Retinal Vessel Segmentation Algorithm Based on Residual Convolution Neural Network

Shuang Xu [1,2]*, Zhiqiang Chen [1,2], Weiyi Cao [3], Feng Zhang [1,2] and Bo Tao [2,3]

[1]Key Laboratory of Metallurgical Equipment and Control Technology, Ministry of Education, Wuhan University of Science and Technology, Wuhan, China, [2]Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan, China, [3]Precision Manufacturing Institute, Wuhan University of Science and Technology, Wuhan, China

Retinal vessels are the only deep micro vessels that can be observed in human body, the accurate identification of which has great significance on the diagnosis of hypertension, diabetes and other diseases. To this end, a retinal vessel segmentation algorithm based on residual convolution neural network is proposed according to the characteristics of the retinal vessels on fundus images. Improved residual attention module and deep supervision module are utilized, in which the low-level and high-level feature graphs are joined to construct the encoder-decoder network structure, and atrous convolution is introduced to the pyramid pooling. The experiments result on the fundus image data set DRIVE and STARE show that this algorithm can obtain complete retinal vessel segmentation as well as connected vessel stems and terminals. The average accuracy on DRIVE and STARE reaches 95.90 and 96.88%, and the average specificity is 98.85 and 97.85%, which shows superior performance compared to other methods. This algorithm is verified feasible and effective for retinal vessel segmentation of fundus images and has the ability to detect more capillaries.

Keywords: retinal vessel segmentation, convolution neural network (CNN), residual network, fundus image, attentional mechanism, deep supervision

1 INTRODUCTION

The deep neural network is a typical bio-inspired intelligence computation technique, which based on the principles of biological processes that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Due to the incredible abilities to solve complex problems, it has attracted much attention from many scholars and have been successful applied to solve complex real-world problems (Sun et al., 2020b; Sun et al., 2020c; Chen et al., 2021a; Jiang et al., 2021a). Retinal vascular occlusion, hypertensive arteriosclerosis and diabetic retinopathy are the most common diseases in retinal diseases and also the main cause of blinding in the world (Horton et al., 2016). It is estimated that the number of people with vision loss will double by 2050 (Varma et al., 2016). Early detection and treatment can preserve 90% vision and also help the auxiliary medical management departments to formulate preventive measures to reduce the number of newly diagnosed cases and reduce the medical related economic burden (Das, 2016; Guo et al., 2018; Sun et al., 2020a). The morphological structure of retinal vessels has important reference value for the diagnosis of the diseases. Accurate and rapid segmentation of retinal vessels is necessary for the treatment. However, the segmentation requires manual labeling by experts at present, which is not

FIGURE 1 | Schematic diagram of residual module.
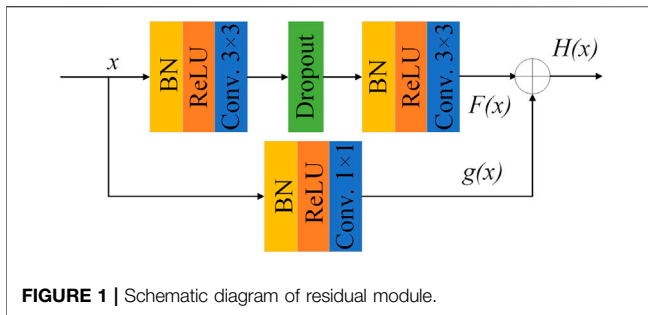
only time-consuming and laborious, but also not accurate enough to carry out large-scale labeling and segmentation. It is significant to develop retinal vessel segmentation algorithms to improve the intelligence of computers in the aspect of disease diagnosis and health screening. Due to the excellent performance of bio-inspired computation method (Deng et al., 2020; Zhao et al., 2020; Tao et al., 2021a; Jiang et al., 2021b), researchers tried various types of algorithms to improve retinal vessel recognition, focusing on the segmentation and extraction of detailed information of connected vessel stems and terminals.

In addition, retina is one of the most reliable, stable and hard-forged information among all the biological features used for identification. As early as the 1930s, some foreign scholars proposed the unique theory of the distribution of retinal vessels. Subsequent studies have shown that the distribution of retinal vessels is different, even for twins (Cao et al., 2017; Zhang 2020). Aside from changes in retinal features due to trauma and disease, the shape of the retinal vessels remains stable throughout life, making it ideal for identification. In the foreseeable future, retina recognition technology has a great hope to be applied to online payment, access control, automatic withdrawal and other civil fields with high security requirements. Therefore, the research on retina recognition technology has great value and good prospects.

In this paper, we proposed a novel retinal vessel segmentation algorithm based on residual convolution neural network, which involves three major steps: 1) the residual learning is introduced in the network structure; 2) the atrous spatial pyramid pooling is built to learn the feature information of different receptive fields; 3) the residual attention module and deep supervision module are applied to improve the accuracy of identifying the capillaries.

The main contribution of this work is to propose the novel algorithm for segmenting retinal vessel from fundus images. It outperforms many recent works, including several methods using deep learning. The proposed algorithm can obtain complete retinal vessel segmentation, including connected vessel stems and terminals, especially the capillaries, and is fast and easily scalable to any fundus image size. Three more specific contributions are also worth mentioning. Firstly, an improved residual attention module is built and combined with a designed deep supervision module, that successfully solves the problem of gradient disappearance and gradient explosion caused by the depth of convolutional neural network. Secondly, an encoder-

decoder network structure is constructed, in which the low-level and high-level feature graphs are joined together. It effectively avoids inefficient learning and sharing in training. Thirdly, atrous spatial pyramid pooling is constructed by introducing the atrous convolution, that effectively enlarges the receptive field while reducing the number of training parameters. Our proposed algorithm can be used to help doctors diagnose retinal disease, and could also support future computer-assisted diagnosis, health screening, and retina identification.

The rest of this paper is organized as follows: *Related Work* discusses the related work contributed by researchers, followed by the *Network Structure*, in which each block has been described in detail. *Experiment and analysis* indicate the data acquisition, experimental results, comparison and analysis. *Conclusion and future work* conclude the paper with a summary and future research directions.

2 RELATED WORK

Retina is made up of complex blood vessels surrounding the tiny nervous system at the back of the eyeball, which contains a large number of features. Researchers have proposed many retinal vessel segmentation algorithms over the years, including matching filtering method (Singh and Srivastava, 2016; Roy et al., 2019), vascular tracking method (Pal et al., 2019; Alaguselvi and Murugan, 2021), image morphology processing and deep learning method (Grewal et al., 2018; Soomro et al., 2019).

Traditional image segmentation methods, as matching filtering method and vascular tracking method, focus on the various filters design and the image morphology process to achieve the purpose of retinal vessel segmentation (Li G. et al., 2019; Li J. et al., 2021). In the reference (Pachade et al., 2020), a unique combination of morphological operations, background estimation, and iterative thresholds was applied to achieve the retinal vessel segmentation. Girard et al. (2019) defined branches by nodes and combined with graph propagation to do the segmentation and classification, and Li et al., 2020b used a deep forest-based segmentation algorithm for retinal vessels. However, in some lesion areas and optic disc edges, these methods may incorrectly detect points as blood vessels. Fan et al. (2019) combined the matched filters and morphological process based on different Gaussian filters in different directions and vector field divergence. In the meantime, multi-scale wavelet transforms (Tian et al., 2021) was used to fuse feature images, and the maximum value of each pixel was calculated to obtain retinal vessel detection images. But the interference of optic disc will lead to the degradation of segmentation performance. Dharmawan et al., 2019a designed a retinal vessel segmentation method based on adaptive filter. Rodrigues and Marengoni, 2017 proposed a method based on morphology and wavelet transform, and Aguirre-Ramos et al. (2018) enhanced vascular contour through Gabor filter and Gaussian fractional derivative. Lu et al. (2016) used multi-scale filtering algorithm to preprocess images and Li et al., 2018 incorporated phase features to segment the retinal vessels. These traditional segmentation methods
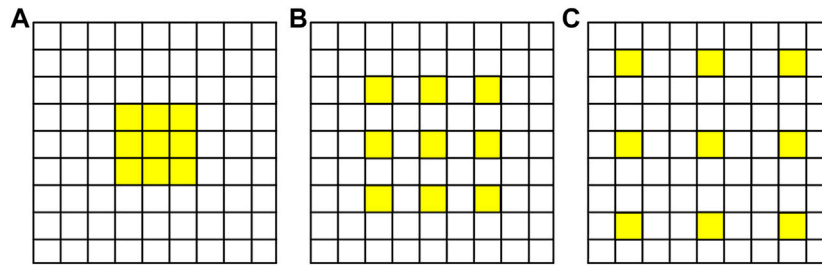
FIGURE 2 | Schematic diagram of atrous convolution with different expansion rates as **(A)** $r = 1$, **(B)** $r = 2$, and **(C)** $r = 3$.
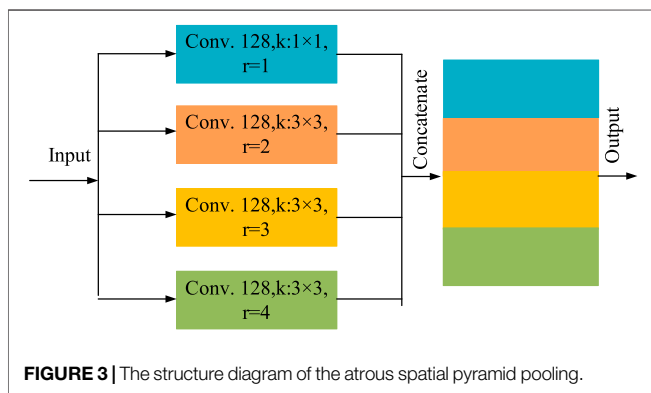

FIGURE 3 | The structure diagram of the atrous spatial pyramid pooling.

FIGURE 4 | The schematic diagram of residual attention module.

preliminarily achieved retinal vessel segmentation, but it still needs manual extraction of the image features due to the low accuracy in details, which may omit key details and fail to achieve end-to-end segmentation.

Deep learning based segmentation method developed rapidly in recent years. Yan et al. (2019) used a three-stage network model to segment the thin and thick vessels respectively, and then segmented the pixels by fusing the vessels. Liu (2021) proposed a method based on an optimized BP neural network, in which image features are extracted by adaptive histogram, matched filter and Hessian matrix. But the algorithm can only be applied on lesions in small area, and the large-scale lesion interference cannot be effectively avoided. Lu et al. (2021) segmented retinal vessels in fundus images through an attentional mechanism and conditional generative adversarial network, and Tang and Yu, 2021 adopted a BP neural network. In Reference (Dharmawan et al., 2019a), a hybrid algorithm was proposed by using a directional sensitive enhancement method and U-NET convolution network to train the enhanced image. But the algorithm is not optimized for multi-scale image segmentation, and the segmentation performance of vessels with lesions still needs improvement. These deep learning based methods have improved the accuracy of retinal vessel segmentation, but the segmentation performance still needs to be significantly improved in order to be widely used in machine-assisted health screening and identification in the future.

3 NETWORK STRUCTURE

Residual Module

To avoid the inevitable problems of deep neural networks as gradient disappearance and gradient explosion, residual learning (Chen et al., 2020; Feng et al., 2020; Yang et al., 2021) is introduced into the network. The parallel method is applied
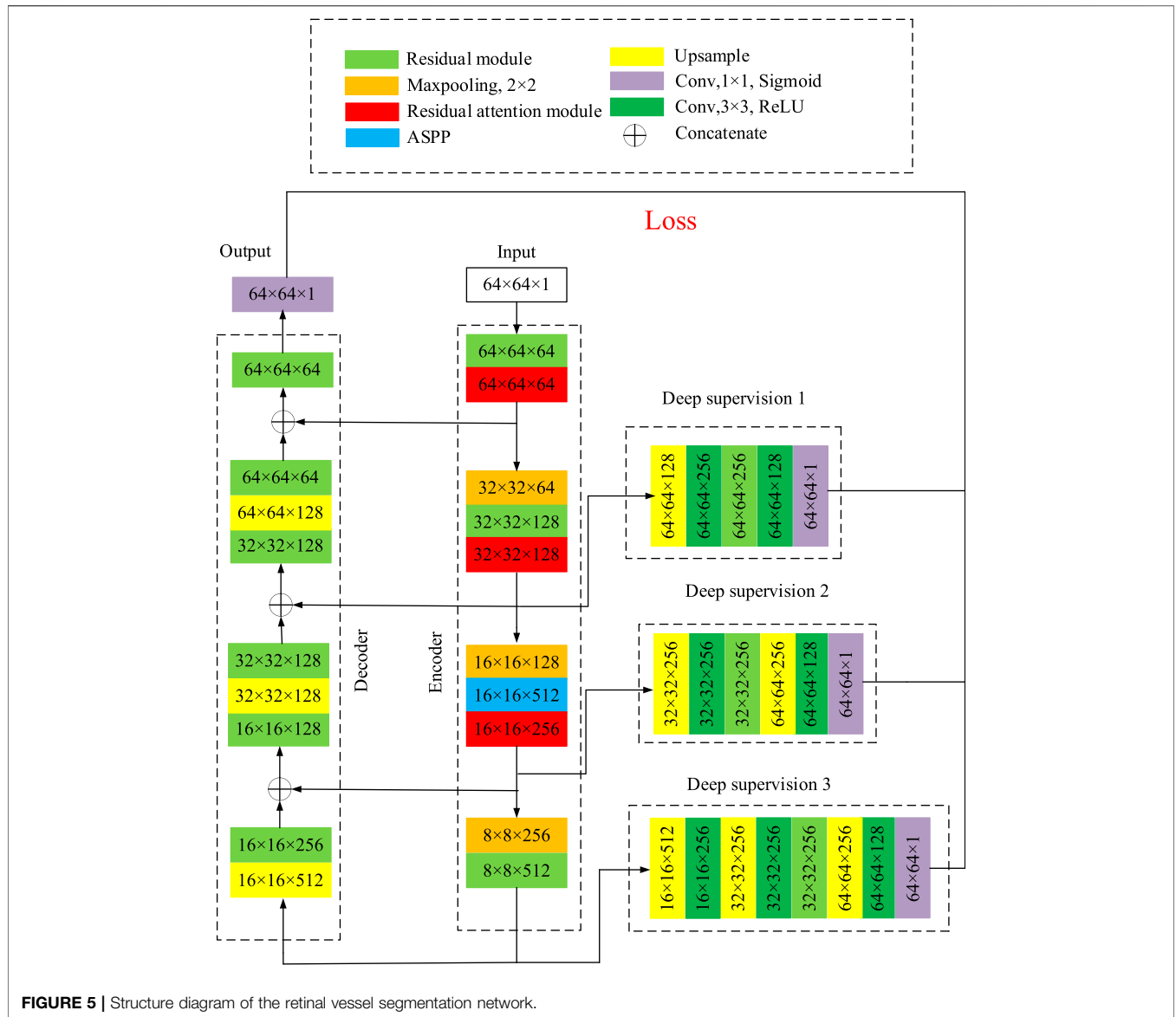
FIGURE 5 | Structure diagram of the retinal vessel segmentation network.

and the identity mapping is added to the output of the stacked convolution layer. It can effectively improve the feature extraction ability of the network. The function can be expressed as

$$F(x) = H(x) - x \qquad (1)$$

where x represents the input; $F(x)$ represents the output of the jagged edge. If $F(x) = 0$, it becomes an identity mapping, while the input and output of the residual module are equal. $H(x)$ represents the final output of the residual module. The structure of the proposed residual module is shown in **Figure 1**. Different from the ordinary residual module, the 1×1 convolution is added to the identity mapping to adjust the number of channels, whose function can be expressed as

$$H(x) = F(x) + g(x) \qquad (2)$$

Here $g(x)$ represents the output of the convolution on the identity map.

The residual mapping of the residual module contains two 3×3 convolution layers and each convolution layer is processed by batch normalization (BN) (Awais et al., 2020) to accelerate the network convergence. Modified linear unit (ReLU) is used as the activation function, and L2 regularization is introduced to avoid network over-fitting. A dropout layer (Wang et al., 2019) (with random inactivation rate = 0.2) is added between the two stacked convolutional layers in the residual module to randomly discard some neurons during training, in order to prevent over-fitting and enhance generalization performance of the network.

Atrous Spatial Pyramid Pooling

The receptive field of the convolution layer is related to the size of the convolution sum. Larger size means larger receptive field and

stronger feature extraction capability of the network, but it also means more parameters to be trained. By utilizing atrous convolution, also called expansion convolution or extended convolution, firstly proposed by Chen et al., 2018, the receptive field can be enlarged without increasing the number of training parameters of the network. Compared with traditional convolution, atrous convolution introduced a hyperparameter called expansion rate r. The larger the value of r is, the larger the receptive field will be. When $r = 1$, the atrous convolution is equal to the ordinary convolution. **Figure 2** is the schematic diagram of atrous convolution with different expansion rates.

In order to enlarge the receptive field of the network without adding too many parameters, so as to increase the feature extraction capability, we utilize the outputs of atrous convolution with different expansion rates and stack them to form the atrous spatial pyramid pooling (ASPP) (Li M. et al., 2021; Lian et al., 2021), The structure is shown as **Figure 3**.

This module is mainly composed of four parallel atrous convolution. including a 1×1 ordinary convolution ($r = 1$) and three 3×3 atrous convolution with expansion rate $r = 2, 3$ and 4 respectively. The number of convolution kernels of each convolution layers is set as 128. The outputs of the four convolution layers are concatenated as the total output of the ASPP. The multi-scale characteristic information of different receptive fields can be learned with different expansion rates, which can increase the recognition ability of small vessels. In addition, it also reduces the parameters that need to be trained and increases the training speed of the network.

Residual Attention Module

With the rapid development of deep learning, it has become particularly important to add attention mechanism into the network in recent years. The attention mechanism of image recognition is mostly constructed by masks. Essentially, it is to train the weight of another layer to identify the key information in the image, so as to increase the sensitivity of the network to the key information. The network can be trained to notice the key areas of each image to generate attention. Spatial transformer network (STN) model (Xu et al., 2021b) uses the attention mechanism to transform the spatial information of the original picture into another space while retaining the key information. In the meantime, Reference (Hu et al., 2020) proposed a SENet model, whose core idea is to learn the weight of each channel through the attention module, in order to generate attention in the channel domain.

In this paper, we propose a residual attention module, whose network structure is shown as **Figure 4**. This module is mainly composed of two parts, mask branch and trunk branch, whose output is represented by $M(x)$ and $T(x)$ respectively. Batch normalization and Rectified Linear Unit (ReLU) activation functions are used by default for all 3×3 convolutional layers. The trunk branch is composed of two 3×3 convolution layers and jump connections, which is used to extract feature information. In the mask branch, two down sampling and two up-sampling are carried out. After the process of Sigmoid activation function, a mask $M(x)$ with the
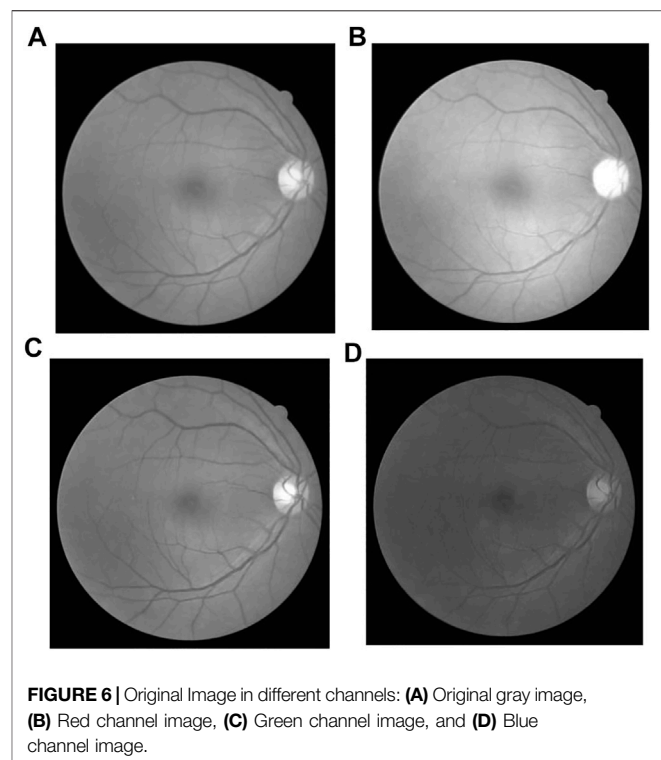

FIGURE 6 | Original Image in different channels: **(A)** Original gray image, **(B)** Red channel image, **(C)** Green channel image, and **(D)** Blue channel image.

same size as the output of the trunk branch is obtained as the weight of the trunk output $T(x)$. It can be expressed as

$$H(\text{x}) = (1 + M(x))T(x) \qquad (3)$$

where x represents the input, $M(x)$ represents the output of the mask branch, $T(x)$ represents the output of the trunk branch, and $H(x)$ represents the output of the residual attention module.

Retinal Vessel Segmentation Network

The structure of the segmentation network proposed in this paper is shown as **Figure 5**, which is mainly composed of encoder, decoder and three deep supervision modules.

The large size of the input image may affect the segmentation, so it is necessary to cut apart the fundus image. The image is cut by windowing of 64×64 pixels, and the sliding step is set to be 16. The dataset can be expanded by randomly flipping, rotation and cropping the images. In **Figure 5**, the width and height of the input image of the network is 64, and the number of channels is 1, which is written as $64 \times 64 \times 1$. During each step, the size of the image and the number of channels will change. The number represents the width $\times$ height $\times$ channel number, as shown in **Figure 5**. In the encoder network, the residual module and the ASPP, together with the residual attention module, will not change the width and height of the image, but increase the number of channels. The 2×2 maxpooling layer reduce the width and height of the image by half and keeps the number of channels constant. In the decoder network, up-sample block doubles the width and height of the image, and the residual
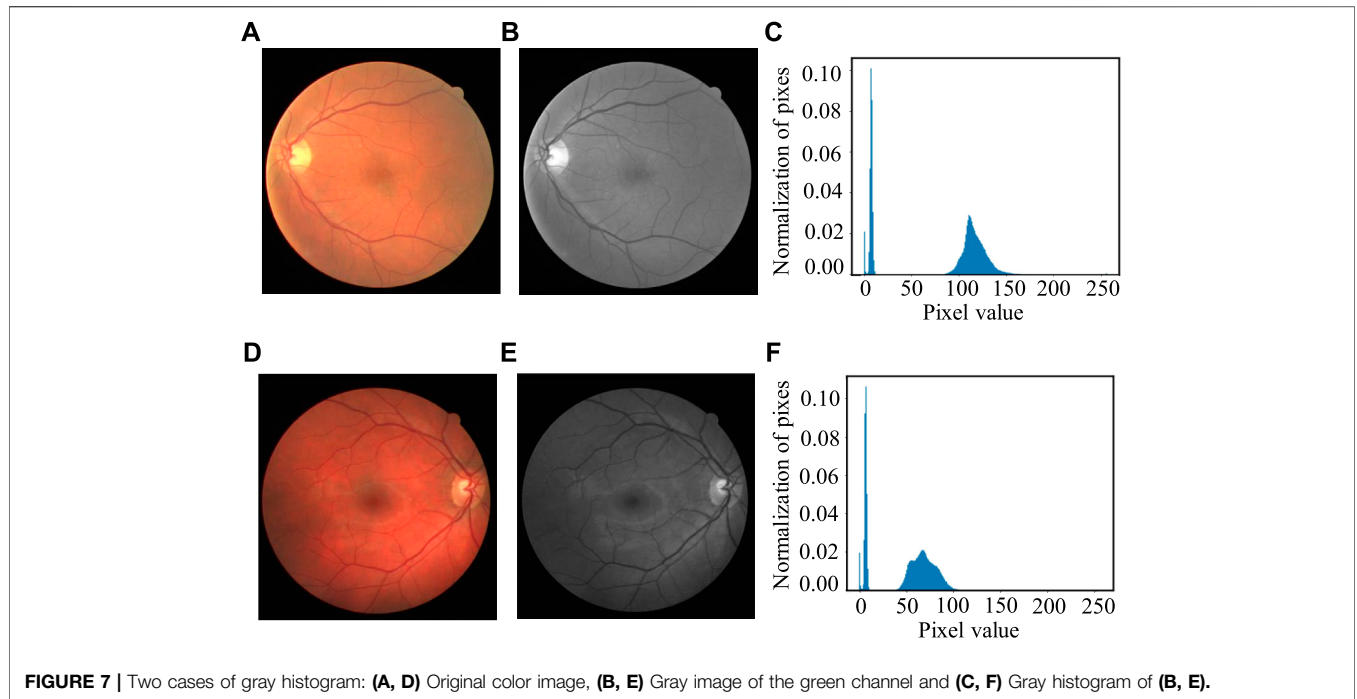
FIGURE 7 | Two cases of gray histogram: **(A, D)** Original color image, **(B, E)** Gray image of the green channel and **(C, F)** Gray histogram of **(B, E).**

module adjusts the number of channels. After each stage of up-sampling and residual modules, stack with the branches on the encoder. The dimension of the image after decoder network is 64 × 64×64. Finally, after the residual module and the convolution layer with kernel size of 1 × 1, the width and height of the image became 64 and the number of channels is 1. As a result, the dimension of the output image is 64 × 64 × 1.

Three branches are introduced in the encoder as the input of the three deep supervision modules with different layers and blocks. The output of the three deep supervision modules and the output of the decoder are used together to calculate the loss function and update the weight parameters. Up-sampling is carried out by means of transposed convolution, and a symmetric coding-decoding network is constructed by splicing the low-level features of encoder and high-level features of decoder with the method of jump connection. A 1 × 1 convolution is used at the last layer of the decoder to adjust the number of channels to 1. Finally, the Sigmoid activation function is used to scale the output to the range from 0 to 1, which can be expressed as

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

where x represents the input of the activation function, $f(x)$ represents the output of the activation function, whose value is normalized.

The loss function of the network is composed of the loss generated by three depth supervision modules and the trunk network. The function can be written as

$$Loss = \beta Loss1 + \beta Loss2 + \beta Loss3 + \beta Loss4 \tag{5}$$

Here, $Loss$ represents the total loss of the network, $Loss1$, $Loss2$ and $Loss3$ represent the loss generated by the three deep supervision modules respectively, and $Loss4$ represents the loss generated by the output layer of the decoder. $\beta = 1 - \frac{\widehat{epochs}}{epochs}$, $\widehat{epochs}$ represents the iteration times of the current network, and $epochs$ represents the total iteration times of the network. It can be seen that as the iteration of the network approaches 0, the weight of loss β generated by the deep supervision module decreases gradually.

Since the pixel number of vascular and non-vascular may can be quite different, the binary cross entropy loss function with weight coefficient (Li X. et al., 2019; Jamin and Humeau-Heurtier, 2020) is adopted to reduce the uneven distribution of positive and negative samples, and its mathematical expression is defined as

$$Loss(n) = -1/m \sum_{i=1}^{m} \partial y_i log\left(\widehat{y_i}\right) + (1 - \partial)(1 - y_i)log\left(1 - \widehat{y_i}\right),$$
$$n = 1, 2, 3, 4 \tag{6}$$

Here, $Loss(n)$ represents the $Loss1$, $Loss2$, $Loss3$ and $Loss4$ in **Eq. 5**, with $n = 1,2,3,4$ respectively; m represents the total number of pixels in the input image; y_i represents the label with value of 0 or 1, 0 represents background, and 1 represents blood vessel. $\widehat{y_i}$ represents the output of the network, $\partial = X_-/m$, $1 - \partial = X_+/m$, where X_- and X_+ represent the number of non-vessel pixels and the number of vessel pixels respectively.
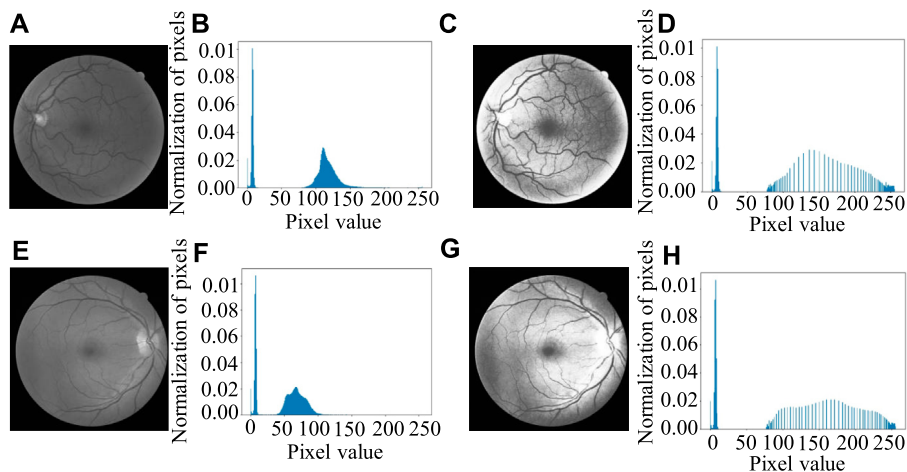
FIGURE 8 | Histogram equalization results: **(A, E)** Gray image of green channel, **(B, F)** Original gray histogram of **(A, E)** separately, **(C, G)** Image after local histogram equalization, **(D, H)** Gray histogram after local histogram equalization.

The batch size of the training is 20, the network iterates for 500 times, the learning rate is set at 0.001, and Adam is used as the optimizer.

4 EXPERIMENT AND ANALYSIS

Database and Training Environment

The network model is established on deep learning framework based on TensorFlow. The hardware configuration of the experiment is i5-1100K and GTX1030, and the software runs on the Win10 system. The database used for the experiment includes Digital Retina Images for Vessel Extraction (DRIVE) (Staal et al., 2004) and Structured Analysis of the Retina (STARE) (Hoover et al., 2000). The DRIVE data set has 20 training images and 20 test images with the resolution of 584 × 584 pixels. The STARE data set contains 20 images with the resolution of 605 × 700 pixels, in which ten images are used as the training set and the other ten as the test set.

Image Preprocessing

Environmental factors, such as illumination, interference and background, will affect the image segmentation, which leads to unsatisfactory result. Therefore, in order to further improve the accuracy of vessel segmentation, appropriate preprocessing operations are required for fundus images. Firstly, the image is transformed into grayscale image, and secondly, Gaussian filter is carried out to eliminate the noise. Thirdly, local histogram equalization (Dhal et al., 2020; Shi, 2021) and Gamma transformation are performed to adjust the contrast of the image.

4.1.1 Channel Separation

In the fundus examination, an RGB color image is obtained by the camera, composed of three separate channels: red, green and blue. The images are shown as **Figure 6**, in which (**Figure 6A**) is the original gray image, that is actually the
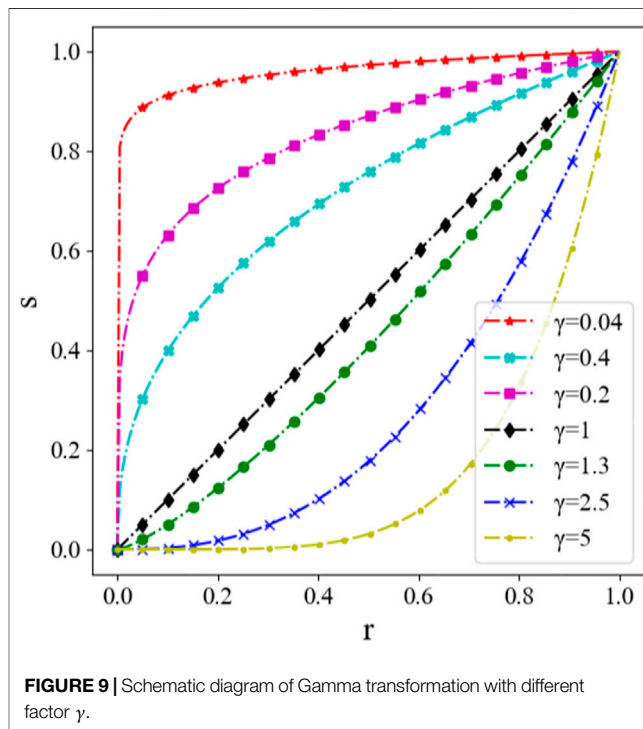

FIGURE 9 | Schematic diagram of Gamma transformation with different factor γ.

average grayscale image of the three channels, and (**Figures 6B,C,D** are the original images in red, green, blue channel separately. The three-channel color image is too large and contains much useless data for subsequent processing. Single channel image information is enough for all the required information, which greatly reduce the amount of data processing and improve the computational efficiency. Channel compression and conversion is usually carried out by the weighted average of three images, or an optimal channel selection. To the human eye, the sensitivity to green is much
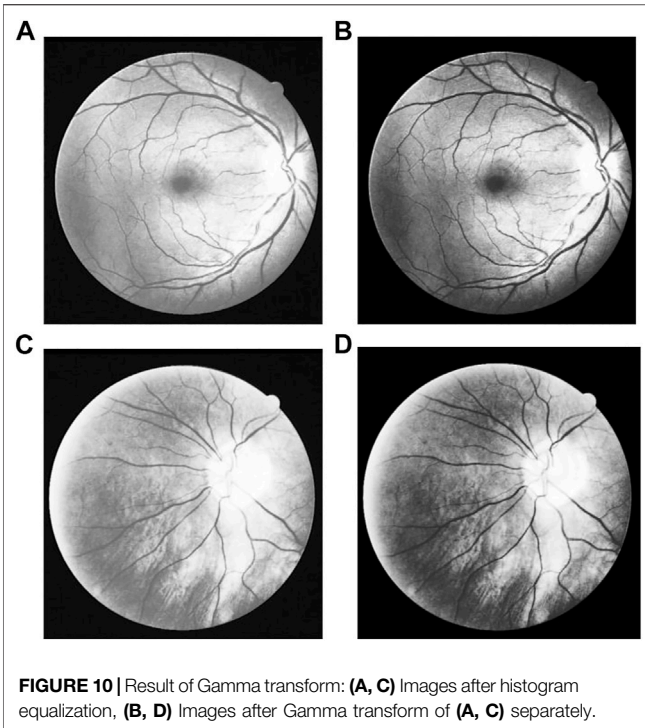
FIGURE 10 | Result of Gamma transform: **(A, C)** Images after histogram equalization, **(B, D)** Images after Gamma transform of **(A, C)** separately.

higher than the other two colors (Ricci and Perfetti, 2007; Marin et al., 2011), as the contrast shown in **Figure 6C**. Therefore, the image in green channel is chosen instead of the original three-cannel image for subsequent image processing.

4.1.2 Gray Histogram Equalization

Histogram (Li et al., 2020a; Sulewski 2020), also known as mass distribution map, is used for statistical reports, in which multiple bars with unequal heights are utilized to represent the distribution of data. During the image process here, the gray histogram represents the image distribution of gray level in the range [0, L-1]. The discrete function can be described as

$$p(r_k) = n_k/n \qquad (7)$$

Here, n is the total number of images; n_k refers to the total number of pixels at the kth gray level; r_k refers to the kth gray level, and $k = 0, 1, 2... L$-1. In our experiment, L is set as 256. The gray histogram represents the frequency of pixels at each gray level shows in the image, with the gray level as the x-axis and the numbers of the pixels as the y-axis. **Figure 7** shows two cases of the gray histogram of fundus images, with **(Figures 7A,D)** the original color images, **(Figures 7B,E)** the gray image after channel separation; **(Figures 7C,F)** the gray histogram of **(Figures 7B,E)** with the vertical axis as the normalized probability distribution of the pixels.

The goal of histogram equalization is to transform the original image from a certain concentration range of pixel values to a wider range so that the contrast between the similar gray value increase. The specific way is utilizing nonlinear stretch to

transform the certain histogram to uniform distribution in a certain range, which can be described as

$$s = T(r), \ 0 \le r \le L - 1 \qquad (8)$$

Here, r represents the pixel gray value before the transformation, s represents the pixel gray value after the transformation, and $T(r)$ represents the transformation function. $T(r)$ is a monotone increasing function in the range of $0 \le T(r) \le L$-1, which ensures that the variables r and s are one-to-one corresponding. $0 \le r \le L$-1 and $0 \le T(r) \le L$-1 can ensure the range of gray value after transformation will not exceed the original one.

We use $p_r(r)$ and $p_s(s)$ to represent the probability density function corresponding to gray level r and s respectively. The inverse transformation from s to r can be described as

$$r = T^{-1}(s), \ 0 \le s \le L - 1 \qquad (9)$$

The inverse transformation function $T^{-1}(s)$ also meets the monotonically increasing condition in the range of $0 \le s \le L$-1. According to the theory, if $p_r(r)$ and $T(r)$ are known and $r = T^{-1}(s)$ is a monotonically increasing function, the gray level probability density function $p_s(s)$ of the image after transformation is shown as follows:

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right| \qquad (10)$$

The transformation function of the histogram equalization function can be written as

$$s = T(r) = (L - 1)\int_0^r p_r(w)dw \qquad (11)$$

Here, w is the integral variable. It can be seen from the right side of the **Eq. 11** that the integral is the area below the function curve, so it meets the condition of monotone increasing. As the integral of $p_r(r)$, the probability density function of r, on the range $[0, L$-1$]$ is 1, and s takes the maximum value L-1, which does not exceed the gray range of r, we can get the following formula as
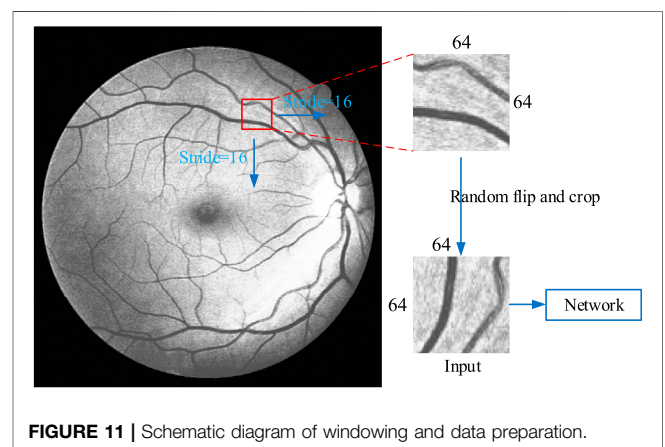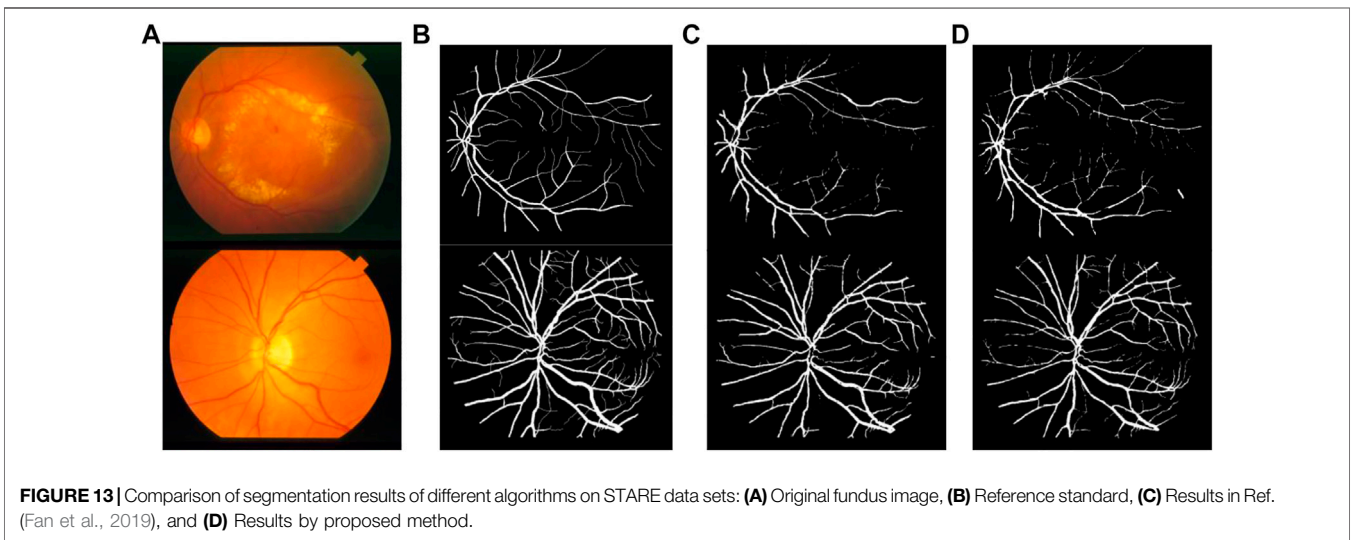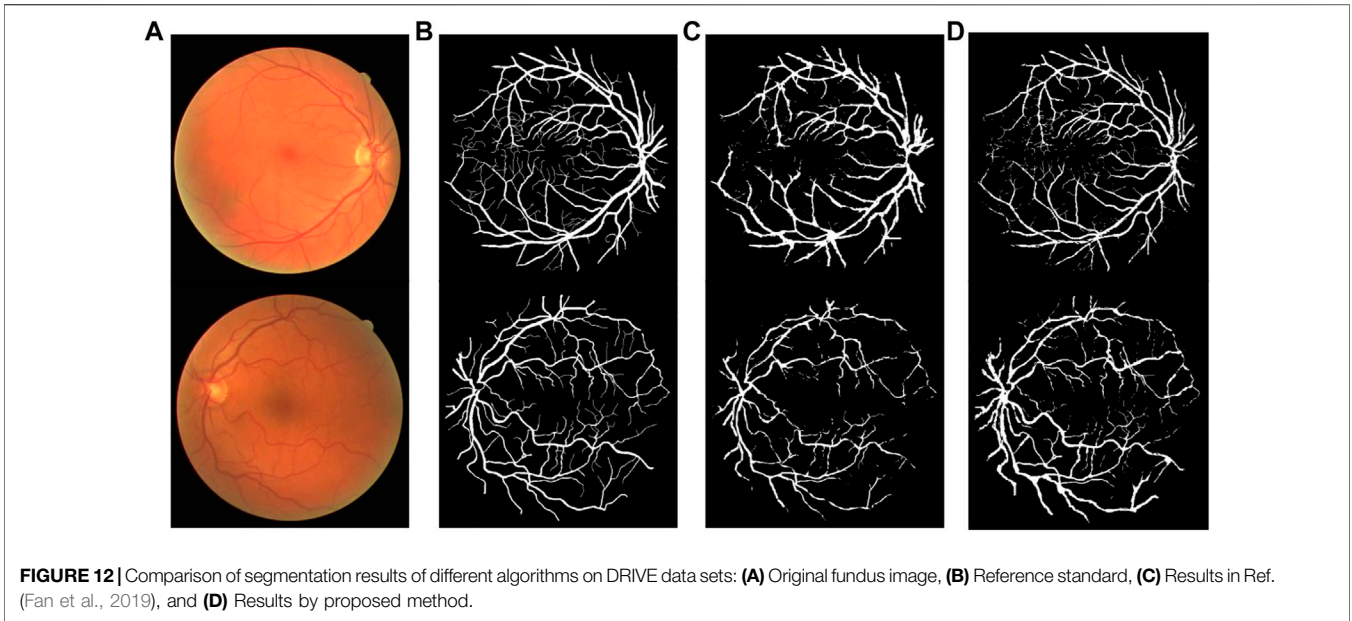

FIGURE 11 | Schematic diagram of windowing and data preparation.

FIGURE 12 | Comparison of segmentation results of different algorithms on DRIVE data sets: **(A)** Original fundus image, **(B)** Reference standard, **(C)** Results in Ref. (Fan et al., 2019), and **(D)** Results by proposed method.

FIGURE 13 | Comparison of segmentation results of different algorithms on STARE data sets: **(A)** Original fundus image, **(B)** Reference standard, **(C)** Results in Ref. (Fan et al., 2019), and **(D)** Results by proposed method.

$$\frac{ds}{dr} = \frac{dT(r)}{dr} (L-1) \frac{d}{dr} \left[\int_0^r p_r(w)dw \right] = (L-1)p_r(r) \quad (12)$$

Substitute **Eq. 12** into **Eq. 10**, the probability density function $p_s(s)$ can be written as

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| = p_r(r) \left| \frac{1}{(L-1)p_r(r)} \right| = \left| \frac{1}{L-1} \right|, 0 \le s \le L-1 \quad (13)$$

Through the above equation, $p_s(s)$ is proved uniformly distributed, which indicates that we can obtain a gray image with more uniform gray distribution and higher contrast by histogram equalization transformation.

For digital images, the gray level is discrete, and the sum of the probability tensity function should be used instead of the integral. The probability of occurrence of gray level r_k is approximately transformed into

$$p_k(r_k) = \frac{n_k}{n}, \quad k = 0, 1, \dots, L-1 \quad (14)$$

Here, n is the number of all pixels in the image, n_k is the number of pixels on the gray-level r_k, L is the total number of

FIGURE 14 | Comparison of local segmentation effects on DRIVE data sets: **(A)** Original fundus image; **(B)** Original local image; **(C)** Local reference standard image; **(D)** Local segmentation results in Ref. (Fan et al., 2019); and **(E)** Local results by proposed method.
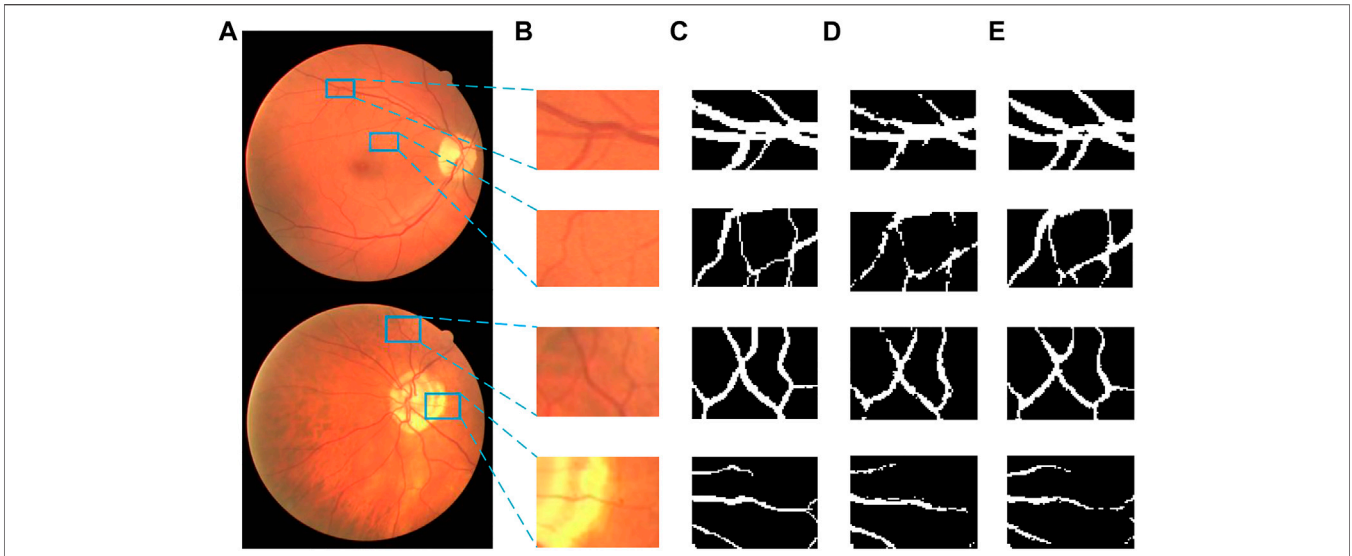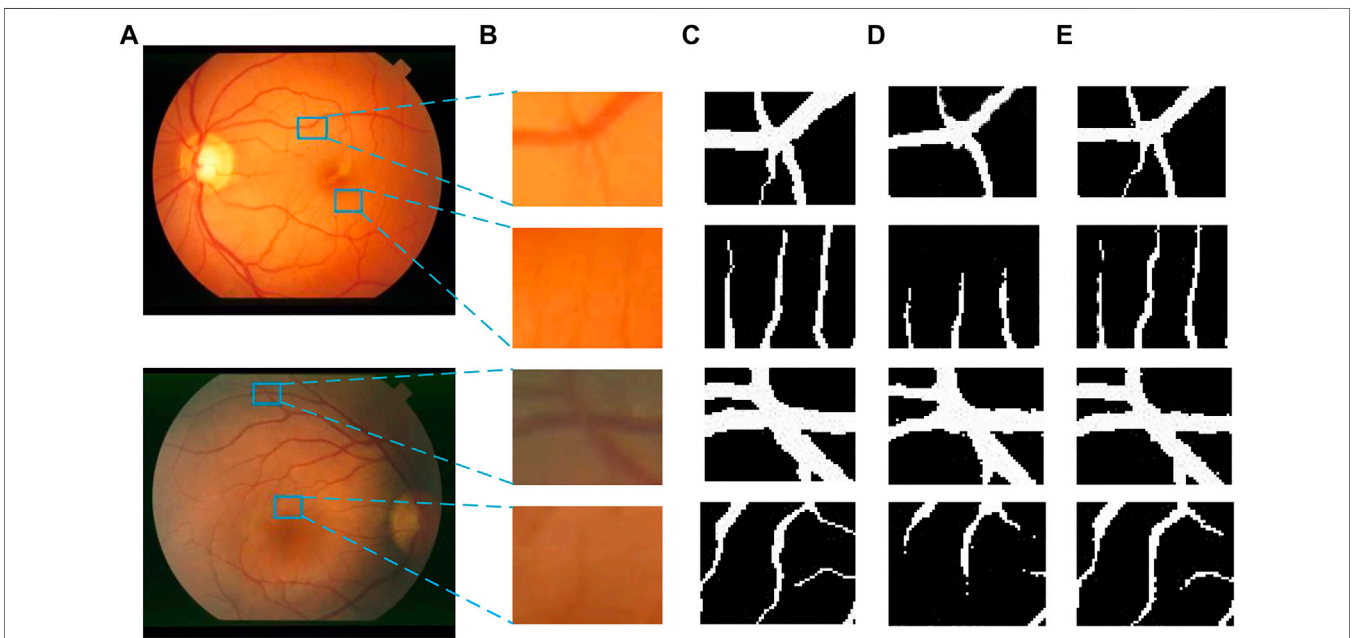

FIGURE 15 | Comparison of local segmentation effects on STARE data sets: **(A)** Original fundus image, **(B)** Original local image, **(C)** Local reference standard image, **(D)** Local segmentation results in Ref. (Fan et al., 2019), and **(E)** Local results by proposed method.

possible gray-level in the image, so the discrete form of the transformation function is

$$s_k = T(r_k) = (L-1) \sum_{j=0}^{k} \frac{n_j}{n} \quad k = 0, 1, \dots, L-1 \quad (15)$$

The operation of discrete histogram equalization is to map each pixel in the input image of gray-level r_k to the pixel in the output image of gray-level s_k through **Eq. 15**. Different from the continuous form, the new image generated is not necessarily completely evenly distributed, but the image tends to be uniform as well with a higher contrast and a larger grayscale after transformation.

For the fundus images, the background will affect the histogram equalization, so the local histogram equalization

TABLE 1 | Segmentation performance comparison of different algorithms on DRIVE.

Method	R_{Acc}	R_{Se}	R_{Sp}	R_{AUC}	Years
Rodrigues et al.	0.9465	0.7165	0.9801	-	2017
Li et al.	0.9574	-	-	0.9702	2018
Aguirre-Ramos et al.	0.9503	0.7854	0.9662	-	2018
Girard et al.	0.9480	0.7490	0.9770	-	2019
Fan et al.	0.9531	0.7035	0.9763	-	2019
Li Z Q et al.	0.9375	0.6945	0.9729	0.9320	2020
Pachade et al.	0.9405	0.7514	0.9676	-	2020
Lu et al.	0.9559	0.8288	0.9745	0.9786	2021
Liu	0.9582	0.8137	0.9055	-	2021
Proposed method	**0.9590**	0.8320	0.9885	0.9713	2021

R_{Acc} represents the accuracy; R_{Se} represents the sensitivity; R_{Sp} represents the specificity; R_{AUC} represents the area under the ROC curve.

TABLE 2 | Segmentation performance comparison of different algorithms on STARE.

Method	R_{Acc}	R_{Se}	R_{Sp}	R_{AUC}	Years
Li Z Q et al.	0.9461	0.7456	0.9693	0.9524	2020
Pachade et al.	0.9543	0.7769	0.9688	-	2020
Lu et al.	0.9683	**0.8432**	0.9775	0.9813	2021
Proposed method	**0.9688**	0.8295	**0.9785**	**0.9820**	2021

R_{Acc} represents the accuracy; R_{Se} represents the sensitivity; R_{Sp} represents the specificity; R_{AUC} represents the area under the ROC curve.

(Lai et al., 2015) transformation for the eyeball part is used instead of global histogram equalization. **Figure 8** shows the result of local histogram equalization. It can be seen from **Figure 8** that the value of background pixels does not change as close to 0, while the histogram of the eyeball part is obviously stretched and the contrast improves obviously.

4.1.3 Gamma Transformation

Gamma transformation (Hoo et al., 2017; Wang et al., 2018; Wang et al., 2021) can be used to adjust the contrast of gray images which are overexposed or underexposed. Nonlinear transformation is utilized to enhance the gray value of the dark area and reduce the gray value of the overexposed area, so that the overall detail of the image will be enhanced. The formula of Gamma transformation can be written as

$$s = cr^{\gamma} \qquad (16)$$

Here, r is the input value of the gray image, with the range of [0,1]; s is the output value after Gamma transformation; c is the gray scale coefficient, usually equals to 1; γ is the Gamma factor, which controls the stretch of the entire transformation. The transformation with different Gamma factor γ is shown in **Figure 9**, where the abscissa and the ordinate represent the gray value before and after Gamma transformation for a certain pixel individually. It can be seen that the result differs with the factor γ.

Gamma transformation is used to process the histogram equalization image, in order to further enhance the contrast. It is drawn from the experiment that the best contrast enhancement effect is with $\gamma = 1.3$. The Gamma transformation results with $\gamma = 1.3$ are shown in **Figure 10**.

4.1.4 Training Data Preparation

As the pixel number of fundus images collected by different instruments is not the same, and also the size of the whole image is relatively large, the segmentation accuracy will be affected and small retinal vessels that contains vital information cannot be extracted when using the full-size images as the direct training data. Therefore, we adopt a sliding window on the image to
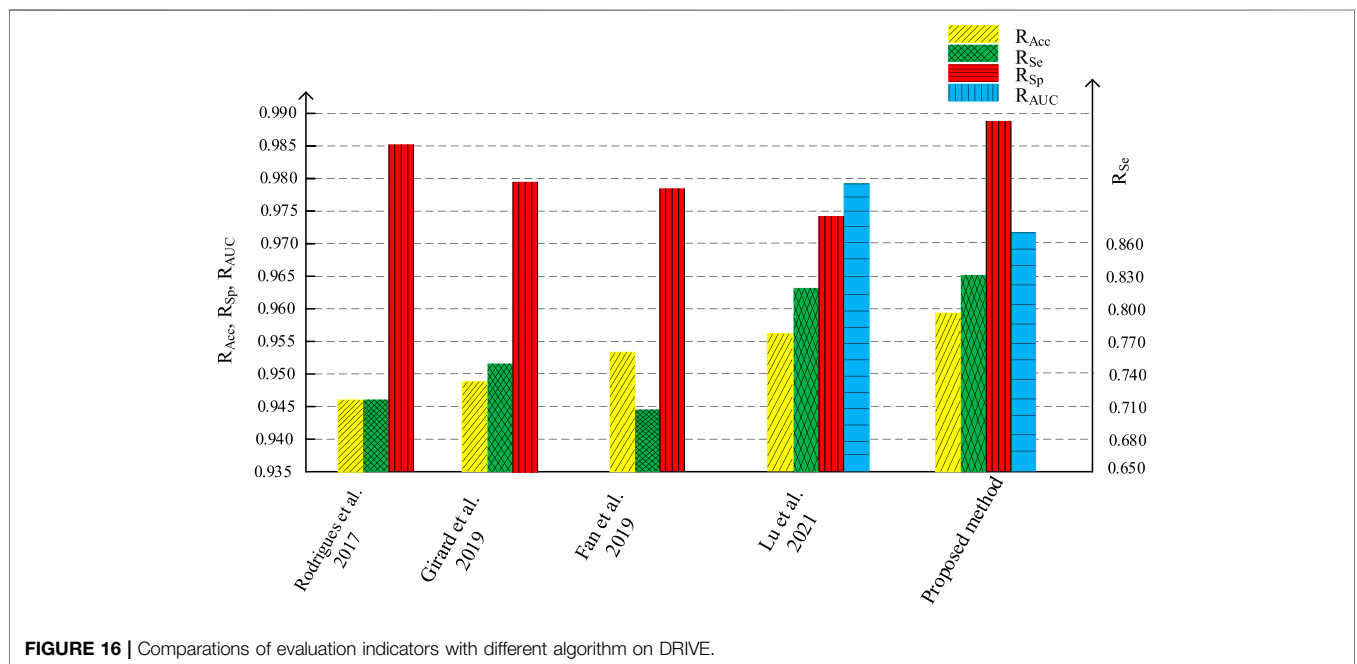

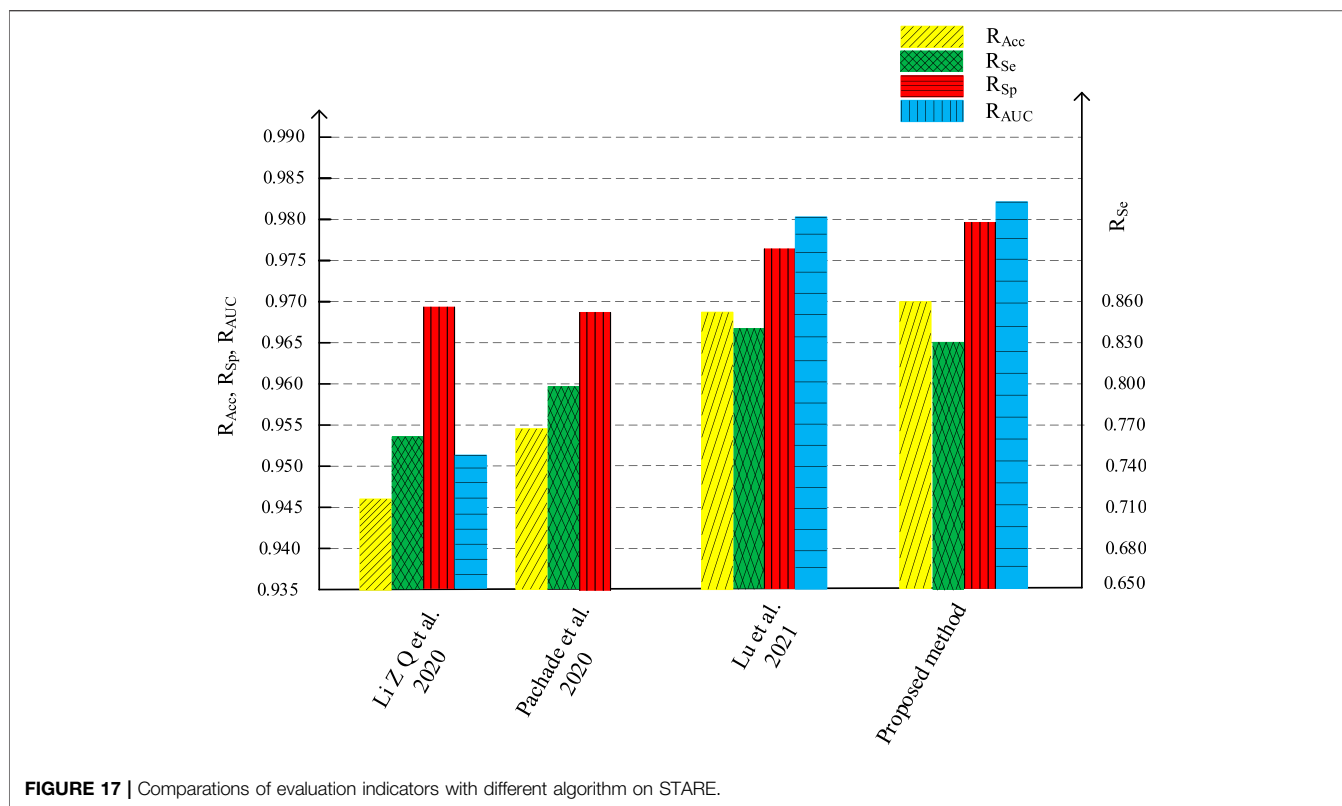**FIGURE 16 |** Comparisons of evaluation indicators with different algorithm on DRIVE.

FIGURE 17 | Comparisons of evaluation indicators with different algorithm on STARE.

capture certain areas as input, as shown in **Figure 11**. Firstly, fill the background pixel with the value of 0 to widen the height and the width of the image as integer multiples of 64. Secondly, a 64×64 window is used to slide from the upper left corner of the image, from left to right and top to bottom, with a sliding step as 16. The sliding window of 64×64 is fixed in this network, but the sliding step size can be change. The smaller the sliding step size is, the more data can be obtained, and larger slide step means less data obtained. The maximum sliding step size should not be greater than 64, as greater than 64 results in some aeras with information will be missed. In the meantime, the sliding step should not be too small, in order not to get a lot of overlapping areas. The sliding step is set as 16, as a total of 27380 64×64 training data can be obtained from DRAVE data set, and 15170 64×64 training data from STARE data set. The obtained data are randomly flipped and clipped in each epoch to ensure that the training data we sent to the network was different in each epoch.

Evaluation Indexes of Model Performance

Retinal vessel segmentation is to classify pixels and determine whether each pixel belongs to blood vessels. Four evaluation methods are used to evaluate the effect of vessel segmentation, which are accuracy R_{Acc}, sensitivity R_{Se}, specificity R_{Sp} and ROC curve. The function of the former three are as follows:

$$R_{Acc} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{17}$$

$$R_{Se} = \frac{T_P}{T_P + F_N} \tag{18}$$

$$R_{Sp} = \frac{T_N}{T_N + F_P} \tag{19}$$

Here T_P represents true positive, whose value equals to the number of accurately segmented vessel pixels; T_N represents the true negative, and its value equals to the number of correctly segmented background pixels; F_P represents false positive, with the value equals to the number of incorrectly segmented vessel pixels; F_N represents the false negative, and its value equals to the number of wrongly segmented background pixels. A certain curve can be draw by the false positive rate (1-R_{Sp}) as the horizontal coordinate and the true positive rate (R_{Se}) as the vertical coordinate, which is the ROC curve (Hoo et al., 2017; Michael et al., 2019). The area under the curve is defined as R_{AUC} (Janssens and Martens, 2020; Muschelli, 2020), and the closer its value to 1, the better the segmentation effect is.

Experimental Results

Figures 12, **13** show the segmentation result on the DRIVE data set and STARE data set individually, in which (**Figures 12A**, **13A**) is the original fundus image; (**Figures 12B**, **13B**) is the image of retinal vessel manually segmented by experts, which is used as the standard reference; (**Figures 12C**, **13C**) is the segmentation result of the typical traditional method in Reference (Fan et al., 2019); and (**Figures 12D**, **13D**) is the segmentation result based on the proposed method. It can be observed from **Figures 12**, **13** that the segmentation image of the

network is mostly consistent with the expert labeled image, and the capillaries are much more detailed than the traditional method. The vast majority of characteristic information for retinopathy recognition and identification is contained in these details.

Figures 14, 15 show the segmentation results of some local areas on the DRIVE data set and STARE data set separately, in which (Figures 14A, 15A) is the original fundus image with the selected local areas by the blue boxes; (Figures 14B, 15B) is the original image of selected areas; (Figures 14C, 15C) is the reference standard images of the selected areas; (Figures 14D, 15D) is the segmentation result based on method in Reference (Fan et al., 2019); and (Figures 14E, 15E) is the result based on the proposed method. The chosen local areas include relatively thick blood vessels, intersecting blood vessels, and thin blood vessels. It can be observed from Figures 14, 15 that the vessel pixels division has a higher accuracy in both thick and thin vessels, resulting in better continuity of blood vessels and fewer broken vessels, which is more conducive to extract key feature information.

We compare the segmentation performance of the proposed algorithm and the methods in the references, applied on DRIVE data set and STARE data set, as shown in Tables 1, 2 respectively. The accuracy of the proposed method on DRIVE data set and STARE data set is 0.9590 and 0.9688 respectively. The sensitivity is 0.8320 and 0.8432, while the specificity is 0.9885 and 0.9785 severally. The accuracy, sensitivity, and specificity on both data sets are generally improved compared to other algorithms. The ROC curve index of the algorithm in this paper is 0.9786 and 0.9820 individually. Though the R_{AUC} is 0.0073 lower than the latest Reference (Lu et al., 2021), the specificity is significantly improved by the algorithm proposed in this paper, and the overall performance is improved in return. As for the STARE data set, the sensitivity of the proposed algorithm is 0.0137 lower than that in Reference (Dharmawan et al., 2019b), but all the other evaluation indicators improved. Therefore, the segmentation performance of the proposed method is effectively improved.

Figures 16, 17 are the histogram of the evaluation indexes of the proposed algorithm and other typical algorithms in DRIVE data set and STARE data set respectively. It can be seen more intuitively from the figure that the proposed algorithm has better segmentation performance and smaller error than other algorithms.

5 CONCLUSION AND FUTURE WORK

The proposed residual convolution neural network based retinal vessel segmentation algorithm has been proved as an effective way to extract the blood vessel from the fundus images. The accuracy of the proposed method on DRIVE data set and STARE data set reaches 0.9590 and 0.9688 respectively, with the sensitivity and specificity on both data sets generally improved compared to other algorithms. It has a better segmentation performance and smaller error than the existing methods, especially in the details of capillaries, which contains most of

the key feature information of diagnose and identification. In addition, by introducing the ASPP module, the receptive field is enlarged and the number of training parameters is reduced, which means a great potential for sharply increasing the volume of the identifications data and shortening the recognition time.

On the basis of this work, it is still necessary to do further research on obtaining more detailed capillary features, and further extract features from the fundus image data of different pathologic conditions, different disease courses, or different healthy people. Besides, specific criteria are needed to evaluate the details of the segmentation quantitively. Continuous research will provide more forceful support for the realization of computer-assisted retinal disease screening and retina identification in the future.

In addition, the network structure and construction method proposed in this paper are of great reference significance to many other applications. Especially the proposed improved residual attention module combined with deep supervision module successfully overcome the gradient disappearance and explosion in the convolution neural network. The encoder-decoder network structure effectively avoids inefficient learning and sharing in training, and the atrous spatial pyramid pooling significantly enlarges the receptive field while reducing the number of training parameters. These contributions have potential implications for other applications of biological heuristic algorithms, not limited to image processing problem (Tao et al., 2021b), but can even be applied in public opinion dissemination (Chen et al., 2021b; Chen et al., 2021c) and behavior analysis (Xu et al., 2021a; Xiang et al., 2021). Further extended research will provide broader support for future applications in other aspects.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SX and ZC built the research plan and drafted the article; ZC conducted the experiments and interpreted the results; WC provided help in data analysis, BT and FZ revised the article.

FUNDING

REFERENCES

Aguirre-Ramos, H., Avina-Cervantes, J. G., Cruz-Aceves, I., Ruiz-Pinales, J., and Ledesma, S. (2018). Blood Vessel Segmentation in Retinal Fundus Images Using Gabor Filters, Fractional Derivatives, and Expectation Maximization. *Appl. Mathematics Comput.* 339, 568–587. doi:10.1016/j.amc.2018.07.057

Alaguselvi, R., and Murugan, K. (2021). Performance Analysis of Automated Lesion Detection of Diabetic Retinopathy Using Morphological Operation. *Signal. Image Video Process. Volume* 15, 797–805. doi:10.1007/s11760-020-01798-x

Awais, M., Bin Iqbal, M. T., and Bae, S.-H. (2021). Revisiting Internal Covariate Shift for Batch Normalization. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 5082–5092. doi:10.1109/TNNLS.2020.3026784

Cao, X., Xue, L., Lin, J., and Yu, L. (2017). Retinal Vessel Segmentation Method Based on Matched Filtering and Automatic Threshold. *J. Electron. Meas. Instrumentation* 1, 51–57. doi:10.13382/j.jemi.2017.01.008

Chen, D., Hu, F., Nian, G., and Yang, T. (2020). Deep Residual Learning for Nonlinear Regression. *Entropy* 22 (2), 193. doi:10.3390/e22020193

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4), 834–848. doi:10.1109/tpami.2017.2699184

Chen, T., Peng, L., Yang, J., and Cong, G. (2021a). Analysis of User Needs on Downloading Behavior of English Vocabulary APPs Based on Data Mining for Online Comments. *Mathematics* 9 (12), 1341. doi:10.3390/math9121341

Chen, T., Rong, J., Yang, J., Cong, G., and Li, G. (2021b). Combining Public Opinion Dissemination with Polarization Process Considering Individual Heterogeneity. *Healthcare* 9 (2), 176. doi:10.3390/healthcare9020176

Chen, T., Yin, X., Peng, L., Rong, J., Yang, J., and Cong, G. (2021c). Monitoring and Recognizing enterprise Public Opinion from High-Risk Users Based on User Portrait and Random forest Algorithm. *Axioms* 10 (2), 106. doi:10.3390/axioms10020106

Das, A. (2016). Diabetic Retinopathy: Battling the Global Epidemic. *Indian J. Ophthalmol.* 64 (1), 2–3. doi:10.4103/0301-4738.178155

Deng, W., Xu, J., Song, Y., and Zhao, H. (2020). An Effective Improved Co-evolution Ant colony Optimisation Algorithm with Multi-Strategies and its Application. *Ijbic* 16 (3), 158. doi:10.1504/IJBIC.2020.111267

Dhal, K. G., Das, A., Ray, S., Gálvez, J., and Das, S. (2020). Histogram Equalization Variants as Optimization Problems: A Review. *Arch. Computat Methods Eng.* 28, 1471–1496. doi:10.1007/s11831-020-09425-1

Dharmawan, D. A., Li, D., Ng, B. P., and Rahardja, S. (2019a). A New Hybrid Algorithm for Retinal Vessels Segmentation on Fundus Images. *IEEE Access* 7, 41885–41896. doi:10.1109/ACCESS.2019.2906344

Dharmawan, D. A., Ng, B. P., and Borijindargoon, N. (2019b). Design of Optimal Adaptive Filters for Two-Dimensional Filamentary Structures Segmentation. *IEEE Signal. Process. Lett.* 26 (10), 1511–1515. doi:10.1109/LSP.2019.2938631

Fan, L., Cheng, Y., and Tian, X. (2019). Retinal Blood Vessel Segmentation Based on Matched Filtering and Morphological Processing. *Computer Knowledge Technology* 15 (17), 188–190. doi:10.14004/j.cnki.ckt.2019.1884

Feng, M., Lu, H., and Yu, Y. (2020). Residual Learning for Salient Object Detection. *IEEE Trans. Image Process.* 29, 4696–4708. doi:10.1109/TIP.2020.2975919

Girard, F., Kavalec, C., and Cheriet, F. (2019). Joint Segmentation and Classification of Retinal Arteries/veins from Fundus Images. *Artif. Intelligence Med.* 94, 96–109. doi:10.1016/j.artmed.2019.02.004

Grewal, P. S., Oloumi, F., Rubin, U., and Tennant, M. T. S. (2018). Deep Learning in Ophthalmology: a Review. *Can. J. Ophthalmol.* 53 (4), 309–313. doi:10.1016/j.jcjo.2018.04.019

Guo, Y., Budak, Ü., Vespa, L. J., Khorasani, E., and Şengür, A. (2018). A Retinal Vessel Detection Approach Using Convolution Neural Network with Reinforcement Sample Learning Strategy. *Measurement* 125, 586–591. doi:10.1016/j.measurement.2018.05.003

Hoo, Z. H., Candlish, J., and Teare, D. (2017). What Is an ROC Curve? *Emerg. Med. J.* 34 (6), 357–359. doi:10.1136/emermed-2017-206735

Hoover, A. D., Kouznetsova, V., and Goldbaum, M. (2000). Locating Blood Vessels in Retinal Images by Piecewise Threshold Probing of a Matched Filter Response. *IEEE Trans. Med. Imaging* 19 (3), 203–210. doi:10.1109/42.845178

Horton, M. B., Silva, P. S., Cavallerano, J. D., and Aiello, L. P. (2016). Clinical Components of Telemedicine Programs for Diabetic Retinopathy. *Curr. Diab Rep.* 16 (12), 129. doi:10.1007/s11892-016-0813-8

Hu, J., Shen, L., Albanie, S., Sun, G., and Wu, E. (2020). Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 42 (8), 2011–2023. doi:10.1109/TPAMI.2019.2913372

Jamin, A., and Humeau-Heurtier, A. (2020). (Multiscale) Cross-Entropy Methods: A Review. *Entropy* 22 (1), 45. doi:10.3390/e22010045

Janssens, A. C. J. W., and Martens, F. K. (2020). Reflection on Modern Methods: Revisiting the Area under the ROC Curve. *Int. J. Epidemiol.* 49 (4), 1397–1403. doi:10.1093/ije/dyz274

Jiang, D., Li, G., Sun, Y., Hu, J., Yun, J., and Liu, Y. (2021a). Manipulator Grabbing Position Detection with Information Fusion of Color Image and Depth Image Using Deep Learning. *J. Ambient Intell. Hum. Comput* 12 (12), 10809–10822. doi:10.1007/s12652-020-02843-w

Jiang, D., Li, G., Tan, C., Huang, L., Sun, Y., Kong, J., et al. (2021b). Semantic Segmentation for Multiscale Target Based on Object Recognition Using the Improved Faster-RCNN Model. *Future Generation Computer Syst.* 123, 94–104. doi:10.1016/j.future.2021.04.019

Lai, Y.-R., Tsai, P.-C., Yao, C.-Y., and Ruan, S.-J. (2015). Improved Local Histogram Equalization with Gradient-Based Weighting Process for Edge Preservation. *Multimed Tools Appl.* 76, 1585–1613. doi:10.1007/s11042-015-3147-7

Li, G., Jiang, D., Zhou, Y., Jiang, G., Kong, J., and Manogaran, G. (2019a). Human Lesion Detection Method Based on Image Information and Brain Signal. *IEEE Access* 7, 11533–11542. doi:10.1109/ACCESS.2019.2891749

Li, H., Munk, A., Sieling, H., and Walther, G. (2020a). The Essential Histogram. *Biometrika* 107 (2), 347–364. doi:10.1093/biomet/asz081

Li, J., Cao, F., Cheng, H., and Qian, Y. (2021a). Learning the Number of Filters in Convolutional Neural Networks. *Ijbic* 17 (2), 75–84. doi:10.1504/IJBIC.2021.114101

Li, M., Wang, Y., and Wang, C. (2021b). Recursive Residual Atrous Spatial Pyramid Pooling Network for Single Image Deraining. *Signal. Processing: Image Commun.* 99, 116430. doi:10.1016/j.image.2021.116430

Li, X., Chang, D., Tian, T., and Cao, J. (2019b). Large-Margin Regularized Softmax Cross-Entropy Loss. *IEEE Access* 7, 19572–19578. doi:10.1109/ACCESS.2019.2897692

Li, Y., Cai, Y., and Gao, X. (2018). Retinal Vessel Segmentation Algorithm Based on Hybrid Phase Feature. *J. Computer Appl.* 38 (07), 2083–2088. doi:10.11772/j.issn.1001-9081.2017123045

Li, Z., Yang, X., and Wu, C. (2020b). Retinal Vessel Segmentation Based on Deep forest. *J. Yunnan Nationalities Univ. (Natural Sci. Edition)* 29 (06), 601–606.

Lian, X., Pang, Y., Han, J., and Pan, J. (2021). Cascaded Hierarchical Atrous Spatial Pyramid Pooling Module for Semantic Segmentation. *Pattern Recognition* 110, 107622. doi:10.1016/j.patcog.2020.107622

Liu, Z. (2021). Construction and Verification of Color Fundus Image Retinal Vessels Segmentation Algorithm under BP Neural Network. *J. Supercomput* 77, 7171–7183. doi:10.1007/s11227-020-03551-0

Lu, N., Li, X., Xiao, J., and Zhou, Y. (2021). Segmentation of Retinal Blood Vessels Combining Attention and Conditional Generative Adversarial Network. *J. Chin. Computer Syst.* 42 (06), 1288–1292. doi:10.3969/j.issn.1672-8513.2020.06.012

Lu, P., Lei, W., Li, Z., and Zhou, S. (2016). A Universal Vessel Segmentation Method Based on Multi-Scale Filtering and Statistical Mixture Model. *Chin. J. Biomed. Eng.* 35 (5), 519–525. doi:10.3969/j.issn.0258-8021.2016.05.002

Marín, D., Aquino, A., Gegundez-Arias, M. E., and Bravo, J. M. (2011). A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features. *IEEE Trans. Med. Imaging* 30 (1), 146–158. doi:10.1109/TMI.2010.2064333

Michael, H., Tian, L., and Ghebremichael, M. (2019). The ROC Curve for Regularly Measured Longitudinal Biomarkers. *Biostatistics* 20 (3), 433–451. doi:10.1093/biostatistics/kxy010

Muschelli, J. (2020). ROC and AUC with a Binary Predictor: a Potentially Misleading Metric. *J. Classif* 37, 696–708. doi:10.1007/s00357-019-09345-1

Pachade, S., Porwal, P., Kokare, M., Giancardo, L., and Meriaudeau, F. (2020). Retinal Vasculature Segmentation and Measurement Framework for Color

Fundus and SLO Images. *Biocybernetics Biomed. Eng.* 40 (3), 865–900. doi:10.1016/j.bbe.2020.03.001

Pal, S., Chatterjee, S., Dey, D., and Munshi, S. (2019). Morphological Operations with Iterative Rotation of Structuring Elements for Segmentation of Retinal Vessel Structures. *Multidim Syst. Sign Process.* 30, 373–389. doi:10.1007/s11045-018-0561-9

Ricci, E., and Perfetti, R. (2007). Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification. *IEEE Trans. Med. Imaging* 26 (10), 1357–1365. doi:10.1109/TMI.2007.898551

Rodrigues, L. C., and Marengoni, M. (2017). Segmentation of Optic Disc and Blood Vessels in Retinal Images Using Wavelets, Mathematical Morphology and Hessian-Based Multi-Scale Filtering. *Biomed. Signal Process. Control.* 36, 39–49. doi:10.1016/j.bspc.2017.03.014

Roy, S., Mitra, A., Roy, S., and Setua, S. K. (2019). Blood Vessel Segmentation of Retinal Image Using Clifford Matched Filter and Clifford Convolution. *Multimed Tools Appl.* 78, 34839–34865. doi:10.1007/s11042-019-08111-0

Shi, L. H. (2021). Synthetic Aperture Radar Lmage Enhancement Based on Adaptive Histogram Equalization. *Radio Eng.* 51 (09), 921–926. doi:10.3969/j.issn.1003-3106.2021.09.012

Singh, N. P., and Srivastava, R. (2016). Retinal Blood Vessels Segmentation by Using Gumbel Probability Distribution Function Based Matched Filter. *Computer Methods Programs Biomed.* 129, 40–50. doi:10.1016/j.cmpb.2016.03.001

Soomro, T. A., Afifi, A. J., ZhengSoomro, L. S., Soomro, S., Gao, J., Hellwich, O., et al. (2019). Deep Learning Models for Retinal Blood Vessels Segmentation: A Review. *IEEE Access* 7, 71696–71717. doi:10.1109/ACCESS.2019.2920616

Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and van Ginneken, B. (2004). Ridge-based Vessel Segmentation in Color Images of the Retina. *IEEE Trans. Med. Imaging* 23 (4), 501–509. doi:10.1109/tmi.2004.825627

Sulewski, P. (2020). Equal-bin-width Histogram versus Equal-Bin-Count Histogram. *J. Appl. Stat.* 48 (12), 2092–2111. doi:10.1080/02664763.2020.1784853

Sun, Y., Hu, J., Li, G., Jiang, G., Xiong, H., Tao, B., et al. (2020a). Gear Reducer Optimal Design Based on Computer Multimedia Simulation. *J. Supercomput* 76 (6), 4132–4148. doi:10.1007/s11227-018-2255-3

Sun, Y., Weng, Y., Luo, B., Li, G., Tao, B., Jiang, D., et al. (2020b). Gesture Recognition Algorithm Based on Multi-scale Feature Fusion in RGB-D Images. *IET image process* 14 (15), 3662–3668. doi:10.1049/iet-ipr.2020.0148

Sun, Y., Xu, C., Li, G., Xu, W., Kong, J., Jiang, D., et al. (2020c). Intelligent Human Computer Interaction Based on Non Redundant EMG Signal. *Alexandria Eng. J.* 59 (3), 1149–1157. doi:10.1016/j.aej.2020.01.015

Tang, S., and Yu, F. (2021). Construction and Verification of Retinal Vessel Segmentation Algorithm for Color Fundus Image under BP Neural Network Model. *J. Supercomput* 77, 3870–3884. doi:10.1007/s11227-020-03422-8

Tao, B., Huang, L., Zhao, H., Li, G., and Tong, X. (2021a). A Time Sequence Images Matching Method Based on the Siamese Network. *Sensors* 21 (17), 5900. doi:10.3390/s21175900

Tao, B., Liu, Y., Huang, L., Chen, G., and Chen, B. (2021b). 3D Reconstruction Based on Photoelastic Fringes. *Concurrency Computat Pract. Exper*, e6481. doi:10.1002/CPE.6481

Tian, F., Li, Y., and Wang, J. (2021). Retinal Blood Vessel Segmentation Based on Multi-Scale Wavelet Transform Fusion. *Acta Optica Sinica* 41 (04), 82–92. doi:10.3788/aos202141.0410001

Varma, R., Vajaranant, T. S., Burkemper, B., Wu, S., TorresHsu, M. C., Hsu, C., et al. (2016). Visual Impairment and Blindness in Adults in the United States. *JAMA Ophthalmol.* 134 (7), 802–809. doi:10.1001/jamaophthalmol.2016.1284

Wang, H., Yang, W., Zhao, Z., Luo, T., Wang, J., and Tang, Y. (2019). Rademacher Dropout: An Adaptive Dropout for Deep Neural Network via Optimizing Generalization gap. *Neurocomputing* 357, 177–187. doi:10.1016/j.neucom.2019.05.008

Wang, P., Liu, F., Yang, C., and Luo, X. (2018). Blind Forensics of Image Gamma Transformation and its Application in Splicing Detection. *J. Vis. Commun. Image Representation* 55, 80–90. doi:10.1016/j.jvcir.2018.05.020

Wang, W., Yuan, X., Chen, Z., Wu, X., and Gao, Z. (2021). Weak-Light Image Enhancement Method Based on Adaptive Local Gamma Transform and Color Compensation. *J. Sensors* 2021, 1–18. doi:10.1155/2021/5563698

Xiang, K., Xu, C., and Wang, J. (2021). Understanding the Relationship between Tourists' Consumption Behavior and Their Consumption Substitution Willingness under Unusual Environment. *Psychol. Res. Behav. Manag.* 14, 483–500. doi:10.2147/PRBM.S303239

Xu, C., Ding, A. S., and Zhao, K. (2021a). A Novel POI Recommendation Method Based on Trust Relationship and Spatial-Temporal Factors. *Electron. Commerce Res. Appl.* 48, 101060. doi:10.1016/j.elerap.2021.101060

Xu, C., Makihara, Y., Li, X., Yagi, Y., and Lu, J. (2021b). Cross-View Gait Recognition Using Pairwise Spatial Transformer Networks. *IEEE Trans. Circuits Syst. Video Technol.* 31 (1), 260–274. doi:10.1109/tcsvt.2020.2975671

Yan, Z., Yang, X., and Cheng, K.-T. (2019). A Three-Stage Deep Learning Model for Accurate Retinal Vessel Segmentation. *IEEE J. Biomed. Health Inform.* 23 (4), 1427–1436. doi:10.1109/JBHI.2018.2872813

Yang, Z., Jiang, D., Sun, Y., Tao, B., Tong, X., Jiang, G., et al. (2021). Dynamic Gesture Recognition Using Surface EMG Signals Based on Multi-Stream Residual Network. *Front. Bioeng. Biotechnol.* 9, 779353. doi:10.3389/fbioe.2021.779353

Zhang, Y. (2020). *The Retinal Identification Based on Principal Component Analysis*. Shenyang: China Medical University. Dissertation.

Zhao, J., Tang, J., Shi, A., Fan, T., and Xu, L. (2020). Improved Density Peaks Clustering Based on Firefly Algorithm. *Int. J. Bio-Inspired Comput.* 15 (1), 24–42. doi:10.1504/IJBIC.2020.105899