*. Retinal vessels have complex bifurcated structure similar to trees, such as bifurcations, crossovers, closely parallel vessels, and junctions, which are difficult to maintain.
- *Influence of abnormal area*. Some retinal images are with abnormal lesions, such as microaneurysms and exudates, which increase the difficulty of the segmentation task.
- *Illumination*. Poor or overexposed illumination caused by the light source of camera will reduce image contrast, which results in the non-sharp boundary of retinal vessels.

To deal with these challenges, existing learning-based approaches usually apply an global image-level pre-processing operation on the whole retina image to enhance the image

- *S. Lian, L. Li, X. Xiao, and S. Li are with the Cognitive Science Department, Xiamen University, Xiamen, Fujian Sheng 361005, China. E-mail: {lancerlian, lilei}@stu.xmu.edu.cn, 6487888499@qq.com, szlig@xmu.edu.cn.*
- *G. Lian is with the College of Mathematics and Informatics, Fujian Normal University, Fuzhou Shi, Fujian Sheng 350100, China. E-mail: lz567@fjnu.edu.cn.*
- *Z. Luo is with the Postdoc Center of Information and Communication Engineering, Xiamen University, Xiamen, Fujian Sheng 361005, China, and also with the Key Laboratory of Cognitive Computing and Intelligent Information Processing, Fujian Education Institutions, Wuyi University, Wuyishan 354300, China. E-mail: zhiming.luo@xmu.edu.cn.*

Manuscript received 14 Dec. 2018; revised 20 Apr. 2019; accepted 12 May 2019. Date of publication 16 May 2019; date of current version 3 June 2021. (Corresponding author: Zhiming Luo.)
Digital Object Identifier no. 10.1109/TCBB.2019.2917188

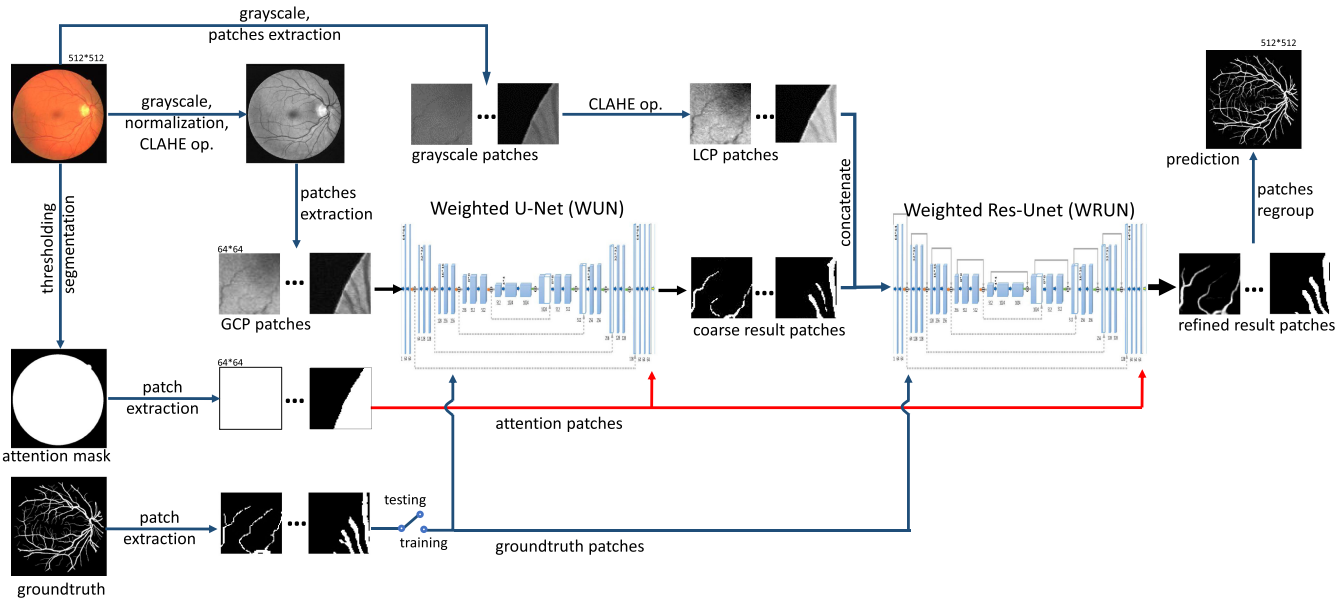

Fig. 1. The overall pipeline of our proposed model which contains two parts. A weighted U-Net (WUN) takes the globally enhanced patches as input to produce a coarse segmentation, and a weighted residual U-Net (WRUN) with locally enhanced patches along with previous segmentation results as input to perform a refinement. GCP and LCP indicate Global CLAHE pre-processing and Local CLAHE pre-processing, respectively.

contrast and randomly crop patches for learning a segmentation model. However, only considering the image-level statistical information for pre-processing will lose sight of the local information which are important for recognizing low-contrast local small vessels.

Retina only appears in specific area, in which pixels outside this area are useless for segmentation, then it's essential to get rid of the irrelevant noisy background. Meanwhile, global image-level and local patch-level pre-processing operations are complementary schemes for enhancing the retinal image contrast. In this paper, we propose a **G**lobal and **L**ocal enhanced residual **U**-n**E**t (*GLUE*) model with a cascaded refinement structure to tackle this challenging retinal vessel segmentation task. Our model as shown in Fig. 1 is composed of two parts, a weighted U-Net (WUN) and a weighted residual U-Net (WRUN). First, the WUN produces a coarse segmentation map from globally enhanced patches. Then the WRUN refines this coarse segmentation by integrating locally enhanced patches, whose parameters are learned automatically instead of manually fine-tuned. In addition, we also add residual connections in the second part of our model (WRUN) for learning more discriminative features. Compared to previous UNet-like structure [9], [10], we also utilize a cascaded U-Net structure which can benefit from both local and global enhanced retinal images. After all these improvements, our model can learn more discriminative features for distinguishing vessel and non-vessel pixels and have a better maintaining of the retinal vessel tree structure.

To evaluate the performance, we conduct experiments on two representative datasets, DRIVE dataset and STARE dataset. Experimental results demonstrate that our proposed method outperforms previous representative approaches. The comparison between the coarse and refined segmentation results shows the effectiveness of our proposed cascade refinement scheme. Besides, the visual example results in the zoomed-in view indicate that our model can deal with challenging situations. This paper is an extension to our previous ITME2018 conference paper [11]. In this version, we make several significant improvements by incorporating globally and locally enhancement on retinal images, and construct our model in a cascaded manner to do a refinement.

The rest of the paper is organized as follows. In Section 2, we review representative retinal vessel segmentation methods. In Section 3, we describe the proposed method in detail. In Section 4, we reported the experimental results. And we conclude our paper in Section 5.

2 Related Work

Realizing the importance of retinal vessel segmentation task, different segmentation methods have been proposed. Excellent surveys of the existing methods for retinal vessel segmentation can be referred in [12], [13]. In this section, we briefly introduce some of the most representative work. In general, the retinal blood vessel segmentation methods can mainly be divided into unsupervised and supervised.

Unsupervised methods attempt to find inherent patterns of retinal vessels without any manual annotation. Most of these approaches are rule-based techniques, including vessel tracking [14], [15], [16], matched filtering [17], [18], [19], morphological processing [20], [21], [22], thresholding [23], [24], [25], etc.

Yin et al. [14] exploit the fact that retinal vessels have connected branch-like structure, and proposed a vessel tracking-based segmentation method. In this method, a Bayesian method is used to detect vessel edge points by maximizing the posterior as criterion. Given these initial seed points, the entire vessel tree is tracked by following the vessel centerline based on local information. With the assumption that retinal vessel's intensity can be modeled as a Gaussian-shaped curve, Wang et al. [18] proposed a Matched Filtering-based method with multiwavelet kernels and multiscale hierarchical decomposition. In this study, vessels are enhanced using matched filtering with multiwavelet kernels. Knowing that vessels are linear and connected structures in retina, mathematical

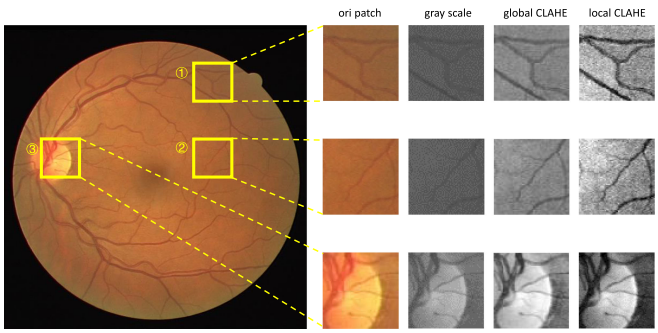

Fig. 2. Three groups of result patches after locally and globally CLAHE processing. Here, we plot examples of original patches, gray scaled patches, global CLAHE processed (GCP) patches, and local CLAHE processed (LCP) patches in zoom-in view.

morphology can be adopted. Fraz et al. [20] used the first order derivative of a Gaussian filter to identify vessel centerlines, followed by a multi-directional morphological top-hat transform to segment the vessels. Moreover, one possible way to delineate retinal vessel pixels is to segment by local thresholding. For example, Jiang et al. [23] proposed an adaptive local thresholding framework on a verification-based multi-threshold probing scheme to segment retinal vessels. Generally, compared to supervised methods, unsupervised methods have a higher segmentation speed, lower computational complexity but lower segment accuracy.

Supervised methods segment retinal vessels by learning a model from a training set annotated by experienced ophthalmologists. All of the supervised methods try to select the most discriminative set of feature vectors from training images to better classifying vessel and non-vessel pixels. The supervised classifiers in such methods range from k-Nearest-Neighbor (kNNs) [26], [27] and support vector machines (SVM) [28], [29] to more complex models such as deep neural networks [30], [31], etc. Niemeijer et al. [26] extract only the green plane of retinal images and conduct k-NN algorithm as classifier to determine the pixels' class. Based on the assumption that image ridges coincide approximately with vessels' centerline, Staal et al. [27] proposed a ridge based vessel segmentation methodology with k-NN classifier. Ricci et al. [28] adopt line operation as feature vector and use SVM for pixel classification. In this method, a line detector based on the evaluation of the average gray level alone lines is applied for first stage unsupervised classification. Then, two orthogonal line detectors are adopted to construct feature vectors for supervised classification using an SVM. Nekovei et al. [30] applied a multilayer perceptron neural network, for which the inputs were derived from a principal component analysis (PCA), to identify retinal vessel pixels. To our best knowledge, this paper is the first study to use artificial neural networks-based method for retinal vessel segmentation.

Benefit from the emergence of large-scale data, the rapid development of computing power and the presentation of various artificial neural network models, deep learning-based methods boost dramatically in the field of computer vision [32], [33], [34], [35]. Accompanied by these developments, Convolutional Neural Networks (CNN) has become an effective approach for analysing medical image [36], [37], [38], solving medical image segmentation problems [39], [40], [41] and can reach the state-of-the-art performance in the task of retinal vessel segmentation [42], [43], [44]. In this part, we will review several CNNs based retinal vessel segmentation methods. Wang et al. [45] proposed a segmentation method which uses CNN as a feature extractor and random forests as the final classifier. Wu et al. [46] first use the CNN to extract binary mask and then use a generalized particle filtering technique to extract retinal vessel tree under a probabilistic tracking framework. Later, Fu et al. [47] developed a multi-scale and multi-level CNN model to do the segmentation and used a Conditional Random Field (CRF) to consider the long-range interactions between pixels. Dasgupta et al. [48] utilized the fully convolutional neural networks and the multi-label inference to do structure predictions of the blood vessel. Son et al. [49] presented a method that adopts the generative adversarial training to improve the segmentation performance. Recently, Zhang et al. [50] proposed an architecture to sufficient use multi-level features and added atrous convolution to get effective multi-scale features for retinal blood vessel segmentation task. Oliveira et al. [44] combined the multiscale analysis provided by the Stationary Wavelet Transform with a multiscale Fully Convolutional Neural Network to cope with the varying width and direction of the vessel structure in the retina. Memari et al. [43] utilized a genetic algorithm enhanced spatial fuzzy c-means method for extracting an initial blood vessel network, with the segmentation further refined by using an integrated level set approach.

In a supervised method, the classification criteria are determined by the ground truth data based on given features. However, data annotated by professional doctors is usually expansive and difficult to obtain. As supervised methods are designed based on pre-classified data, their performance is usually better than that of unsupervised ones and can produce pretty good results for healthy retinal images. Although supervised methods, especially CNN-based methods have achieved satisfactory segmentation results in many scenarios, there are still many challenging issues of dealing with small vessels, poor illumination, etc.

3 Proposed Method

In this section, we describe the proposed **G**lobal and **L**ocal enhanced residual **U**-n**E**t (*GLUE*) model for tackling the retinal vessel segmentation task in detail, and the overall pipeline is shown in Fig. 1. The proposed model is a cascaded CNN model which consists of 2 parts, a WUN and a WRUN. The WUN takes the globally enhanced patches as input to produce a coarse segmentation, and the WRUN takes the locally enhanced patches along with previous segmentation results as input to do a refinement. The remainder of this section is organized as follows. Section 3.1 introduces the retinal image pre-processing steps used in our approach. Section 3.2 describes the detailed architecture of our proposed GLUE model. Finally, the scheme of overlapped prediction patches regrouping is discussed in Section 3.3.

3.1 Retinal Image Pre-Processing

Globally and Locally Contrast Enhancement. As can be seen from Fig. 2, the original retinal image is with extremely low contrast. Suitable pre-processing steps can increase the contrast which alleviate the learning difficulty of CNNs and get better performance. In the following, we introduce the pre-processing steps that we applied in this study. For each retinal image,

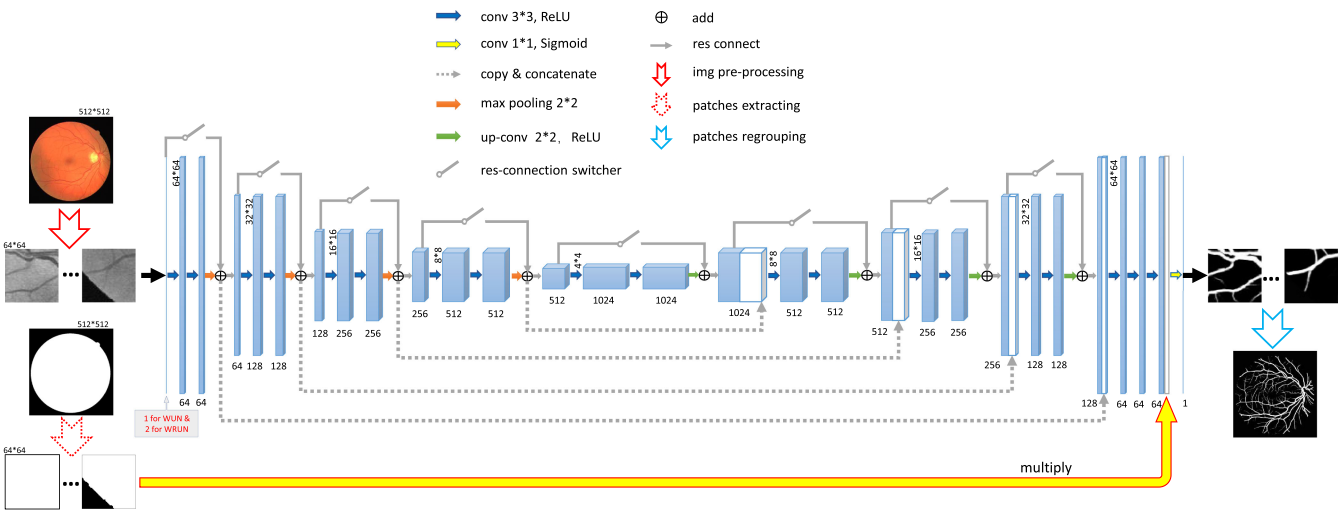

Fig. 3. The brief architecture of proposed GLUE model. The meaning of all types of signs is marked on the top of the picture. In particular, when all the res-connection switchers are disconnected, the model indicates WUN, which acts as the first part of our approach. When all the res-connection switchers are connected, the model indicates WRUN, which serve as the second part of our approach. Due to the different inputs as introduced in Section 4.3, the number of input channel for the first part and second part is 1 and 2, respectively, which is marked in red color in this figure.

we resize it to $512 * 512$ and convert it into grayscale. To increase the retinal images' contrast, the contrast limited adaptive histogram equalization (CLAHE) [51] operation is applied on retinal images. Here, we apply CLAHE operation both globally on entire retinal image, and locally on extracted retinal image patch.

Global CLAHE processing (GCP) and local CLAHE processing (LCP) are complementary with each other. Applying CLAHE operation on the entire image can maintain global information and avoid the degeneration effects of local noise like light reflection and diseased area, but GCP does not work well for some extremely low contrast areas. On the other hand, performing CLAHE operation in a local area can boost its local contrast information, especially in local areas with very thin capillary blood vessels.

As shown in Fig. 2, no matter applying CLAHE operation globally or locally, all the patches' contrasts are enhanced. The two schemes have different advantages. For patch ① and ② where the local contrasts are rather low, LCP performs obviously better than GCP, and the vessels are with more contrast enhancement. While for patch ③ where light reflection or other local anomaly occurs, GCP is more robust than LCP. As such, we propose to use a cascade deep model to fully utilize the GCP and LCP as shown in Fig. 1. For the first stage, we apply CLAHE operation on entire images, then randomly crop globally enhanced image patches for training. For the second refining stage, we first extract images to patches and then apply a local CLAHE operation to increase the local contrast. By doing so, the proposed model can learn more discriminative features by considering the retinal images' global and local statistical information.

3.2 The Architecture of the Proposed GLUE Model

In order to take full utilize of the global and local CLAHE operation, as well as model the dependencies among adjacent pixels and thus enforce complex bifurcated structure of the retinal vessels, we implemented a cascaded CNN model. The architecture of our proposed GLUE model is shown in Fig. 1 which contains 2 parts, a WUN for coarse retinal vessel segmentation and a WRUN to refine the segmentation results.

The overall architecture of the proposed WUN and WRUN is shown in Fig. 3. In particular, when all the res-connection switchers (marked as gray switch symbols in the picture) are disconnected, the model indicates WUN, which serves as the first part of our approach. While when all the res-connection switchers are connected, the model indicates WRUN, which serve as the second part of our approach. Similar to the original U-Net model [52], both the WUN and WRUN have an encoder-decoder architecture. While beyond such architecture, we made several significant improvements by adding a weighted attention mechanism on both parts, and the skip connection scheme on the second part as introduced in [34]. All the parameters of the model are given in Fig. 3. In particular, due to the different inputs, the number of input channel for the first part and second part is 1 and 2, respectively. The first part of our model (WUN) contains 23.49 M parameters, and the second part of our model (WRUN) has 32.43 M parameters.

3.2.1 The Weighted Attention Mechanism

The retinal images in DRIVE [27] and STARE [53] dataset are with circular-like region of interest (ROI) and dark background. In order to model this geometric constraint, we add an attention scheme, like our former iris segmentation study [54]. A circular template ROI mask $\mathcal{M}$ is generated to estimate the potential area where the fundus is most likely to appear. And the generated attention mask $\mathcal{M}$ will be used as a weighted attention scheme, which is shown in Fig. 3 in the bottom as large yellow arrow. The attention mechanism is implemented by multiplying the model's second last layer's feature maps with an attention mask. This operation can be expressed as

$$\mathcal{I}(x,y) = \begin{cases} \mathcal{F}(x,y) * 1.0 & (x,y) \in \mathcal{M} \\ \mathcal{F}(x,y) * 0.0 & (x,y) \notin \mathcal{M} \end{cases}, \tag{1}$$

in which $\mathcal{F}(x,y)$ represents the features at position (x,y). Due to the generate attention masks are very accurate, so we simply set the attention weight to be 1.0. But this does not mean that our attention mechanism is invalid. When the retinal

TABLE 1
The Details of the Two Datasets Used for Evaluation

Dataset	DRIVE	STARE
Sensor	Canon CR5 non-mydriatic 3CCD camera	TopCon TRV-50 fundus camera
Resolution	565*584	700*605
Format	tif	ppm
Color	RGB	RGB
# of training	20	10
# of testing	20	10

images are extracted into patches, the attention mechanism can help our model better locate the target area in each patch.

By using this weighted attention mechanism, our model will only pay attention at the target ROI area and discard the irrelevant noisy background. For the DRIVE dataset, we directly use the provided mask of fundus area as the weighted attention mask. While for STARE dataset, we computed the attention mask by a simple processing step which we convert the retinal image into grayscale, apply Gaussian filtering and then separate the fundus regions by doing binary thresholding at the value 40. On the DRIVE dataset, our mask extraction method can reach an accuracy of 99.2 percent. For more general cases, we may not get a perfect mask from noisy background, as such we can get the attention mask by using soft weight such as 0.1 instead of directly ignore them.

As indicated in Fig. 3, both the proposed WUN (first part of our approach) and WRUN (second part of our approach) adopt this weighted attention mechanism for better modeling the geometric constraint of fundus.

3.2.2 The Skip Connection Scheme

As demonstrated in [34], adding skip connection can increase the depth and improve the accuracy of deep CNNs. Inspired by this work, we also added skip connections into the second part of our model as indicated by the solid gray arrow in Fig. 3. For each convolutional block in WRUN, the skip connection scheme is implemented by using the following equation

$$\boldsymbol{y} = F(\boldsymbol{x}, \{\boldsymbol{w}_i\}) + H(\boldsymbol{x}), \tag{2}$$

where F consists of two convolution operations and one max-pooling or one up-sampling operation, H is either the identical mapping or a convolution operation to keep the input with the same feature dimensions as F.

Notice that, as introduced in Section 3.2, the first part of our model is the WUN without residual connection. In this way, in coarse segmentation step, our model can utilize the weight of network pre-trained on ImageNet [55].

3.2.3 Loss Function

In order to train the proposed model, we choose the binary cross entropy as segmentation loss function for both the WUN and the WRUN, which goes as

$$\mathcal{L}(p, q) = -\frac{1}{n}\sum_{k=1}^{n} q_k \log p_k + (1 - q_k) \log (1 - p_k), \tag{3}$$

where n represents the total number of training pixels, p and q represent predicted probability and its corresponding groundtruth *(0 for background pixels and 1 for blood vessel pixels)*.

3.3 Patch Regrouping

During the testing phase, instead of doing a random overlapping patch cropping as in the training phase, we extract 64×64 image patches in sequence with an overlapping stride of 8. For each 512×512 testing retinal image, the number of overall extracted patches is $((512 - 64)/8 + 1) * ((512 - 64)/8 + 1) = 3249$. After getting the prediction of each patch, we can easily get the final segmentation result of the whole retinal image by regrouping the results from all the patches according to their positions. Specifically, for all the overlapped pixels, our computed the probability map by averaging the values of all the overlapped pixels' predictions.

4 Experiments

4.1 Datasets

We evaluate the performance of our method on two publicly available benchmark datasets: DRIVE [27] and STARE [53]. Both the two datasets are representative dataset in the task of retinal vessel segmentation. The details of the two datasets are listed in Table 1.

The *DRIVE*[1] dataset contains 20 training RGB images and 20 testing RGB images with the resolution of 768×584 pixels. There are 7 images shown signs of mild early diabetic retinopathy in this dataset.

The *STARE*[2] dataset consists of 20 retinal fundus slides captured by a TopCon TRV-50 fundus camera. Half of the dataset comprises images of healthy subjects, and the rest contains the pathological cases which make the segmentation more challenging.

For both the DRIVE and STARE dataset, there are two groups of manual segmentation masks available annotated by two independent human experts. The manual annotations of the first expert were used as the ground truth for training and evaluation.

4.2 Evaluation Metrics

Accuracy is a widely used evaluation metric for the task of binary segmentation, which computes the percentage of correctly classified pixels in the whole image. Eq. (4) is used for calculating the Accuracy on test set

$$Acc = \frac{TP + TN}{TP + FN + TN + FP}, \tag{4}$$

where TP, TN, FP and FN represent the number of true positive, true negatives, false positives and false negatives, respectively.

Sensitivity (also referred as recall and true positive rate) is another commonly used statistical measures of the performance of binary segmentation task. It measures the proportion of actual positives that are correctly classified as such. The equation goes as

1. https://www.isi.uu.nl/Research/Databases/DRIVE/
2. http://cecas.clemson.edu/ ahoover/stare/

TABLE 2
Accuracy, Sensitivity, Specificity, and Precision Comparisons on Proposed GLUE Model and Other Promising Approaches

Method	Method	Year	DRIVE				STARE			
			Acc	Sen	Spec	Prec	Acc	Sen	Spec	Prec
2nd human expert		-	0.9725	0.7760	0.9725	0.8123	0.9346	0.8956	0.9381	0.6361
Unsupervised methods										
Yin et al. [14]	BMP	2013	0.9267	0.6522	0.9710	-	0.9420	0.7034	0.9668	-
Wang et al. [18]	MF + AT	2013	0.9461	-	-	-	0.9682	-	-	-
Fraz et al. [20]	MTHT + BPS + RG	2013	0.9422	0.7302	0.9742	0.8112	0.9423	0.7318	0.9660	0.7294
Odstrcilik et al. [56]	MF + MET	2013	0.9340	0.7060	0.9693	-	0.9341	0.7847	0.9512	-
Emary et al. [57]	FCM + CS	2014	0.9376	0.6316	0.9838	-	0.9448	0.5864	0.9871	-
Imani et al. [21]	MCA	2015	0.9523	0.7524	0.9753	-	0.9590	0.7502	0.9745	-
Mapayi et al. [25]	AT + GLCM	2015	0.9461	0.7632	0.9634	-	0.9510	0.7626	0.9657	-
Kumar et al. [58]	MF+LoG+CLAHE	2016	0.9626	0.7006	-	-	0.9637	0.7675	-	-
Neto et al. [59]	GS + MTHT	2017	-	0.7942	0.9631	-	-	0.7695	0.9537	-
Khan et al. [60]	GLP + MPT	2017	0.9600	0.7470	0.980	-	0.9510	0.7780	0.9660	-
Khan et al. [61]	LD + HT	2018	0.9506	0.7696	0.9651	-	0.9513	0.7521	0.9812	-
Memari et al. [43]	MF + FCM	2018	0.9610	0.7610	0.9810	-	0.9510	0.7820	0.9650	-
Supervised methods										
Ricci et al. [28]	LD + SVM	2007	0.9595	-	-	-	0.9646	-	-	-
Fraz et al. [62]	GOA+MT+GFR+DT	2012	0.9480	0.7406	0.9807	0.8532	0.9534	0.7548	0.9763	0.7956
Wang et al. [45]	CNN + RF	2015	0.9767	0.8173	0.9733	-	**0.9813**	0.8104	0.9791	-
Roychowdhury et al. [63]	8 HCFs + GMM	2015	0.9519	0.7249	0.9830	-	0.9515	0.7719	0.9726	-
Liskowski et al. [64]	CNN	2016	0.9495	0.7763	0.9768	-	0.9566	0.7867	0.9754	-
Fu et al. [47]	CNN + CRF	2016	0.9523	0.7603	-	-	0.9585	0.7412	-	-
Dasgupta et al. [48]	FCN	2017	0.9533	0.7691	0.9801	0.8498	-	-	-	-
Orlando et al. [42]	CRF + SOSVM	2017	0.9454	0.7897	0.9684	0.7854	0.9571	0.7773	0.9789	0.7740
Oliveira et al. [44]	SWT + FCN	2018	**0.9821**	0.8039	0.9804	-	0.9694	0.8315	0.9858	-
U-Net		-	0.9594	0.7698	0.9798	0.8210	0.9669	0.7536	0.9827	0.8662
Ours coarse (from 1st part WUN)		-	0.9625	0.7536	0.9737	0.8392	0.9681	0.7461	0.9825	0.8620
GLUE (w/o CLAHE)		-	0.9605	0.8029	0.9746	0.8412	0.9623	0.8210	0.9802	0.8723
GLUE in LF		-	0.9668	0.8170	0.9826	0.8619	0.9703	0.8102	0.9815	0.8717
Ours GLUE		-	0.9692	**0.8278**	**0.9861**	**0.8637**	0.9740	**0.8342**	**0.9916**	**0.8823**

Herewe list the main methods used in each paper. The acronyms for the algorithms stand for: Bayesian method with a maximized posterior (BMP), matched filtering (MF), adaptive thresholding (AT), morphological top-hat transform (MTHT), bit plane slicing (BPS), region growing (RG), minimum error thresholding (MET), fuzzy c-means (FCM), cuckoo search (CS), morphological component analysis (MCA), adaptive thresholding (AT), gray level co-occurence matrix (GLCM), Laplacian of Gaussian (LoG), contrast limited adaptive histogram equalization (CLAHE), Gaussian smoothing (GS), generalized linear model (GLM), moment-preserving thresholding (MPT), line detector (LD), hysteresis thresholding (HT), support vector machine (SVM), gradient orientation analysis (GOA), morphological transformation (MT), Gabor filter response (GFR), decision trees (DT), convolutional neural networks (CNN), random forest (RF), hand-craft features (HCF), Gaussian mixture model (GMM), conditional random fields (CRF), fully convolutional networks (FCN), structured output support vector machine (SOSVM), stationary wavelet transform (SWT), and late fusion (LF).

$$Sen = \frac{TP}{TP + FN}, \tag{5}$$

Specificity (also called as true negative rate) measures the proportion of actual negatives that are correctly identified as such. The equation for computing the Specificity is

$$Spec = \frac{TN}{TN + FP}, \tag{6}$$

Precision (also known as positive predictive value (PPV)) quantifies the ratio of pixels classified as retinal vessel that are correctly identified. The equation goes as

$$Prec = \frac{TP}{TP + FP}, \tag{7}$$

We apply these four metrics on each testing image and then report the final average value on the testing set for comparison.

4.3 Implementation

We implement our model in by using the Keras with TensorFlow backend. The outputs of the first part concatenate with LCP patches directly, and together with attention masks, will act as input to the second part. The Adam optimizer is adopted to train our model with an initial learning rate of $5 * 10^{-5}$. The contracting paths in both parts of our proposed model have the same structure as VGG16 [33] (without fully connected layers). We then initialize the weights of the first part (WUN) by using the pre-trained weights on ImageNet provided by Keras.

The proposed model is trained and tested following the pipeline discussed in Section 3 on a machine with Intel i7-7700K CPU and an NVIDIA 1080Ti GPU. For each training image, we randomly extracted 500 overlapped $64 * 64$ patches to construct the training set. We adopt the widely used data augmentation operations in each extracted patch, such as horizontal flip, width shift range, etc. We first train the parameters of the first WUN for 30 epochs, and then train the second WRUN for another 30 epochs by fixing the weights of the first part. The training batch size for both models is 128.

4.4 Experiment Results

4.4.1 Quantitative Evaluation

In this section, we evaluate the segmentation performance of the proposed GLUE model. The detailed evaluation protocols are as follows.

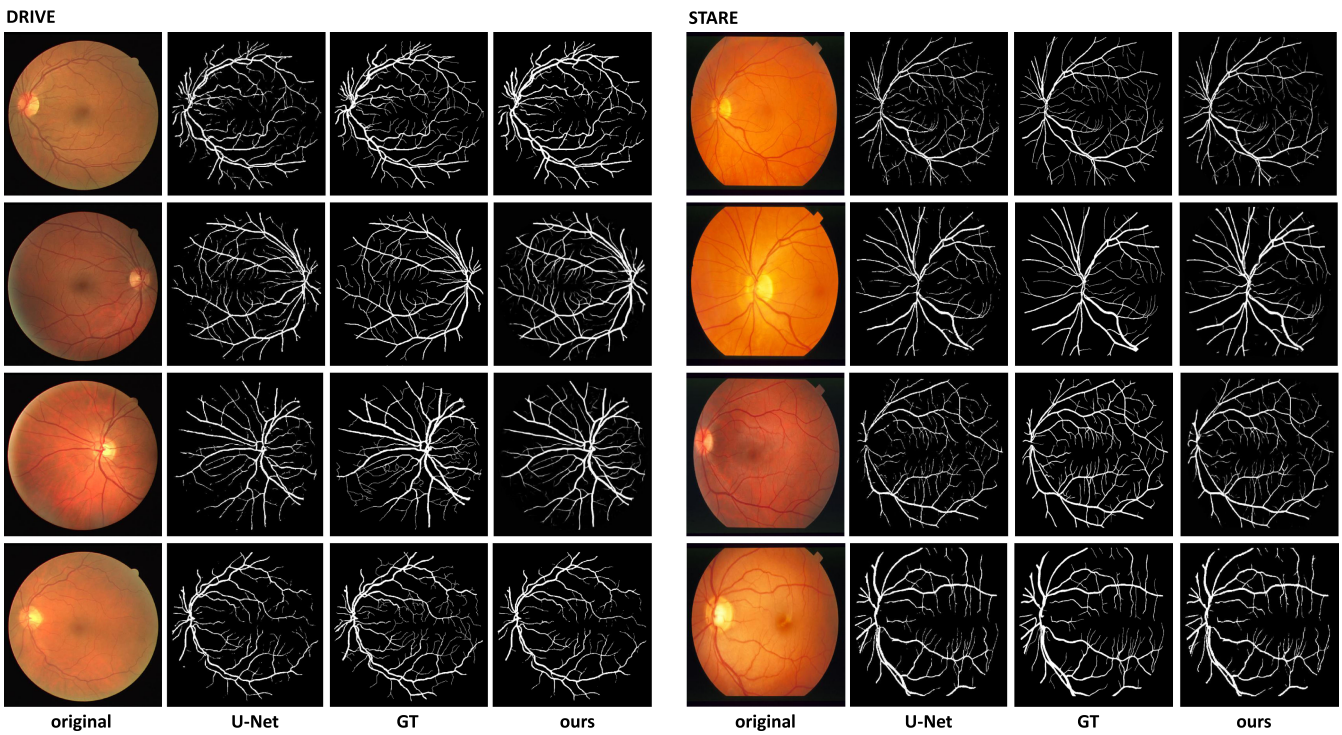

Fig. 4. Examples of retinal vessels segmentation result from DRIVE dataset (listed in the left) and STARE dataset (listed in the right). Results show that our model can deal well with the problems of low illumination, hard area of the optic disk. Moreover, our model shows better blood vessel connectivity than U-Net. Detailed results in zoom-in view of extracted patches are displayed in Fig. 5.

- We implement an U-Net model on both datasets as the baseline model. Both our proposed model and the baseline U-Net model adopt the backbone of VGG-16, and initialize with the weights pre-trained on ImageNet.
- To evaluate the influence of applying local and global CLAHE pre-processing operation, we trained a GLUE model using the original gray-scale retinal image as input.
- Following the idea of late fusion [65], [66], we design a paralle two-path GLUE model in late fusion manner. Instead of using a cascaded structure, the final conv layer's (conv9-3) output of both WUN and WRUN are concatenated, followed by a conv layer to fuse the their outputs.
- Furthermore, we compared with other representative state-of-the-art methods, including unsupervised methods and supervised methods.

The detailed evaluation metrics of our method and other representative methods are listed in Table 2. From the table, we find that our proposed GLUE model can significantly surpass the baseline U-Net model on accuracy, sensitivity, specificity and precision performance. The results of the proposed approach is close to manual segmentations provided by the second human observer. Comparing the GLUE with and without CLAHE operation, we can find that using the CLAHE operation can significantly boost the performance. Besides, unlike late fusion methods that fuse the predictions from multiple features, we leverage the merit of two different preprocessing operations and combine them by using a cascaded refinement, and the results show that our cascaded structure has a considerable improvement over late fusion structure. Also, the proposed model is in the leading position on each metric compared to other methods, except that [44] and [18] have higher accuracy on DRIVE and STARE, respectively.

Specifically, the comparison between the control groups in the bottom four rows illustrates the effectiveness of our global and local CLAHE operations, and the effectiveness of GLUE's cascade refinement structure. One thing we need to note is that many approaches (including ours) surpass the manual segmentation results by the second observer on some metrics, that is because retinal vessel segmentation is a hard task and the procedure of manual segmentation can be considered highly subjective (examples can be seen in Fig. 5).

4.4.2 Qualitative Segmentation Results

In Fig. 4, we plot some segmentation results produced by the proposed model and the baseline U-Net model. Results show that our model can cope well with the difficulties, including low illumination, hard area of the optic disk, etc. Besides, results from our proposed model show better vessel connectivity and small vessel sensitivity than the ones from U-Net. For better displaying the details of segmentation results, we also plot extracted patches' segmentation results in zoom-in view in Fig. 5. The selected examples indicate that compared to U-Net, our model can accurately segment tiny and indistinct vessels, and maintain the geometric connection of retinal vessels. Specificity, as is indicated by red arrows in No.④ result of STARE, our model can overcome the effects of lesions and abnormal areas in retinal images. As mentioned above, retinal vessel segmentation is a challenging task and can be considered highly subjective. Due to the low contrast, the classification of many areas is ambiguous. In fact, the misclassified area marked out by green arrows (in No.②⑥ in DRIVE and No.⑤ in STARE, which regarded as false positive) should actually be true positive. This shows that after a

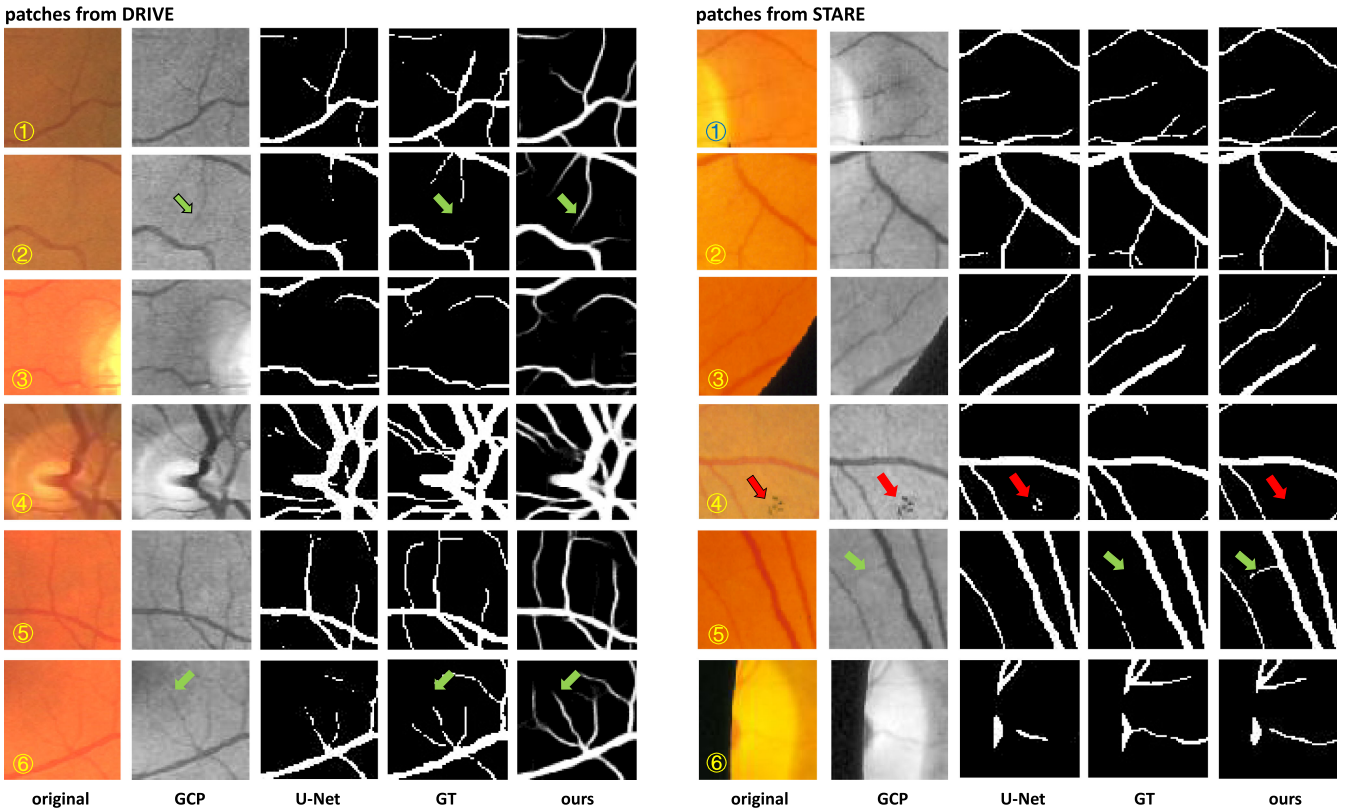

Fig. 5. Examples of extracted patches' segmentation result from DRIVE dataset (listed in the left) and STARE dataset (listed in the right). The results show that compared with U-Net, our model is able to find small blood vessels and maintain better vessel connectivity.

series of effective processing and training, in some cases, our model has a higher sensitivity for small and indistinguishable vessels than manual segmentation.

4.4.3 *Efficiency Analysis*

At last, we compute the time cost of our model during the training and testing. It takes about 91 and 65 minutes for training our GLUE model on the DRIVE and the STARE dataset respectively. For the testing phase, it only need 6.2 second for segmenting 20 retinal images (about 0.31 s / image).

5 Conclusion

High-performance retinal vessel segmentation is a key step for medical diagnosing, including chronic eye disease, cardiovascular disease and diabetic retinopathy, etc. While problems such as poor illumination, missing of small vessels, complex vessel geometry, make retinal vessel segmentation a challenging task. Former learning-based approaches generally apply an image-level pre-processing operation on whole retina image for segmentation model learning. However, retina region's appearance is a unitary area, and pixels outside such area is of no use for segmentation. Moreover, as introduced in Section 3.1, applying CLAHE operation globally and locally are complementary schemes with each other. So, in this paper, we proposed the GLUE model for addressing the challenging retinal vessel segmentation problem. Our model contains a WUN for coarse segmentation and a WRUN for segmentation refinement. By applying CLAHE operation globally and locally, our model can benefit from both the GCP and LCP patches' information. And by adding attention scheme and residual connection scheme, and applying cascaded refinement structure, our model can focus only on retinal region in each patch, avoid false segmentation caused by irrelevant noisy background, learn more discriminative features and have a better maintaining of the retinal vessel tree structure.

We adopt accuracy, sensitivity, specificity, and precision for measuring the accuracy of predicted segmentation results. We evaluate our method on the widely used DRIVE and the STARE benchmark dataset, and compare it with U-Net model as baseline model, and many other representative methods. Experimental results demonstrate that our model can achieve high-performance segmentation results. Also, the performance comparison of GLUE with GLUE's coarse segmentation, GLUE in late fusion manner, and GLUE without CLAHE operation shows the validity of our structure. Detailed segmentation results in the zoom-in view show that our model can not only accurately segment tiny and indistinct vessels, but also maintain the geometric connection of retinal vessels. Also, our model can deal well with abnormal and noisy regions like lesion areas, edge areas and optic disk areas. Since the procedure of manual segmentation can be considered highly subjective, we expect larger datasets and more accurate manual annotation for further improving the performance of our approach.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61876159, No. 61806172, No. 61572409, No. U1705286 & 61571188), the China Postdoctoral Science

Foundation Grant (No. 2019M652257), the Fujian Province 2011 Collaborative Innovation Center of TCM Health Management, the Collaborative Innovation Center of Chinese Oolong Tea Industry-Collaborative Innovation Center (2011) of Fujian Province, the Fund for "Integration of Cloud Computing and Big Data, Innovation of Science and Education (No. 2017A11032), and the Open Project Program of the Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian Education Institutions (Wuyi University).

REFERENCES

[1] J. W. Y. Yau, S. L. Rogers, R. Kawasaki, E. L. Lamoureux, J. W. Kowalski, T. Bek, S.-J. Chen, J. M. Dekker, A. Fletcher, J. Grauslund, et al., "Global prevalence and major risk factors of diabetic retinopathy," *Diabetes Care*, vol. 35, 2012, Art. no. DC_111909.

[2] W. L. Wong, X. Su, X. Li, C. M. G. Cheung, R. Klein, C.-Y. Cheng, and T. Y. Wong, "Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis," *Lancet Global Health*, vol. 2, no. 2, pp. e106–e116, 2014.

[3] Y.-C. Tham, X. Li, T. Y. Wong, H. A. Quigley, T. Aung, and C.-Y. Cheng, "Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis," *Ophthalmology*, vol. 121, no. 11, pp. 2081–2090, 2014.

[4] S. H. Hardarson and E. Stefánsson, "Retinal oxygen saturation is altered in diabetic retinopathy," *Brit. J. Ophthalmology*, vol. 96, no. 4, pp. 560–563, 2012.

[5] T. Y. Wong and R. McIntosh, "Hypertensive retinopathy signs as risk indicators of cardiovascular morbidity and mortality," *Brit. Med. Bulletin*, vol. 73, no. 1, pp. 57–70, 2005.

[6] A. V. Stanton, B. Wasan, A. Cerutti, S. Ford, R. Marsh, P. P. Sever, S. A. Thom, and A. D. Hughes, "Vascular network changes in the retina with age and hypertension," *J. Hypertension*, vol. 13, no. 12 Pt 2, pp. 1724–1728, 1995.

[7] B. Bowling, *Kanski's Clinical Ophthalmology E-Book: A Systematic Approach*. Amsterdam, The Netherlands: Elsevier, 2015.

[8] D. Maji, A. Santara, P. Mitra, and D. Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," *arXiv:1603.04833*, 2016.

[9] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 483–499.

[10] X. Dong, Y. Yan, W. Ouyang, and Y. Yang, "Style aggregated network for facial landmark detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 379–388.

[11] X. Xiao, S. Lian, Z. Luo, and S. Li, "Weighted Res-UNet for high-quality retina vessel segmentation," in *Proc. IEEE Int. Conf. Inf. Technol. Med. Edu.*, 2018, pp. 327–331.

[12] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "Blood vessel segmentation methodologies in retinal images–A survey," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 407–433, 2012.

[13] C. L. Srinidhi, P. Aparna, and J. Rajan, "Recent advancements in retinal vessel segmentation," *J. Med. Syst.*, vol. 41, no. 4, 2017, Art. no. 70.

[14] Y. Yin, M. Adel, and S. Bourennane, "Automatic segmentation and measurement of vasculature in retinal fundus images using probabilistic formulation," *Comput. Math. Methods Med.*, vol. 2013, 2013, Art. no. 260410.

[15] J. Zhang, H. Li, Q. Nie, and L. Cheng, "A retinal vessel boundary tracking method based on Bayesian theory and multi-scale line detection," *Comput. Med. Imag. Graph.*, vol. 38, no. 6, pp. 517–525, 2014.

[16] J. De, H. Li, and L. Cheng, "Tracing retinal vessel trees by transductive inference," *BMC Bioinf.*, vol. 15, no. 1, 2014, Art. no. 20.

[17] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–269, Sep. 1989.

[18] Y. Wang, G. Ji, P. Lin, and E. Trucco, "Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition," *Pattern Recognit.*, vol. 46, no. 8, pp. 2117–2133, 2013.

[19] M. Krause, R. M. Alles, B. Burgeth, and J. Weickert, "Fast retinal vessel analysis," *J. Real-Time Image Process.*, vol. 11, no. 2, pp. 413–422, 2016.

[20] M. M. Fraz, A. Basit, and S. A. Barman, "Application of morphological bit planes in retinal blood vessel extraction," *J. Digit. Imag.*, vol. 26, no. 2, pp. 274–286, 2013.

[21] E. Imani, M. Javidi, and H.-R. Pourreza, "Improvement of retinal blood vessel detection using morphological component analysis," *Comput. Methods Programs Biomed.*, vol. 118, no. 3, pp. 263–279, 2015.

[22] E. M. Sigursson, S. Valero, J. A. Benediktsson, J. Chanussot, H. Talbot, and E. Stefánsson, "Automatic retinal vessel extraction based on directional mathematical morphology and fuzzy classification," *Pattern Recognit. Lett.*, vol. 47, pp. 164–171, 2014.

[23] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 1, pp. 131–137, Jan. 2003.

[24] M. U. Akram and S. A. Khan, "Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy," *Eng. Comput.*, vol. 29, no. 2, pp. 165–173, 2013.

[25] T. Mapayi, S. Viriri, and J.-R. Tapamo, "Adaptive thresholding technique for retinal vessel segmentation based on GLCM-energy information," *Comput. Math. Methods Med.*, vol. 2015, 2015, Art. no. 597475.

[26] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M. D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *Proc. SPIE Med. Imag.: Image Process.*, 2004, pp. 648–657.

[27] 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 Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.

[28] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1357–1365, Oct. 2007.

[29] L. Xu and S. Luo, "A novel method for blood vessel detection from retinal images," *Biomed. Eng. Online*, vol. 9, no. 1, 2010, Art. no. 14.

[30] R. Nekovei and Y. Sun, "Back-propagation network and its configuration for blood vessel detection in angiograms," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 64–72, Jan. 1995.

[31] 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," *Brit. J. Ophthalmology*, vol. 83, no. 8, pp. 902–910, 1999.

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

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

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[35] Z. Zhong, L. Zheng, Z. Zheng, S. Li, and Y. Yang, "CamStyle: A novel data augmentation method for person re-identification," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1176–1190, Mar. 2019.

[36] Y. Feng, F. Yang, X. Zhou, Y. Guo, F. Tang, F. Ren, J. Guo, and S. Ji, "A deep learning approach for targeted contrast-enhanced ultrasound based prostate cancer detection," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, to be published, doi: 10.1109/TCBB.2018.2835444.

[37] R. Ju, C. Hu, Q. Li, et al., "Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 16, no. 1, pp. 244–257, Jan./Feb. 2019.

[38] Y. Feng, L. Zhang, and J. Mo, "Deep manifold preserving autoencoder for classifying breast cancer histopathological images," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, to be published, doi: 10.1109/TCBB.2018.2858763.

[39] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, 2017.

[40] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Med. Image Anal.*, vol. 30, pp. 108–119, 2016.

[41] C. Zotti, Z. Luo, A. Lalande, and P.-M. Jodoin, "Convolutional neural network with shape prior applied to cardiac MRI segmentation," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 3, pp. 1119–1128, May 2019.

[42] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 16–27, Jan. 2017.
[43] N. Memari, A. R. Ramli, M. I. Bin Saripan, S. Mashohor, and M. Moghbel, "Retinal blood vessel segmentation by using matched filtering and fuzzy C-means clustering with integrated level set method for diabetic retinopathy assessment," *J. Med. Biological Eng.*, pp. 1–19, 2018.
[44] A. F. M. Oliveira, S. R. M. Pereira, and C. A. B. Silva, "Retinal vessel segmentation based on fully convolutional neural networks," *Expert Syst. Appl.*, vol. 112, pp. 229–242, 2018.
[45] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *Neurocomput.*, vol. 149, pp. 708–717, 2015.
[46] A. Wu, Z. Xu, M. Gao, M. Buty, and D. J. Mollura, "Deep vessel tracking: A generalized probabilistic approach via deep learning," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2016, pp. 1363–1367.
[47] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "DeepVessel: Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2016, pp. 132–139.
[48] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2017, pp. 248–251.
[49] J. Son, S. J. Park, and K.-H. Jung, "Retinal vessel segmentation in fundoscopic images with generative adversarial networks," *arXiv:1706.09318*, 2017.
[50] B. Zhang, S. Huang, and S. Hu, "Multi-scale neural networks for retinal blood vessels segmentation," *arXiv:1804.04206*, 2018.
[51] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*. San Diego, CA, USA: Academic, 1994, pp. 474–485.
[52] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241.
[53] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.
[54] S. Lian, Z. Luo, Z. Zhong, X. Lin, S. Su, and S. Li, "Attention guided U-Net for accurate iris segmentation," *J. Visual Commun. Image Representation*, vol. 56, pp. 296–304, 2018.
[55] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 248–255.
[56] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "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.
[57] E. Emary, H. M. Zawbaa, A. E. Hassanien, G. Schaefer, and A. T. Azar, "Retinal vessel segmentation based on possibilistic fuzzy c-means clustering optimised with cuckoo search," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2014, pp. 1792–1796.
[58] D. Kumar, A. Pramanik, S. S. Kar, and S. P. Maity, "Retinal blood vessel segmentation using matched filter and Laplacian of Gaussian," in *Proc. IEEE Int. Conf. Signal Process. Commun.*, 2016, pp. 1–5.
[59] L. C. Neto, G. L. B. Ramalho, J. F. S. R. Neto, R. M. S. Veras, and F. N. S. Medeiros, "An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images," *Expert Syst. Appl.*, vol. 78, pp. 182–192, 2017.
[60] K. B. Khan, A. A. Khaliq, and M. Shahid, "A novel fast GLM approach for retinal vascular segmentation and denoising," *J. Inf. Sci. Eng.*, vol. 33, no. 6, pp. 1611–1627, 2017.
[61] M. A. U. Khan, T. M. Khan, D. G. Bailey, and T. A. Soomro, "A generalized multi-scale line-detection method to boost retinal vessel segmentation sensitivity," *Pattern Anal. Appl.*, pp. 1–20, 2018.
[62] 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 Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, Sep. 2012.
[63] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Blood vessel segmentation of fundus images by major vessel extraction and subimage classification," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 3, pp. 1118–1128, May 2015.
[64] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
[65] G. Ye, D. Liu, I.-H. Jhuo, and S.-F. Chang, "Robust late fusion with rank minimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 3021–3028.
[66] X. Dong, Y. Yan, M. Tan, Y. Yang, and I. W. Tsang, "Late fusion via subspace search with consistency preservation," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 518–528, Jan. 2019.

Sheng Lian received the bachelor's degree in computer science from the Huazhong University of Science and Technology, China. He is currently working toward the PHD degree at Xiamen University, China. His research interests include medical image analysis, computer vision, and machine learning.

Lei Li is currently working toward the master's degree at Xiamen University, China. His research interests include object detection, medical image analysis, computer vision, and marching learning.

Guiren Lian is currently serving as the associate professor of the College of Mathematics and Informatics, Fujian Normal University, China. His research interests include medical imaging and computer vision.

Xiao Xiao is currently working toward the master's degree at Xiamen University, China. Her research interests include computer vision and machine learning.

Zhiming Luo received the PhD degree in computer science from Xiamen University, China, and the University of Sherbrooke, Canada. He is currently a postdoctoral researcher with Xiamen University, China. His research interests include traffic surveillance video analytics, medical image analysis, computer vision, and machine learning.

Shaozi Li received the BS degree from Hunan University, the MS degree from Xi'an Jiaotong University, and the PhD degree from the National University of Defense Technology. He currently serves as the chair and professor of the Cognitive Science Department, Xiamen University, the vice director of Technical Committee on Collaborative Computing of CCF, and the vice director of the Fujian Association of Artificial Intelligence. His research interests cover artificial intelligence and its applications, moving objects detection and recognition, machine learning, computer vision, multimedia information retrieval, etc. He has directed and completed more than 20 research projects, including several National 863 Programs, National Nature Science Foundation of China, and PhD Programs Foundation of Ministry of Education of China. He is a senior member of the IEEE, ACM, and China Computer Federation (CCF).