Research Article

Automatic Retinal Vessel Segmentation Based on an Improved U-Net Approach

Zihe Huang,1 Ying Fang,1 He Huang,1 Xiaomei Xu,1 Jiwei Wang,2 and Xiaobo Lai1

1School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou, China
2Department of Information, The 73th Group Army Hospital of P. L. A, Xiamen, China

Correspondence should be addressed to Jiwei Wang; 772073190@qq.com and Xiaobo Lai; dmia_lab@zcmu.edu.cn

Received 13 February 2021; Revised 30 March 2021; Accepted 8 April 2021; Published 24 April 2021

Academic Editor: Chenxi Huang

Copyright ©2021 Zihe Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Retinal blood vessels are the only deep microvessels in the blood circulation system that can be observed directly and noninvasively, providing us with a means of observing vascular pathologies. Cardiovascular and cerebrovascular diseases, such as glaucoma and diabetes, can cause structural changes in the retinal microvascular network. Therefore, the study of effective retinal vessel segmentation methods is of great significance for the early diagnosis of cardiovascular diseases and the vascular network’s quantitative results. This paper proposes an automatic retinal vessel segmentation method based on an improved U-Net network. Firstly, the image patches are rotated to amplify the image data, and then, the RGB fundus image is preprocessed by normalization. Secondly, after the improved U-Net model is constructed with 23 convolutional layers, 4 pooling layers, 4 upsampling layers, 2 dropout layers, and Squeeze and Excitation (SE) block, the extracted image patches are utilized for training the model. Finally, the fundus images are segmented through the trained model to achieve precise extraction of retinal blood vessels. According to experimental results, the accuracy of 0.9701, 0.9683, and 0.9698, sensitivity of 0.8011, 0.6329, and 0.7478, specificity of 0.9849, 0.9967, and 0.9895, F1-Score of 0.8099, 0.8049, and 0.8013, and area under the curve (AUC) of 0.8895, 0.8845, and 0.8686 were achieved on DRIVE, STARE, and HRF databases, respectively, which is better than most classical algorithms.

1. Introduction

The retinal blood vessels are essential parts of the whole microcirculation system with extremely wealthy blood vessel information. They are the only deep microvessels in the blood circulation system that can be observed directly and noninvasively. Most ophthalmic diseases are attributed to fundus retinopathy [1], glaucoma, diabetic retinopathy, and age-related macular degeneration, the four primary reasons. Therefore, early detection, diagnosis, and treatment can effectively reduce the degree of patients affected by the disease or slow down its progress. The retinal blood vessels are also the main anatomical structures visible from color fundus images [2]. Color fundus images can directly detect retinal vascular microaneurysms, fundus hemorrhage, exudate, and other diseases. They are essential bases for ophthalmologists to diagnose and treat diseases. Accurate segmented images play an essential role in the preliminary screening, subsequent diagnosis, and treatment of fundus diseases. The manual segmentation is tedious, time consuming, and requires skilled techniques. When different ophthalmologists segment the same blood vessel image in clinical practice, the results obtained are different, which will cause errors in the final segmentation results [3]. The automatic segmentation technology of retinal vessels can relieve ophthalmologists’ diagnostic pressure and effectively solve inexperienced ophthalmologists’ problems. Therefore, the automatic segmentation of retinal vessels is important for the clinical diagnosis and treatment of ophthalmic diseases.

The highly complex structure of retinal vessels, low contrast between the retinal target vessel and the background image, and noise interference when acquiring retinal images can significantly increase the difficulty of fundus retinal blood vessel segmentation. Although there are many barriers to realizing the accurate segmentation of retinal vessels, it is of considerable significance to auxiliary clinical
diagnosis and treatment. In recent years, it has attracted many scientific experts to explore the field and achieved fruitful results [4]. According to the learning mode, the existing retinal vascular segmentation methods can be roughly classified as unsupervised methods [5–16] and supervised [17–19].

The unsupervised method can be subdivided into different methods, such as matched filtering, vascular tracking, morphological processing, and model-based approaches. Fraz et al. [5] combined vascular centerline detection with morphological bit plane section technology to extract blood vessels from retinal images. To detect the blood-vessel-like pattern in a noisy environment, Zana and Klein [6] implemented an algorithm about mathematical morphology and curvature evaluation. Niemeijer et al. compared many vascular segmentation algorithms [7]. According to the research, the highest accuracy of these algorithms is 0.9416. Mendonça and Campilho [8] combined the strength and partial morphological features of the vascular structure with the different filters for centerline extraction and the morphological operator for filling vascular segments to gain the final segmentation results of fundus retinal vessels. Garg et al. [9] proposed a curvature-based method combined with an improved region growth method to extract complete vascular structures from color retinal images. Chutatape et al. [10] used Gaussian filtering and Kalman filtering to make the best linear estimation of subsequent blood vessels’ position on the known centerline of blood vessels. Gang et al. [11] obtained a two-dimensional Gaussian filtering method through mathematical analysis and simulation experiment process, which significantly affects amplitude correction and improves the detection success rate. Vlachos and Dermatas [12] presented a retinal vessel image extraction algorithm whose basic principle is iterative line tracking and morphological postprocessing, with an average accuracy of 0.929. Hoover et al. put forward a threshold detection algorithm based on matched filter [13]. To eliminate the optic disc and background noise in the fundus image, Zardad et al. [14] enhanced the blood vessels in all directions and used the adaptive threshold as an unsupervised classifier for pixel classification and perform morphological postprocessing at last. Srinidhi et al. [15] used local and global information to decompose the vascular tree into multiple subtrees and divide the vascular subtrees according to arteries and veins, respectively. Upadhyay et al. [16] implemented local directional-wavelet transform and global curvelet transforms to enhance vessel and segmentation. The supervised method uses the truth map as the training sample to train the model. Aslani and Sarnel come up with mixed feature vectors from different algorithms to represent pixels [17] and mixed feature vectors to train the random forest (RF) classifier to classify vascular/nonvascular pixels. Marin et al. employed a neural network scheme to classify pixels and calculate 7-dimensional vectors composed of grayscale and moment-invariant features to represent pixels [18]. Jiang et al. [19] proposed a D-Net model for retinal blood vessels, which can obtain denser feature information and reduce the excessive loss of tiny blood vessels’ feature information.

Traditional supervision methods are usually defined by experience and require manual intervention, which may result in deviation. Therefore, an automatic and effective feature extractor deep learning is necessary to achieve higher efficiency. Wang et al. combined two advanced classifiers named the convolution neural network (CNN) and random forest (RF) for segmentation [20]. Maji et al. processed color fundus images and detected blood vessels on the foundation of the ConvNet integrated algorithm framework [21]. Che et al. utilized supervised blood vessel segmentation based on an artificial neural network to estimate the effect of aging on the performance [22]. The results illustrate that people of different ages have an impact on various aspects of the subdivision results. Lisowski and Krawiec presented deep neural network segmentation architecture with an extensive training dataset strengthened by global contrast normalization, zero-phase preprocessing, geometric transformation, and gamma correction [23]. Tan et al. took advantage of a 10-layer CNN segment and distinguished exudates, bleeding, and microaneurysms simultaneously [24]. The network structure consists of two stages to improve the performance, segment the fundus image’s pathological features with superior accuracy, and demonstrate the strong feature extraction ability of deep learning. Wu et al. [25] integrated residual learning with DenseNet, which uses each layer of vascular feature map information effectively so that the model can obtain more robust morphological structure information according to the gold standard. However, due to the excessive use of dense connection modules, the amount of memory will be expanded, and the computational complexity will be added. Fu et al. [26] used a deep learning framework based on fully convolutional networks (FCN) to figure out the problem of vascular edge detection. Lin et al. [27] proposed an automatic retinal blood vessel segmentation network that combines globally nested edge detectors with global smoothness regularization from conditional random fields. Khan et al. [28] extended the FCN variant, which significantly reduced the number of adjustable hyperparameters and decreased the computational overhead of the training and testing phases.

Although the automatic segmentation algorithm of retinal vessels can effectively segment the retinal vascular network, the automatic segmentation method’s performance still needs to be improved. There are some limitations, such as insufficient segmentation degree and poor continuity of microvascular vessels, which cannot meet clinical diagnosis needs. The U-Net model has been applied to the segmentation of medical images for years, and many improved structures based on the U-Net model [29–36] have achieved good segmentation results. In this paper, we explore an automatic segmentation algorithm of retinal vessels based on improved U-Net. The proposed scheme of the work is shown in Figure 1. It contains training and testing stages. In the training stage, the retinal vessel training set images and the gold standard images are preprocessed, and the image blocks are extracted to train the improved U-Net model. The parameters are learned by training data. In the test stage, the same preprocessing method is used for the test images, and the test set image blocks are tested and extracted through the training model.
This article’s remainder structure is arranged as follows: Section 2 introduces the improved automatic segmentation method of retinal vessels in detail. Section 3 will compare and analyze our experimental results. Finally, Section 4 draws conclusions about this work.

2. Methods

In order to improve algorithm performance, we combine partial advantages of the unsupervised and supervised methods. Firstly, the RGB fundus image is preprocessed. Secondly, the U-Net model is constructed with a SE block, and the extracted patches are applied to train the improved U-Net model. Finally, the trained model is tested to obtain the eventual results. The overview of the segmentation framework is shown in Figure 1, which consists of three steps:

Step 1: fundus image preprocessing: we extract the training set’s patches and the gold standard of fundus image, and the patches are rotated to amplify the data and normalized subsequently.

Step 2: building and training the model: we construct the U-Net model with an SE block and set up the parameters, and the extracted patches are applied to train the model.

Step 3: testing and segmenting retinal vessels. Firstly, the test set’s fundus patches are extracted and normalized; then, we test and extract retinal vessels automatically through the trained model.

2.1. Materials. For the sake of evaluating the effectiveness of the algorithm, the fundus image datasets DRIVE (http://www.isi.uu.nl/Research/Databases/DRIVE/), STARE (http://cecas.clemson.edu/~ahoover/stare/), and HRF (https://www5.cs.fau.de/research/data/fundus-images/) are used to segment the retinal vessels. DRIVE is a fundus image dataset usually referred to for evaluating the performance of retinal vascular segmentation methods in other experiments. DRIVE is concluded from the Netherlands DR screening Project. It contains 40 color fundus images with a resolution of $768 \times 584$ pixels. The STARE dataset consists of 20 normal and diseased fundus images with $605 \times 700$ pixels. The full name of HRF is High-Resolution Fundus Image Database. HRF consists of 15 images of healthy patients, 15 images of patients with diabetic retinopathy and 15 images of patients with glaucoma, all collected by using clinical fundus cameras with a resolution of $3504 \times 3504$.

2.2. Preprocessing. Our model is prone to overfitting during the training process due to the relatively small number of training samples in our DRIVE, STARE, and HRF databases. Simultaneously, the collected retinal images are susceptible to mutual interference. Preprocessing that contains the amplification and normalization of the fundus image data is necessary to deal with these problems. Figure 2 shows the results of the image normalization, where (a) is the original image and (b) is the normalized image. The process is briefly described as follows:

1. Image data amplification: patch extraction and geometric transformation are used to amplify the image data. Taking each voxel as the center, the two-dimensional region with the size of $n \times n$ pixels is extracted as the corresponding patches of the voxel. The window size of the model input is $48 \times 48$.
2. The images of the training set and the patches extracted by the gold standard are rotated. Figure 2,
(a) is the extracted image patches, and Figures 3(b)–3(d) show visualizations of the image patches rotated by 90°, 180°, and 270°, respectively.

(2) Image normalization: there is inevitable noise interference in the fundus image. The features are normalized so that they have a mean of 0 and a variance of 1. The mean and variance are standardized as follows:

\[ x^*_k = \frac{x_k - \mu}{\sigma}, \]

where \( x_k \) denotes the value of the \( k \)th pixel before standardization, \( x^*_k \) represents the intensity of the \( k \)th voxel after standardization, \( \mu \) is the average intensity of all pixels, and \( \sigma \) is the standard deviation of the intensity of all pixels.

2.3. Improved U-Net Architecture Model. The U-Net model was first put forward by Ronneberger et al. [37] in 2015 and applied to the cell segmentation of medical images. In recent years, Alom et al. [38] applied U-Net to retinal vascular segmentation, combined with a cyclic neural network, to further improve retinal vascular segmentation accuracy. The U-Net neural network is an end-to-end training supervised image segmentation network based on FCN. It has excellent performance under fewer datasets and can be widely used in medical image segmentation research. This paper optimizes the basic structure of U-Net.

2.3.1. SE Block. SE block is a network substructure in SENet. It can be easily integrated into almost any network architecture. As for the traditional convolution process, its channel dependence is not fully utilized. Moreover, the filters learned by the channel are all performed in the local receptive field, resulting in each feature map being unable to use the context information outside the region. The channel dimension of the SE block in the convolutional layer adds the attention mechanism. It aims to adopt a feature recalibration strategy to learn the importance of each feature from the global information through the network loss function. Then, the valuable features are promoted according to their importance, and the less valuable features are suppressed. This training method improves the performance of the network. The addition of the SE block will only add a small amount of calculation and parameters to the original network, which can be wholly accepted compared with the improvement of its effect.

SE block is roughly divided into 3 parts. For an input \( x \), supposing the feature channels value is \( c_1 \), it would get a particular feature after some transformations such as convolution. Moreover, this particular feature has the value of feature channels \( c_2 \). Unlike the traditional CNN, the features previously obtained need to be recalibrated next. Firstly, in the squeeze stage, the feature map should go through a process which is named global average pooling. The purpose of this process is to convert two-dimensional feature channels into an actual number. In a sense, there is a global receptive field for each actual number so that the low convolutional layer could also use information obtained by the global receptive field. Secondly, the next stage is the excitation operation. It is analogous to the gate mechanism in the recurrent neural network. Each feature channel should get a weight identified by the parameter \( W \), and the parameter \( W \) is regarded as the correlation between feature channels. Finally, the reweight operation always recounts a weight as the importance of each feature channel. Usually, it is the weight of the output of the excitation and then weights the previous feature channel by channel through multiplication to achieve the original feature in the channel dimension recalibration operation.

2.3.2. Improved U-Net Network Architecture. The U-Net network model has an encoder and a decoder, and it includes two parts: the contraction path on the left and the expansion path over there. As Figure 4 shows, the network
contains a $3 \times 3$ convolutional layer, the copy and cropping, and the max-pooling layer. The deconvolution layer has a step size of two. Rectified linear unit (Relu) is a nonlinear activation function used for the entire network structure. According to the feature recalibration, the network could embed the SE block. Then, the interdependence between feature channels is modelled explicitly. It is worth noting that each feature channel would be obtained automatically according to the learning stage. Then, the block promotes useful features and suppresses the less valuable features for the current task.

The contraction path of the network is an encoder that adopts the structure of a convolutional neural network. Two $3 \times 3$ convolution layers are used for feature extraction, and the activation function of Relu is applied to improve the U-net network’s expression ability. Next is the channel attention mechanism used by the 1-layer SE module to add feature maps. Then, a $2 \times 2$ max-pooling layer is used to reduce the resolution of the input signal. The extension path is the decoder, which restores the image size through upsampling. It improves the subdivision capability and contour prediction level of the model through the data features extracted by the joint downsampling process (the visualization result of downsampling and upsampling is exhibited in Figure 5). Finally, a $1 \times 1$ convolutional layer is applied to convert the final output channel number into the number of divided categories.

Our model mainly sets up 23 convolution layers, 4 pooling layers, 8 SE layers, 4 upsampling layers, and 2 dropout layers. Each layer’s structure consists of an input layer, 4 contraction paths, a bottom layer, 4 extension paths, and an output layer. The activation layer is the Relu activation function.

2.3.3. Training Improved the U-Net Model. To make the experiment more rigorous, the segmentation results of retinal vessels should be kept accurate and stable. Compared with the traditional U-Net, batch standardization (BN) is applied to improve the training stage’s stability and reduce the possibility of gradient disappearance. The segmentation performance is optimized, and the effect has remarkable universality. Besides, BN helps the model converge. The rational formula is as follows:
where $x$ represents the input feature, $y$ indicates the standardized feature, $\varepsilon$ is close to zero, and $\gamma$ and $\beta$ denote the training parameters, which can be iteratively updated during the training process.

The cross-entropy loss function is applied in the process of model training. The optimization of the parameters of the whole network would be achieved owing to the Adam algorithm. The parameter update process of the algorithm is as follows:

$$x_n = x_0 - \eta \cdot \frac{m_n}{\sqrt{\text{Var}[x] + \varepsilon}}$$

$$m_n = a_1 \times m_0 + (1 - a_1) f' (x_0),$$

$$v_n = a_2 \times v_0 + (1 - a_2) [f' (x_0)]^2,$$

$$v_n = a_2 \times v_0 + (1 - a_2) [f' (x_0)]^2,$$

where $a_1$, $a_2$ represent the loss rate, $\eta$ is the learning rate, $v_n$, $v_0$ denote the new and old parameters, and $m$ represents the momentum. The algorithm can determine the range of learning rates in an iteration to ensure the stability of parameters and high computational efficiency. In the model, the epoch is set to 10, and an iteration (epoch) is a complete transmission of all training samples. The algorithm learning rate is adjusted to every 5 epochs, and the learning rate is reduced to 0.96. At the same time, the dropout is set to 0.5. Half of the training data are randomly discarded to prevent overfitting.

3. Results and Discussion

3.1. Parameter Configuration. 10 iterations can achieve the ideal training accuracy in this paper. By optimizing the cross-entropy loss function, the pixel segmentation error rate is minimized. This paper uses the Adam algorithm to optimize the loss function, and the learning rate is set to 0.001. The cross-entropy function is defined as follows:

$$H(p, q) = - \sum_x p(x) \log q(x).$$

The cross entropy represents the distance between the two probability distributions. The difficulty of the probability distribution $p(x)$ is represented by the probability distribution $q(x)$, in which $p$ is the correct answer and $q$ is the predicted value. The smaller the cross entropy, the closer the distribution of the two probabilities. The software applied in this experiment are Pycharm (python3.6), Keras, and its TensorFlow port. The running processor is Intel (R) Core (TM) i7-7700HQ CPU @ 2.81 GHz, with 16 GB memory. The GPU graphics card is NVIDIA GeForce GTX 1050, and the operating system is 64 bit Windows 10.

3.2. Evaluation Metrics. We use three commonly used metrics to evaluate the performance objectively, including accuracy (ACC), sensitivity (TPR), specificity (TNR), F1-score, and area under the curve (AUC). Different metrics are calculated as follows:

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$\text{TPR} = \frac{TP}{TP + FN},$$

$$\text{TNR} = \frac{TN}{FP + TN},$$

$$\text{F1-score} = \frac{2 \times TP}{2 \times TP + FP + FN},$$

where ACC represents the proportion of the sum of the correctly segmented pixels to the sum of all pixels, TPR is the ratio of the correctly segmented retinal vessel pixels to the pixels in the gold standard image, and TNR represents the proportion of the correctly segmented background pixels to the real background. TP in the formula is true positive, indicating the number of retinal vessels correctly segmented. TN, which represents the number of background pixels correctly segmented, is a true negative value. FP means false positive, suggesting that the number of pixels segmented into blood vessels by mistake. FN is a false negative, showing the number of pixels erroneously segmented into the background. AUC represents the area under the ROC curve and is also adopted as an essential evaluation metric for vessel segmentation. The values of AUC and F1-score are 1 for a perfect classifier. Figure 6 shows ROC curves of the proposed model training and testing on DRIVE, STARE, and HRF.

3.3. Results. The improved U-Net model is used to extract retinal vessels from the databases of DRIVE, STARE, and HRF color fundus images. The test results are shown in
Figures 7–9, representing the segmentation results of retinal vessels from DRIVE, STARE, and HRF databases. The first column is the original image, and the second column indicates the improved U-Net segmentation result; then, the third column represents the result of expert manual segmentation.

Table 1 shows the quantitative results of ACC, TPR, and TNR obtained by automatic segmentation of retinal vessels from the DRIVE fundus image database. Table 2 exhibits the segmentation of 20 images from the STARE database. Table 3 shows the average values (AVG) and standard deviation (STD) of ACC, TPR, and TNR from the DRIVE, STARE, and HRF color fundus image databases.

3.4. Discussion. The segmentation results using the improved U-Net model are visually compared with the segmentation results of Zhang, Jiang, and Zana’s three algorithms. It can be seen from the segmentation diagram in Figure 10 that the proposed algorithm has better results than the other three algorithms. Compared with Jiang and Zhang’s algorithm, our algorithm has higher integrity and a more significant advantage in vascular continuity. In the segmentation background, our segmentation effect is closer to the manual segmentation result than Zhang and Zana’s algorithms. The vascular trunk and ending connectivity is excellent, but most of the microretinal vessels are better segmented. Therefore, we can provide a valuable reference for clinical diagnosis.

We compared our improved U-Net with several state-of-the-art methods on the HRF database, and the segmentation results are shown in Figure 11. It can be seen subjectively and objectively that AG-UNet [40] and M-GAN [41] have limitations in capturing those thin blood vessels. IterNet [42] seems to miss the vessels around the optic discs on images. NMF+3D U-Net [43] and our method have shown better performance, but our method is more sensitive to small blood vessels and produces more distinct segmentation results. The proposed method is able to detect weak vessels that may be lost in AG-UNet and M-GAN. Thus, it is more powerful to preserve considerable details.

Tables 4 and 5 show the quantitative comparison of the ACC, TPR, and TNR results obtained by the proposed algorithm and partial unsupervised and supervised method for retinal vascular segmentation of DRIVE and STARE fundus image database images, respectively. In the experiment, the result of manual segmentation by the first expert in the DRIVE and STARE fundus image database is used as the ground truth. As Table 4 shows, this algorithm has achieved a higher ACC than most vascular segmentations, except for that of Wang et al. [20] and Khowaja et al. [52], whose methods are based on hierarchical classification. Wang introduced CNN and random forest into hierarchical classification. Thus, the model has strong generalization capability. Khowaja adopts the hierarchical feature extractor instead of the hybrid feature set with subspace learning methods to improve the segmentation performance. It is noteworthy that the TNR performance of this algorithm is excellent, and its performance is better than that of other algorithms. As for TPR, our algorithm is better than that of Zana and Klein [6], Lin et al. [27], and other unsupervised methods. The TPR of the generative adversarial network (GAN) is based on conditional patches as high as 0.7746, second only to the algorithm we proposed. GAN proposed by Khan et al. [28] utilizes a generator network and a patch-based discriminator network conditioned on the sample data with an additional loss function to learn both thin and thick vessels. The model is possible to probe the different patch sizes so that GAN has a certain level of competitiveness with current advanced techniques. Table 5 shows that although the number of vascular images in the training set is not large, our algorithm is still effective in the test set and can obtain ideal segmentation results with excellent robustness and generalization ability. It could be observed from Table 6 that
all methods achieve a similar level of specificity where M-GAN is 0.0036 and NMF+3D U-Net is 0.0008, slightly higher than ours. Comparing to the listed methods, AG-UNet is slightly more sensitive under the HRF database while IterNet is second only to AG-UNet. Our approach and M-GAN both reach preferable performance where the accuracy of M-GAN is 0.0002 higher than ours.

In recent years, many improved structures based on the U-Net model have achieved good segmentation results. However, our improved model is based on the new attention mechanism module proposed in 2020cvpr. ©+_here is no article applying the Squeeze-Attention module to the field of medical image segmentation, so the improved U-Net model we proposed still has a certain degree of innovation and novelty.

To prove the effectiveness of the proposed model, we compared the results of the algorithm with other state-of-the-art modified U-Net as follows. It can be noticed from Table 7 that our algorithm has an advantage over other modified U-Net models. However, the sensitivity of our method on the STARE database is the lowest. One possible reason is that the number of images from the STARE database is too small to achieve the original effect. Therefore, we can learn from other excellent segmentation methods to improve preprocessing and finally achieve the best results.

Figure 7: Visualization of the random samples results for the DRIVE image dataset. (a) Original image, (b) U-Net, and (c) manual.
Figure 8: Visualization of the random samples results for the STARE image dataset. (a) Original image, (b) U-Net, and (c) manual.

Figure 9: Visualization of the random samples results for the HRF image dataset. (a) Original image, (b) U-Net, and (c) manual.

Table 1: Quantitative results of the proposed method for the DRIVE fundus database.

| Image | ACC  | TPR  | TNR  |
|-------|------|------|------|
| 01    | 0.9646 | 0.8386 | 0.9779 |
| 02    | 0.9676 | 0.8001 | 0.9879 |
| 03    | 0.9577 | 0.6764 | 0.9900 |
| 04    | 0.9677 | 0.7663 | 0.9891 |
Table 1: Continued.

| Image | ACC  | TPR  | TNR  |
|-------|------|------|------|
| 05    | 0.9621 | 0.6464 | 0.9947 |
| 06    | 0.9601 | 0.6831 | 0.9911 |
| 07    | 0.9630 | 0.6973 | 0.9938 |
| 08    | 0.9578 | 0.6093 | 0.9943 |
| 09    | 0.9654 | 0.6630 | 0.9938 |
| 10    | 0.9680 | 0.7683 | 0.9877 |
| 11    | 0.9650 | 0.7594 | 0.9882 |
| 12    | 0.9676 | 0.7592 | 0.9901 |
| 13    | 0.9619 | 0.7703 | 0.9838 |
| 14    | 0.9690 | 0.7661 | 0.9886 |
| 15    | 0.9715 | 0.7547 | 0.9897 |
| 16    | 0.9668 | 0.7636 | 0.9880 |
| 17    | 0.9633 | 0.6837 | 0.9909 |
| 18    | 0.9673 | 0.8114 | 0.9816 |
| 19    | 0.9729 | 0.8923 | 0.9811 |
| 20    | 0.9692 | 0.8035 | 0.9831 |

Table 2: Quantitative results of the proposed method for the STARE fundus database.

| Image   | ACC  | TPR  | TNR  |
|---------|------|------|------|
| im0162  | 0.9691 | 0.8231 | 0.9811 |
| im0163  | 0.9727 | 0.8659 | 0.9817 |
| im0235  | 0.9691 | 0.8073 | 0.9849 |
| im0236  | 0.9683 | 0.7868 | 0.9864 |
| im0239  | 0.9515 | 0.5499 | 0.9914 |
| im0240  | 0.9423 | 0.4981 | 0.9940 |
| im0255  | 0.9648 | 0.6624 | 0.9955 |
| im0291  | 0.9709 | 0.4812 | 0.9975 |
| im0319  | 0.9748 | 0.4856 | 0.9973 |
| im0324  | 0.9523 | 0.3371 | 0.9970 |

Table 3: AVG and STD of the evaluation index of the DRIVE, STARE, and HRF fundus database.

| Database | ACC AVG | STD AVG | TPR AVG | ATD AVG | TNR AVG | STD AVG |
|----------|---------|---------|---------|---------|---------|---------|
| DRIVE    | 0.9701  | 0.0287  | 0.8011  | 0.3817  | 0.9849  | 0.0194  |
| STARE    | 0.9683  | 0.0315  | 0.6329  | 0.3987  | 0.9967  | 0.0208  |
| HRF      | 0.9698  | 0.0231  | 0.7478  | 0.3029  | 0.9895  | 0.0115  |

Figure 10: Comparison of different segmentation methods for the DRIVE fundus database. (a) Original image, (b) ground truth, (c) Zhang [31], (d) Jiang and Mojon [39], (e) Zana and Klein [6], and (f) our algorithm.
Figure 11: Vessel segmentation results on the HRF dataset, from up to down is input images, ground truth, AG-UNet, IterNet, M-GAN, NMF+3D U-Net, and ours, respectively.

Table 4: Performance comparison of different methods for the DRIVE fundus database.

| Method                        | Algorithm                  | ACC  | TPR  | TNR  |
|-------------------------------|----------------------------|------|------|------|
| Supervised method             |                            |      |      |      |
| Vlachos and Dermatas [12]     | 0.9290                     | 0.7470| —    | —    |
| Zana and Klein [6]            | 0.9377                     | 0.6971| —    | —    |
| Mendonça and Campilho [8]     | 0.9452                     | 0.7344| 0.9764| —    |
| Jiang and Mojon [39]          | 0.9212                     | 0.6399| —    | —    |
| You et al. [44]               | 0.9434                     | 0.7410| —    | —    |
| Chaudhuri et al. [45]         | 0.8773                     | 0.2716| 0.9785| —    |
| Zhang et al. [46]             | 0.9476                     | 0.7743| 0.9725| —    |
| Zhao et al. [47]              | 0.9540                     | 0.7420| 0.9820| —    |

Unsupervised method
Table 4: Continued.

| Method            | Algorithm | ACC   | TPR   | TNR   |
|-------------------|-----------|-------|-------|-------|
| Supervised method | Fraz et al. [5] | 0.9480 | 0.7406 | 0.9807 |
|                   | Lin et al. [27] | 0.9536 | 0.7632 | —     |
|                   | Fu et al. [26] | 0.9470 | 0.7294 | —     |
|                   | Niemeijer et al. [7] | 0.9416 | — | —     |
|                   | Maji et al. [21] | 0.9470 | — | —     |
|                   | Soares et al. [48] | 0.9466 | 0.7332 | 0.9782 |
|                   | Staal et al. [49] | 0.9442 | 0.7322 | 0.9646 |
|                   | Zhu et al. [50] | 0.9607 | 0.7528 | 0.9820 |
|                   | Marin et al. [18] | 0.9452 | 0.7067 | 0.9801 |
|                   | GAN [51] | 0.9562 | 0.7746 | 0.9753 |
|                   | Wang et al. [20] | 0.9767 | 0.8173 | 0.9733 |
|                   | Khowaja et al. [52] | 0.9753 | 0.8176 | 0.9709 |
|                   | Our algorithm | 0.9701 | 0.8011 | 0.9849 |

Table 5: Performance comparison of different methods for the STARE fundus database.

| Method            | Algorithm | ACC   | TPR   | TNR   |
|-------------------|-----------|-------|-------|-------|
| Unsupervised method | Zhang et al. [46] | 0.9554 | 0.7791 | 0.9758 |
|                   | Hoover et al. [13] | 0.9264 | 0.6747 | 0.9565 |
|                   | Mendonça and Campilho[8] | 0.9440 | 0.6996 | 0.9730 |
|                   | Jiang and Mojon [39] | 0.9009 | — | —     |
|                   | You et al. [44] | 0.9497 | 0.7260 | —     |
|                   | Zhao et al. [47] | 0.9560 | 0.7800 | 0.9780 |
| Supervised method | Fraz et al. [5] | 0.9534 | 0.7548 | 0.9763 |
|                   | Lin et al. [27] | 0.9603 | 0.7423 | —     |
|                   | Fu et al. [26] | 0.9545 | 0.7140 | —     |
|                   | Niemeijer et al. [7] | 0.9534 | 0.7548 | 0.9763 |
|                   | Soares et al. [48] | 0.9480 | 0.7207 | 0.9747 |
|                   | Staal et al. [49] | 0.9516 | — | —     |
|                   | Marin et al. [18] | 0.9526 | 0.6944 | 0.9819 |
|                   | GAN [51] | 0.9647 | 0.7940 | 0.9869 |
|                   | Wang et al. [20] | 0.9813 | 0.8104 | 0.9791 |
|                   | Khowaja et al. [52] | 0.9751 | 0.8239 | 0.9749 |
|                   | Our algorithm | 0.9683 | 0.6032 | 0.9967 |

Table 6: Performance comparison of different methods for the HRF fundus database.

| Method            | ACC   | TPR   | TNR   |
|-------------------|-------|-------|-------|
| AG-UNet [40]      | 0.9600 | 0.8297 | 0.9875 |
| IterNet [42]      | 0.9623 | 0.7524 | 0.9743 |
| M-GAN [41]        | 0.9700 | 0.6948 | 0.9931 |
| NMF+3D U-Net [43] | 0.9688 | 0.7451 | 0.9903 |
| Our algorithm     | 0.9698 | 0.7478 | 0.9895 |

Table 7: Performance analysis of different modified U-Nets on DRIVE and STARE databases with respect to the measuring metrics.

| Method            | DRIVE ACC | TPR | TNR | STARE ACC | TPR | TNR |
|-------------------|-----------|-----|-----|-----------|-----|-----|
| MSFFU-Net [32]    | 0.9694 | 0.7762 | 0.9835 | 0.9537 | 0.7721 | 0.9885 |
| Dense U-net [34]  | 0.9511 | 0.7986 | 0.9736 | 0.9538 | 0.7914 | 0.9722 |
| AA-UNet [53]      | 0.9558 | 0.7941 | 0.9798 | 0.9640 | 0.7598 | 0.9878 |
| DUNet [54]        | 0.9566 | 0.7963 | 0.9800 | 0.9641 | 0.7595 | 0.9878 |
| EEA U-net [55]    | 0.9577 | 0.7918 | 0.9708 | 0.9445 | 0.8021 | 0.9561 |
| Our algorithm     | 0.9701 | 0.8011 | 0.9849 | 0.9683 | 0.6032 | 0.9967 |
4. Conclusions
The segmentation of retinal vascular images automatically plays a vital role in computer-aided fundus diagnosis and disease screening. There are some differences between retinal vessels and background. Nevertheless, the contrast is not strong enough, which makes the accurate segmentation of retinal vessels harder [56]. Therefore, the segmentation of fundus retinal vessels automatically is a complex and challenging mission. To deal with the problem that the accuracy is not enough in retinal blood vessel segmentation, a new algorithm that could segment retinal blood vessel automatically comes up. The algorithm is based on an improved U-Net segmentation model in this paper. Firstly, the RGB fundus image is preprocessed; secondly, the U-Net model is constructed. The extracted image block could be used during the training stage in the U-Net model; finally, the trained U-Net model is tested to obtain the final results to achieve automatic fundus image segmentation. The experimental results of retinal blood vessel segmentation on the DRIVE, STARE, and HRF color fundus image database show that the algorithm has high accuracy in extracting retinal vessels and can retain more microscopic vessels and complete network. This algorithm can provide necessary theoretical and technical support for ophthalmologists to track the development of fundus lesions and reveal the pathogenesis with a particular clinical reference value.

Although the method of the U-Net network proposed in our paper has made some achievements in the segmentation of retinal vessels, there are still some problems and challenges. First of all, the number of retinal fundus images is limited. Due to ethical issues such as protecting patients’ privacy, it is not easy to obtain fundus images, and the amount of data is not abundant. Simultaneously, the ground truth corresponding to RGB fundus images relies on expert observation and manual segmentation, so the database arrangement is tedious. The amount of data in the existing public fundus image database is small, which has some limitations to the training model. Secondly, the RGB fundus image in the image set is affected by uneven illumination and low image resolution, which cannot guarantee the effective recognition of blood vessels. Given the shortcomings, our research’s subsequent work is to integrate or postprocess the results to further optimize the segmentation results based on the U-Net network combined with effective technologies and methods.

Data Availability
The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
No conflicts of interest are declared by the authors.

Acknowledgments
This work was funded in part by the National Natural Science Foundation of China (Grant no. 62072413) and also supported by the Natural Science Foundation of Zhejiang Province of China (Grant no. LY16F010008).

References
[1] N. I. Kurysheva, “Selective elective ivy national natural science foundation of China,” Vestnik Ofthalmologii, vol. 135, no. 2, pp. 144–150, 2019.
[2] S. S. Kar and S. P. Maity, “Automatic detection of retinal lesions for screening of diabetic retinopathy,” IEEE Transactions on Biomedical Engineering, vol. 65, no. 3, pp. 608–618, 2018.
[3] C. Bai, L. Huang, X. Pan, J. Zheng, and S. Chen, “Optimization of deep convolutional neural network for large scale image retrieval,” Neurocomputing, vol. 303, no. 16, pp. 60–67, 2018.
[4] P. Zhou, C. Bai, J. Xia, and S. Chen, “CMRDF: a real-time food alerting system based on multimodal data,” IEEE Internet of Things Journal, vol. 99, p. 1, 2020.
[5] M. M. Fraz, P. Remagnino, A. Hoppe et al., “An ensemble classification-based approach applied to retinal blood vessel segmentation,” IEEE Transactions on Biomedical Engineering, vol. 59, no. 9, pp. 2538–2548, 2012.
[6] F. Zana and J.-C. Klein, “Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation,” IEEE Transactions on Image Processing, vol. 10, no. 7, pp. 1010–1019, 2001.
[7] M. Niemeijer, J. Staal, B. van Ginneken et al., “Comparative study of retinal vessel segmentation methods on a new publicly available database,” Proceedings of SPIE the International Society for Optical Engineering, vol. 5370, pp. 648–656, 2004.
[8] A. M. Mendonça and A. Campilho, “Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction,” IEEE Transactions on Medical Imaging, vol. 25, no. 9, pp. 1200–1213, 2006.
[9] S. Garg, S. Jayanthi, and S. Chandra, “Unsupervised curvature-based retinal vessel segmentation,” in Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro, pp. 344–347, IEEE, Washington, DC, USA, April 2007.
[10] O. Chutatape, Z. Liu, and S. M. Krishnan, “Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters,” in Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society, vol. 6, pp. 3144–3149, Hong Kong Sar, China, November 1998.
[11] L. Gang, O. Chutatape, and S. M. Krishnan, “Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter,” IEEE Transactions on Biomedical Engineering, vol. 49, no. 2, pp. 168–172, 2002.
[12] M. Vlachos and E. Dermatas, “Multi-scale retinal vessel segmentation using line tracking,” Computerized Medical Imaging and Graphics, vol. 34, no. 3, pp. 213–227, 2010.
[13] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, “Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response,” IEEE Transactions on Medical Imaging, vol. 19, no. 3, pp. 203–210, 2000.
[14] M. Zardadi, N. Mehrshad, and S. M. Razavi, “Unsupervised segmentation of retinal blood vessels using the human visual system line detection model,” Information Systems & Telecommunication, vol. 4, pp. 125–133, 2016.
[15] C. L. Srinidhi, P. Aparna, and J. Rajan, “Automated method for retinal artery/vein separation via graph search
active contour model with hybrid region information with application to retinal images,” *IEEE Transactions on Medical Imaging*, vol. 34, no. 9, pp. 1797–1807, 2015.

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

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

[50] C. Zhu, Y. Xiang, B. Zou et al., “Retinal vessel segmentation in fundus images using CART and AdaBoost,” *Journal of Computer-Aided Design and Computer Graphics*, vol. 26, no. 3, pp. 445–451, 2014.

[51] W. Abbas, M. H. Shakeel, N. Khurshid et al., “Patch-based generative adversarial network towards retinal vessel segmentation,” in *Proceedings of the Neural Information Processing*, Vancouver, Canada, December 2019.

[52] S. A. Khowaja, P. Khuwaja, and I. A. Ismaili, “A framework for retinal vessel segmentation from fundus images using hybrid feature set and hierarchical classification,” *Signal, Image and Video Processing*, vol. 13, no. 2, pp. 379–387, 2019.

[53] Y. Lv, H. Ma, J. Li, and S. Liu, “Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation,” *IEEE Access*, vol. 8, pp. 32826–32839, 2020.

[54] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, “DUNet: a deformable network for retinal vessel segmentation,” *Knowledge-Based Systems*, vol. 178, no. 15, pp. 149–162, 2019.

[55] V. Sathananthavathi and G. Indumathi, “Encoder enhanced atrous (EEA) unet architecture for retinal blood vessel segmentation,” *Cognitive Systems Research*, vol. 67, no. 4, pp. 84–95, 2021.

[56] Z. Gao, Y. Li, Y. Sun et al., “Motion tracking of the carotid artery wall from ultrasound image sequences: a nonlinear state-space approach,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 1, pp. 273–283, 2017.