An Enhanced GAN for Retinal Image Generation Segmentation

Xiao Yang[1,a], Yirong Guo[1,b], Lihuang She[1,c] and Shi Zhang[1,d]

[1]*School of Information & Engineering Northeastern University Shenyang, China*
a. 1801694@stu.neu.edu.cn, b. 1701720@stu.neu.edu.cn,
c. shelihuang@ise.neu.edu.cn, d. zhangshi@cse.neu.edu.cn

Keywords: Retinal vasculature, segmentation, CNN, GAN.

Abstract: Automated segmentation of the retinal vasculature is a critical task for precise geometric analysis and automated diagnostics. In recent years, convolutional neural networks have shown excellent performance compared to traditional segmentation methods. In this paper, we choose the Generative Adversarial Network with game ideas in the field of computer vision to segment the retinal images, and improve the network framework on the basis of the original version. This paper mainly improves the network of generators by combining the Generative Adversarial Network with residual learning, dense connection and U-Net, and increases the utilization of image features and solves the problem of less training data. Through experimental evaluation, the performance gain of the Generative Adversarial Network when applied to the segmentation task is generated. The results show that the network can finely segment small blood vessels and achieve statistically significant improvements in the analysis of accuracy, sensitivity, specificity, and ROC curves on two common data sets, DRIVE and STARE.

1. Introduction

Analysis of the retinal vascular network can provide a wealth of information for the general systemic state of the eye and body. For example, an ophthalmologist can detect early signs of increased systemic vascular load and associated retinal diseases caused by diabetes and hypertension based on the information[1]. Retinal vascular segmentation technology is becoming a basic and important component of clinical medical disease screening and diagnosis[2],so automatic retinal vascular segmentation methods have been widely studied by many scholars.

The current retinal segmentation algorithm can automatically segment the vascular network, but there is still some influence for segmentation caused by different contrasts of the blood vessels and other stuff in the different retinal databases and illumination and pathology on the retina. Due to the importance of retinal vascular segmentation, researchers have proposed a series of methods to solve this challenging task.

In the existing retinal blood vessel segmentation method, the highest segmentation accuracy is based on the machine learning method[3]. It is mainly divided into two categories: one is based on methods of supervised learning and another is unsupervised learning. In the method of supervised learning, the pixel points need to marked to Fig out whether it belong to the blood vessel point or the background point. Fraz.[4] proposed a retinal vessel segmentation supervised learning method

for children's fundus images. Gabor filtering, morphological transformation, etc. are applied to construct a multi-dimensional feature vector. Finally, the integrated classifier was used to classify the segmentation results of blood vessels. The method with no post-processing and pre-processing was tested on a new database and achieved high segmentation accuracy. Wang et al[5] proposed a retinal vessel segmentation method including preprocessing, hierarchical feature extraction and integrated classification, which was tested and displayed in a common database.

The unsupervised method does not need the training sample an expertly calibrated, and it can directly segment the retinal blood vessels of the fundus image. Salazar-Gonzalez et al.[6]proposed the partitioning method which apply the appropriate amount of flow to segment, and verified the effectiveness of the algorithm on the STARE and DRIVE databases. Fu H et al[7] combined multi-level convolutional neural networks and conditional random fields into an integrated deep network, simulating remote interaction between pixels, and implementing the most advanced retinal vascular segmentation performance on DRIVE, STARE and CHASE_DB1 data sets. And the runtime is valid.

In fact, blood vessel segmentation can be seen as an image conversion task that can generate an image of a blood vessel segmentation probability map from the input fundus image. If the output is limited to the results marked by a human expert, a clearer and more accurate blood vessel map can be obtained. Generating a confrontation network[8] is a network framework that can create and utilize actual output as a gold standard. It consists of two subnetworks: a generator and a discriminator. The discriminator attempts to distinguish the output of the generator from the gold standard image, while the generator produces as realistic output as possible so that the discriminator cannot distinguish between gold standards and outputs.

This paper will improve the generation of the confrontation network on the basis of the original network, and combine the generation of the network with the residual network, the dense connection and the U-Net, and propose a retinal blood vessel segmentation method based on the improved generation of the anti-network. The method of this paper can not only extract accurate and clear retinal blood vessels, but also produce fewer false positives compared with the existing methods, segmenting more elaborate blood vessels, and also showing good performance on two common databases of DRIVE and STARE. Statistical effect.

2. Network Structure

The framework herein is generally similar to the traditional GAN framework: giving the generator a retinal fundus image and generating a probability map of the retinal blood vessels which is of the same size as the input image. The range of values in the probability map is from 0 to 1, indicating the probability of being a pixel of a blood vessel. Then, the discriminator obtains a pair of fundus images and probability maps, and judges whether the probability map is a gold standard image manually labeled by humans or an output of the generator. The purpose of the generator is to output a probability map that the discriminator cannot distinguish from the gold standard. The purpose of the discriminator is to correctly determine the source of the probability map, in such a way that the constrainer generator produces a more realistic output. The overall framework of the vascular GAN is generally shown in Figure 1.

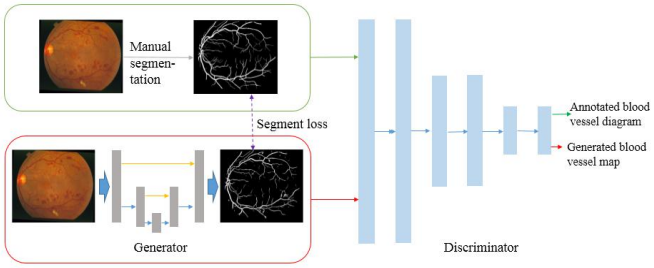


Figure 1: Vascular-type GAN network.

2.1. U-Net

U-Net[12] is a network suitable for medical image processing based on the full convolutional neural network, which mainly includes two parts of symmetry: contraction path and extension path. The shrink path is mainly to capture the context information in the picture, and the extended path is in the picture.

U-Net[12] is a network suitable for medical image processing based on the full convolutional neural network, which mainly includes two parts of symmetry: contraction path and extension path. The shrinking path is mainly to capture the context information in the picture, and the extended path is to accurately locate the part of the picture that needs to be segmented. This network uses data enhancement to perform end-to-end training on data with fewer samples, which is suitable for clinically segmentation of retinal images with fewer samples. Therefore, this paper applies this structure to the generator to enable it to be integrated with the GAN network. At the same time, it can combine the underlying information to improve the training precision and high-level information to facilitate the extraction of complex features. It is easy to get more features of retinal blood vessels and better segment the retinal images.

2.2. Resnet

In general, the inputs and can get the correct solution to accurately predict the classification by applying the network to fit. Introduce the residual function in ResNet (that is, the deviation between the target value and the input value), and fit by training, and then get from.

ResNet[13] is similar to the principle of differential amplifiers. It adds direct-connected channels to the network, allowing raw input information to be passed directly to subsequent layers. This allows simple equivalent mapping without additional parameters and computational complexity, and the entire network can still carry out end-to-end backpropagation training, which can greatly accelerate the training of neural networks. The specific module idea is shown below.

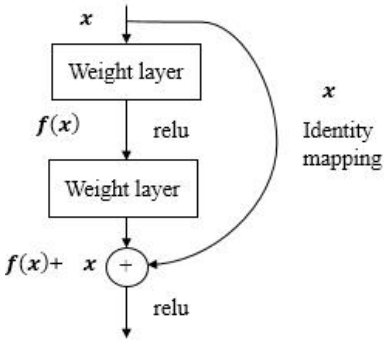
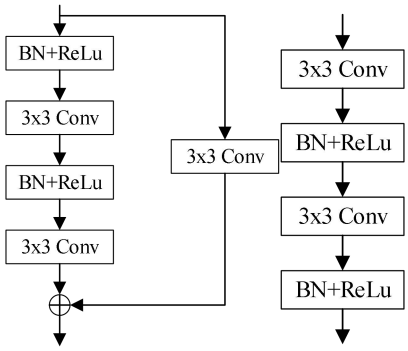


Figure 2: ResNet's residual learning block.

Traditional convolutional neural networks or fully connected networks have more or less information loss in the process of information transmission. At the same time, they may cause gradients to disappear or gradient explosions, resulting in deep network untrainable. ResNet can effectively solve this problem to a certain extent. By directly transferring the input information to the output, the integrity of the information can be protected. The entire network only needs to learn the difference between input and output, which can simplify the learning objectives and difficulty, and solve the problem at the same time. The problem of degradation caused by the increase in network age is easier to optimize and the convergence speed is faster. The ResNet module used in this article is shown in the Figure below.



(a)ResNet block (b)Deconvlution block

Figure 3: Generate specific block.

2.3. Densenet

The dense connection module of DenseNet[14] establishes the dense connection between all the layers in the previous layer and the back layer, combines the feature information of the front layer and the layer to the next layer as the output, and realizes the feature reuse through the connection of the features on the channel, shortening The distance between the front and back layers, improving efficiency, reducing network parameters and layers combined with bypass settings, reducing algorithm complexity, while alleviating the problem of gradient disappearance and model collapse, thus making the network easier to train and having a certain regular effect.

If the input picture is X_0 , it will pass through an L -layer neural network. The nonlinear transformation of the i -layer is denoted as $H_i()$, which can be the cumulative operation of various functions such as BN, ReLu, and Pooling, and the characteristic output of the i -layer is X_i . The i -th layer output of the traditional convolutional neural network will be used as the input to the next layer, written as $X_i = H_i(X_{i-1})$. As shown in the Figure, the output of the i -layer is not only related to the previous layer, but also related to the output of the previous layer, which is denoted as $X_l = H_l([X_0, X_1, ..., X_{l-1}])$,where [] represents the concatenate stitching. The nonlinear transformation is generally a combination of $BN + \mathrm{Re}Lu + Conv(3\times 3)$.

This article mainly uses DenseNet's splicing concept to make full use of image features. The generator structure of this paper after combination is shown in Figure 5.

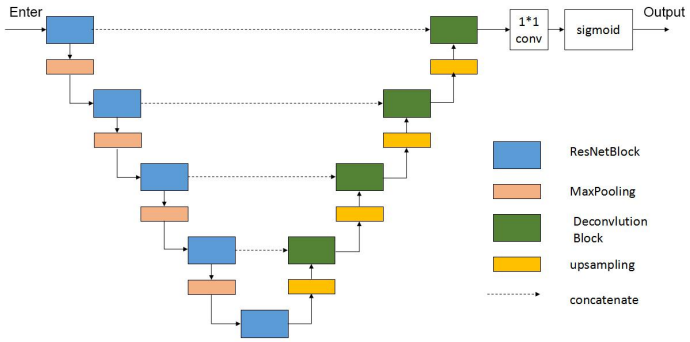


Figure 4: Generator Network.

2.4. Discriminator Network

The most important thing in the generation of confrontation networks other than generators is the discriminator that is used to play with them. The main function of the generator is to model and learn the distribution law of real data, and the main function of the discriminator is to identify the input data and finally output the type of data. The specific structure of the discriminator used in this paper is shown in the Figure below.

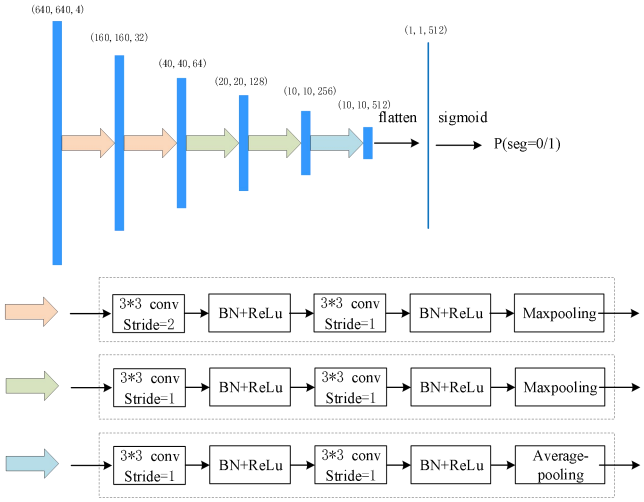


Figure 5: Discriminator Network.

2.5. Loss Function

The generator G can be regarded as a mapping of the fundus image x to the probability map y, and the discriminator D maps (x, y) to $\{0,1\}^N$, where 0 represents the output of the generator and 1 represents the gold standard of the manual annotation. N indicates the number of decisions. It should be noted that in the GAN with the image as the decision criterion, N=1, and in the GAN network with the pixel as the decision unit $N = W \times H$, where the size of the image is $W \times H$ [9].

The objective function of the GAN network for the segmentation problem can be expressed as:

$$L_{GAN}(G,D) = E_{x,y\sim pdata(x,y)}[\log D(x,y)] + E_{x\sim pdata(x)}[\log(1 - D(x,G(x)))] \quad (1)$$

The architecture optimization problem of the GAN network is solved by the following formula:

$$G^* = \arg\min_G \left[\max_D E_{x,y\sim pdata(x,y)}[\log D(x,y)] + E_{x\sim pdata(x)}[\log(1-D(x,G(x)))] \right] \quad (2)$$

Training the discriminator to make the correct judgment requires $D(x,y)$ to be maximized while $D(x,G(x))$ is minimized. On the other hand, the generator needs to prevent the discriminator from making the correct judgment by generating an output that is difficult to distinguish from real data. Since the ultimate goal is to get the actual output from the generator, the objective function is defined as the minimum maximum of the target.

In fact, you can also use the gold standard image by adding a loss function, such as binary cross entropy, as follows:

$$L_{SEG}(G) = E_{x,y\sim pdata(x,y)}[-y\cdot\log G(x) - (1-y)\cdot\log(1-G(x))] \quad (3)$$

By summarizing the GAN target and the segmentation loss, the objective function is expressed as:

$$G^* = \arg\min_G [\max_D L_{GAN}(G,D)] + \lambda L_{SEG}(G) \quad (4)$$

Among them, use λ to balance the two objective functions.

3. Experiment

3.1. Database

The method in this paper is evaluated on two public databases, DRIVE[10] and STARE[11]. The DRIVE database consists of the fundus image of the corresponding mask image and the artificial segmentation results of the blood vessels. There are 40 groups of images, which are divided into 20 training sets and 20 test sets. In the training set, a single hand-divided image is provided for each fundus image, and two segmented images manually labeled by an experienced ophthalmologist are provided for each fundus image in the test set. For testing and training, it is customary to use the first annotated blood vessel segmentation image as the gold standard and compare the result to the resulting image of the second image.

The STARE data set contains 20 sets of fundus push and vessel segmentation images from two independent annotator markers. This article uses the first person's mark as the gold standard and the other as the reference point for human observation. Since the data set does not give a mask for the corresponding fundus image, we generate this part of the image ourselves. The first 10 images in the data set are used for training and the rest are tested to compare algorithm performance.

To train the network, these data sets are not enough. Therefore, in this paper, the fundus image of the training sample and the corresponding artificially segmented image label are respectively rotated by 3 degrees, so that one training sample becomes 120 sheets, thereby enhancing the data of the training samples.

3.2. Evaluation Indicators

In order to systematically analyze the segmentation performance of the algorithm, the following three indicators are used as the measurement criteria:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

$$Sn = \frac{TP}{TP + FN} \tag{6}$$

$$Sp = \frac{TN}{TN + FP} \tag{7}$$

Among them, TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively. Accuracy (Acc) indicates the ratio of the pixel that divides the correct pixel to the pixel of the whole image. Sensitivity (Sn) indicates the ratio of the correct segment of the blood vessel to the pixel. The specificity (Sp) indicates that the correct background point is the pixel. ratio. The closer the values of the three are to 1, the better the performance of the segmentation algorithm.

The ROC curve is also an important indicator to measure the results of vascular segmentation. The horizontal axis is the false positive rate, the vertical axis is the true positive rate, and the response is different under different thresholds. The closer the curve is to the vertical axis, the more robust the segmentation algorithm is. In general, the area under the curve AUC is used to quantify the representation. The closer this parameter is to 1, the better the segmentation performance of the algorithm.

The dice coefficients represent the similarity coefficients between the standard image and the predicted image. The specific calculation formula is as follows:

$$dice = \frac{2|X \cap Y|}{|X| + |Y|} \tag{8}$$

Among them, it represents a standard image and represents a predicted image. The closer the parameter value is to 1 the better the performance of the algorithm.

3.3. Experimental Setup

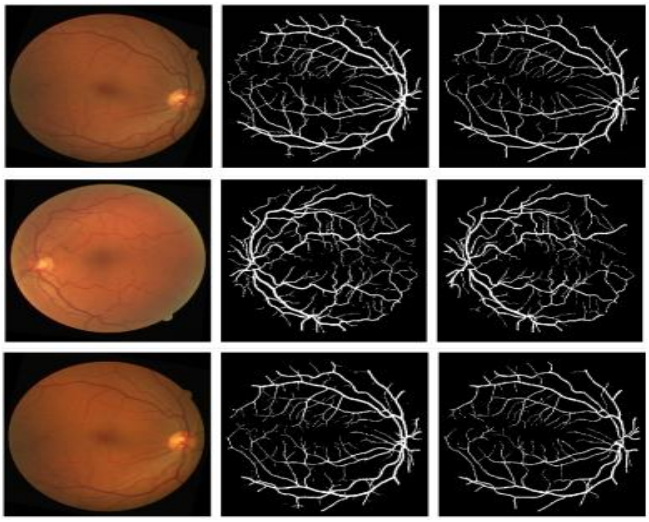
The simulation platform for this experiment is PyCharm, using keras and its TensorFlow port. The computer is conFigd as Intel(R) CORETM i7-7700CPU@3.6GHz*8, 16GB RAM, Nvidia GeForce GTX 1070 GPU, and 64-bit Ubuntu 16.04 LTS. This article maintains the same experimental setup, including the use of uniform hyperparameters during training and the application of the same image pre-processing procedures to eliminate external factors that may have an impact on the assessment.

3.4. Results

In order to verify the comprehensive performance of the algorithm, the images in the commonly used DRIVE and STARE databases were tested.

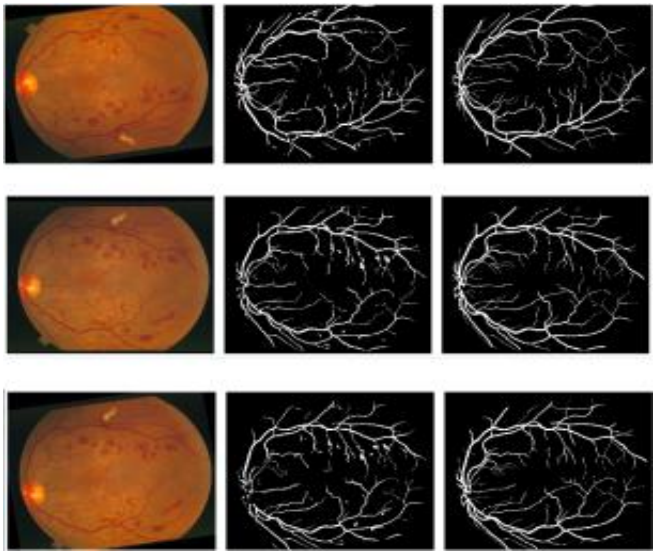
Figure 6 shows the test results of the algorithm in the DRIVE database. The first column is the original retinal color image, the second column is the result of the algorithm segmentation, and the

third column is the standard result of the expert manual segmentation. Each row is different images selected from the database.



(a) Original retina color image (b) Our results (c) Experts manually segment standard results

Figure 6: Segmentation results of the algorithm in the DRIVE database.



(a) Original retina color image (b) Our results (c) Experts manually segment standard results

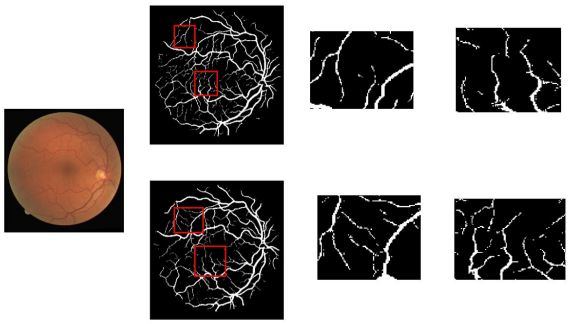
Figure 7: Segmentation results of the algorithm in the STARE database.

Figure 7 shows the test results of the algorithm in the STARE database. The first column is the original retinal color image, the second column is the standard segmentation result for the algorithm, and the third column expert manually divides the standard result. Each row is a different image selected from the STARE database.

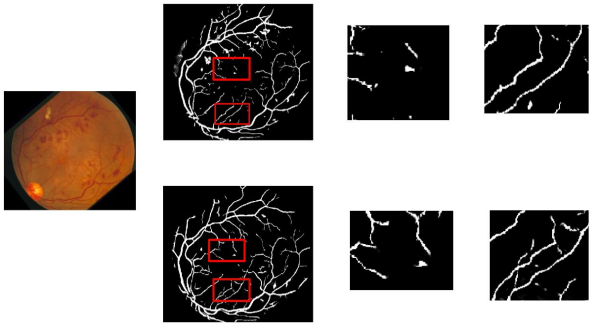
It can be seen from the segmentation results in the above two databases that the proposed retinal segmentation algorithm based on improved generation of anti-network can segment the vascular network of color retinal fundus images, and the vascular network has good integrity and coherence. The algorithm as a whole has a good segmentation of both vascular and non-vascular points, and achieved good segmentation results.

This paper also compares the segmentation results with the pre-implementation generation confrontation network. The left side is the original retinal fundus image, the first line indicates the

pre-improvement generation against the network results, and the second line is the result of this paper. The red squares are marked with the areas highlighted in this article. They are enlarged in order from top to bottom and arranged in order from left to right, as shown in Fig 8. It can be seen from the Fig that the pre-improvement generation confrontation network can segment the main part of the blood vessel, but the details are lost, but the method of this paper can not only divide the general network of blood vessels, but also achieve good results in small blood vessels.



(a)DRIVE data



(b)STARE data

Figure 8: DRIVE and STARE data sets small blood vessel segmentation enlarged view.

Table 1 and Table 2 show the algorithms and literatures[15], [16],and[17], and the accuracy (Acc), sensitivity (Sn), and specificity (Sp) in the DRIVE and STARE data sets, respectively. Several indicators compared the performance of retinal vascular segmentation. The algorithm in this paper has good accuracy on both datasets. On the DRIVE and STARE datasets, the accuracy of the algorithm is 0.74% and 0.13% higher than that of the Wavelets method.

Table 1: Comparison of Algorithm Test Results (Drive).

data	DRIVE				
	Acc	Sn	Sp	AUC	Dice
CRFs[15]	0.9438	0.7829	0.9672	-	0.7797
Kernel-Boost[16]	0.9456	0.7563	0.9732	-	0.7795
Wavelets[17]	0.9387	0.7628	0.9644	0.9436	0.7599
2nd manual	0.9473	0.7746	0.9725	-	0.7889
Our algorithm	0.9575	0.7656	0.9857	0.9598	0.7832

Table 2: Comparison of Algorithm Test results (Stare).

data	STARE				
	Acc	Sn	Sp	AUC	Dice
CRFs[15]	0.9648	0.8380	0.9796	-	0.8323
Kernel-Boost[16]	-	-	-	-	-
Wavelets[17]	0.9529	0.7817	0.9728	0.9694	0.7756
2nd manual	0.9380	0.9445	0.9373	-	0.7605
Our algorithm	0.9542	0.7363	0.9842	0.9696	0.7701

4. Conclusion

In this paper, the adversarial network is applied to the retinal blood vessel segmentation, and the original network structure is improved. By combining with the residual network and the densely connected network, the utilization of image features is improved. The experimental results show that the game theory of the discriminator and generator is conducive to clear and accurate segmentation of the vascular network. In addition, fewer false positives appear in the segmentation of small blood vessels, and more detailed information is drawn with clearer lines. The improved network structure also yielded good test results in statistical measurements of the database regarding accuracy, sensitivity, specificity, and the like. In the follow-up work, we will further improve the segmentation quality of retinal vessels by combining the anatomical knowledge of vascular structures.

Acknowledgments

This work financially supported by the funding of basic research business fees in Chinese universities fund(project batch number is N181604007).

References

[1] Son J , Park S J , Jung K H . Retinal Vessel Segmentation in Fundoscopic Images with Generative Adversarial Networks[J]. 2017.

[2] Yi L , Honggong Z , Guang H . Automatic Retinal Vessel Segmentation via Deeply Supervised and Smoothly Regularized Network[J]. IEEE Access, 2018:1-1.

[3] Zhu C, Zou B, Xiang Y, et al. Research Progress of Retinal Vessel Segmentation Methods in Color Fundus Images[J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27(11): 2046-2057.

[4] Fraz M M, Rudnicka A R, Owen C G, et al. Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification[J]. International Journal of Computer Assisted Radiology and Surgery, 2014, 9(5): 795-81.

[5] Wang S, Yin Y, Cao G, et al. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning[J]. Neurocomputing, 2015, 149(B): 708-717.

[6] Salazar-Gonzalez, Kaba D, Li Y M, et al. Segmentation of the blood vessels and optic disk in retinal images[J]. IEEE Journal of Biomedical and Health Informatics, 2014, 18(6):1874-1886.

[7] Fu H, Xu Y, Lin S, Wong DWKW, Liu J: Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. In: International conference on medical image computing and computer-assisted intervention. Springer, 2016, pp 132–139.

[8] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y: Generative adversarial nets.In: Advances in neural information processing systems, 2014,pp 2672–2680.

[9] Son J , Park S J , Jung K H . Towards Accurate Segmentation of Retinal Vessels and the Optic Disc in Fundoscopic Images with Generative Adversarial Networks[J]. Journal of Digital Imaging, 2018.

[10] Staal J, Abramoff MD, Niemeijer M, Viergever MA, Ginneken BV: Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging 23(4):501–509, 2004.

[11] Hoover AD, Kouznetsova V, Goldbaum M: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans Med Imaging 19(3):203–210,2000.

[12] Ronneberger O , Fischer P , Brox T . U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. 2015.

[13] Zhang Q , Cui Z , Niu X , et al. Image Segmentation with Pyramid Dilated Convolution Based on ResNet and U-Net[J]. 2017.

[14] Iandola F, Moskewicz M, Karayev S, et al. DenseNet: Implementing Efficient ConvNet Descriptor Pyramids[J]. Eprint Arxiv, 2014.

[15] Orlando JI, Blaschko M: Learning fully-connected crfs for blood vessel segmentation in retinal images. In: International conference on medical image computing and computer-assisted intervention. Springer, 2014, pp 634–641.

[16] Becker C, Rigamonti R, Lepetit V, Fua P: Supervised feature learning for curvilinear structure segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer, 2013, pp 526–533.

[17] Soares JVB, Leandro JJG, Cesar RM, Jelinek HF, Cree MJ:Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. IEEE Trans Med Imaging 25(9):1214–1222, 2006.