CSU-Net: A Context Spatial U-Net for Accurate Blood Vessel Segmentation in Fundus Images

Bo Wang, Shengpei Wang, Shuang Qiu, Wei Wei, Haibao Wang, and Huiguang He, *Senior Member, IEEE*

***Abstract*—Blood vessel segmentation in fundus images is a critical procedure in the diagnosis of ophthalmic diseases. Recent deep learning methods achieve high accuracy in vessel segmentation but still face the challenge to segment the microvascular and detect the vessel boundary. This is due to the fact that common Convolutional Neural Networks (CNN) are unable to preserve rich spatial information and a large receptive field simultaneously. Besides, CNN models for vessel segmentation usually are trained by equal pixel level cross-entropy loss, which tend to miss fine vessel structures. In this paper, we propose a novel Context Spatial U-Net (CSU-Net) for blood vessel segmentation. Compared with the other U-Net based models, we design a two-channel encoder: a context channel with multi-scale convolution to capture more receptive field and a spatial channel with large kernel to retain spatial information. Also, to combine and strengthen the features extracted from two paths, we introduce a feature fusion module (FFM) and an attention skip module (ASM). Furthermore, we propose a structure loss, which adds a spatial weight to cross-entropy loss and guide the network to focus more on the thin vessels and boundaries. We evaluated this model on three public datasets: DRIVE, CHASE-DB1 and STARE. The results show that the CSU-Net achieves higher segmentation accuracy than the current state-of-the-art methods.**

***Index Terms*—Fundus images, blood vessel segmentation, CSU-Net, feature fusion, structure loss.**

Manuscript received February 3, 2020; revised May 17, 2020 and June 20, 2020; accepted July 18, 2020. Date of publication July 22, 2020; date of current version April 5, 2021. This work was supported in part by The National Key Research and Development Program of China under Grant 2017YFB1302704, in part by the National Natural Science Foundation of China under Grants 61976209 and 81701785, in part by CAS International Collaboration Key Project (No. 173211KYSB20190024), and in part by the Strategic Priority Research Programof CAS under Grant XDB32040200. *(Corresponding author: Huiguang He.)*

Bo Wang is with the School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China, and also with Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China (e-mail: wangbo2018@ia.ac.cn).

Shengpei Wang, Wei Wei, and Haibao Wang are with Research Center for Brain-Inspired Intelligence, the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China (e-mail: wangshengpei2014@ia.ac.cn; weiwei2018@ia.ac.cn; haibaow@hotmail.com).

Shuang Qiu is with Research Center for Brain-Inspired Intelligence, the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China (e-mail: shuang.qiu@ia.ac.cn).

Huiguang He is with Research Center for Brain-Inspired Intelligence, the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China, and with Center for Excellence in Brain Science and Intelligence Technology, the Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China (e-mail: huiguang.he@ia.ac.cn).

Digital Object Identifier 10.1109/JBHI.2020.3011178

I. Introduction

ANALYSIS of blood vessel structure from fundus image has been widely used in diagnosing ophthalmic diseases caused by arteriosclerosis, diabetic retinopathy, glaucoma or stroke [1]. To analyze the morphology of the blood vessel, it is essential to segment vessel accurately from fundus images, which can completely reflect the variations in vessel width, vessel density and vessel branching patterns. Since manual annotation of vessels is a tricky and time-consuming work, automatic segmentation of vessels from fundus images become particularly important.

As fully convolutional neural network (FCN) [2] has achieved great success in segmentation tasks on natural images, various FCN variants, such as PSPNet [3] and DeepLab [4], make the most advanced performance in the benchmark semantic segmentation tasks. Of all the variants, U-Net [5] becomes the most widely used model and provides excellent performance in medical image analysis. Using encoder to extract the feature representations and decoder to increase the feature resolution, U-Net is designed to capture the precise localization by combining low-level and high-level features, which can achieve better results from small datasets. In recent years, various U-Nets have also been proposed. Ozan *et al.* [6] added an attention module in U-Net to determine the importance of the skip map. A recurrent U-Net was proposed by Alom *et al.* [7], where encoder layers are replaced by the recurrent convolution operation. However, all these variants use the backbone designed for classification as the encoder, which makes the model unable to preserve rich spatial information and large receptive field in segmentation task simultaneously. Besides, all these U-Net models are trained by equal pixel-level cross-entropy loss, which guides models to emphasize more on thick vessels.

There are two main contributions in our paper. First, we propose a Context Spatial U-Net (CSU-Net) for vessel segmentation. The modified U-Net has a two-channel encoder: a context channel with multi-scale convolution to capture more receptive field and a spatial channel with large kernel to retain spatial

2168-2194 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

TABLE I
SUMMARY OF THE AUTOMATIC RETINAL VESSEL SEGMENTATION METHODS

Reference	Year	Vessel segmentation technique	Dataset	Evaluation metrics	Categories
Chaudhuri et al. [8]	1989	Gaussian matched filtering	DRIVE	AC, AUC	Image processing methods
Hoover et al. [9]	2000	Matched filtering and threshold probing	DRIVE	SP, SE, AC	
Al-Rawi et al. [10]	2007	Improved gaussian matched filter	DRIVE	AUC, AC	
Fraz et al. [11]	2012	Morphology-based method	DRIVE, STARE	SE, SP, AC	
Zardadi et al. [12]	2016	Analysis of the Hessian matrix	STARE	SE, SP, AUC	
Jiang et al. [13]	2003	Multi-threshold probing strategy	DRIVE	SE, AC	
Al-Diri et al. [14]	2009	Active contour model	DRIVE, STARE	SE, SP	
Staal et al. [15]	2004	k-NN classifier	DRIVE, STARE	AC, AUC	Shallow learning methods
You et al. [16]	2011	Support vector machine	DRIVE, STARE	SE, SP, AC	
Soares et al. [17]	2006	Gaussian mixture model	DRIVE, STARE	AC, AUC	
Lupascu et al. [18]	2010	AdaBoost classifier	DRIVE	SE, AC, AUC	
Marin et al. [19]	2010	Neural network classifier	DRIVE, STARE	SE, SP, AC, AUC	
Aslani et al. [20]	2016	Random forest classifier	DRIVE, STARE	SE, SP, AC, AUC	
Zhu et al. [21]	2016	Classification and regression tree classifier	DRIVE	SE, SP, AC	
Liskowski et al. [22]	2016	Deep learning model	DRIVE, STARE	SE, SP, AC, AUC	Deep learning methods
Fu et al. [23]	2016	Convolutional neural networks and CRF	DRIVE, STARE	SE, AC	
Gao et al. [24]	2017	U-Net	DRIVE	SE, SP, AC, PR	
Xiao et al. [25]	2017	Weighted Res-UNet	DRIVE, STARE	SE, AC	
Wu et al. [26]	2018	MS-NFN model	DRIVE, CHASEDB1	SE, SP, AC, AUC	
Yan et al. [27]	2018	U-Net with jointed loss	DRIVE, CHASEDB1, STARE	SE, SP, AC, PR	
Zhuang et al. [28]	2018	LadderNet	DRIVE, CHASEDB1	SE, SP, AC, AUC, F1	
Jin et al. [29]	2019	Deformable U-Net	DRIVE, CHASEDB1, STARE	AC, AUC	
Alom et al. [7]	2019	Recurrent U-Net	DRIVE, CHASEDB1, STARE	SE, SP, AC, AUC	

information. To combine the features from two channels, we build a feature fusion module (FFM). Also, an attention skip module (ASM) is designed to strengthen the feature representation and reduce the redundancy. The second major contribution is that we introduce a new structure loss, which assigns different weights to different positions. It can mine better structural information contained in the features and make the model focus more on the thin vessels and boundaries.

The remaining sections of this paper are organized as follows. The existing automated vascular segmentation algorithms are reviewed in Section II. In Section III, we introduce the datasets and the proposed CSU-Net model. In Section IV, we present the experimental results and analyze the performance of the CSU-Net in vessel segmentation. Finally, Section V conducts the conclusions.

II. RELATED WORK

As the continuous advancement of computer vision technology, many automatic vessel segmentation methods in fundus images have been proposed. We divide these algorithms into two broad classifications: image processing methods and machine learning methods. All the reviewed papers are summarized in Table I. The dataset and performance metrics in the table will be illustrated with Section III and IV.

A. Image Processing Methods

In the image processing methods, algorithms that follow classical computer vision approach, like matched filtering (MF), morphology-based and model-based algorithms. The MF for improving blood vessel segmentation was first adopted in [8]. With a matched filtering method, 12 different templates are used for the fundus image. Then, Hoover and Kouznetsova [9] combined the threshold probing technique and local region-based properties for vessel segmentation. Al-Rawi *et al.* [10] proposed an all-purpose optimization technique to ascertain better filter parameters, which improved the performance of a Gaussian MF. The morphology-based method presented in [11] segment vessels from fundus image by combing vessel centerlines regression and morphological bit plane slicing. Zardadi *et al.* [12] conducted classification on pixel through a threshold method after that executed morphological post-processing. In [13], a model-based method was applied, which adopt local threshold and integrated vessel information into the verification stage. Al-Diri and Mojon [14] used the active contour model, which has two pairs of contours to capture each vessel edge. Image processing methods are applicable to a large amounts of unlabeled data in retinal vessel task. However, because of the sensitivity to abnormal retina image, image processing methods suffer from low accuracy and generalization.

B. Machine Learning Methods

Methods based on machine learning that take advantage of ground truth data onto vessel segmentation can be classified: shallow machine learning method and deep learning-based method.

1) Shallow Machine Learning Methods: Normally, machine learning methods extract features by handcraft and utilize machine learning algorithms like kNN [15] and SVM [16] for vessel segmentation. For instance, Soares *et al.* [17] exploited pixel luminosity together with 2-dimensional Gabor wavelet transformation to extract feature vector and used Bayesian classifier for vessel segmentation. Lupascu *et al.* [18] extracted a 41-dimensional feature vector in each pixel and trained an AdaBoost classifier for vessel segmentation. Marin *et al.* [19]

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

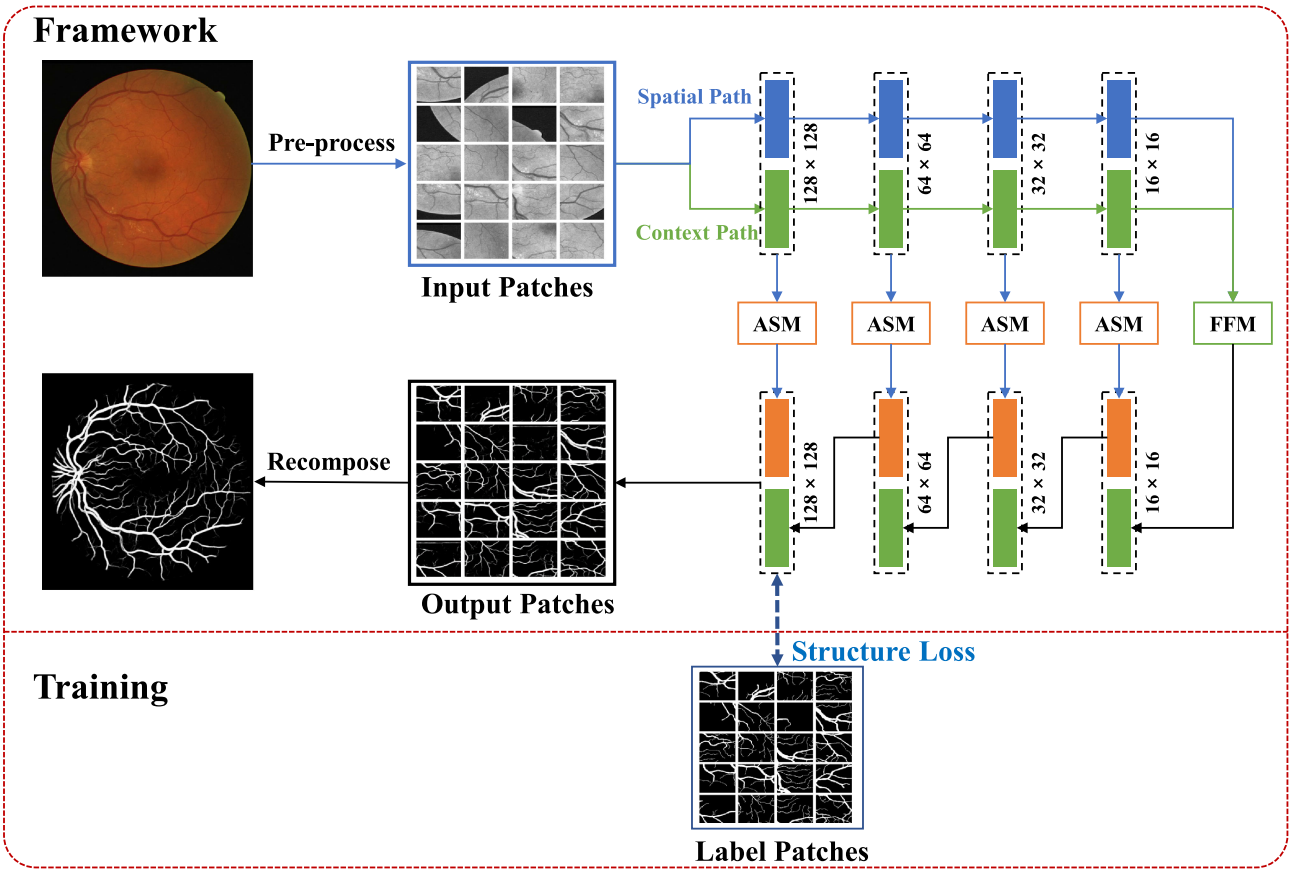


Fig. 1. The overview of the proposed Context Spatial U-Net framework. It consists of three cascaded stages : 1) fundus image pre-processing and patches extraction. 2) feeding patches into the CSU-Net for predication. 3) predication results reconstruction. The proposed model has encoder and decoder, which is a U-Net. In encoder, a context channel with multi-scale convolution to capture more receptive field and a spatial channel having large kernel to retain spatial information. Besides, FFM and ASM are applied to strengthen the feature representation and reduce the redundancy. In training stage, structure loss is using to update the network weights.

adopted a 7-dimensional feature vector to train neural network. Aslani *et al.* [20] used different extractors to calculate hybrid feature vector on pixels, which was classified by random forest classifier. Zhu *et al.* [21] extracted a 36-dimensional feature vector of each pixel in Green channel, and classify vessels pixels with classification and regression tree.

2) Deep Learning Methods: In the past few years, many researches have focus more on deep learning-based models to solve the problem of blood vessel segmentation. Liskowski *et al.* [22] proposed a deep neural network (DNN) for vessel segmentation and achieved high accuracy on the DRIVE dataset. Fu *et al.* [23] combined DNN and a fully connected Conditional Random Fields (CRFs) and regarded the vessel segmentation problem as a boundary detection task. U-Net [5] has upsampling operators to enlarge the resolution of output and combines low-level and high-level features to capture precise localization. The encoder to decoder architecture with skip connections enables efficient information flow and presents better performance in medical image analyses. There is some effective improvements based on U-Net in [7], [24], [25]. A multi-scale network followed network (MS-NFN) was proposed [26] to segment blood vessel. Each submodel of MS-NFN includes two identical U-Net. Yan *et al.* [27] proposed a new loss function, which jointed the segment-level and pixel-wise to train a U-Net and improved the ability to segment thin vessels. Zhuang *et al.* [28] proposed LadderNet, a multi-branch CNN for semantic segmentation, which has multiple paths of information flows. Jin and Meng [29] proposed a Deformable U-Net for blood vessel segmentation, and a U-shape structure was employed to exploit the local features of retinal vessels. In our work, we focus on deep learning models for accurate blood vessel segmentation in the fundus image.

III. Methodology

In this paper, a CSU-Net model is proposed for blood vessel segmentation. The procedure consists of the following steps: (1) fundus image pre-processing and patches extraction (2) feeding patches into CSU-Net for predication, (3) predication results reconstruction. The overview of the framework of CSU-Net is shown in Fig. 1.

A. Materials

The proposed CSU-Net is trained and tested on three datasets: DRIVE [30], CHASE [31], and STARE [9]. The DRIVE includes 40 color fundus images. Seven images have mild signs of

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

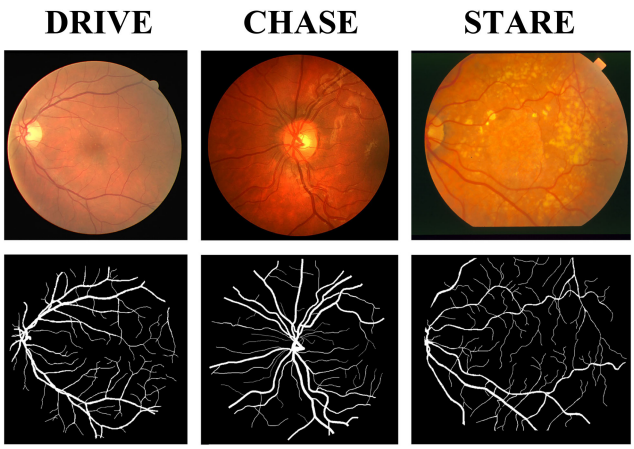


Fig. 2. An example of fundus images and labels (annotated by the first expert) from three datasets.

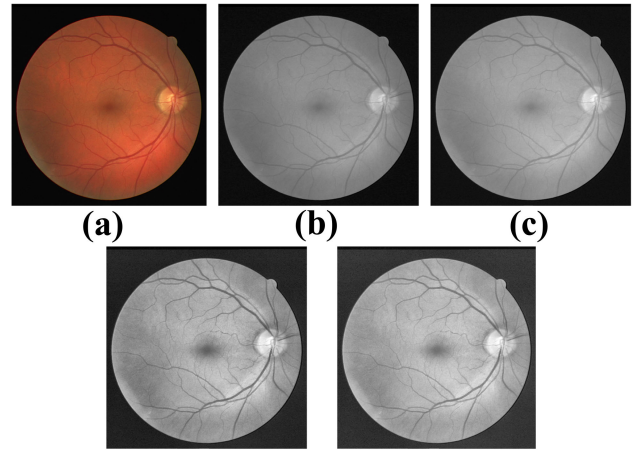


Fig. 3. Sample retinal images in the preprocessing stage. (a) the original fundus image. (b) grayscale. (c) normalization. (d) CLAHE. (e) gamma adjustment.

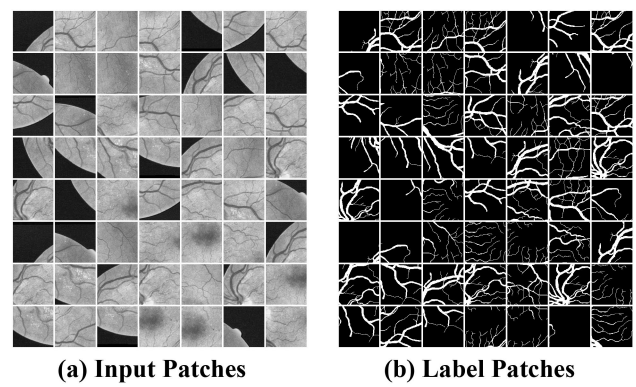


Fig. 4. Example of the input patches are shown in (a) and the corresponding label patches are shown in (b).

diabetic retinopathy. The other images are from healthy subjects. We train on the last 20 images (including 3 abnormal images) and test on the first 20 images (including 4 abnormal images). Each image has 584×565 pixels. The CHASE dataset has 28 color fundus images. The first 20 images are used for training, and the others are for testing, whose size is 999×960. STARE contains 20 fundus images, where 10 of which represent patients with retinal abnormalities, and the remaining ones represent normal retina. The size of each image is 605×700. Empirically, we divide the STARE randomly into training and testing images in the ratio of 10/10. Training part has four abnormal images and the other ones exist in testing set. Each test sample of all datasets is annotated by two experts. We follow the practice of using the results of the first expert as label and verifying the effectiveness of our method compared with the second expert results [27]. Fig. 2 presents an example of fundus images and labels annotated by the first expert from three datasets.

B. Fundus Image Pre-processing and Patch Extraction

The fundus image contains noise and inhomogeneous illumination, so it's essential to make pre-processing. Firstly, we transformed the fundus image from color to gray [32]. The formula is:

$$X_{\text{gray}} = 0.299 \times R + 0.578 \times G + 0.114 \times B \tag{1}$$

where R, G, B represent the three color channels. The gray image is preserved in Fig. 3(b). The second step is data normalization. For every image X, we should make each pixel of the X with zero-mean and unit-variance.

$$X_{\text{norm}} = \frac{X - \mu}{\sigma} \tag{2}$$

where μ and σ represent the mean value and standard variance of X. Then, image data must be scaled to [0,255].

$$X = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \times 255 \tag{3}$$

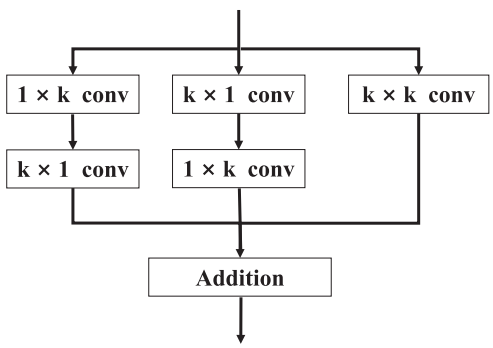
where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of all pixels in the image X. The normalized image is presented in Fig. 3(c). Finally, we use contrast limited adaptive histogram equalization (CLAHE) [33] and gamma adjustment to enhance the image's foreground and background contrast, which is shown in Fig. 3(d)(e). Finally, the intensity values of the image are mapped to a range of 0 to 1.

In the training stage, we randomly extract 1000 patches of size 128×128 from each fundus image, so that the training dataset contains 20 000 patches of DRIVE, 20 000 patches of CHASE and 10 000 patches of STARE. For testing, we sample patches with the stride of 15 from each test fundus image, which can make the balance of performance and speed. The example of the input patches and the corresponding label patches are shown in Fig. 4.

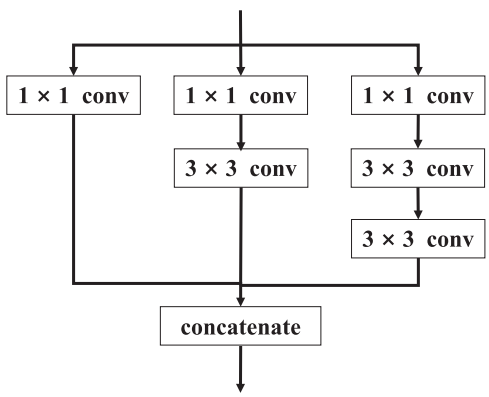
C. The Proposed Model

To segment vessel more accurately, we proposed a CSU-Net for blood vessel segmentation in fundus images. Fig. 1 clarifies the architecture of the model. The CSU-Net has an encoder and a decoder, so it is also a variant of the U-shaped model. In the encoder, a context channel is designed to capture more receptive field and a spatial channel is adopted to retain spatial information. Besides, spatial attention mechanism and feature

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.



(a) Baseblock in Spatial Channel



(b) Baseblock in Context Channel

Fig. 5. Baseblock of spatial channel and context channel in CSU-Net. (a) Baseblock in Spatial Channel. (b) Baseblock in Context Channel.

fusion idea are adopted to strengthen the feature representation and reduce the redundancy.

1) Spatial Channel: In semantic tasks, it is important for the model whether it can capture rich spatial information and large receptive field. Common CNN models use the backbone designed for classification as encoder, which unable to satisfy the two requirements simultaneously. According to the Global Convolutional Network [34], a spatial channel using large kernel is designed especially for preserve spatial information on original picture. The spatial channel contains four layers. Each layer consists of a baseblock, as shown in Fig. 5(a), and one max pooling layer. The baseblock has three kinds convolutions: $k \times k$, $1 \times k$ followed by $k \times 1$, and $k \times 1$ followed by $1 \times k$, and each convolution also followed by batch normalization and relu. Since too large convolution kernels will increase the computational complexity, we set the convolution kernels of the four layers to 7, 7, 5, and 5 in sequence.

2) Context Channel: To capture a larger receptive field, we build a context channel to get more semantic information. The spatial channel also contains four layers. The baseblock of a layer is presented in Fig. 5(b). Unlike spatial channel, we combine small kernels like 1×1 and 3×3 to capture more small scale features, and it can increase the semantic capture ability of the network and also reduce the parameters.

3) Attention Skip Module: In skip connection, we design an Attention Skip Module (ASM) with self-spatial attention mechanism to strengthen the original features. As illustrated in Fig. 6, we first estimate the spatial and the context feature maps from two encoder paths. Secondly, we apply the element-wise addition on two features to fuse spatial and context information. Then, we send the fusion feature to 3×3 convolution $\rightarrow$ BatchNormalization $\rightarrow$ Sigmoid to learn an attention map for all positions. Their channels are decreased to 1 but the same height and width compared with the input size. Third, we aggregate the original spatial and context features of all positions by element-wise sum with the weights defined by the attention map. ASM is based on spatial attention mechanism, which can strengthen the feature representation and reduce redundancy.

4) Feature Fusion Module: Since the context channel capture more semantic information and spatial channel preserve more spatial information, the level of feature from two channels is different. For extracting a better feature from the encoder, we designed a feature fusion module (FFM) to fuse features from two channels, which is shown in Fig. 6. Firstly, 3×3 convolution is applied to spatial and context features respectively to adapt them for follow-up processing. Then, these features are fused by element-wise multiplication and the fused features are sent to other 3×3 convolutional layers for learning a refined representations. Finally, the refined representations are added to the original spatial and context features. This module can fuse two paths and avoid adding too much redundant information.

D. The Proposed Structure Loss

The pixel proportion of the thin and thick vessels is highly imbalanced in fundus images. Common CNN models are trained by equal pixel level cross-entropy loss, which treat all vessel pixels with equal importance and tend to miss fine vessels structures. To increase the ability of models to segment thin vessels accurately and provide clear local boundary details, a new structure loss is introduced to train the model. Unlike common cross-entropy loss, structure loss will first compute weight and put greater attention on the position of the thin vessels and boundaries.

Let us denote by W_{ce} a spatial weight map for the cross-entropy loss for image X, computed by the corresponding ground truth map Y_{mask}. Based on the manual ground truth map Y_{mask}, we first apply the average-pooling operation $P^{avg}(\cdot)$ with kernel size K to obtain the corresponding attention map A. The formula is:

$$A = P_K^{avg}(Y_{\text{mask}}) - Y_{\text{mask}} \tag{4}$$

where attention map A has the same size as Y_{mask}. $A_{i,j}$ represents the attention weight at pixel (i, j), and $A_{i,j} \in [0, 1]$ for all pixels. If the value of $A_{i,j}$ is large, the pixel at (i, j) is very different from other pixels around it and deserves more attention. Fig. 7 shows the attention map on three datasets with different K values. We observe that pixels located at the thin vessels and boundaries have larger attention weight on all datasets. Besides, the kernel size K controls the boundary sensitivity of the attention map. When K equals to 3, the boundary is too narrow, and when $K = 7$, the boundary becomes blurry. In our experiments, we set $K = 5$ according to the boundary sensitivity

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

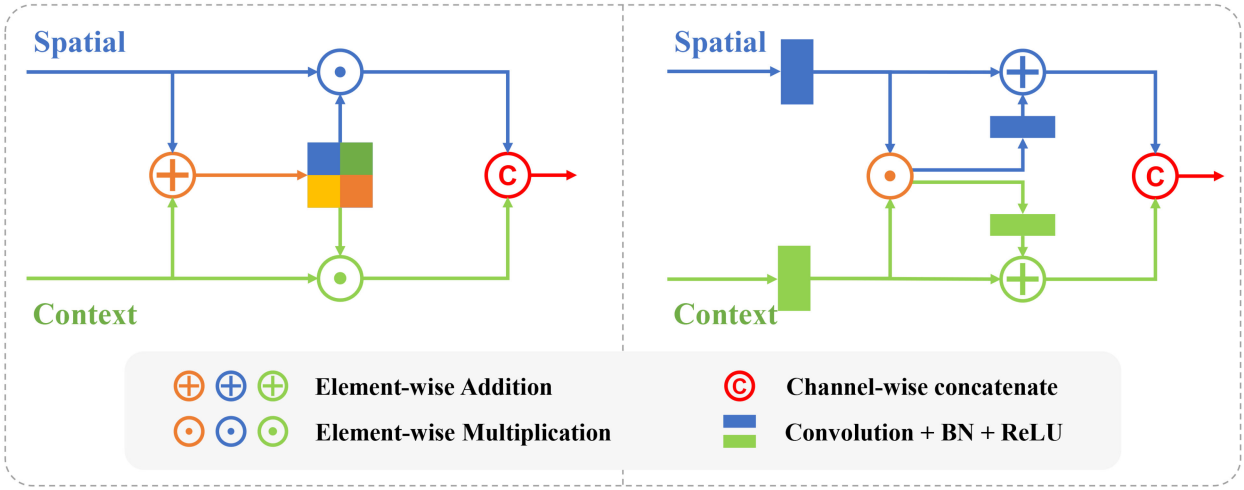


Fig. 6. Structure of ASM (left) and FFM (right) in CSU-Net. ASM is designed to reduce redundancy. FFM is designed to fusion two path's information and make a better representation.

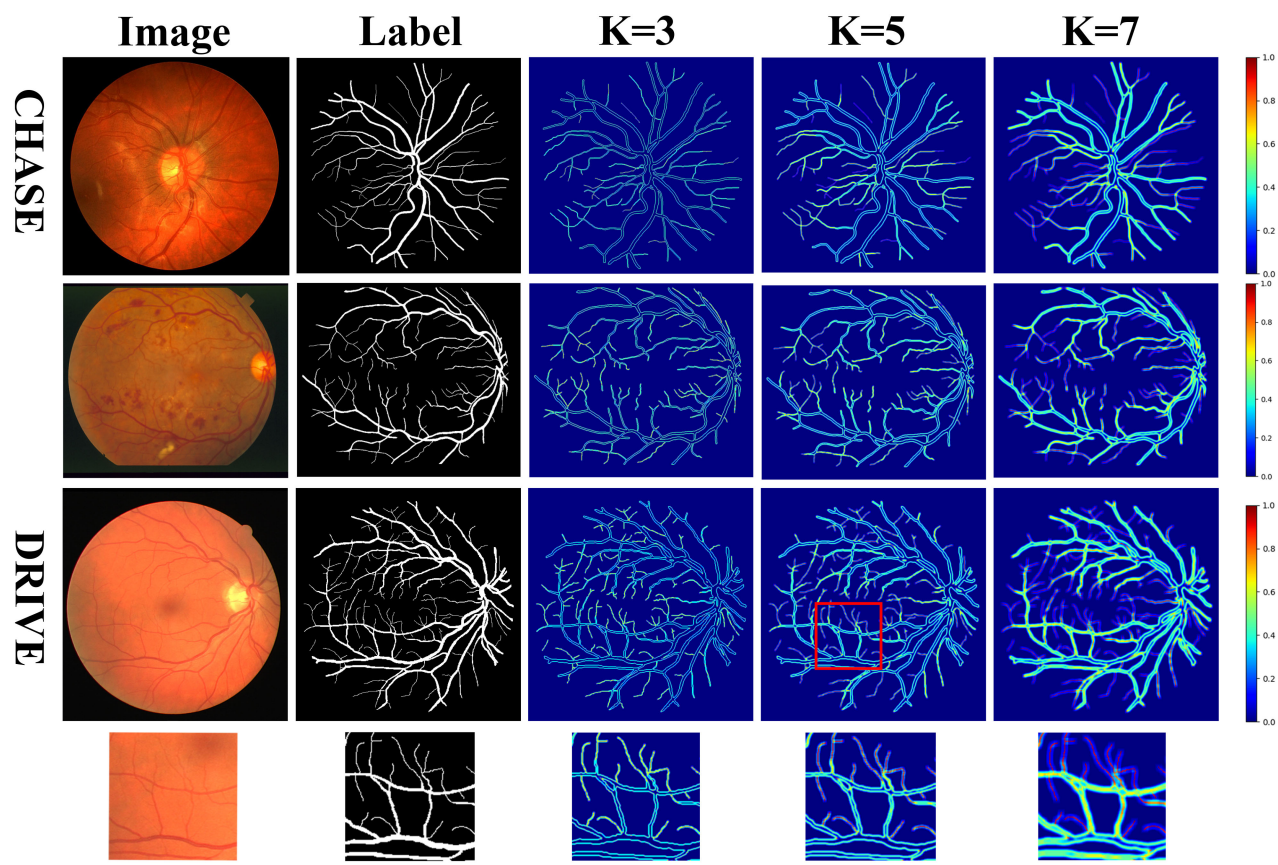


Fig. 7. Some examples of the attention map A. From left to right: the original fundus image, mask, the attention map ($K = 3$), the attention map ($K = 5$), the attention map ($K = 7$). We can see that pixels located at thin vessels and boundaries have larger attention weight.

and the validation performance. Then the spatial weight map for the cross-entropy loss W_{ce} can be formulated as:

$$W_{ce} = |1 + \beta \times A| \tag{5}$$

where $|\cdot|$ is the absolute value function. β is a hyperparameter for balancing the proportion of the attention map and the background in spatial weight map. We use $\beta = 5$ according to the validation performance. With the learned attention information, we build strong attention for pixels belonging to the thin vessels and boundaries. During propagation, more information is passed via strong weights around the boundaries and thin vessels. In this way, local structural information has been encoded

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

into cross-entropy loss, which can help the model focus on the important.

Finally, the structure loss function can be expressed as,

$$Loss_{str}(Y,\hat{Y}) = -\sum W_{ce}^{i} \times \left[Y_i \log \hat{Y}_i + (1-Y_i)\log(1-\hat{Y}_i)\right] \quad (6)$$

where Y_i is the label, and $\hat{Y}_i$ is predicted vector. The impact of hyperparameters will be analyzed in Section IV.

IV. Results and Discussion

In this section, we systematically analyze the performance of CSU-Net in the following content. First of all, we introduce the evaluation metrics and the setup in the training stage. Secondly, we show the performance of the model as a whole and present the segmentation results of the test set. Then, we compare CSU-Net with some other methods, most of which are based on deep leading algorithm. Finally, We conduct the ablation experiments on the CSU-Net to prove the effectiveness of module designs.

A. Experimental Setup

1) Evaluation Metrics: Compared with the label, we divide the pixels in the predication probability picture into false positive (FP), true positive (TP), true negative (TN) and false negative (FN). Then, we can calculate F1-score, sensitivity (SE), specificity (SP) and accuracy (AC) to evaluate the performance of CSU-Net. Besides, the area under receiver operating characteristics curve (AUC) is also presented. In addition, Cohen's kappa coefficient [35] is used to measure the agreement between proposed algorithm results and the annotations by medical experts. Metrics are calculated as follows:

$$AC = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$SE = Recall = \frac{TP}{TP+FN} \quad (8)$$

$$SP = \frac{TN}{TN+FP} \quad (9)$$

$$Precision = \frac{TP}{TP+FP} \quad (10)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

$$Kappa = \frac{P_o - P_e}{1-P_e} \quad (12)$$

where P_o is observed accuracy and P_e is expected accuracy. The formulas are:

$$P_o = AC = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

$$P_e = \frac{(TP+FP)(TP+FN)+(FN+TN)(FP+TN)}{(TP+FP+FN+TN)^2} \quad (14)$$

2) Training Details: Our model is implemented with Pytorch framework on ubuntu 16.04 (64 Bit) platform, equipped with NVIDIA GTX 2080Ti. In the training stage, the model is trained by adam optimizer [36] with default parameters and minibatch size is 16. Besides, we use the plateau strategy in learning rate, which set 0.01, 0.001, 0.0001 on epochs 0, 25, and 50, respectively. The total learning epochs is 200.

TABLE II
Results of Normal and Abnormal Images on DRIVE and STARE Datasets

Dataset	Image	F1-score	SE	SP	AC	AUC
DRIVE	Normal	0.8263	0.8088	0.9783	0.9568	0.9798
	Abnormal	0.8204	0.8000	0.9778	0.9551	0.9813
	All images	0.8251	0.8071	0.9782	0.9565	0.9801
STARE	Normal	0.8696	0.8467	0.9889	0.9748	0.9828
	Abnormal	0.8262	0.8374	0.9778	0.9631	0.9816
	All images	0.8516	0.8432	0.9845	0.9702	0.9825

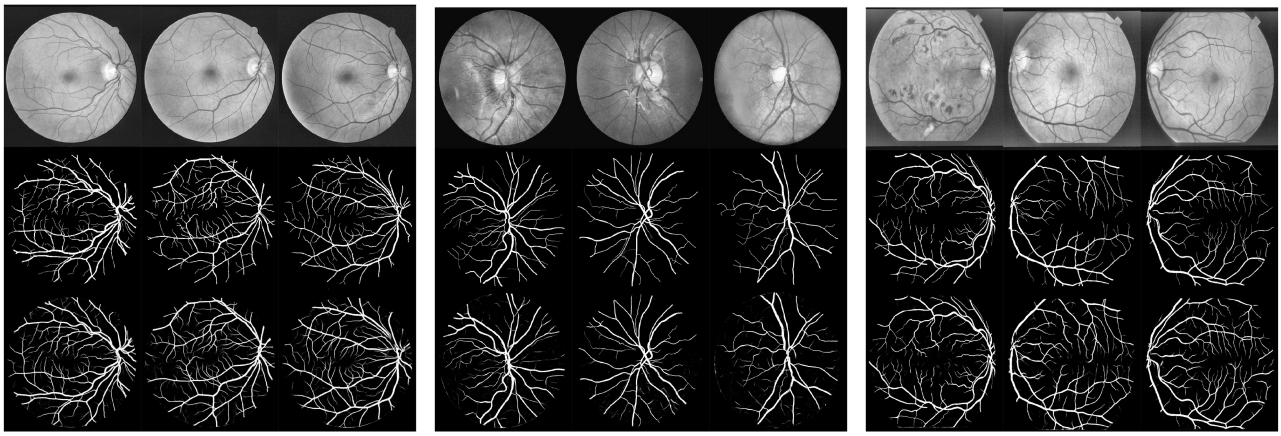
B. Overall Performance

In order to estimate the proposed CSU-Net, a serious experiments were conducted on the DRIVE, CHASE and STARE datasets. Part of the segmentation results of CSU-Net for all three datasets are shown in Fig. 8(a), 8(b) and 8(c). To further observe the performance of CSU-Net, we visualize the F1-score of each test images on three datasets, which is shown in Fig. 9. Although the CSU-Net was trained by using the annotations from the first observer, the model achieves improvement in F1-score against the second expert results, which means the gap between the model result and the label is smaller than inter-observers differences. Besides, CSU-Net achieves the highest mean values of F1-score on all three datasets, which indicates that the proposed model has high generalization ability and performs well in different kinds of fundus images on the blood vessel segmentation task.

Blood vessels may become more difficult to segment in the presence of disease like hemorrhages. For further analysis of the robustness of CSU-Net in such cases, we separate the test data from DRIVE and STARE into healthy and pathological group based on their diagnosis (CHASE does not contain diagnosis). Table II presents the performance comparison between the normal and abnormal cases. Compared with the normal images, most of indicators for abnormal case only drop a tiny fraction of percent on two datasets and even some indicators are higher than the normal. The result proves that the proposed CSU-Net performs high robustness on the pathological images.

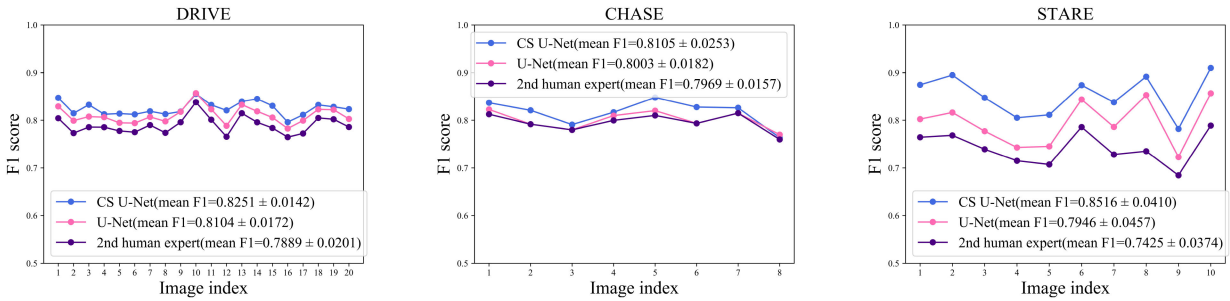
In order to demonstrate the agreement of proposed algorithm results with the annotations by medical experts, the Consistency of Diagnostic Test is performed. Since vessel segmentation is a binary classification task by distinguishing vessels from background, we use Cohen's kappa coefficient for evaluation. Generally, $0 \leq kappa \leq 1$, and two results have very high consistency when $kappa \geq 0.8$. The scores of kappa on the three data sets are: DRIVE (0.8103), CHASEDB1 (0.8244), STARE (0.8350), which indicates that the proposed model is highly consistent with the gold standard.

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.



(a) (b) (c)

Fig. 8. Test results in three datasets: (a) DRIVE; (b) CHASE; (c) STARE. From top to bottom: preprocessing image, label and the prediction of CSU-Net.



(a) (b) (c)

Fig. 9. The predication of the F1-score for each test images in three datasets: (a) DRIVE; (b) CHASE; (c) STARE.

TABLE III
COMPARISON OF CSU-NET WITH OTHER MODELS ON DRIVE

Approaches	Time	F1-score	SE	SP	AC	AUC
Second expert annotation	NaN	0.7889	0.7743	**0.9819**	**0.9637**	0.8781
You et al. [16]	2011	NaN	0.7410	0.9751	0.9434	0.9737
Fraz et al. [37]	2012	NaN	0.7152	0.9759	0.9430	NaN
Azzopardi et al. [38]	2015	NaN	0.7655	0.9704	0.9442	0.9614
Liskowsk et al. [22]	2016	NaN	0.7763	0.9768	0.9495	0.9720
Orlando et al. [39]	2016	0.7857	0.7897	0.9684	NaN	NaN
Dasgupta et al. [40]	2017	0.8074	0.7691	0.9801	0.9533	0.9744
Zhang et al. [41]	2017	0.7953	0.7861	0.9712	0.9466	0.9703
Xia et al. [42]	2018	NaN	0.7740	0.9800	0.9540	NaN
Zhuang et al. [28]	2018	0.8202	0.7856	0.9810	0.9561	0.9793
Alom et al. [7]	2019	0.8155	0.7751	0.9816	0.9556	0.9782
Context Spatial U-Net	2019	**0.8251**	**0.8071**	0.9782	0.9565	**0.9801**

TABLE IV
COMPARISON OF CSU-NET WITH OTHER MODELS ON CHASE DATASET

Approaches	Time	F1-score	SE	SP	AC	AUC
Second expert annotation	NaN	0.7969	0.8313	0.9829	**0.9733**	0.9071
Lam et al. [43]	2010	NaN	0.7585	0.9587	0.9387	0.9487
Li et al. [44]	2015	NaN	0.7507	0.9793	0.9581	0.9793
Ronneberger et al. [5]	2015	0.7783	0.8288	0.9701	0.9578	0.9772
Liskowsk et al. [22]	2016	NaN	0.7867	0.9754	0.9566	0.9785
Zhang et al. [41]	2017	0.7581	0.7644	0.9716	0.9502	0.9706
Zhuang et al. [28]	2018	0.8031	0.7978	0.9818	0.9656	**0.9839**
Jin et al. [29]	2019	0.7883	0.8155	0.9752	0.9610	0.9804
Alom et al. [7]	2019	0.7800	0.7726	0.9820	0.9553	0.9779
Context Spatial U-Net	2019	**0.8105**	**0.8427**	**0.9836**	0.9706	0.9824

C. Comparison With the Existing Methods

In our experiments, we compare the proposed CSU-Net with several other advanced approaches on DRIVE, CHASE, and STARE datasets, respectively. The quantitative results is presented in Tables III, IV and V. From the table, the CSU-Net achieves the highest F1-score on all three datasets. Because F1-score is the comprehensive embodiment of recall and precision, the results indicate that CSU-Net can segment the background and vessels accurately. The pixel ratio of vessels and background in fundus images is unbalanced, the high sensitivity is very important. Our model achieves the best sensitivity on DRIVE and CHASE and is a little lower than the second expert results in STARE, which indicates that CSU-Net has strong ability to detect the retinal vessels and not adding more false pixels. Meanwhile, the CSU-Net also achieves the highest specificity on CHASE and STARE, which means the CSU-Net can better classify the background. Besides, the CSU-Net obtains the best performances on DRIVE (AUC: 0.9801, AC: 0.8565), STARE

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

TABLE V
COMPARISON OF CSU-NET WITH OTHER MODELS ON STARE DATASET

Approaches	Time	F1-score	SE	SP	AC	AUC
Second expert annotation	NaN	0.7417	**0.9017**	0.9564	0.9522	0.9291
Lam et al. [43]	2010	NaN	NaN	NaN	0.9567	0.9739
Fraz et al. [37]	2012	NaN	0.7311	0.9680	0.9442	NaN
Azzopardi et al. [38]	2015	NaN	0.7716	0.9701	0.9563	0.9497
Liskowsk et al. [22]	2016	NaN	0.7867	0.9754	0.9566	0.9785
Orlando et al. [39]	2016	0.7701	0.7680	0.9738	NaN	NaN
Son et al. [45]	2017	0.8353	0.8350	NaN	0.9657	0.9777
Zhang et al. [41]	2017	0.7815	0.7882	0.9729	0.9547	0.9740
Xia et al. [42]	2018	NaN	0.7670	0.9770	0.9530	NaN
Lu et al. [46]	2018	NaN	0.8090	0.9770	0.9628	0.9801
Context Spatial U-Net	2019	**0.8516**	0.8432	**0.9845**	**0.9702**	**0.9825**

TABLE VI
ABLATION STUDY OF TWO HYPERPARAMETERS ON DRIVE FOR CSU-NET. WE VARY THE VALUE OF K AND β, FINDING $K = 5$, $\beta = 5$ GIVES THE BEST PERFORMANCE ON DRIVE DATASET

AC/SE	K=3	K=5	K=7
$\beta = 2$	0.9553/0.8058	0.9561/0.8063	0.9551/0.8053
$\beta = 5$	0.9559/0.8067	**0.9565/0.8071**	0.9554/0.8063
$\beta = 7$	0.9557/0.8065	0.9562/0.8069	0.9554/0.8061

TABLE VII
RESULTS OF U-NET WITH DIFFERENT LOSSES ON DRIVE DATASET

Models	F1-score	SE	SP	AC	AUC
CE U-Net	0.8094	0.7709	0.9783	0.9532	0.9695
ST U-Net	**0.8105**	**0.7773**	**0.9794**	**0.9537**	**0.9716**

(AUC: 0.9825, AC: 0.9702) and achieves high score on CHASE dataset (AUC: 0.9824, AC:0.9706), indicates that the CSU-Net exhibits advanced performance compared with the other methods. Besides, compared with the second expert results, almost all indicators prove that the CSU-Net can match the human experts.

D. Ablation Analysis

We firstly demonstrate effectiveness of the proposed structure loss. There are two hyperparameters (i.e., K and β) to be determined. We conduct the experiment of CSU-Net on DRIVE dataset. Table VI lists the scores of AC and SE, when hyperparameters are in different values. As can be seen, when K equals 5, these indicators reach the highest score, which also meets our guess that the attention map has good boundary sensitivity. In addition, β represents the ratio of thin vessels and boundaries to background in the loss function. We increase β gradually from 2 to 7 and measure the corresponding scores of the above metrics. When $\beta = 5$, the model achieves the best performance. Secondly, we compare the U-Net trained by cross-entropy loss (named CE U-Net) and the U-Net trained by structure loss (named ST U-Net). Referring to Table VII, the result shows that ST U-Net is more accurate than CE U-Net. Besides, we compare the CSU-Net trained by these two different losses. Visualizations of the segmented vessel probability map in Fig. 10, the proposed structure loss increases the power of models to segment the thin vessels accurately and provides clear local boundary details.

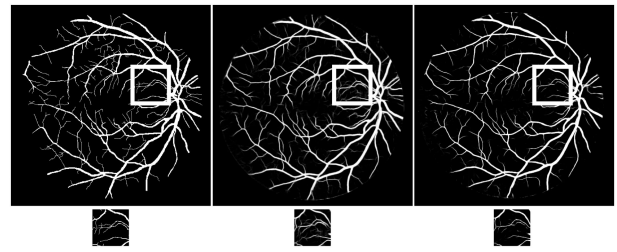


Fig. 10. Test results on the CSU-Net trained by two different losses. From left to right: label, the CSU-Net trained by structure loss and the CSU-Net trained by cross-entropy loss.

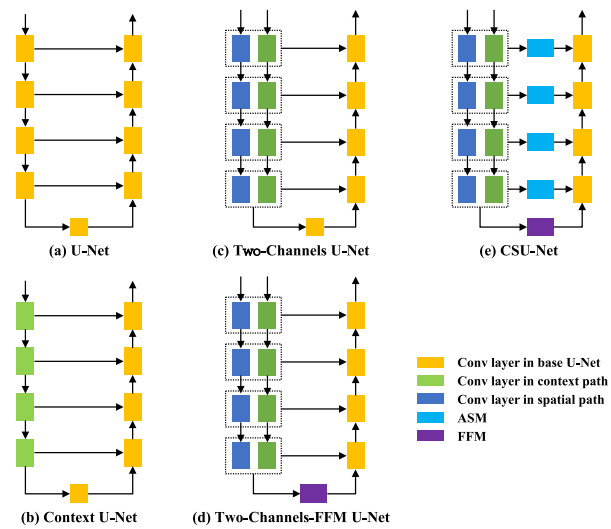


Fig. 11. Simplified illustration of different models tested in our experiments. (a) U-Net. (b) context U-Net, we replace convolution block of the U-Net by the baseblock of context path. (c) two-channels U-Net, expanding the encoder with two channels. (d) two-channels-FFM U-Net, adding FFM to two-channels U-Net. (e) CSU-Net, our proposed model.

TABLE VIII
RESULTS WITH DIFFERENT IMPROVED MODELS ON DRIVE DATASET

Models	F1-score	SE	SP	AC	AUC
U-Net	0.8105	0.7773	**0.9794**	0.9537	0.9716
Context U-Net	0.8135	0.7961	0.9778	0.9547	0.9737
Two-Channels U-Net	0.8178	0.7916	0.9790	0.9551	0.9777
Two-Channels-FFM U-Net	0.8204	0.8069	0.9766	0.9555	0.9782
Context Spatial U-Net	**0.8251**	**0.8071**	0.9782	**0.9565**	**0.9801**

Compared to the basic U-Net with CSU-Net, CSU-Net mainly has the following improvements: (1) expanding to two paths encoder; (2) adding skip attention module; (3) using fusion feature module to combine the different level features. For analyzing the effects of these improvements, we perform experiments on five different architectures, namely U-Net, Context U-Net, Two-Channels U-Net, Two-Channels-FFM U-Net and Context Spatial U-Net, which are presented in Fig. 11. Then, we evaluate all the models on DRIVE test data and the result is shown in Table VIII.

To investigate the impact of a module designed, we conduct various experiments in the DRIVE for different architectures trained by structure loss. The results in Table VIII show that

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

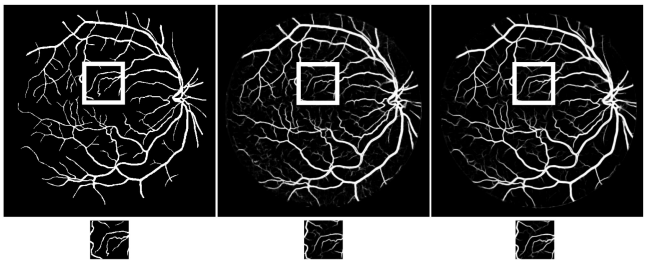


Fig. 12. Test results on UNet and the CSU-Net trained by structure loss. From left to right: label, the UNet result and the CSU-Net result.

(1) CSU-Net performs better than other improved U-Nets in F1-score, SE, SP, AC and AUC on DRIVE dataset, which indicates the effectiveness of module designed; (2) the performance with the U-Net having ASM and FFM is better than non-FFM or non-ASM models, which indicates that ASM and FFM structure can fuse the different level of feature representation and reduce the redundancy; (3) decoder with context and spatial channel can preserve both the spatial and semantic information, which makes performance improvement. Fig. 12 shows that the proposed CSU-Net produces more accurate results and also preserve more details than U-Net.

E. Limitation

Although the proposed CSU-Net has achieved good vessel segmentation accuracy, the computational complexity is still high. The average processing time for inferencing a testing image of the DRIVE, CHASE and the STARE datasets are 6,83 s, 12.61 s, and 8.13 s, respectively. The main reason is that the steps of patch extraction and predication results reconstruction need to consume a lot of memory and operations.

V. Conclusion

We propose a Context Spatial U-Net to improve the accuracy of blood vessel segmentation in fundus images. The CSU-Net has two-channel encoder, which can enhance the ability of the U-Net to capturing both spatial and context information in vessel segmentation with an end-to-end manner. Moreover, a new structure loss is introduced to distribute the different weights to the different positions, which can mine the structure information contained in the features and help the network focus more on the thin vessels and boundaries. The proposed CSU-Net is evaluated on three medical datasets. The results indicates that the CSU-Net performs better in vessel segmentation than the other advanced methods on all datasets. In future work, we will test our model on other medical image tasks such as vessel classification (arteries vs. veins) and auxiliary diagnosis of ophthalmic diseases.

References

[1] J. J. Kanski and B. Bowling, *Clinical Ophthalmology: A Systematic Approach*. Amsterdam, The Netherlands: Elsevier Health Sciences, 2011.

[2] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.

[3] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2881–2890.

[4] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2017.

[5] 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.

[6] O. Oktay *et al.*, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.

[7] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," *J. Med. Imag.*, vol. 6, no. 1, 2019, Art. no. 014006.

[8] 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.

[9] A. 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.

[10] M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Comput. Biol. Med.*, vol. 37, no. 2, pp. 262–267, 2007.

[11] M. M. Fraz *et al.*, "An approach to localize the retinal blood vessels using bit planes and centerline detection," *Comput. Methods Programs Biomedicine*, vol. 108, no. 2, pp. 600–616, 2012.

[12] M. Zardadi, N. Mehrshad, and S. M. Razavi, "Unsupervised segmentation of retinal blood vessels using the human visual system line detection model," *J. Inf. Syst. Telecomm.*, vol. 4, no. 2, pp. 125–133, 2016.

[13] 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.

[14] B. Al-Diri, A. Hunter, and D. Steel, "An active contour model for segmenting and measuring retinal vessels," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1488–1497, Sep. 2009.

[15] 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.

[16] X. You, Q. Peng, Y. Yuan, Y.-m. Cheung, and J. Lei, "Segmentation of retinal blood vessels using the radial projection and semi-supervised approach," *Pattern Recognit.*, vol. 44, no. 10-11, pp. 2314–2324, 2011.

[17] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1214–1222, Sep. 2006.

[18] C. A. Lupascu, D. Tegolo, and E. Trucco, "FABC: Retinal vessel segmentation using AdaBoost," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1267–1274, Sep. 2010.

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

[20] S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomed. Signal Process. Control*, vol. 30, pp. 1–12, 2016.

[21] C. Zhu, B. Zou, Y. Xiang, J. Cui, and H. Wu, "An ensemble retinal vessel segmentation based on supervised learning in fundus images," *Chin. J. Electron.*, vol. 25, no. 3, pp. 503–511, 2016.

[22] 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.

[23] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in *Proc. IEEE 13th Int. Symp. Biomed. Imag.*, 2016, pp. 698–701.

[24] X. Gao, Y. Cai, C. Qiu, and Y. Cui, "Retinal blood vessel segmentation based on the gaussian matched filter and U-Net," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat.*, 2017, pp. 1–5.

[25] X. Xiao, S. Lian, Z. Luo, and S. Li, "Weighted Res-UNet for high-quality retina vessel segmentation," in *Proc. 9th Int. Conf. Inf. Technol. Med. Edu.*, 2018, pp. 327–331.

[26] Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, "Multiscale network followed network model for retinal vessel segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2018, pp. 119–126.

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.

[27] Z. Yan, X. Yang, and K.-T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 1912–1923, Sep. 2018.

[28] J. Zhuang, "LadderNet: Multi-path networks based on U-Net for medical image segmentation," 2018, *arXiv:1810.07810*.

[29] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "DuNet: A deformable network for retinal vessel segmentation," *Knowledge-Based Syst.*, vol. 178, pp. 149–162, 2019.

[30] J. Staal, M. Abramoff, M. Niemeijer, M. 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.

[31] C. G. Owen *et al.*, "Measuring retinal vessel tortuosity in 10-year-old children: Validation of the computer-assisted image analysis of the retina (caiar) program," *Investigative Ophthalmol. Vis. Sci.*, vol. 50, no. 5, pp. 2004–2010, 2009.

[32] B. Payette, "Color Space Converter: R'G'B' to Y'CbCr," *Xilinx*, vol. 1, no. 1, pp. 637–643, Sep. 2002.

[33] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vis., Graphics, Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.

[34] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters–Improve semantic segmentation by global convolutional network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4353–4361.

[35] J. Cohen, "A coefficient of agreement for nominal scales," *Educational Psychological Meas.*, vol. 20, no. 1, pp. 37–46, 1960.

[36] D. P. Kingma and J. Ba, "ADAM: A method for stochastic optimization," in *3rd Int. Conf. Learn. Represent.*, 2015. [Online]. Available: http://arxiv.org/abs/1412.6980

[37] M. M. Fraz *et al.*, "Blood vessel segmentation methodologies in retinal images–A survey," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 407–433, 2012.

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

[39] 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. 2016.

[40] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *Proc. IEEE 14th Int. Symp. Biomed. Imag.*, 2017, pp. 248–251.

[41] J. Zhang, Y. Chen, E. Bekkers, M. Wang, B. Dashtbozorg, and B. M. ter Haar Romeny, "Retinal vessel delineation using a brain-inspired wavelet transform and random forest," *Pattern Recognit.*, vol. 69, pp. 107–123, 2017.

[42] H. Xia, F. Jiang, S. Deng, J. Xin, and R. Doss, "Mapping functions driven robust retinal vessel segmentation via training patches," *IEEE Access*, vol. 6, pp. 61 973–61 982, 2018.

[43] B. S. Lam, Y. Gao, and A. W.-C. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling," *IEEE Trans. Med. Imag.*, vol. 29, no. 7, pp. 1369–1381, Jul. 2010.

[44] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 109–118, Jan. 2015.

[45] J. Son, S. J. Park, and K.-H. Jung, "Retinal vessel segmentation in fundoscopic images with generative adversarial networks," 2017, *arXiv: 1706.09318*.

[46] J. Lu, Y. Xu, M. Chen, and Y. Luo, "A coarse-to-fine fully convolutional neural network for fundus vessel segmentation," *Symmetry*, vol. 10, no. 11, pp. 607–616, 2018.

Authorized licensed use limited to: Lanzhou University. Downloaded on December 09,2021 at 13:14:36 UTC from IEEE Xplore. Restrictions apply.