



Dual-path multi-scale context dense aggregation network for retinal vessel segmentation

Wei Zhou^a, Weiqi Bai^a, Jianhang Ji^a, Yugen Yi^{b,*}, Ningyi Zhang^b, Wei Cui^{c,**}

^a College of Computer Science, Shenyang Aerospace University, Shenyang, China

^b School of Software, Jiangxi Normal University, Nanchang, China

^c Institute for Infocomm Research, The Agency for Science, Technology and Research (A*STAR), Singapore

ARTICLE INFO

Keywords:

Fundus image
Vessel segmentation
Context information
Multi-scale fusion
Dual-path fusion

ABSTRACT

There has been steady progress in the field of deep learning-based blood vessel segmentation. However, several challenging issues still continue to limit its progress, including inadequate sample sizes, the neglect of contextual information, and the loss of microvascular details. To address these limitations, we propose a dual-path deep learning framework for blood vessel segmentation. In our framework, the fundus images are divided into concentric patches with different scales to alleviate the overfitting problem. Then, a Multi-scale Context Dense Aggregation Network (MCDAU-Net) is proposed to accurately extract the blood vessel boundaries from these patches. In MCDAU-Net, a Cascaded Dilated Spatial Pyramid Pooling (CDSPP) module is designed and incorporated into intermediate layers of the model, enhancing the receptive field and producing feature maps enriched with contextual information. To improve segmentation performance for low-contrast vessels, we propose an InceptionConv (IConv) module, which can explore deeper semantic features and suppress the propagation of non-vessel information. Furthermore, we design a Multi-scale Adaptive Feature Aggregation (MAFA) module to fuse the multi-scale feature by assigning adaptive weight coefficients to different feature maps through skip connections. Finally, to explore the complementary contextual information and enhance the continuity of microvascular structures, a fusion module is designed to combine the segmentation results obtained from patches of different sizes, achieving fine microvascular segmentation performance. In order to assess the effectiveness of our approach, we conducted evaluations on three widely-used public datasets: DRIVE, CHASE-DB1, and STARE. Our findings reveal a remarkable advancement over the current state-of-the-art (SOTA) techniques, with the mean values of *Se* and *F1* scores being an increase of 7.9% and 4.7%, respectively. The code is available at <https://github.com/bai101315/MCDAU-Net>.

1. Introduction

Retinal vessel segmentation is an essential pre-processing step for fundus image analysis [1–3]. In practice, ophthalmologists segment the retinal vessels manually, while this process is a time-consuming and subjective work. Therefore, it is particularly important to design an automatic approach for retinal vessel segmentation [4,5].

However, accurate retinal vessel segmentation faces several major challenges due to the following reasons [6]. First, due to the low contrast accompanying the fundus image, distinguishing the retinal vessel from background in the areas of low contrast in fundus image is difficult, especially for thin vessel, leading to incomplete vessel segmentation, as

illustrated in Fig. 1(a). Second, retinal vessels differ in their size and shape, as depicted in Fig. 1(b). Third, the anatomical of retinal vessel is complicated caused by lesions (exudates, hemorrhage) in retinal, as shown in Fig. 1(c).

Numerous segmentation approaches have been developed to address the aforementioned complex challenges associated with retinal vessel segmentation [7–17]. Notably, U-Net and its variations have demonstrated exceptional performance in this regard. However, these approaches still face several challenges. Firstly, the down-sampling operation may result in the loss of spatial information, particularly for low-contrast thin vessels. Secondly, the skip connection operation may introduce irrelevant regions such as background noise or non-vessel

* Corresponding author.

** Corresponding author.

E-mail addresses: yi510@jxnu.edu.cn (Y. Yi), Cui_Wei@i2r.a-star.edu.sg (W. Cui).

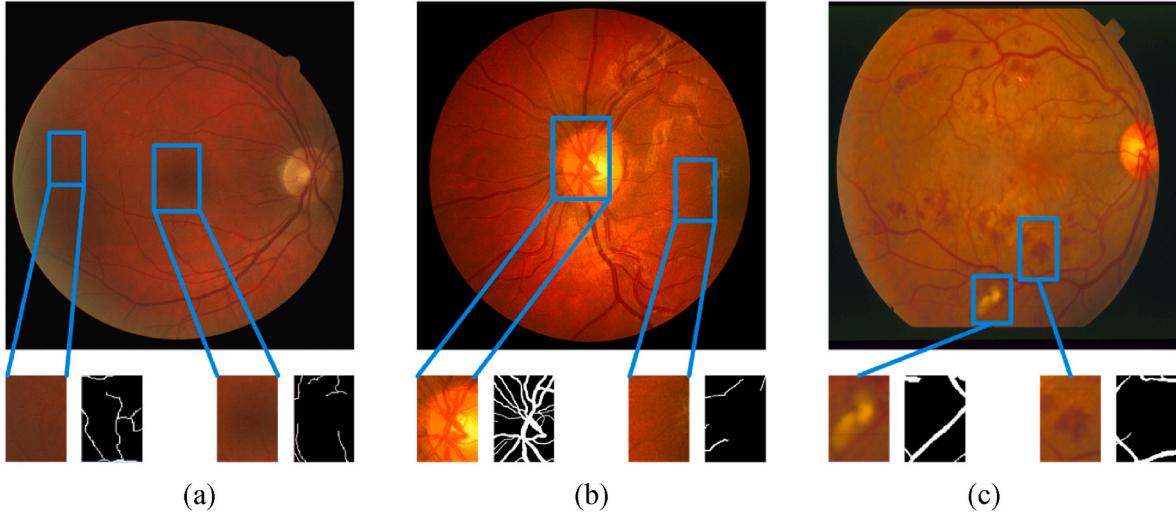


Fig. 1. Challenging cases of retinal vessel segmentation. (a) Low contrast, (b) large-scale variations in vascular tree, (c) lesions (exudates, hemorrhage). Each pair of small square boxes denote detailed view and the corresponding ground truths.

regions, thereby weakening the segmentation performance. Thirdly, although patch-based segmentation approaches have been proposed to mitigate over-fitting caused by small training datasets, they often ignore important contextual information surrounding the extracted patches, which can significantly affect the shape and location of vessels.

In order to tackle the aforementioned challenging issues, we introduce a novel dual-path deep learning framework specifically designed for retinal blood vessel segmentation. In this framework, the fundus images are firstly divided into concentric patches of varying sizes, effectively mitigating the overfitting challenge commonly associated with limited training datasets. Subsequently, we present a Multi-scale Context Dense Aggregation Network (MCDAU-Net) to address the limitations arising from consecutive sampling operations and simple skip connections. Although MCDAU-Net demonstrates proficiency in extracting blood vessels from individual patches, it neglects important contextual information encompassing the extracted patches, resulting in reduced efficacy in segmenting small blood vessels. To address this issue, we introduce a fusion module that integrates the segmentation results derived from patches of varied sizes. This strategy effectively explores context information, thereby enhancing the continuity of microvascular structures.

The core novelties can be summarized as below:

- (1) Proposed a dual-path deep learning framework for the precise segmentation of blood vessels, exhibiting promising prospects for the advancement of medical image analysis in the field.
- (2) Proposed a MCDAU-Net for enhancing the process of feature propagation and generating highly discriminative feature maps.
- (3) Proposed a fusion module to effectively convey contextual information and enhance the integrity and continuity of microvascular structures.

The remaining of the paper is structured as follows: Section 2 will review current automated segmentation approaches for retinal vessel. Section 3 introduces the proposed approach in detail. Extensive experimental results and analyses are presented in Section 4. Finally, we will conclude this paper in Section 5.

2. Related work

Recently, lots of retinal vessel segmentation approaches using machine learning and artificial intelligence techniques [18–23] have been proposed, containing unsupervised approaches and supervised

approaches [24].

2.1. Unsupervised approaches for retinal vessel segmentation

Unsupervised segmentation approaches often employ morphological features of the retinal vessel to construct a simple model without depending any annotations [25]. Specially, unsupervised segmentation approaches mainly consist of matched filtering based, thresholding based, line detection based. In 1989, Chaudhuri et al. [10] firstly employed Gaussian matched filter and multi-directional templates to segment the retinal vessel. However, their approach can hardly recover the whole vessels, particularly the narrow vessel in the low contrast areas. After that, Cinsdikici et al. [26] introduced the ant colony algorithm to further improve the response of matched filter. Since the shape and size of retinal vessels vary in large scale, one-scale filters can hardly work well. Under this circumstance, Li et al. [27] developed a multiscale matched filter to further enhance segmentation efficiency. Apart from the matched filtering-based segmentation approaches, some thresholding based and line detection-based segmentation approaches have been designed recently. For example, Zardadi et al. [28] presented a three-stage approach by taking multiple morphological operations and adaptive thresholding into consideration. In Ref. [29], Zhou et al. introduced a linear detector and Hidden Markov Model (HMM) to trace the thick and thin vessels, respectively. Generally, although the unsupervised approaches can segment the retinal vessel without training and have less complexity, accurately segmenting the low-contrast thin vessel is hard. Moreover, connectivity of the vessel is relatively poor when compared with supervised segmentation approaches.

2.2. Supervised approaches for retinal vessel segmentation

Supervised approaches always utilize the training images with the gold standard to yield a classifier with excellent performance, including shallow machine learning approaches and deep learning-based approaches.

2.2.1. Shallow machine learning approaches

For shallow machine learning approaches, they first extract the hand-craft features and then apply some classical classifiers such as K-Nearest Neighbors (KNN) [30], Support Vector Machines (SVMs) [31], Bayesian classifier [16], AdaBoost classifier [32] etc., for retinal vessel segmentation. For example, authors employed the KNN classifier to classify each pixel for vessel extraction [30]. In Ref. [31], they presented

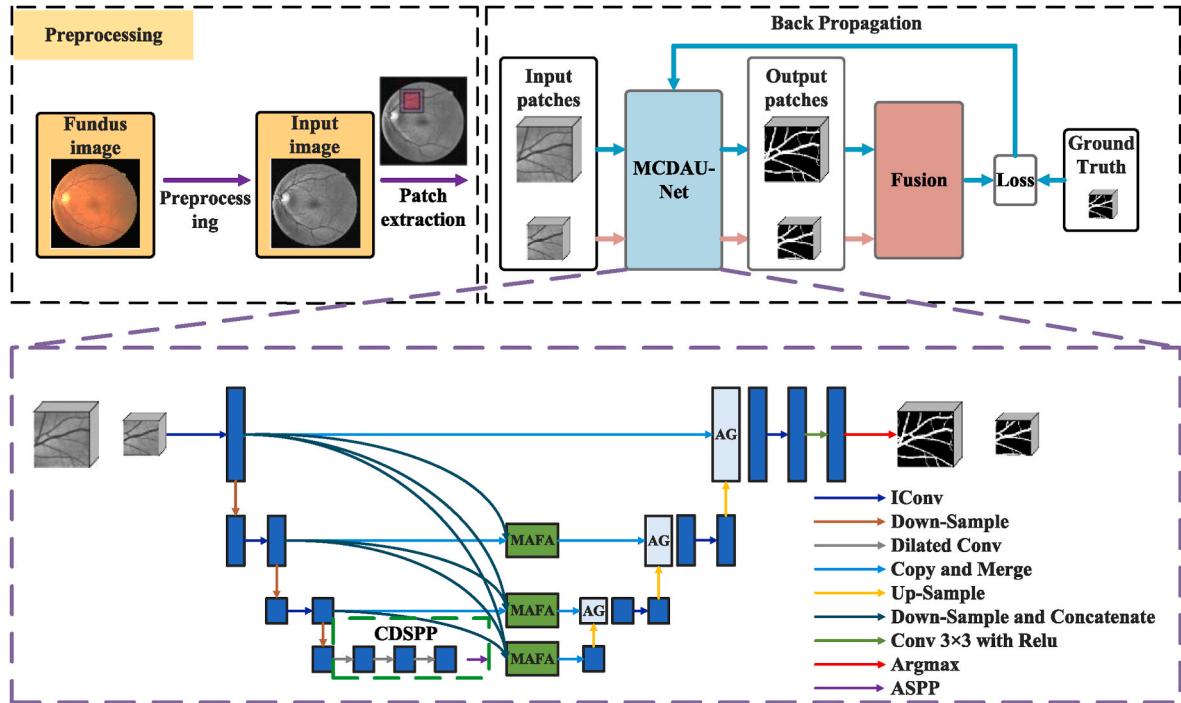


Fig. 2. The proposed framework.

a semi-supervised approach and employed a SVM as the classifier. Similarly, Soares et al. [16] learned the intensity feature from each pixel. Then, a Bayesian classifier was employed to classify the retinal vessel from the non-vessel categories. In Ref. [32], for each pixel, the authors extracted 41 features. Following that, an AdaBoost classifier was utilized to acquire retinal vessel segmentation results. Due to the introduction of the labeled training images, shallow machine learning approaches can perform better than the unsupervised approaches. However, the major disadvantage in them is that the segmentation performance largely relies on the hand-crafted features for target datasets. Therefore, they have poor generalization ability on a new coming dataset.

2.2.2. Deep learning-based approaches

Deep neural networks have strong feature learning ability, which are widely used for retinal vessel segmentation, acquiring excellent segmentation performance. Generally, they consist of two major categories, namely end-to-end approaches and patch-based approaches [33–35].

Since end-to-end approaches are simple, a series of relevant approaches have been put forward. For instance, the authors regarded the vessel segmentation as an edge detection task [36]. Rich hierarchical representations were learned by multi-level network. Then, Conditional Random Field (CRFs) [37] was incorporated, aiming at modeling the remote interactions between pixels. Similarly, reference [38] presented a U-Net network with salience mechanism. To enhance the performance of network, reference [39] replaced the convolutional layers with the pooling layers and put forward a full convolutional neural network. Further, references [40,41] treated vessel segmentation as a retrieving problem. They introduced matching algorithm as well as walking technique into network to enhance the completeness and continuity of microvascular. Later, Wu et al. [6] presented a U-shape context sensitive network, which can efficiently aggregate multi-scale features for retinal vessel segmentation. To precisely extract the tiny blood vessels, reference [42] presented an Encoder Enhanced Atrous (EEA) network, in which the encoder was strengthened via importing the depth concatenation process. In Ref. [43], the authors designed a deep fusion network for multi-source vessel segmentation, in which the tiny vessel can be

exploited through Frangi filter guided. In Ref. [44], the Edge-aware flows were introduced into U-Net for retinal vessel segmentation, which can enhance the representative of the tiny blood vessel. Although end-to-end approaches have many merits and achieved great segmentation performance, they still have the following limitations. On one hand, since the resolution of fundus images is large, end-to-end approaches always deal with the whole image via reducing resolution. Therefore, this process will lose lots of spatial information due to the limited computing resources. On the other hand, data augmentations are necessary to be introduced to overcome the over-fitting issue caused by insufficient sample size.

To address the above limitations, patch-based deep learning approaches have been designed. For instance, the authors suggested a CNN architecture for pixel-level classification, which can automatically learn features from both the raw images and predicted patterns [45]. In 2016, the authors first employed data augmentation techniques to extend the size of dataset [46]. Then, a deep neural network was trained on a large-scale dataset for vessel segmentation. To speed-up the patch-based FCNN training, a patch-based FCNN approach was designed by local entropy sampling and a jump-over network with quasi-equilibrium losses [47]. To achieve the balance of thick and thin vessels, the authors incorporated two level losses, including segment-level and pixel-level, into network model for segmentation [48]. Similarly, reference [49] presented a MS-NFN model for vessel segmentation, which contained both ‘up-pool’ and ‘pool-up’ NFN submodules. Inspired by the concept of “network followed network”, Wu et al. [50] proposed an NFN + model to enhance the segmentation performance furthermore. To accurately segment the details of retinal vessel, reference [51] employed the Nest U-net and patch-learning to construct the segmentation model. Recently, reference [52] presented a simplified version of UNet model by introducing residual blocks and batch normalization in the encoding and decoding stages for tiny blood vessel segmentation.

Although patch-based segmentation approaches have become the mainstream, most of them do not fully consider the contextual features embedded in the surrounding pixels. Hence, distinguishing the retinal vessel from background in the areas of low contrast is hard, especially for thin vessel. Besides, the anatomical of retinal vessel is complicated

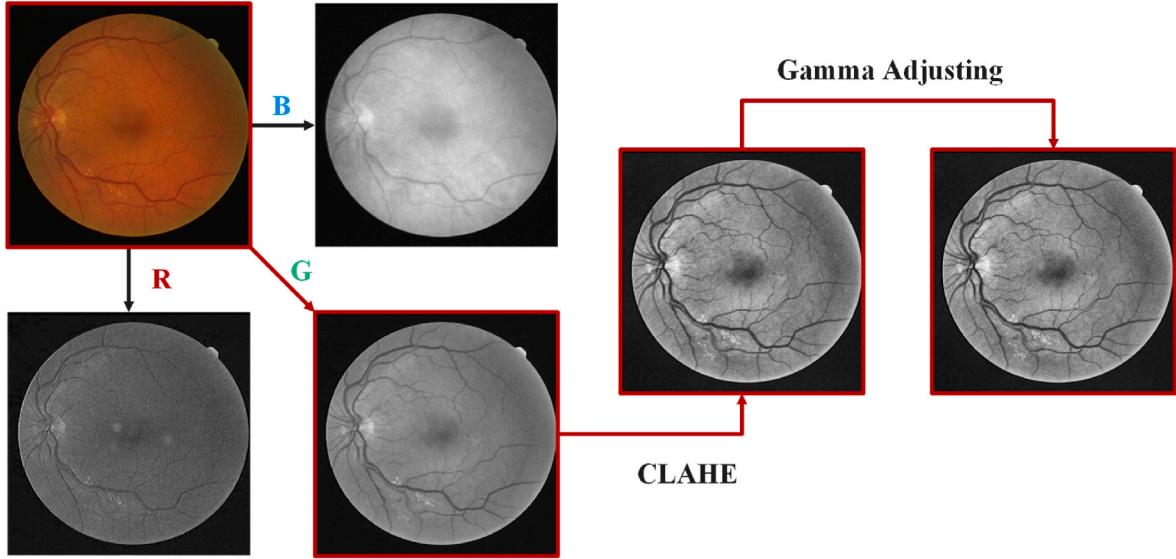


Fig. 3. An example of data preprocessing.

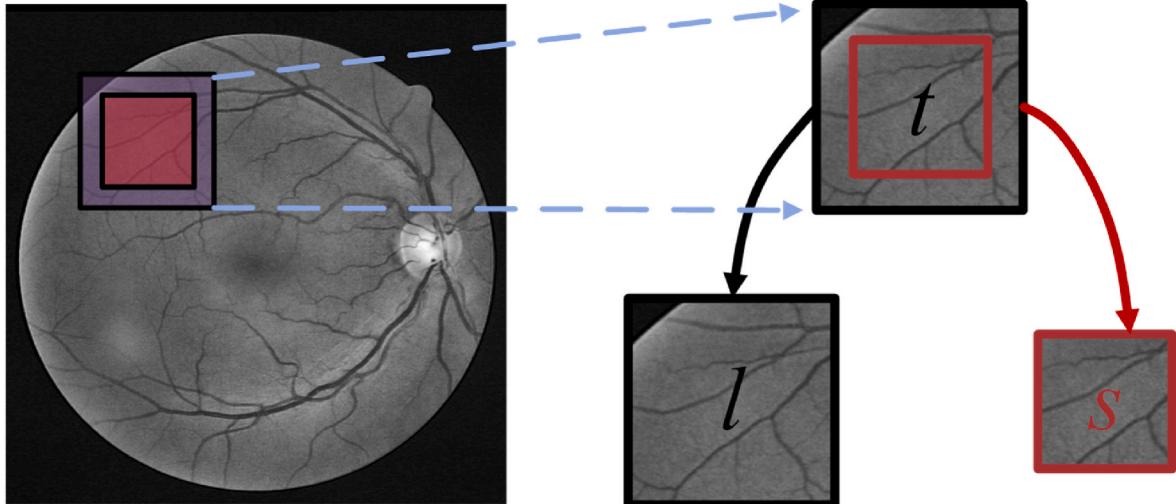


Fig. 4. The descriptions of patch extraction. The red box represents the small patch s and the target t . The black box represents the larger patch l , which is used to compensate for the missing global information.

(e.g., lesions or large-scale variation), simply stacking multi-scale modules will lead to information loss on microvascular or details.

3. Methodology

To address these limitations, we propose a dual-path deep learning framework for blood vessel segmentation. Fig. 2 depicts the overall framework of our approach, which is consisted of Preprocessing, MCDAU-Net and Fusion.

3.1. Preprocessing

The high-level noise and inhomogeneous illumination areas in fundus images will directly affect the effectiveness of the vessel's segmentation. Extensive researches have shown that data preprocessing can improve segmentation results and accelerate the training of the network. Therefore, several preprocessing steps are introduced before modeling. For each color retinal fundus image, the green channel image is first extracted. Then, Contrast Limited Adaptive Histogram Equalization

(CLAHE) [53] is applied to enhance the contrast between vessels and the background. Together with CLAHE, the gamma adjustment can further reduce the effects of inhomogeneous illumination within the image. Finally, image normalization is applied to reduce image dimension and accelerate network convergence. As seen from Fig. 3, the blood vessel information in the green channel image is more prominent. After that, the contrast between the vessels and the background is further enhanced through CLAHE and gamma adjustment. Meanwhile, the problem of inhomogeneous illumination can be alleviated.

Considering that the training of CNN models usually relies on large-scale datasets, the limited number high-resolution fundus images may lead to the overfitting issue. To address this issue, this paper adopts the patch-based deep learning technique to expand the number of images. However, almost all patch-based approaches utilize only the patches that precisely match the desired region, while ignoring the contextual information around the patches. For each patch represented as target region t , we extract two concentric patches with an inclusion relation, as shown in Fig. 4. From this figure, it can be seen that the small patch s exactly overlaps with t , while the large patch l consists of both the entire

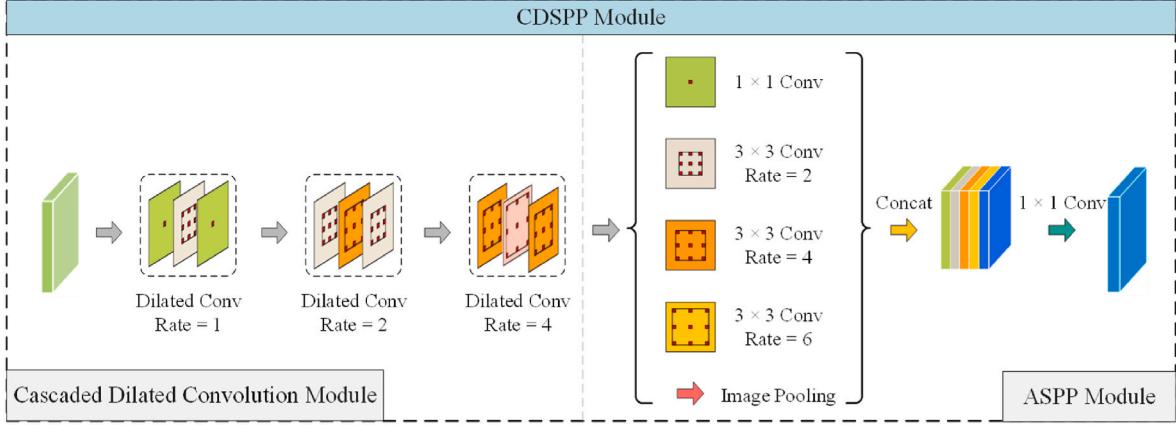
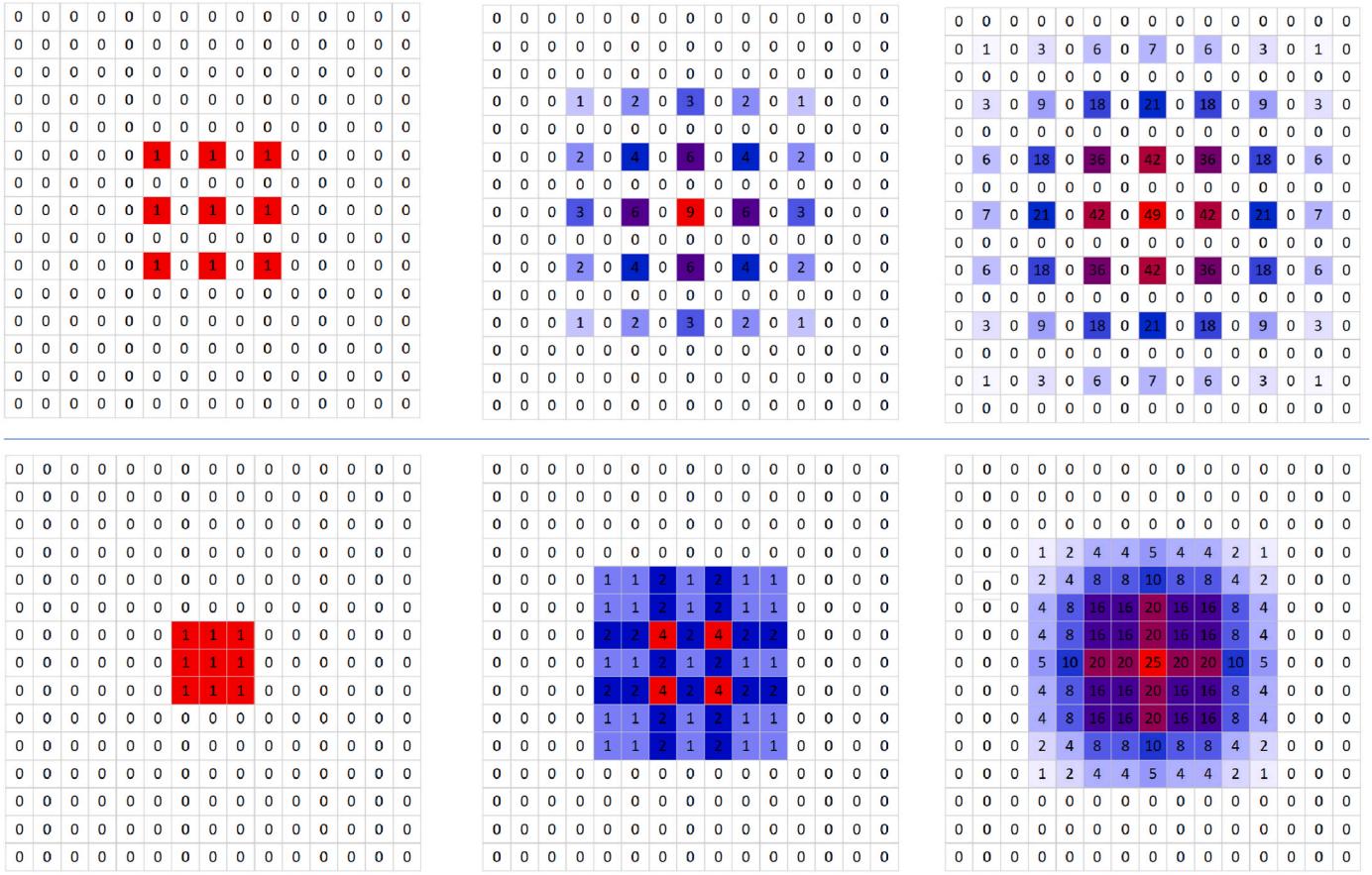


Fig. 5. The CDSPP module.

Fig. 6. The receptive fields of cascaded dilated convolution module. Left to right: the number indicates that the calculation of each pixel through convolution layers with size 3×3 . Line 1: the dilation rate = (2, 2, 2). Line 2: the dilation rate rates = (1, 2, 1).

target region and its surrounding pixels embedded around the target region t . Finally, these patches are combined to form the input sequence $[L, S]$, in which $L = concat [l_1, l_2, l_3, l_4\dots], S = concat [s_1, s_2, s_3, s_4\dots]$.

3.2. MCDAU-net

Fig. 2 depicts the overall structure of MCDAU-Net, which is a U-shaped structure and contains three proposed modules: CDSPP, IConv and MAFA.

3.2.1. Cascaded Dilated Spatial Pyramid Pooling (CDSPP)

Continuous convolution and pooling operations cause significant loss of microscopic details [54]. Moreover, detail-related features that cannot be accurately restored by simple up-sampling operation. To address these issues and enhance the semantic segmentation performance, we propose a CDSPP module, which is consisted of cascaded dilated convolution module and ASPP module as shown in Fig. 5.

The cascaded dilated convolution module uses multiple dilated convolutional layers with different dilation rates in a sequential manner. Each layer captures contextual information at a different scale, allowing the network to capture both local and global features effectively. By

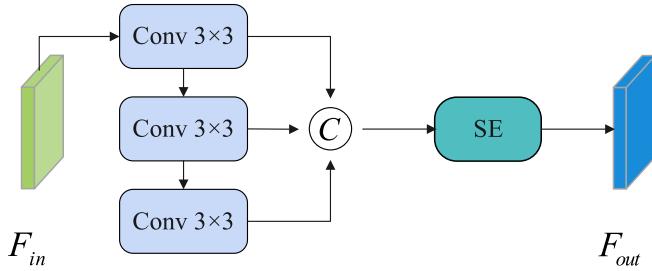


Fig. 7. Schematic diagram of IConv module.

stacking these layers, the network can incorporate information from a broader spatial range.

In practice, in order to achieve a balance between efficiency and computational cost, dilation convolution is usually appropriately applied to down-sampling of the network [55]. Inspired by Ref. [56], the multi-gird rate is set to (1, 2, 1). The final dilation rate is the multiplication of the base rate and the corresponding multi-gird rate. As illustrated in Fig. 6, the values of multi-gird rate in the top row and bottom row are set to (2, 2, 2) and (1, 2, 1).

Atrous Spatial Pyramid Pooling (ASPP) module provides a powerful mechanism for capturing multi-scale contextual information efficiently,

leading to improved semantic segmentation performance in computer vision tasks. Inspired by the merits of ASPP, this paper places it after the cascaded dilated convolution module to construct the CDSPP module. Therefore, more contextual information and global information can be utilized to achieve excellent segmentation results.

3.2.2. InceptionConv (IConv)

Traditional encoder-decoder architectures contain the finite receptive field, which leads to the following two major limitations. For one thing, the extracted local features are limited. For another, the global context information is ignored. Inspired by Szegedy et al. [54], we propose an InceptionConv module (IConv for short), which replaces the single two-convolution structure with the combination of Inception module and standard convolutions together, as depicted in Fig. 7.

To achieve multi-scale feature fusion, we propose a modified Inception module by discarding the multi-branch structure in inception. In this paper, three 3×3 convolutions with different receptive field sizes of 3×3 , 5×5 , and 7×7 are applied to feature extraction, respectively. The output feature maps of each convolution are concatenated along the channel dimension to produce multi-scale feature maps. This approach enhances the model's ability to extract local and global blood vessel features, while also enabling effective detection of vessels with varying shapes and sizes. Moreover, the SE block is incorporated into the IConv

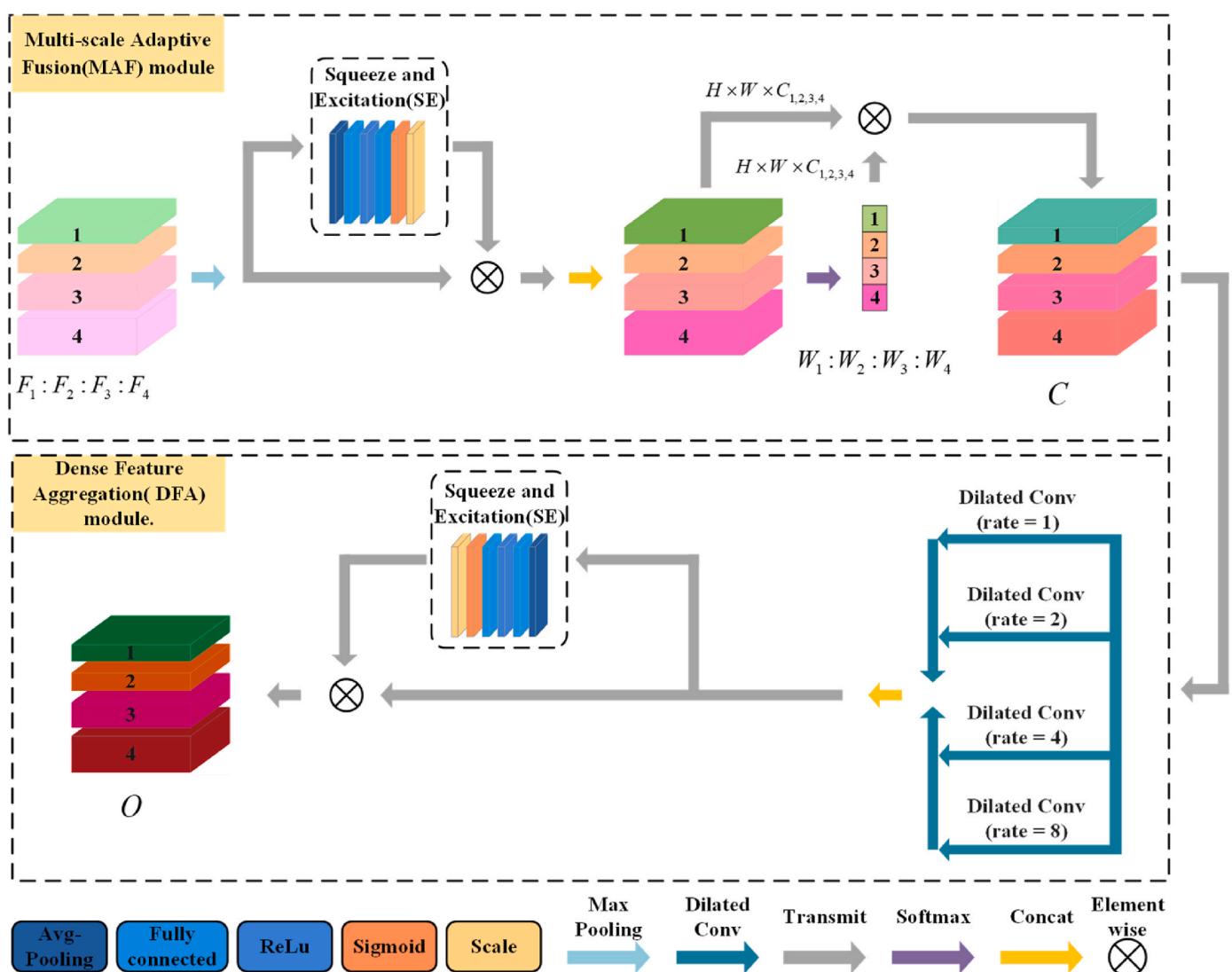


Fig. 8. Schematic diagram of MAFA. (a) Multi-scale Adaptive Fusion (MAF) module, (b) Dense Feature Aggregation (DFA) module.

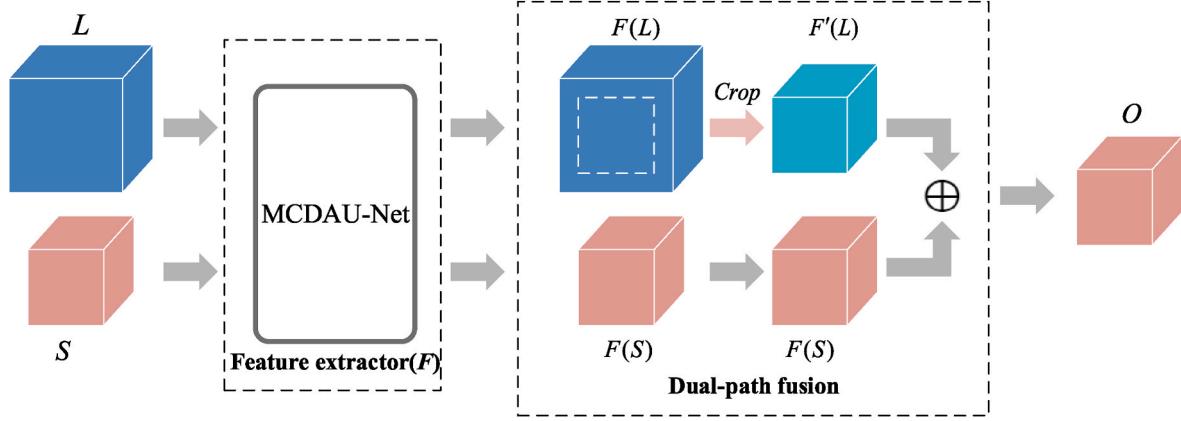


Fig. 9. The proposed fusion module.

module to suppress irrelevant features and enhance feature expression.

The output of IConv module is:

$$\begin{aligned} F_1 &= \text{Conv}_{3 \times 3}(F_{in}) \\ F_2 &= \text{Conv}_{3 \times 3}(F_1) \\ F_3 &= \text{Conv}_{3 \times 3}(F_2) \\ F_{out} &= \text{SE}(\text{Concat}(F_1 : F_2 : F_3)) \end{aligned} \quad (1)$$

where F_{in} represents the input feature maps and F_{out} represents the output feature maps, Concat represents the concatenation operation in the channel dimension and SE is a SE block.

3.2.3. Multi-scale Adaptive Feature Aggregation (MAFA)

The feature maps of different scales always involve various semantic feature information. For example, low-level feature maps have higher resolution with more detailed information, while lower semantic information with more noise. High-level feature maps have stronger semantic information, but lower resolution and relatively poor segmentation results. To fully take advantage of these features, this paper designs a MAFA module consisting of the Multi-scale Adaptive Fusion (MAF) module and the Dense Feature Aggregation (DFA) module, as illustrated in Fig. 8.

In Fig. 8, the MAF module is on the top and the DFA module is on the bottom. In the MAF module, each layer input contains the output feature maps from the current layer encoder unit and all the shallower layer encoder units. Taking the fourth network block as an example, its input feature maps are from the fourth layer encoder unit and the first, second and third-layer encoder units, respectively denoted as $F_1 : F_2 : F_3 : F_4$. The obtained multi-scales feature maps $F_1 : F_2 : F_3 : F_4$ are first performed with max-pooling to generate feature maps with the same size. After that, they are used to obtain the attention vector in channel. By recalibrating the channel attention vectors with the softmax function, we can obtain the multi-scale channel calibration weights $W_1 : W_2 : W_3 : W_4$. Finally, the recalibrated weights are multiplied with their respective feature maps element by element, and then are concatenated together to achieve multi-scale feature fusion. A refined feature map C with rich multi-scale information is regarded as the output of MAF.

The formula of MAF can be expressed as:

$$W_1 : W_2 : W_3 : W_4 = S(\text{SE}(F_M(F_1 : F_2 : F_3 : F_4))) \quad (2)$$

$$C = \text{Concat}((F_1 : F_2 : F_3 : F_4) \otimes (W_1 : W_2 : W_3 : W_4)) \quad (3)$$

where F_M is the down-sampling operation. SE denotes the channel attention module and S represents the Softmax activation function. \otimes is the element-wise product.

After obtaining the multi-scale feature maps by MAF, these maps are regarded as the inputs of the DFA module. In DFA, dilated convolution

with different dilation factors is firstly employed to the feature map C , aiming at better extracting features with contextual information. As shown in Fig. 8, for a feature map C , we use the dilated convolution with different dilation rates for parallel multi-branch feature extraction. Subsequently, the feature maps obtained from each branch are concatenated along the channel dimension, followed by their input into the Squeeze-and-Excitation (SE) blocks. The primary objective of these SE blocks is to effectively suppress the propagation of irrelevant features.

The output of the DFA can be expressed as:

$$C_i = \text{DeConv}_{3 \times 3}(C) \quad (4)$$

$$O = \text{SE}(\text{Concat}(C_1 : C_2 : C_3 : C_4)) \quad (5)$$

where $\text{DeConv}_{3 \times 3}$ represents the dilated convolution with kernel size of 3×3 . $C_1 : C_2 : C_3 : C_4$ is the feature maps after dilated convolution operations, O denotes the output feature map of DFA.

3.3. Fusion

For most of patch-based approaches, they just utilize the patches that precisely match the desired region, ignoring the contextual information around the patches. Facing this situation, we have proposed a fusion module to fuse the contextual information. For given fundus images, after preprocessing, the input sequence $[L, S]$ is fed into the feature extractor MCDAU-Net for generating the probability feature maps $[F(L), F(S)]$. Following that, the proposed fusion module is used to $[F(L), F(S)]$, achieving contextual information fusion, as depicted in Fig. 9. During this process, $F'(L)$ is firstly achieved by center cropping $F(L)$ to the same size as the $F(S)$. Then, $F'(L)$ is merged with $F(S)$ at pixel level to construct the final feature map. Finally, the feature map containing context and edge information is regarded as the output of our proposed approach as below:

$$\begin{aligned} [F(L), F(S)] &= F([L, S]) \\ F(L) &= \text{Crop}(F(L)) \\ O &= F_n([F'(L), F(S)]) \end{aligned} \quad (6)$$

where F is a feature extractor and Crop denotes the operation of cropping from the center. The fusion module is denoted as F_n .

3.4. Loss function

A suitable loss function possesses the ability to accurately represent the disparity between the probability map and the ground truth, thereby providing guidance for training the network. Given that vessel segmentation entails a pixel-level binary classification task, the loss function employed for the model is the pixel-level binary cross-entropy. The mathematical definition of this function is as follows:

Table 1
Descriptions of the datasets.

Dataset	Image Number	Training Set	Test Set	Image Resolution
DRIVE	40	20	20	584 × 565
CHASE-DB1	28	20	8	999 × 960
STARE	20	15	5	700 × 605

$$L = \sum_i^N (y'_i \log(y_i) + (1 - y'_i) \log(1 - y_i)) \quad (7)$$

where y'_i represents the true probability value for pixel i and y_i represents the predicted probability value for pixel i . N represents the total number of pixels.

4. Experiments

4.1. Implementation details

Our work is based on the Pytorch platform, and we conducted our experiments on an NVIDIA TITAN XP graphics card with 24 GB of memory. Prior to training, each image in the dataset underwent a series of pre-processing steps, including green channel extraction, CLAHE, and gamma adjustment with a parameter of 0.9. Subsequently, the images were cropped into patches of size 96×96 , followed by center cropping to obtain patches of size 64×64 , which were used as input to our model. Both the cropped patch and the original patch were fed into the model, and the pixel-level cross-entropy loss function was employed to measure the discrepancy between the predicted probability map and the ground truth. The network's parameters were updated using the Adam optimizer with an initial learning rate of $1e-3$. Furthermore, we adopted the “poly” learning rate, inspired by DeepLabV2 [57], with a power hyperparameter of 0.9, and updates were made every 10 epochs. Our experiment entailed a total of 100 training epochs and a batch size of 20.

4.2. Datasets

To evaluate the performance of the proposed approach, three public retinal fundus vessel segmentation datasets, involving DRIVE [30], CHASE-DB1 [18], and STARE [58] are used. The specifications of each dataset are shown in Table 1 and some typical images are given in Fig. 10.

DRIVE: It contains 40 retinal fundus images, in which 33 images are normal and 7 images are early diabetic retinopathy. There are 20 images in training set and 20 images in test set. The images are captured digitally by a Canon CR5 non-eyedrop 3CCD camera at 45° .

CHASE-DB1: It consists of 28 retinal images from 14 children, of which the first 20 images are applied for model training and the remaining are used for model testing. The images are captured by a handheld Nidek NM-200-D fundus camera at 30° FOV.

STARE: It contains 20 retinal fundus images captured by Topcon camera at 35° FOV and half of them are pathologic images. Since there is no explicitly division for training and test sets, this paper employs the widely used division strategy, i.e., 5 images are randomly chosen as the training set, and the remaining images are regarded as the test set.

In the CHASE-DB1, DRIVE, and STARE datasets, image patch extraction for model training and testing was performed using cropping techniques with specific patch sizes and stride values. Considering the trade-off between time and performance, we choose a suitable patch size for our experiments. For the CHASE-DB1 dataset, 18,000 image patches sized 96×96 were extracted for model training using random cropping with a stride of 32. Similarly, a stride of 64 was employed for sliding cropping in the testing phase, resulting in 1800 image patches of the same size. In the DRIVE dataset, 18200 patches of size 96×96 were obtained for model training, while the STARE dataset utilized 18220 patches. For testing, a stride of 64 was used for sliding cropping, resulting in 1620 patches of size 96×96 for the DRIVE dataset, and 500 patches of the same size for the STARE dataset.

4.3. Evaluation metrics

To objectively evaluate the performance of our model, four vital performance evaluation indicators are employed, including Accuracy (ACC), Specificity (Sp), Sensitivity (Se), and $F1$ score are used as the evaluation metrics, which are defined as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Se = \frac{TP}{TP + FN} \quad (9)$$

$$Sp = \frac{TN}{FP + TN} \quad (10)$$

$$F1 = \frac{2 \times TP}{2 \times TP + FN + FP} \quad (11)$$

where TP and TN represent the number of blood vessel pixels and background pixels that are correctly segmented. FP denotes the number of pixels where background is falsely predicted as blood vessel and FN is the number of pixels where blood vessel is incorrectly predicted as background. ACC is always utilized to describe the segmentation performance. The higher the ACC obtained, the better the segmentation

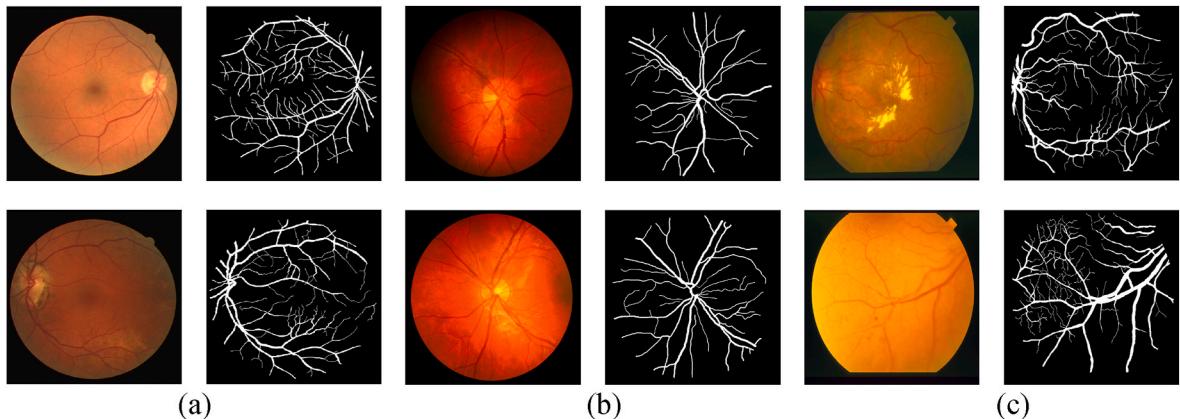


Fig. 10. Challenging cases of vessel segmentation. (a) DRIVE, (b) CHASE-DB1, (c) STARE. The original images are on the left, the corresponding ground truth images are on the right.

Table 2
Ablation experiments on the CHASE-DB1 dataset.

Model	F1	Se	Sp	ACC
Baseline	0.7888	0.8215	0.9742	0.9602
Baseline + CDSPP	0.7922	0.8304	0.9737	0.9606
Baseline + CDSPP + IConv	0.7924	0.8390	0.9724	0.9602
Baseline + CDSPP + IConv + MAFA	0.7938	0.8438	0.9721	0.9604
Ours (Baseline + CDSPP + IConv + MAFA + Fusion)	0.8031	0.8796	0.9694	0.9612

performance of the model. *Se* and *Sp* are utilized to evaluate the accuracy of *TP* pixels and *TN* pixels segmentation. By considering these metrics, we aimed to provide a comprehensive assessment of the model's effectiveness and suitability.

4.4. Results

4.4.1. Ablation studies

Among the utilized three datasets, CHASE-DB1 dataset is the most challenging one due to various illumination changes and some low contrast areas. We carry out the ablation experiments on it. The proposed approach has four major components including CDSPP, IConv, MAFA, and Fusion modules. First, the original U-Net is chosen as the Baseline. Then, the four vital components are added sequentially. The mean *F1*, *Se*, *Sp*, and *ACC* are used to evaluate the segmentation performance. The quantitative and visualization results are given in Table 2 and Fig. 11.

4.4.1.1. Efficacy of the CDSPP module. First of all, the proposed CDSPP module is introduced into the Baseline (Baseline + CDSPP) and the visualization result is shown in Fig. 11(e). Evidently, the segmentation

performance of the Baseline + CDSPP model surpasses that of the baseline network, as demonstrated in Fig. 11(d), owing to its ability to capture a greater number of vessel pixels. The information derived from Table 2 enables us to ascertain that all evaluation indicators are improved to some extent after adding the CDSPP module to the Baseline (Baseline + CDSPP). Further, compared to the Baseline, the *Se/F1* score improved from 82.15%/78.88%–83.04%/79.22%, improving by 0.89%/0.34%, respectively. The main reason is that the cascaded dilation module expands the receptive field while maintaining the resolution of feature maps, avoiding microscopic details loss due to down-sampling operation. Furthermore, the pyramid module with spatial continuity is able to alleviate the challenging issue of false segmentation.

4.4.1.2. Efficacy of the IConv module. Second, we utilize IConv module to replace the standard convolution of the Baseline (Baseline + CDSPP + IConv). As shown in Table 2, the segmentation performance has an improvement in terms of *Se*. More visual segmentation results are also shown in Fig. 11(f). It can be seen that the sixth column achieves excellent segmentation performance for both thick and tiny vessels, especially for thin vessels with large scale variations within low-contrast regions (see the segmentation results in the first row and third row). Since IConv module is designed by using standard convolutions with different kernel sizes, it can exploit multi-scale features and achieve effective feature fusion. Compared with Baseline, the feature extraction ability of the network is strengthened, which can acquire more excellent segmentation performance. Overall, the combination of CDSPP + IConv is effective in our model.

4.4.1.3. Efficacy of the MAFA module. Third, we import the MAFA module into (Baseline + CDSPP + IConv) referred to as (Baseline +

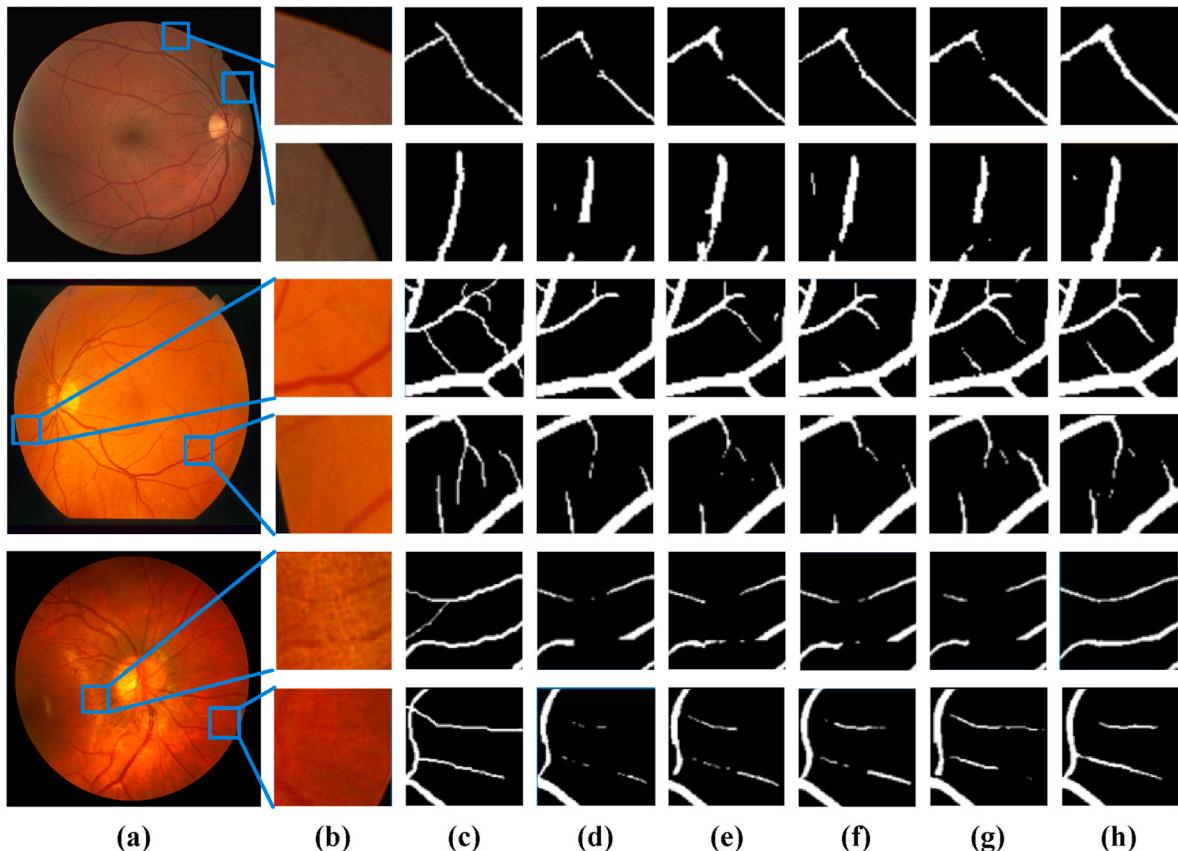


Fig. 11. Visualization results on ablation experiments. (a) Original images, (b) image patches, (c) ground truth, (d) Baseline, (e) Baseline + CDSPP, (f) Baseline + CDSPP + IConv, (g) Baseline + CDSPP + IConv + MAFA and (h) Ours (Baseline + CDSPP + IConv + MAFA + Fusion).

Table 3

Comparative experiments on the DRIVE dataset.

Method	Year	F1	Se	Sp	ACC
U-Net [59]	2015	-	0.7915	0.9808	0.9640
Li et al. [60]	2015	-	0.7569	0.9816	0.9527
AttU-Net [61]	2018	0.7917	0.7580	0.9850	0.9651
M-Net [62]	2018	-	0.7559	0.9835	0.9634
CE-Net [63]	2019	-	0.8309	0.9816	0.9545
VeeselNet [64]	2019	-	0.8038	0.9802	0.9578
Wang et al. [65]	2019	-	0.7287	0.9771	0.9443
Dhevendra et al. [66]	2020	-	0.7375	0.9788	0.9480
Jinzhu Yang et al. [67]	2020	-	0.6817	0.9765	0.9507
Jinzhu Yang et al. [68]	2020	-	0.7181	0.9747	0.9522
Sathananthavathi et al. [42]	2021	-	0.7918	0.9708	0.9577
Wang, zhao and Yu [51]	2021	0.7863	0.8060	0.9869	0.9512
Wu et al. [6]	2021	-	0.8289	0.9838	0.9697
DF-Net [43]	2022	-	0.7733	0.9853	0.9623
Edge-Aware U-Net [44]	2022	0.8021	0.7719	0.9799	0.9701
Human Observer	-	-	0.7760	0.9724	0.9472
Ours	-	0.8129	0.8215	0.9739	0.9589

Table 4

Comparative experiments on the CHASE-DB1 dataset.

Method	Year	F1	Se	Sp	ACC
U-Net [59]	2015	-	0.8055	0.9832	0.9720
Li et al. [60]	2015	-	0.7507	0.9793	0.9581
AttU-Net [61]	2018	-	0.7721	0.9850	0.9726
Zengqiang et al. [69]	2018	-	0.7641	0.9806	0.9607
M-Net [62]	2018	-	0.7606	0.9855	0.9709
CE-Net [63]	2019	0.7872	0.8042	0.9839	0.9723
VeeselNet [64]	2019	-	0.8132	0.9814	0.9661
SD-UNet [70]	2019	0.8030	0.8547	0.9816	0.9736
Sathananthavathi et al. [42]	2021	-	0.6457	0.9653	0.9340
Manuel et al. [52]	2021	-	0.8063	0.9707	0.9571
Wang, zhao and Yu [51]	2021	-	0.8435	0.9782	0.9630
Wu et al. [6]	2021	-	0.8365	0.9839	0.9744
DF-Net [43]	2022	-	0.8316	0.9885	0.9812
Edge-Aware U-Net [44]	2022	0.7662	0.8506	0.9981	0.9811
Human Observer	-	-	0.8105	0.9711	0.9545
Ours	-	0.8031	0.8796	0.9694	0.9612

CDSPP + IConv + MAFA) to verify its effectiveness. Since the Baseline network employs the simple jump connection in each layer to describe local information, the semantic information cannot be fully explored, especially for multi-scale blood vessel structure. As illustrated in Fig. 11 (g), our MAFA module can acquire a more precise and complete blood segmentation results than Baseline. The main reason is that the MAFA module enhances the propagation of relevant features by constructing multi-scale feature maps. Hence, feature maps with different scales in different layers can be concatenated to generate meaningful and discriminative feature maps, which is good for the segmentation of tiny vessels.

4.4.1.4. Efficacy of fusion module. Finally, in order to effectively convey contextual information and enhance the integrity and continuity of microvascular structures, we design a fusion module and introduce it into MCDAU-Net (Baseline + CDSPP + IConv + MAFA) referred to as ours (Baseline + CDSPP + IConv + MAFA + Fusion). As seen from Fig. 11 (h), compared with the Baseline network, our approach is able to capture relatively complete topology and refined segmentation results. Thanks to the dual-path fusion training, our segmentation model can fully compensate the missing global information of the patch extraction and improve the segmentation effect of tiny patch. As shown in Table 2, our approach shows tremendous improvement in the Se and F1 scores, which is 5.47% and 5.8%, respectively, compared to the baseline network. Note that Sp is slightly lower than the Baseline due to a moderate trade-off between Sp and Se. As can be seen from the visual and statistical results, each component in our model is effective, and the

Table 5

Comparative experiments on the STARE dataset.

Method	Year	F1	Se	Sp	ACC
U-Net [59]	2015	-	0.7839	0.9871	0.9716
Li et al. [60]	2015	-	0.7726	0.9844	0.9628
AttU-Net [61]	2018	-	0.7903	0.9853	0.9722
M-Net [62]	2018	-	0.7446	0.9908	0.9701
CE-Net [63]	2019	-	0.7916	0.9853	0.9715
Zhun Fan et al. [71]	2019	-	0.9710	0.9700	0.9570
Henry A. Leopold et al. [72]	2019	-	0.6433	0.9472	0.9045
Sushil et al. [73]	2020	-	0.6149	0.9823	0.9513
Wang, zhao and Yu et al. [51]	2021	0.7947	0.8230	0.9945	0.9641
Wu et al. [6]	2021	-	0.8207	0.9839	0.9736
DF-Net [43]	2022	-	0.8048	0.9899	0.9716
Edge-Aware U-Net [44]	2022	0.7552	0.6912	0.9911	0.9691
Human Observer	-	-	0.8952	0.9384	0.9346
Ours	-	0.8497	0.8505	0.9686	0.9774

best segmentation results can be obtained by combining these components together. Therefore, the proposed approach is well suited for retinal vessel segmentation.

4.4.2. Comparisons with the state-of-the-art methods

To demonstrate the superiority of our approach, we conducted a comparative analysis against the current SOTA methodologies using three publicly available datasets for blood vessel segmentation. Additionally, the segmentation results obtained by second human observer are also listed for comparison. More detailed comparison results are depicted in Tables 3–5.

First, our model is evaluated on the DRIVE dataset. In the low-contrast area located at the center of the images, there exists a large number of vessels ends. Accurate segmentation of such vessels ends is extremely crucial for precise vessel segmentation. As shown in Table 3, our model achieves 81.29%, 82.15%, 97.39%, and 95.89% for the F1, Se, Sp, and ACC metrics, respectively. Particularly, our approach obtains F1 of 81.29% and Se of 82.15%, which indicate that the proposed model can correctly segment vessel ends even in low-contrast regions. By introducing the MAFA module and the fusion module, our model is able to effectively fuse multi-scale features and incorporate a wealth of vessel information. This enhances weak vessel information, which in turn facilitates better feature extraction by the model. Although some metrics are not superior to some methods, our results are comparable to the best on of them within 1% difference.

Second, the experiments are performed on the CHASE-DB1 dataset, whose corresponding comparison results are summarized in Table 4. Seen from this table, we can observe that our model reaches 80.31% and 87.96% for F1 score and Se, where the Se score is much higher than the other approaches. Compared to the recently proposed model [43], although it achieves 98.12% in Sp score, it only achieves 83.16% on F1. The CHASE-DB1 dataset is more challenging, in which the images have various illumination and low contrast areas, making it difficult to detect tiny blood vessels. By introducing the IConv module and the CDSPP module, our model can perform deep feature extraction and generate a denser feature representation map through the fusion of multi-scale features. Higher evaluation metrics indicate that our approach still has a strong ability to detect tiny blood vessels, even faced with various illuminations and weak contrast images.

At last, Table 5 depicts the comparative results of different approaches on the STARE dataset. The STARE dataset contains half of the pathological images, including a series of lesions such as hemorrhage and exudative inflammation. These pathological features may affect network performance. Thanks to the proposed dual-path fusion framework, lesion regions are segmented into numerous patches, which not only reduces the impact of the lesion regions on the segmentation results, but also avoids the edge information loss and blurring caused by image segmentation. From this table, we can see that our approach reaches the top performance in all metrics. Compared to the papers [43,

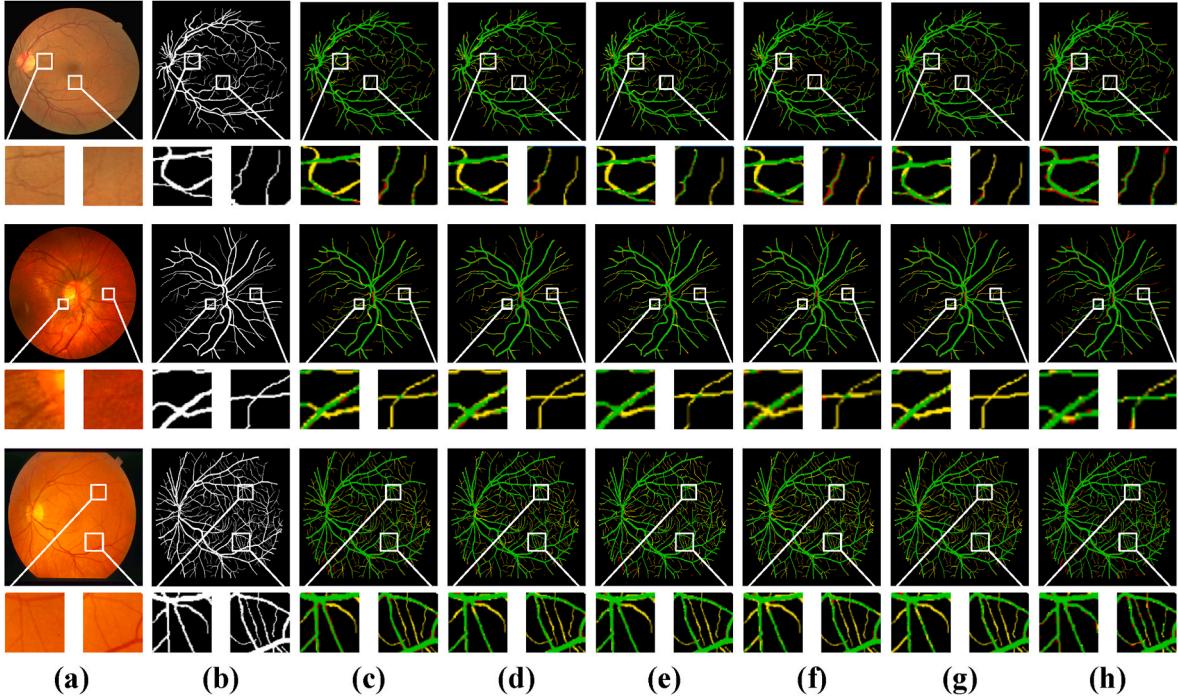


Fig. 12. Some visual segmentation examples of different approaches: DRIVE (top), CHASE-DB1 (middle), STARE (bottom). (a) Original images, (b) Ground truth, (c) U-Net, (d) M-Net, (e) ATTU-Net, (f) CE-Net, (g) SD-UNet and (h) Ours. The pixels in red, yellow and green colors indicate FP, FN, TP, respectively.

[44] published in 2022, our approach is far superior to them in both *Se* and *ACC* metrics, with the F1 score being 9.45% higher than [44]. The results indicate that our approach far outperforms the SOTA methods in the literature.

To intuitively observe the segmentation effect, some typical visualization results of our approach and SOTA approaches are shown in Fig. 12. From this figure, it can be observed that U-Net has a large number of yellow pixels (represented as FN), which cannot achieve continuous segmentation of tiny vessels (see Fig. 12(c)). Inspired by the success of attention, ATTU-Net focuses on the useful salient features by using a gated attention mechanism and achieves a comparable performance against U-Net (see Fig. 12(e)). However, it is clear that the segmentation results of U-Net and ATTU-Net are still unsatisfactory. The main reason is that they are all subjected to the receptive field, which hardly extracts multi-scale features. Under this circumstance, M-Net regarded as a variation of U-Net is developed by combining multi-scale input layer and a side-output layer, to extract effective multi-scale features (see Fig. 12(d)). To further improve the performance of blood vessel segmentation, CE-Net combines the residual multi-kernel pooling module and the dense dilation convolution module together to expand the receptive field. Nevertheless, from Fig. 12(f), it can be seen that CE-

Net has more yellow pixels and red pixels. This means that the segmentation results of CE-Net not only have discontinuous segmentation, but also even have false segmentation. The main reason is that the vessel trees are difficult to extract accurately due to lacking both sufficient semantic information and global contextual information during the up-sampling process. Recently, SD-UNet is designed by introducing the edge information into U-Net to further guide network segmentation, while it has more yellow pixels than other approaches, as shown in Fig. 12(g).

Although ATTU-Net, M-Net, CE-Net, and SD-UNet can segment retinal blood vessels to some extent, they still cannot accurately segment tiny vessels. In contrast, our approach is able to explicitly expand the receptive field of the semantic segmentation network and utilize multi-scale feature maps, which are mainly attributed to the proposed CDSPP and MAFA modules. With the help of them, the network can fully utilize the semantic information at different scales and adaptively assign weight coefficients to different feature maps, aiming at enabling feature fusion at different scales. Meanwhile, thanks to the proposed fusion module and the IConv module, it can efficiently explore the contextual information and deeper semantic of the retinal blood vessels with varying shapes and sizes. In summary, our approach is generally

Table 6

The limitations of current methods and advantages of our method.

Methods	Limitations of current methods	Advantages of our method
U-Net [59]	The spatial information is missing for low-contrast thin vessels caused by the down-sampling operation. Besides, irrelevant regions are introduced, weakening the segmentation performance.	The channel attention mechanism is introduced to suppress the expression of irrelevant information. Besides, IConv and MAFA modules are used to achieve deep feature extraction and feature reuse.
M-Net [62]	The unstructured information, i.e., irregular shapes or rotations, cannot be fully explored.	CDSPP module allows the network to generate feature maps with different sizes, so as to better deal with irregular shapes or rotation problems.
ATTU-Net [61]	It is subjected to the receptive field, which can hardly extract the multi-scale features.	Both CDSPP and IConv modules can dynamically adjust the receptive field to extract the multi-scale features.
CE-Net [63]	The lack of sufficient semantic information and global contextual information during the up-sampling process will lead to capture the incomplete vessel structures.	MAFA module combines the high-level semantic information with the low-level detailed information to improve the accuracy of the vessel tree. Besides, different weights are assigned to feature maps at different scales to better capture the global information.
SD-UNet [70]	The context information and global information are ignored, making it difficult to accurately segment tiny blood vessels.	Dule-path framework and fusion module can deliver the context information, strengthening the completeness and continuity of microvascular.

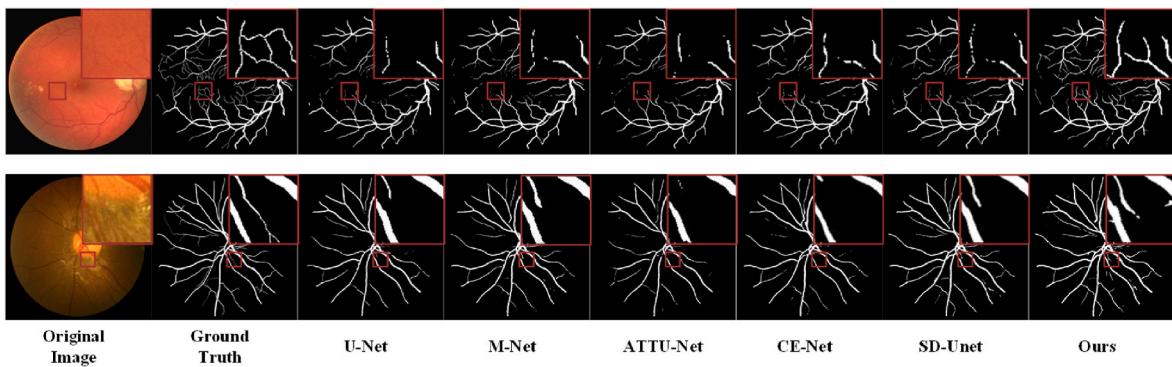


Fig. 13. Some failure cases.

superior to other approaches, as shown in Fig. 12(h).

To enable a comprehensive comparison of our method with existing approaches, we present Table 6, which highlights the limitations of current research and emphasizes the significance of the improvements achieved by our approach.

4.5. Discussion

From Tables 3–5, it can be observed that the *Se* metrics of all methods are relatively lower than other metrics. The main reasons are listed as below: On the one hand, the complex morphological structure of retinal blood vessels makes it difficult to achieve accurate segmentation. On the other hand, striking a balance between sensitivity and specificity in supervised training is hard. The higher value of *Sp* means that the model is more capable of accurately detecting non-vessel pixels in the background accurately, while the probability of vessel pixels could be falsely recognized as non-vessel pixels and will be increased accordingly. Meanwhile, sensitivity and overall accuracy will be affected [74]. The occurrence of this situation is unacceptable in supervised training. Hence, this paper chooses to sacrifice a small portion of the specificity in favor of higher sensitivity. As can be seen from these tables, although some approaches achieve higher results than our approach on certain datasets, they cannot perform better on all datasets. On the contrary, the proposed approach has a much higher sensitivity and relatively stable performance than the SOTA approaches on all datasets.

Despite the impressive performance demonstrated by our blood vessel segmentation model, it remains imperfect in certain extreme scenarios, such as those involving extreme low-contrast or vessels that are obscured by lesions, leading to their invisibility. Fig. 13 depicts a scenario where the model fails to accurately identify the entire vessel structure when segmenting tiny blood vessels within a low-contrast region. Similar to previous models, accurately identifying blood vessels becomes challenging when the vessel pixels are entirely concealed by large lesion regions. While our approach has been successful in identifying pixels that surpass SOTA methods in such extreme conditions, it still has significant room for improvement.

5. Conclusion

In order to tackle the complex challenges associated with inadequate contextual information and information loss concerning microvascular features, a novel dual-path deep learning framework has been proposed. This framework comprises three components: Preprocessing, MCDAU-Net, and Fusion. Initially, the fundus images are partitioned into a collection of patches with varying sizes. Subsequently, these patches are inputted into the MCDAU-Net, enabling the exploration of multi-scale contextual information and the acquisition of discriminative segmentation results. Lastly, a fusion module has been devised to explicitly facilitate the transfer of contextual pixel information surrounding each patch, augmenting the continuity of microvascular structures. Our

approach has undergone qualitative and quantitative analyses using three publicly available datasets, demonstrating a substantial improvement in blood vessel segmentation accuracy, particularly for small blood vessels.

While our segmentation results display promise and excellence, opportunities for enhancement remain. Notably, our model's effectiveness is limited in cases where fundus images exhibit extremely low-contrast or where blood vessels are obscured by lesions, resulting in their invisibility.

In the future, we plan to introduce other modality information, such as fluorescein angiography, to assist vessel segmentation. Moreover, we will explore a lightweight network model that can be embedded in medical devices, making this research more clinically applicable.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with the manuscript submitted to *Computers in Biology and Medicine* and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Nos. 62062040, 62006174, 62102270, 62262030), Natural Science Foundation of Liaoning province (No. 2023-MS-246), the Outstanding Youth Project of Jiangxi Natural Science Foundation (No. 20212ACB212003), the Jiangxi Province Key Subject Academic and Technical Leader Funding Project (No. 20212BCJ23017).

References

- [1] J.J. Kanski, B. Bowling, *Clinical Ophthalmology: a Systematic Approach*, Elsevier Health Sciences, 2011.
- [2] L. Li, M. Verma, Y. Nakashima, H. Nagahara, R. Kawasaki, Iternet: retinal image segmentation utilizing structural redundancy in vessel networks, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020, pp. 3656–3665.
- [3] C. L Srinidhi, P. Aparna, J. Rajan, Recent advancements in retinal vessel segmentation, *J. Med. Syst.* 41 (2017) 1–22.
- [4] C. Wang, M. Oda, Y. Hayashi, Y. Yoshino, T. Yamamoto, A.F. Frangi, K. Mori, Tensor-cut: a tensor-based graph-cut blood vessel segmentation method and its application to renal artery segmentation, *Med. Image Anal.* 60 (2020), 101623.
- [5] Z. Li, X. Zhang, H. Müller, S. Zhang, Large-scale retrieval for medical image analytics: a comprehensive review, *Med. Image Anal.* 43 (2018) 66–84.
- [6] H. Wu, W. Wang, J. Zhong, B. Lei, Z. Wen, J. Qin, SCS-net: a scale and context sensitive network for retinal vessel segmentation, *Med. Image Anal.* 70 (2021), 102025.
- [7] X. Jiang, D. Mojon, Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (1) (2003) 131–137.
- [8] A.M. Mendonca, A. Campilho, Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction, *IEEE Trans. Med. Imag.* 25 (9) (2006) 1200–1213.

- [9] S. Roychowdhury, D.D. Koozekanani, K.K. Parhi, Iterative vessel segmentation of fundus images, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 62 (7) (2015) 1738–1749.
- [10] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, M. Goldbaum, Detection of blood vessels in retinal images using two-dimensional matched filters, *IEEE Trans. Med. Imag.* 8 (3) (1989) 263–269.
- [11] F. Zana, J.C. Klein, Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation, *IEEE Trans. Image Process.* 10 (7) (2001) 1010–1019.
- [12] U.T. Nguyen, A. Bhuiyan, L.A. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, *Pattern Recogn.* 46 (3) (2013) 703–715.
- [13] J.I. Orlando, E. Prokofyeva, M.B. Blaschko, A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 64 (1) (2016) 16–27.
- [14] E. Ricci, R. Perfetti, Retinal blood vessel segmentation using line operators and support vector classification, *IEEE Trans. Med. Imag.* 26 (10) (2007) 1357–1365.
- [15] D. Marín, A. Aquino, M.E. Gegúndez-Arias, J.M. Bravo, A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features, *IEEE Trans. Med. Imag.* 30 (1) (2010) 146–158.
- [16] J.V. Soares, J.J. Leandro, R.M. Cesari, H.F. Jelinek, M.J. Cree, Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification, *IEEE Trans. Med. Imag.* 25 (9) (2006) 1214–1222.
- [17] R. Annunziata, E. Trucco, Accelerating convolutional sparse coding for curvilinear structures segmentation by refining SCIRD-TS filter banks, *IEEE Trans. Med. Imag.* 35 (11) (2016) 2381–2392.
- [18] Z. Wu, S. Shen, X. Lian, X. Su, E. Chen, A dummy-based user privacy protection approach for text information retrieval, *Knowl. Base Syst.* 195 (2020), 105679.
- [19] Z. Wu, S. Xuan, J. Xie, C. Lin, C. Lu, How to ensure the confidentiality of electronic medical records on the cloud: a technical perspective, *Comput. Biol. Med.* 147 (2022), 105726.
- [20] Z. Wu, S. Shen, H. Li, H. Zhou, C. Lu, A basic framework for privacy protection in personalized information retrieval: an effective framework for user privacy protection, *J. Organ. End User Comput.* 33 (6) (2021) 1–26.
- [21] Z. Wu, G. Li, S. Shen, X. Lian, E. Chen, G. Xu, Constructing dummy query sequences to protect location privacy and query privacy in location-based services, *World Wide Web* 24 (2021) 25–49.
- [22] Z. Wu, S. Shen, H. Zhou, H. Li, C. Lu, D. Zou, An effective approach for the protection of user commodity viewing privacy in e-commerce website, *Knowl. Base Syst.* 220 (2021), 106952.
- [23] Z. Wu, J. Xie, S. Shen, C. Lin, G. Xu, E. Chen, A confusion method for the protection of user topic privacy in Chinese keyword based book retrieval, *ACM Transactions on Asian and Low-Resource Language Information Processing* 22 (5) (2023) 1–19.
- [24] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, S.A. Barman, An ensemble classification-based approach applied to retinal blood vessel segmentation, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 59 (9) (2012) 2538–2548.
- [25] C. Liu, P. Gu, Z. Xiao, Multiscale U-net with spatial positional attention for retinal vessel segmentation, *Journal of Healthcare Engineering* (2022), 5188362–5188372.
- [26] M.G. Cinsdikici, D. Aydin, Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm, *Comput. Methods Progr. Biomed.* 96 (2) (2009) 85–95.
- [27] Q. Li, J. You, D. Zhang, Vessel segmentation and width estimation in retinal images using multiscale production of matched filter responses, *Expert Syst. Appl.* 39 (9) (2012) 7600–7610.
- [28] M. Zardadi, N. Mehrshad, S.M. Razavi, Unsupervised segmentation of retinal blood vessels using the human visual system line detection model, *Journal of Information Systems and Telecommunication* 14 (4) (2016) 125–133.
- [29] C. Zhou, X. Zhang, H. Chen, A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model, *Comput. Methods Progr. Biomed.* 187 (2020), 105231.
- [30] J. Staal, M.D. Abràmoff, M. Niemeijer, M.A. Viergever, B. Van Ginneken, Ridge-based vessel segmentation in color images of the retina, *IEEE Trans. Med. Imag.* 23 (4) (2004) 501–509.
- [31] X. You, Q. Peng, Y. Yuan, Y.M. Cheung, J. Lei, Segmentation of retinal blood vessels using the radial projection and semi-supervised approach, *Pattern Recogn.* 44 (10–11) (2011) 2314–2324.
- [32] C.A. Lupascu, D. Tegolo, E. Trucco, FABC: retinal vessel segmentation using AdaBoost, *IEEE Trans. Inf. Technol. Biomed.* 14 (5) (2010) 1267–1274.
- [33] F. Arcadu, F. Benmansour, A. Maunz, J. Willis, Z. Haskova, M. Prunotto, Deep learning algorithm predicts diabetic retinopathy progression in individual patients, *NPJ Digital Medicine* 2 (1) (2019) 92.
- [34] L. Dai, L. Wu, H. Li, C. Cai, Q. Wu, H. Kong, W. Jia, A deep learning system for detecting diabetic retinopathy across the disease spectrum, *Nat. Commun.* 12 (1) (2021) 3242.
- [35] S. Moccia, E. De Momi, S. El Hadji, L.S. Mattos, Blood vessel segmentation algorithms-review of methods, datasets and evaluation metrics, *Comput. Methods Progr. Biomed.* 158 (2018) 71–91.
- [36] H. Fu, Y. Xu, S. Lin, D.W. Kee Wong, J. Liu, Deepvessel: retinal vessel segmentation via deep learning and conditional random field, in: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016: 19th International Conference* vol. 19, Springer International Publishing, Athens, Greece, 2016, pp. 132–139. October 17–21, 2016, Proceedings, Part II.
- [37] H. Fu, Y. Xu, D.W.K. Wong, J. Liu, Retinal vessel segmentation via deep learning network and fully-connected conditional random fields, in: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2016, April, pp. 698–701.
- [38] J. Hu, H. Wang, S. Gao, M. Bao, T. Liu, Y. Wang, J. Zhang, S-unet: a bridge-style u-net framework with a saliency mechanism for retinal vessel segmentation, *IEEE Access* 7 (2019) 174167–174177.
- [39] T.A. Soomro, A.J. Afifi, J. Gao, O. Hellwich, M.A. Khan, M. Paul, L. Zheng, November. Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network, in: *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, IEEE, 2017, pp. 1–8.
- [40] H. Zhao, H. Li, L. Cheng, Improving retinal vessel segmentation with joint local loss by matting, *Pattern Recogn.* 98 (2020), 107068.
- [41] L. Mou, L. Chen, J. Cheng, Z. Gu, Y. Zhao, J. Liu, Dense dilated network with probability regularized walk for vessel detection, *IEEE Trans. Med. Imag.* 39 (5) (2019) 1392–1403.
- [42] V. Sathananthavathi, G. Indumathi, Encoder enhanced atrous (EEA) unet architecture for retinal blood vessel segmentation, *Cognit. Syst. Res.* 67 (2021) 84–95.
- [43] P. Yin, H. Cai, Q. Wu, DF-Net: deep fusion network for multi-source vessel segmentation, *Inf. Fusion* 78 (2022) 199–208.
- [44] Y. Zhang, J. Fang, Y. Chen, L. Jia, Edge-aware U-net with gated convolution for retinal vessel segmentation, *Biomed. Signal Process Control* 73 (2022), 103472.
- [45] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, G. Yang, Hierarchical retinal blood vessel segmentation based on feature and ensemble learning, *Neurocomputing* 149 (2015) 708–717.
- [46] P. Liskowski, K. Krawiec, Segmenting retinal blood vessels with deep neural networks, *IEEE Trans. Med. Imag.* 35 (11) (2016) 2369–2380.
- [47] Z. Feng, J. Yang, L. Yao, Patch-based fully convolutional neural network with skip connections for retinal blood vessel segmentation, in: *2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2017, September, pp. 1742–1746.
- [48] Z. Yan, X. Yang, K.T. Cheng, Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 65 (9) (2018) 1912–1923.
- [49] Y. Wu, Y. Xia, Y. Song, Y. Zhang, W. Cai, Multiscale network followed network model for retinal vessel segmentation, in: *Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference*, Granada, Spain, September 16–20, 2018, Proceedings, Part II vol. 11, Springer International Publishing, 2018, pp. 119–126.
- [50] Y. Wu, Y. Xia, Y. Song, Y. Zhang, W. Cai, NFN+: a novel network followed network for retinal vessel segmentation, *Neural Network* 126 (2020) 153–162.
- [51] C. Wang, Z. Zhao, Y. Yu, Fine retinal vessel segmentation by combining Nest U-net and patch-learning, *Soft Comput.* 25 (2021) 5519–5532.
- [52] M.E. Gegúndez-Arias, D. Marín-Santos, I. Perez-Borrero, M.J. Vasallo-Vazquez, A new deep learning method for blood vessel segmentation in retinal images based on convolutional kernels and modified U-Net model, *Comput. Methods Prog. Biomed.* 205 (2021), 106081.
- [53] S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, K. Zuiderveld, Adaptive histogram equalization and its variations, *Comput. Vis. Graph. Image Process* 39 (3) (1987) 355–368.
- [54] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, A. Rabinovich, Going deeper with convolutions, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
- [55] M. Sun, K. Li, X. Qi, H. Dang, G. Zhang, Contextual information enhanced convolutional neural networks for retinal vessel segmentation in color fundus images, *J. Vis. Commun. Image Represent.* 77 (2021), 103134.
- [56] L.C. Chen, G. Papandreou, F. Schroff, H. Adam, Rethinking Atrous Convolution for Semantic Image Segmentation, 2017 arXiv preprint arXiv:1706.05587.
- [57] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4) (2017) 834–848.
- [58] A.D. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, *IEEE Trans. Med. Imag.* 19 (3) (2000) 203–210.
- [59] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference* vol. 18, Springer International Publishing, Munich, Germany, 2015, pp. 234–241. October 5–9, 2015, Proceedings, Part III.
- [60] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, T. Wang, A cross-modality learning approach for vessel segmentation in retinal images, *IEEE Trans. Med. Imag.* 35 (1) (2015) 109–118.
- [61] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, D. Rueckert, Attention U-Net: Learning where to Look for the Pancreas, 2018 arXiv preprint arXiv:1804.03999.
- [62] H. Fu, J. Cheng, Y. Xu, D.W.K. Wong, J. Liu, X. Cao, Joint optic disc and cup segmentation based on multi-label deep network and polar transformation, *IEEE Trans. Med. Imag.* 37 (7) (2018) 1597–1605.
- [63] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, J. Liu, Ce-net: context encoder network for 2d medical image segmentation, *IEEE Trans. Med. Imag.* 38 (10) (2019) 2281–2292.
- [64] T. Kittrungrotakul, X.H. Han, Y. Iwamoto, L. Lin, A.H. Foruzan, W. Xiong, Y. W. Chen, VesselNet: a deep convolutional neural network with multi pathways for robust hepatic vessel segmentation, *Comput. Med. Imag. Graph.* 75 (2019) 74–83.

- [65] W. Wang, W. Wang, Z. Hu, Retinal vessel segmentation approach based on corrected morphological transformation and fractal dimension, *IET Image Process.* 13 (13) (2019) 2538–2547.
- [66] D.A. Palanivel, S. Natarajan, S. Gopalakrishnan, Retinal vessel segmentation using multifractal characterization, *Appl. Soft Comput.* 94 (2020), 106439.
- [67] J. Yang, C. Lou, J. Fu, C. Feng, Vessel segmentation using multiscale vessel enhancement and a region based level set model, *Comput. Med. Imag. Graph.* 85 (2020), 101783.
- [68] J. Yang, M. Huang, J. Fu, C. Lou, C. Feng, Frangi based multi-scale level sets for retinal vascular segmentation, *Comput. Methods Progr. Biomed.* 197 (2020), 105752.
- [69] Z. Yan, X. Yang, K.T. Cheng, A three-stage deep learning model for accurate retinal vessel segmentation, *IEEE Journal of Biomedical and Health Informatics* 23 (4) (2018) 1427–1436.
- [70] P. Yin, R. Yuan, Y. Cheng, Q. Wu, Deep guidance network for biomedical image segmentation, *IEEE Access* 8 (2020) 116106–116116.
- [71] Z. Fan, J. Lu, C. Wei, H. Huang, X. Cai, X. Chen, A hierarchical image matting model for blood vessel segmentation in fundus images, *IEEE Trans. Image Process.* 28 (5) (2018) 2367–2377.
- [72] H.A. Leopold, J. Orchard, J.S. Zelek, V. Lakshminarayanan, PixelBNN: augmenting the PixelCNN with batch normalization and the presentation of a fast architecture for retinal vessel segmentation, *Journal of Imaging* 5 (2) (2019) 26.
- [73] S.K. Saroj, V. Ratna, R. Kumar, N.P. Singh, Efficient Kernel Based Matched Filter Approach for Segmentation of Retinal Blood Vessels, 2020, 03601 arXiv preprint arXiv:2012.
- [74] T.J. Jebaseeli, C.A.D. Durai, J.D. Peter, Extraction of retinal blood vessels on fundus images by kirsch's template and Fuzzy C-Means, *J. Med. Phys.* 44 (1) (2019) 21.