

A Novel Feature Fusion Model Based on Non-subsampled Shear-wave Transform for Retinal Blood Vessel Segmentation

Feng Lijuan¹ and Zhang Fan²

¹ School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology, Zhengzhou City, 450064, China
857003841@qq.com

² Zhengzhou University of Technology
Zhengzhou 450064, China
fanzhang0225@163.com

Abstract. Background: Fundus image is a projection of the inner surface of the eye, which can be used to analyze and judge the distribution of blood vessels on the retina due to its different shape, bifurcation and elongation. Vascular trees are the most stable features in medical images and can be used for biometrics. Ophthalmologists can effectively screen and determine the ophthalmic conditions of diabetic retinopathy, glaucoma and microaneurysms by the morphology of blood vessels presented in the fundus images. Traditional unsupervised learning methods include matched filtering method, morphological processing method, deformation model method, etc. However, due to the great difference in the feature complexity of different fundus image morphology, the traditional methods are relatively simple in coding, poor in the extraction degree of vascular features, poor in segmentation effect, and unable to meet the needs of practical clinical assistance. Methods: In this paper, we propose a new feature fusion model based on non-subsampled shear-wave transform for retinal blood vessel segmentation. The contrast between blood vessels and background is enhanced by pre-processing. The vascular contour features and detailed features are extracted under the multi-scale framework, and then the image is postprocessed. The fundus images are decomposed into low frequency sub-band and high frequency sub-band by non-subsampled shear-wave transform. The two feature images are fused by regional definition weighting and guided filtering respectively, and the vascular detection image is obtained by calculating the maximum value of the corresponding pixels at each scale. Finally, the Otsu method is used for segmentation. Results: The experimental results on DRIVE data set show that the proposed method can accurately segment the vascular contour while retaining a large number of small vascular branches with high accuracy. Conclusion: The proposed method has a high accuracy and can perform vascular segmentation well on the premise of ensuring sensitivity.

Keywords: Retinal blood vessel segmentation, non-subsampled shear-wave transform, feature fusion, regional definition weighting, guided filtering, Otsu method.

1. Introduction

Retinal vascular disease detection is the main method to assist in the diagnosis of hypertension, diabetes and many other diseases. Regular retinal vascular disease detection

can detect vascular abnormalities in time and help patients find the disease early, thereby preventing the further development and deterioration of the disease. The blood vessels of the retina in the fundus are very complicated, with small blood vessels, fuzzy outlines and intertwined distribution, so it is more difficult to segment blood vessels. Therefore, accurate and rapid segmentation of blood vessels is of great significance for the detection of retinal blood vessels [1-3].

Vessel segmentation methods mainly include supervised and unsupervised methods. The supervised method distinguishes the blood vessel and the background through a trained classifier, and the unsupervised method mainly detects the blood vessel by maximizing the filter response on the gray profile of the blood vessel cross-section. Liang et al. [4] proposed a u-shaped retinal vessel segmentation algorithm with adaptive vascular morphology and scale information, it used u-shaped segmentation model to perform end-to-end training on preprocessed images, and used local information entropy sampling for data enhancement. Wang et al. [5] proposed a multiple vessel segmentation (MVP) algorithm, which segmented retinal vessels into well-constrained subsets, and grouped blood vessel pixels with similar geometric characteristics in the subsets. Parallel training was performed on a group of homogeneous classifiers to form a discriminative decision for each group. Yan et al. [6] proposed a three-stage deep learning model, which separately segmented thick and thin blood vessels, further identified non-vascular pixels through vascular fusion, and improved the results by improving the overall blood vessel thickness consistency. Odstrcilik et al. [7] designed five different kernels based on the typical cross-sectional profile of blood vessels, and considered the five width categories of retinal blood vessels, and rotated the kernels to 12 different directions to cover all directions of the blood vessels. Dharmawan et al. [8] proposed a new retinal vessel segmentation framework based on the best adaptive filter. Rodrigues et al. [9] proposed a new optic disc detection algorithm based on wavelet transform and mathematical morphology, and used the tubular characteristics of blood vessels to segment retinal arteries and veins. Aguirre et al. [10] proposed a method that used Gabor filter and Gaussian fractional derivative to significantly enhance the structure and contour of blood vessels, and applied thresholds and a series of morphological-based decision rules to separate blood vessels.

Most of the existing methods perform well in blood vessel segmentation on the fundus retinal image, but they will lose the contour information of the blood vessel to a certain extent or produce more false positive pixels at the blood vessel contour. It is difficult to be more accurate. The small branches of the blood vessel are segmented, which leads to the decrease of the blood vessel integrity and the segmentation accuracy. This paper considers the segmentation of blood vessel contours and small branches, and proposes a new feature fusion model based on non-subsampled shear-wave transform (NSST) for retinal blood vessel segmentation with fusing blood vessel contour feature information and detailed feature information in a multi-scale framework. The fundus images are decomposed into low frequency sub-band and high frequency sub-band by NSST. The two feature images are fused by regional definition weighting and guided filtering respectively, and the vascular detection image is obtained by calculating the maximum value of the corresponding pixels at each scale. Finally, the Otsu method is used for segmentation, which solves the problem of vessel contour and detail information loss on a single scale, thus effectively improving the accuracy of vessel segmentation.

The remaining paper is settled as follows: Section 2 presents the related works in de-

tail. Section 3 discusses the proposed method. Then, experiments are provided in Section 4. Finally, the paper is concluded in Section 5.

2. Related Works

Blood vessels are one of the most important components of the retina. Retinal vascular segmentation and the division of vascular morphological attributes, such as length, width, tortuosity and angle, can be used for the diagnosis, screening, treatment and evaluation of a variety of cardiovascular and ophthalmic diseases, such as diabetes, glaucoma, hard exudate and hypertension and other systemic diseases [11].

However, the segmentation of retinal blood vessels by an ophthalmologist is time-poor and lacks accuracy. Therefore, the research on the automatic segmentation of blood vessels and automatic recognition [12] of blood vessel morphological attributes becomes a crucial step to assist medical diagnosis.

Existing retinal segmentation methods can be roughly divided into unsupervised and supervised methods [13]. The unsupervised learning retinal blood vessel segmentation methods can be divided into matching filtering, mathematical morphology, blood vessel tracking and clustering. Reference [14] proposed an improved Top-Hat transformation, which was applied to retinal blood vessel segmentation, it used circular structural elements of different radii to detect blood vessels of different widths, further improving the segmentation rate of tiny blood vessels, but it was more sensitive to noise. Reference [15] proposed a blood vessel tracking method based on Bayesian probability, which combined the information of blood vessel connectivity and gray level to fit the blood vessel structure. This method could better solve the problem of segmentation and fracture at the intersection of blood vessels, but had low segmentation accuracy for small blood vessels. In reference [16], principal component analysis (PCA) was used to extract features of blood vessels, and threshold value was used for segmentation. It had low segmentation of small blood vessels and segmentation rupture at the intersection of blood vessels.

Supervised learning requires experts to provide gold standard retinal images with tags, use a set of features based on local or global images to train the classifier, act as prior knowledge and guide training. Reference [17] proposed a random forest classifier integrating multiple features, which could well solve the situation of segmentation and fracture at the intersection of blood vessels, but there was a phenomenon of mis-segmentation of optic disc into blood vessels. In reference [18], the vascular segmentation algorithm integrating phase characteristics better solved the problem of insufficient detection of vascular phase consistency features, but there was still a problem of insufficient microvascular segmentation. In reference [19], blood vessel segmentation was regarded as a binary classification problem, and a hybrid 5D feature based on support vector machine (SVM) was proposed to distinguish blood vessel pixels from non-blood vessel pixels, which had strong adaptability to noise and solved the problem of optic disc being misdivided into blood vessels, but there was the problem of small blood vessels being easily broken. Therefore, this paper proposes a new feature fusion model based on non-subsampled shear-wave transform (NSST) for retinal blood vessel segmentation with fusing blood vessel contour feature information and detailed feature information in a multi-scale framework.

3. Proposed Retinal Blood Vessel Segmentation Method

The retinal blood vessel segmentation method based on NSST is mainly divided into three stages: image preprocessing, blood vessel feature information extraction, and blood vessel detection and segmentation. The contrast-limited adaptive histogram equalization (CLAHE) method is used to enhance the contrast of the green channel image of the original image and reduce noise. The contour feature information of the blood vessel is obtained by calculating the gradient amplitude of the retinal image, and the detailed feature information of the blood vessel is obtained by calculating the maximum principal curvature. The contour and the detailed information are processed for feature image post-processing respectively. Finally, the NSST is used to fuse the image containing the contour and detail information of the blood vessel to obtain the blood vessel detection information at multiple scales, and the Otsu method is used for segmentation. The flowchart of multi-scale blood vessel segmentation is shown in figure 1.

3.1. Image Pre-processing

In the fundus retinal image, the contrast of the green channel image is relatively high, so the green channel image is selected as the initial image for subsequent operations. The retinal blood vessel area is a low-contrast dark area, and the CLAHE algorithm can enhance the image while suppressing noise. The CLAHE is simple to calculate, only the amplitude limiting parameter is determined, so the CLAHE method can be used to enhance the retinal image [20]. The preprocessing of the original image is shown in figure 2.

3.2. Vessel Feature Information Extraction

The blood vessels in the retinal image of fundus show a network structure, and the diameter of the blood vessels is different, so it is difficult to accurately detect the blood vessel information using the information of a single scale. In this paper, vascular information is extracted and detected in a multi-scale framework. In general, the original image $I(x, y)$ and Gaussian kernel $G(x, y; \sigma)$ with variance σ^2 to define information at different scales [12-14]:

$$I_{\sigma}(x, y, \sigma) = I(x, y) \times G(x, y; \sigma) \quad (1)$$

Where $G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} \exp(-\frac{x^2+y^2}{2\sigma^2})$.

The derivative of the image can be approximated numerically as the convolution of the image with the derivative of the scale-normalized Gaussian kernel, so the first and second derivatives of the image can be expressed as:

$$\partial I_{\sigma}(x, y; \sigma) = I(x, y) \times \sigma \partial G(x, y; \sigma) \quad (2)$$

$$\partial^2 I_{\sigma}(x, y; \sigma) = I(x, y) \times \sigma^2 \partial^2 G(x, y; \sigma) \quad (3)$$

Where $I(x, y)$ denotes the original image. σ is the scale factor.

A. Contour information extraction

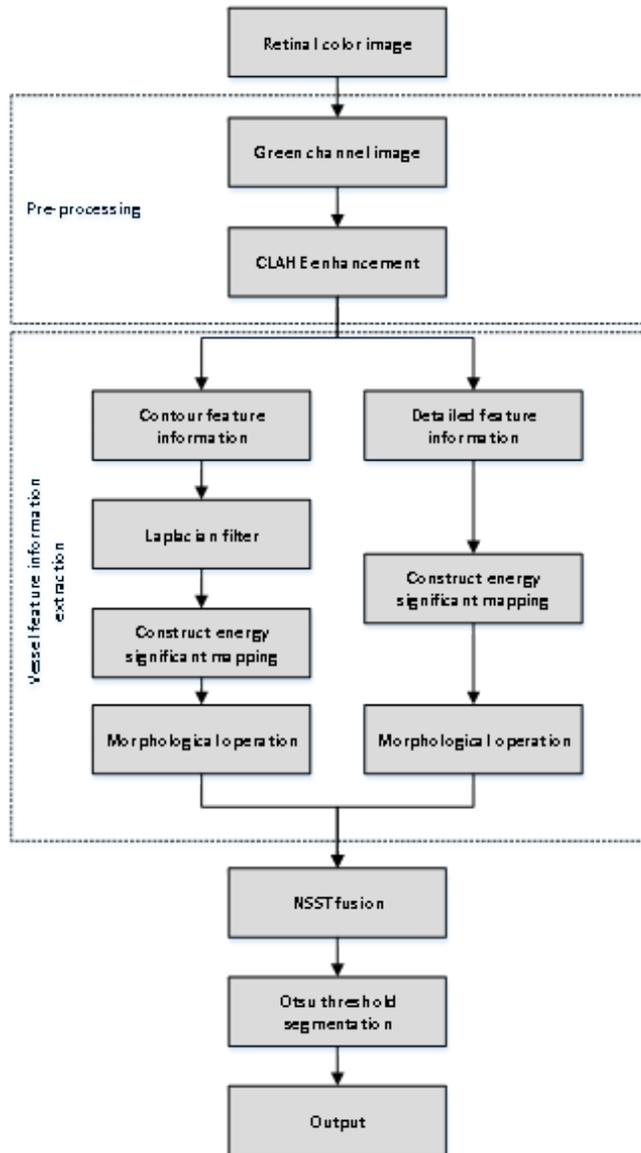


Fig. 1. Flow chart of proposed blood vessel segmentation

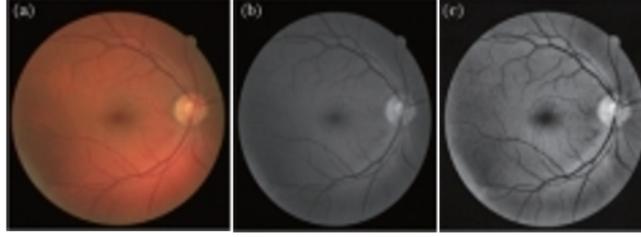


Fig. 2. Preprocessed images. (a) Original color image; (b) green channel image; (c) contrast enhanced image by CLAHE

The gradient amplitude describes the change in image intensity near the pixels. The image obtained by convolution of the original image $I(x, y)$ with Gaussian kernel $G(x, y; \sigma)$ is $I_\sigma(x, y; \sigma)$, and the gradient at point (x, y) can be defined as a vector:

$$\text{grad}[I(x, y)] = [\partial_x I_\sigma \quad \partial_y I_\sigma]^T \quad (4)$$

We can see from the definition of gradient that: 1) The vector $\text{grad}[I(x, y)]$ points to the direction with the maximum increase rate of $I(x, y)$, and the direction of the image edge is perpendicular to this gradient direction; 2) If $G[I(x, y)]$ is used to represent the amplitude of $\text{grad}[I(x, y)]$, then there is:

$$G[I(x, y)] = \sqrt{(\partial_x I_\sigma)^2 + (\partial_y I_\sigma)^2} \quad (5)$$

In the multi-scale framework, $G[I(x, y)]$ represents the slope of image intensity at a specific scale σ . The gradient amplitude images at scale $\sigma=1, 2$, and 4 pixels are shown in figure 3.

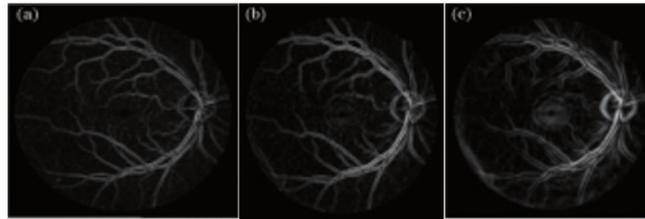


Fig. 3. Gradient amplitudes values. (a) $\sigma=1$ (b) $\sigma=2$ (c) $\sigma=4$

When scale σ is 1, 2, and 4 pixels respectively, most of the vascular contour information in the retinal images can be obtained by extracting the gradient amplitude information of the images. Convolution with Gaussian for image will suppress most structures whose feature length is smaller than scale σ , and the image will become blurred gradually with the increase of scale.

B. Detail information extraction

The second order directional derivative describes the change of image gradient intensity near the pixel point. The main curvature information of blood vessels in image $I(x, y)$ can be obtained by Hessian matrix. The Hessian matrix can be expressed as:

$$H = [\partial_{xx}I_\sigma \quad \partial_{xy}I_\sigma, \partial_{yx}I_\sigma \quad \partial_{yy}I_\sigma]^T \quad (6)$$

In a two-dimensional image, the Hessian matrix is a two-dimensional positive definite matrix with two eigenvalues, each of which has a corresponding eigenvector. When blood vessel relative to the background of low dark tubular structure, the blood vessels in pixel Hessian matrix has a larger eigenvalue λ_1 and a smaller eigenvalue λ_2 , $\lambda_1 \geq \lambda_2$. The maximum eigenvalue λ_1 corresponds to the maximum principal curvature of the Hessian matrix, so if $\lambda_1 \geq 1$, the pixels belonging to the vascular region in the image will be weighted as vascular pixels. Figure 4 shows the maximum principal curvature image when scale σ is 1, 2, and 4 pixels respectively.

The results of figure 4 show that with the increase of the scale factor, vessels with a radius similar to the scale factor are more obvious, while the information of vessels with a radius smaller than the scale factor is suppressed. With small scale, the maximum principal curvature image can well reflect the information of each small branch of the vessel. With large scale, the maximum principal curvature image can fully reflect the main artery information.

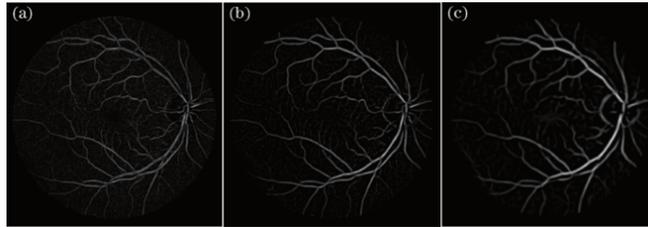


Fig. 4. Maximum principal curvature images. (a) $\sigma=1$ (b) $\sigma=2$ (c) $\sigma=4$

C. Feature image post-processing

The gradient amplitude image describes the contour information in the original image, while the maximum principal curvature image captures the details in the original image. Therefore, before image fusion, the two images are post-processed respectively to achieve better fusion effect.

1. Contour feature image post-processing

Assuming that the fundus retinal image is $I(x, y)$, its gradient amplitude image can be expressed as $G[I(x, y)]$. First, The Gradient amplitude image is processed by Laplacian filter, and the high-pass image is obtained as shown in figure 5(a):

$$U^G = G \times L_{ap} \quad (7)$$

Where L_{ap} is Laplace filter with size 3×3 .

Then the local average value of the absolute value of the high-pass image is taken to construct the energy significant mapping image [21].

$$S^G = |U^G \times A_{vg}| \quad (8)$$

Where A_{vg} is the average filter of size 3×3 .

The energy significant mapping image S^G is shown in figure 5(b), which contains the contour information of small vessels, but there are many noises around them. Therefore, the mathematical morphology is used to remove these noises. Since the blood vessels in retinal images are tubular structures, linear structures are selected as structural elements. The structure element takes two parameters: length and Angle. First, the length of the linear structural element is increased from the minimum diameter of the vessel to the maximum diameter with stride=1 pixel based on the diameter of the vessel. The angles of linear structural elements are selected from 0 to 170, the interval is 10 [22], and 198 linear structural element templates are obtained. The result of morphological open operation can be calculated with 198 element templates. Assume that the gray value of the open operation result of the l -th template at pixel point (x, y) is $I_{open_l}(x, y)$, then the gray value of each pixel after final morphological processing should be the maximum value of the corresponding position among 198 operation results, namely,

$$I_R = \max_l I_{open_l}(x, y) \quad (9)$$

Where I_R is the image processed by mathematical morphology, and the result is shown in figure 5(c).

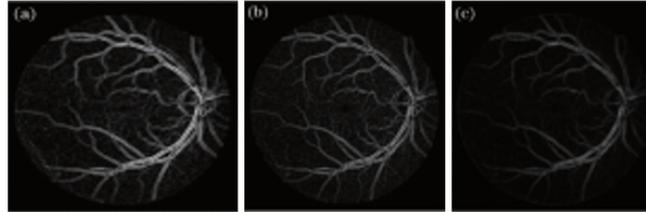


Fig. 5. Results of post-processing for contour feature image. (a) Laplacian filtered image; (b) energy significant mapping image; (c) mathematical morphology processing

2. Detail feature image post-processing

The detailed feature image is $F[I(x, y)]$, and a similar approach is adopted for post-processing. Since the detail feature image inherits the high frequency information of the original image, it is unnecessary to use Laplace filter to obtain the high pass image. The absolute value of the detail feature image $F[I(x, y)]$ is directly averaged locally to construct the energy significant mapping image $S^F = |F \times A_{vg}|$, and then the noise around the small vessels is removed by mathematical morphology processing. The post-processing results of detail feature images are shown in figure 6.

The results of figure 5(c) and figure 6(b) show that after the post-processing of contour feature image and detail feature image respectively, the contour and detail information in the image are clearer, and the post-processing effectively reduces the noise, which provides a good foundation for the subsequent fusion.

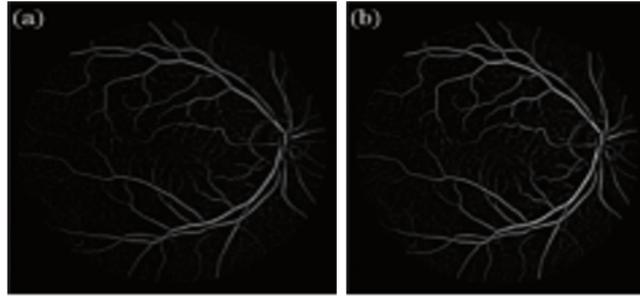


Fig. 6. Results of post-processing for detail feature image. (a) Energy significant mapping image; (b) image obtained by mathematical morphology processing

3.3. NSST-based image fusions

After post-processing, the noise in contour feature image and detail feature image is significantly reduced. The contour information and detail information contained in these two images are more prominent respectively. Therefore, NSST is used to fuse these two images [23-26]. The process of using NSST to fuse contour feature images and detail feature images is shown in figure 7. Wavelet coefficients are fused by regional definition weighting for low frequency and guided filtering for high frequency.

A. Low frequency component fusion rule

The purpose of low frequency sub-band fusion of fundus vascular image is to obtain a clearer fundus image. Logically, image blur is caused by the fact that the contour of the object in the image is not obvious, the grayscale change of the contour edge is not strong, and the sense of hierarchy is not strong. On the contrary, if the gray level of the contour edge changes significantly and the hierarchy is strong, the image will be clear. For an image, we can regard the image as a two-dimensional function $f(x, y)$. According to calculus, the sharpness of the image can be expressed by the gradient of the image.

For low frequency components, a fusion strategy based on region definition weighting is designed. As an evaluation index of clarity, energy of gradient (EOG) has clear physical meaning in spatial domain, good time performance and high sensitivity [27,28]. EOG uses the difference between adjacent points to calculate a gradient at a point, as defined below:

$$E = \sum_x \sum_y [f(x+1, y) - f(x, y)]^2 + [f(x, y+1) - f(x, y)]^2 \quad (10)$$

c_1 , c_2 and c are used to represent the low-frequency coefficients of image A, image B and fusion image F in the same sub-band, the same direction and the same position, respectively. r_1 and r_2 represent the EOG of low-frequency sub-band coefficients of image

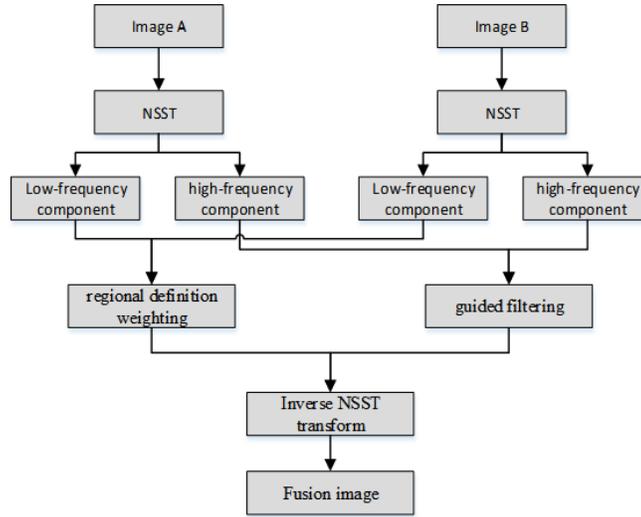


Fig. 7. Fusion process of NSST

A and image B in the same sub-band, direction and position in the NSST region, then the fusion strategy can be expressed as:

$$c = w_1c_1 + w_2c_2 \tag{11}$$

Where $w_1 = \frac{r_1}{r_1+r_2}$, $w_2 = \frac{r_2}{r_2+r_2}$

B. High-frequency component fusion rule

For high-frequency sub-band fusion, the aim is to preserve as much spatial detail, edge, and contour as possible. The high frequency sub-band of fundus vascular image A after NSST decomposition is taken as the input image, the high frequency sub-band of fundus vascular image B after NSST decomposition is taken as the guide image, and the output image is the high frequency sub-band image after fusion. In addition, the input image and the guided image must be on the same level and in the same direction. The fusion strategy can be expressed as:

$$H_F^{l,k} = G_{r,\varepsilon}(H_A^{l,k}, H_B^{l,k}) \tag{12}$$

$H_A^{l,k}$ and $H_B^{l,k}$ represent the high frequency sub-band of image A and image B in the k direction of the $l - th$ decomposition layer respectively, while $H_F^{l,k}$ represents the high frequency fusion sub-band in the k direction of the $l - th$ decomposition layer.

The results of NSST fusion for contour feature image and detail feature image are shown in figure 8. Figure 8 shows that the NSST fuses the image detail and contour information, which significantly improves the image quality. The outline of blood vessels and the information of thicker blood vessels are retained completely. At the same time, small vessels are retained to a large extent and non-vascular pixel interference around them is reduced, but some vascular information smaller than the current scale is lost in this process.

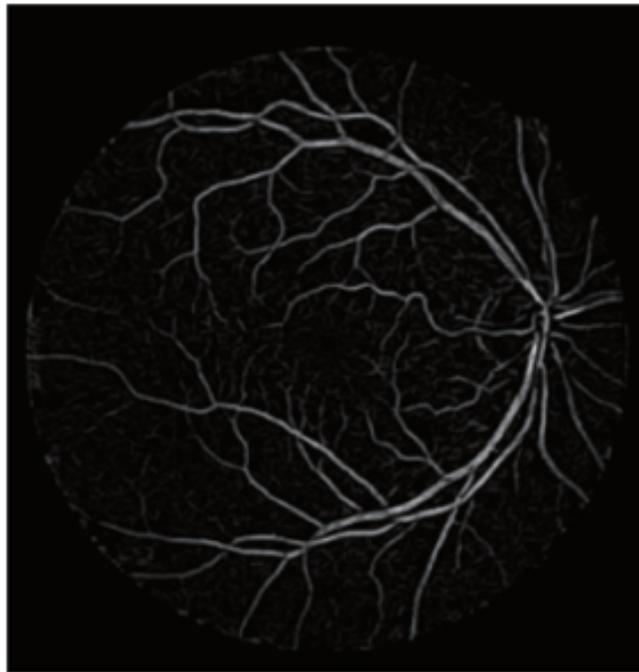


Fig. 8. Fusion image with NSST when $\sigma = 1$

3.4. Multi-scale Vascular Detection

Vessel detection at a single scale σ can only ensure that the vessel information corresponding to the radius of the current scale can be completely detected, while the vessel information smaller than this scale may be lost. So it needs to compute all feature information within a certain range $\sigma(\sigma_{min} \leq \sigma \leq \sigma_{max})$, where σ_{min} and σ_{max} are the minimum and maximum radius of blood vessels in retinal images to be detected. These two parameters must be known in advance and depend on the pixel resolution of the original image and the field of view of the fundus camera. In DRIVE data sets, σ_{min} and σ_{max} are usually set as 1 and 6 respectively [29-31].

First, at every scale $\sigma(\sigma_{min} \leq \sigma \leq \sigma_{max})$, the contour feature images G_k and detail feature images F_k are obtained respectively:

$$G_k = \sqrt{(\partial_x I_{\sigma_k})^2 + (\partial_y I_{\sigma_k})^2} \quad (13)$$

$$F_k = f(p_0, \sigma_k) = \lambda_1 \quad (14)$$

Then, G_k and F_k are post-processed respectively.

- The High pass image $U_k^G = G_k \times L_{ap}$ is obtained by Laplace transform of gradient amplitude image G_k . Reconstruction of energy significant mapping image $S_k^G = U_k^G \times A_{vg}$. Finally, the mathematical morphology processing image $I_{R_k}^G$ is obtained.
- The energy significant mapping image $S_k^F = F_k \times A_{vg}$ of detail feature image F_k is directly constructed, and then the image $I_{R_k}^F$ is obtained by mathematical morphology processing.

Finally, NSST is used to fuse the post-processing images of contour features and detail features at each scale, and the fusion image $R_k = fusion(I_{R_k}^G, I_{R_k}^F)$ at each scale is obtained. Then, it is corresponding to the multi-scale, and the maximum value of pixels corresponding to each scale is calculated to obtain the multi-scale vascular detection image:

$$M_k = \max_{\sigma_{min} \leq \sigma_k \leq \sigma_{max}} R_k \quad (15)$$

The images obtained by multiscale vascular detection are shown in figure 9. Figure 10 shows the contrast of image details obtained from multi-scale vascular detection.

Figure 10 shows that the contour feature image can retain the edge contour information of the image well and accurately describe the boundary of blood vessels (figure 10(b)). The detailed feature image retains the information of the central area of the vessels and the small vessels (figure 10(c)). As shown in figure 10(d), fusion of the two images can preserve as much as possible the information of small branches of blood vessels on the premise of accurately describing image boundaries. However, in a single scale, the limitation of scale will lead to thinning of blood vessels, and part of blood vessel information not corresponding to the current scale will be lost in the image. The corresponding scale of figure 10(d) is $\sigma=1$ pixel, figure 10(e) is $\sigma=2$ pixel, and figure 10(f) is $\sigma=4$ pixel. When the scale is $\sigma=1$ pixel, many small branching vessels are retained, but the main and branching vessels are thinner than the original vessels. When the scale gradually increasing, as shown in figure 10(e) and (f), the width of main vessels gradually becomes normal,



Fig. 9. Image obtained by multi-scale vascular detection

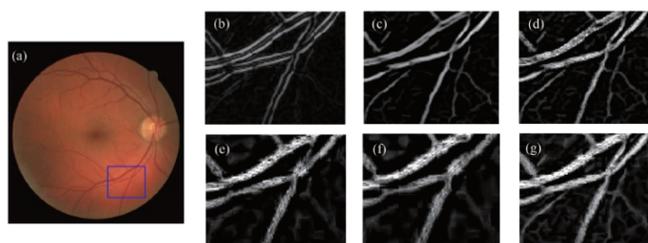


Fig. 10. Comparison of detail images obtained by multi-scale blood vessel detection. (a) Original color image; (b) contour feature; (c) detail feature, and (d) detail image of wavelet transform fusion at $\sigma=1$ pixel; (e) detail images obtained by blood vessel detection when $\sigma=2$ pixel and (f) $\sigma=4$ pixel; (g) detail image obtained by multi-scale blood vessel detection

but the information of small vessel branches will be lost. Therefore, this paper adopts the method of vascular detection under the multi-scale framework to ensure the equality and consistency of all pixels and all scales in the detection image. As can be seen from figures 8-10, the vascular information in the image is relatively intact. The contour of all the larger vessels is clear, and the vascular information at the edge is not lost. Most of the fine vessels are also well preserved, and the interference of surrounding non-vascular pixels is reduced.

4. Experiments and Analysis

4.1. Data Set

Color fundus retinal images from DRIVE dataset are used in this paper. The dataset consists of 20 training images and 20 testing images. These images are taken with a Canon CR5 non-dilated camera with a field of 45° . The field of view (FOV) of each image is round, with a diameter of about 540 pixels. Each image has a corresponding mask image depicting FOV and two expert manual segmentation results. Image size is 768×584 pixels and an 8-bit color channel is used in the RGB model.

4.2. Evaluation Index

In order to quantitatively evaluate the segmentation results of blood vessels, three indicators are used to evaluate the segmentation results including accuracy (ACC), sensitivity (SE) and specificity (SP) [32]. The calculation formulas are shown in formulas (16)-(18). ACC represents the ratio of all correctly classified pixels to all the pixels in the retinal images. SE represents the proportion of vascular pixels correctly identified in the segmentation result. SP represents the proportion of non-vascular pixels correctly identified in the segmentation result. TP represents the number of vessel pixels judged as vessels in the results. FP represents the number of background pixels judged as blood vessels in the results. FN represents the number of blood vessel pixels judged as background in the algorithm results. TN represents the number of background pixels judged as background in the algorithm result.

$$ACC = (TP + TN)/(TP + FP + TN + FN) \quad (16)$$

$$SE = TP/(TP + FN) \quad (17)$$

$$SP = TN/(TN + FP) \quad (18)$$

4.3. Results and Analysis

The edge of the field of view of fundus camera can be misjudged as vascular pixel by vascular detection in fundus retinal images. Therefore, in order to ensure the segmentation results and the accuracy of evaluation index calculation, the given mask is first used to remove the interference of the edge, and then the classical Otsu algorithm is used to

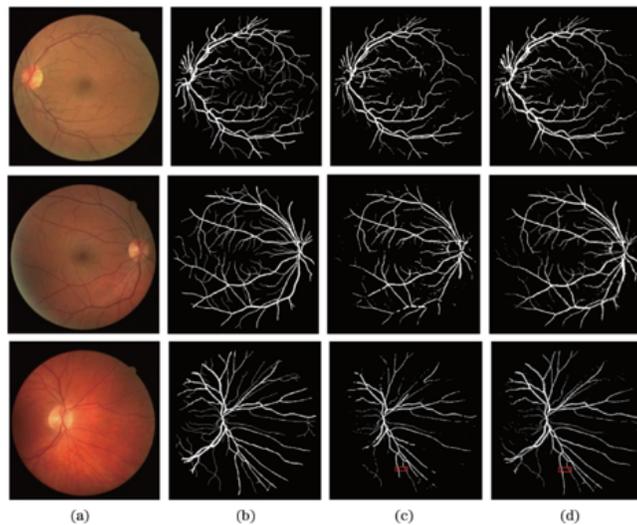


Fig. 11. Comparison of retinal vascular segmentation.(a) Original color fundus retinal images; (b) Gold standard images; (c) Results of vascular segmentation with single scale; (d) Vascular segmentation results with multiple scales.

perform segmentation directly. DRIVE test set is used to test the proposed method, and comparative analysis is conducted with gold standard images. The comparison of retinal vascular segmentation effect is shown in figure 11.

Figure 11 compares the segmentation effects of the two methods on three randomly selected retinal images in the DRIVE test set. Figure 11(a) is the original color retinal fundus image, figure 11(b) is the gold standard image, figure 11(c) is the result of blood vessel segmentation with single scale, and figure 11(d) is the result of blood vessel segmentation with the multi-scale framework. The third image in figure 11 shows the 1D cross section of the middle row marking the sub-area as shown in figure 12. The comparison results show that the proposed method can effectively reduce the influence of optic disc on vascular segmentation and reduce the possibility of many false positive pixels around blood vessels. The proposed method can ensure the accuracy of vascular segmentation, while retaining most of the small vessels, and has good segmentation performance.

The detail segmentation results with fusing vascular contour information and detail information under multi-scale and single-scale frameworks is shown in figure 13. Figure 13 is the obtained conclusion by comparing the segmentation results at different scales. In figure 13, only the corresponding good results are listed in figure 13(c) (i.e., the corresponding scale is 4 pixel). By comparing figure 13 (b) (d), it can be found that a large number of small vessel structures are lost in the vessel segmentation result of fusion under the single-scale framework, and the information of some main vessels is also lost in the image of the second row, resulting in incomplete vascular structure and poor connectivity. The vessel segmentation obtained by using the multi-scale framework is shown in figure 13(d). This new method retains more information about small vessels while ensuring the

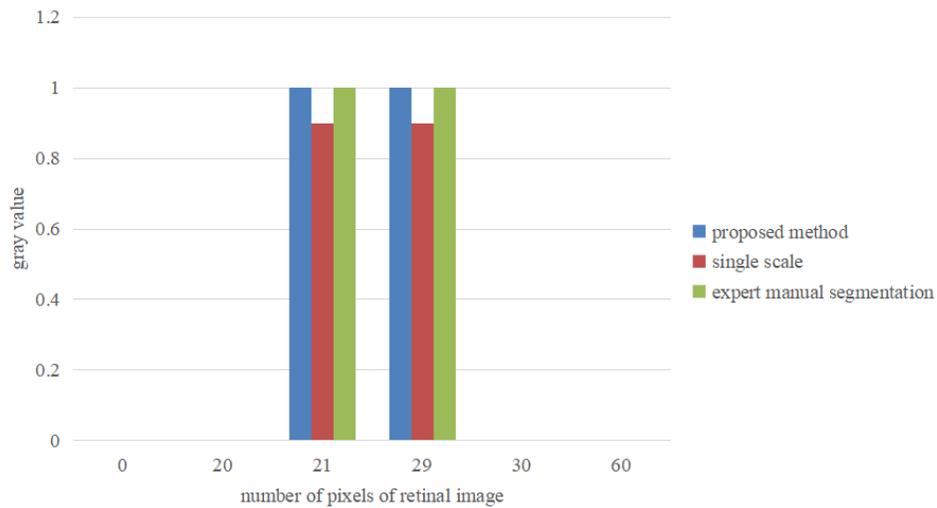


Fig. 12. The 1D cross section of the middle row of the marked area in the third image in figures 11(c) and (d)

integrity of the main vessel, and the structure of vessels is relatively complete and the connectivity is good.

In order to reflect the fusion effect of the vascular wheel information and detail information with NSST transform on the vascular contour area in retinal images, the fusion results of the vascular contour information and detail information and the unfusion results of the vascular contour information and detail information are compared and analyzed in a multi-scale framework. Vascular segmentation results of fused and unfused contour information and detail information at multiple scales are shown in figure 14.

Figure 14 compares the fusion results of vascular contour information and detail information with those of unfusion contour information and detail information using NSST in a multi-scale framework. It can be seen that, when vascular contour information and detailed information are not fused, contour features are very vague in the segmentation results where vascular branches and multiple vessels are intertwined, which leads to a lot of vascular information that cannot be accurately segmented. The fusion of vascular contour information and detailed information can ensure the accuracy and integrity of vascular contour at the intersection of vascular branches and multiple vessels.

4.4. Discussion

In order to more intuitively compare the performance of the proposed algorithm with other classical algorithms, the ACC, SE and SP are compared and analyzed. The performance comparison is shown in Table 1.

The sensitivity of M-GAN in Table 1 is the highest, slightly higher than 0.0103 of the proposed algorithm in this paper. However, the accuracy of the proposed algorithm is 0.0119 higher than that of M-GAN. Compared with other methods, the proposed method has the highest accuracy and it is 0.0117 higher than other methods. The sensitivity of

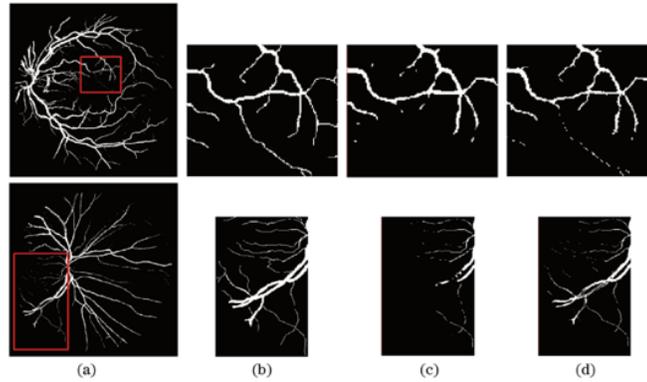


Fig. 13. Vascular segmentation results with contour information and detail information fusion in multi-scale and single-scale frames. (a) Fusion result with multi-scale framework; (b) Gold standard image details; (c) Detailed diagram of fusion results in a single-scale framework; (d) Detailed diagram of fusion results in a multi-scale framework.

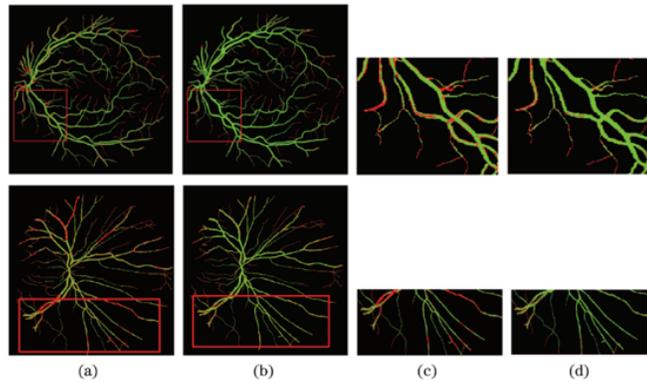


Fig. 14. Vessel segmentation results of fused and unfused contour information and detail information at multiple scales. (a) Unfused vascular contour information and detailed information; (b) Fusion of vascular contour information and detailed information; (c) Marked detailed diagram in figure 14(a); (d) Marked detailed diagram in figure 14 (b).

Table 1. Performance comparison of retinal vessel segmentation methods

Method	ACC	SE	SP
Sine-Net [33]	0.9351	0.7071	0.9704
UNMU [34]	0.9431	0.6781	0.9821
M-GAN [35]	0.9474	0.7199	0.9804
EOA [36]	0.9476	0.7176	0.9812
Proposed	0.9593	0.7097	0.9817

the proposed method in this paper is higher than that of Sine-Net and UNMU, and only slightly lower than that of M-GAN and EOA. The accuracy and sensitivity of the proposed method are higher than the average values of other methods. The above performance comparison results show that the proposed method has a high accuracy and can perform vascular segmentation well on the premise of ensuring sensitivity.

5. Conclusion

A multi-scale NSST transform fusion method for retinal vessel segmentation is proposed. The proposed method combines the contour and detail features of retinal vessels with multiple scales, and keeps the fine branches of retinal vessels well while ensuring the accurate contour of retinal vessels. The average accuracy, sensitivity and specificity were 0.9593, 0.7097 and 0.9817 respectively. On the premise of ensuring the sensitivity, the proposed method has high accuracy and good overall performance, while keeping a certain balance between sensitivity and accuracy indexes. However, in order to further improve the accuracy of vascular segmentation, the interference of optic disc on vascular segmentation needs to be solved. In the future, this proposed method will be utilized in the clinical trials.

Acknowledgments. The author is very grateful to the anonymous expert review, which provides valuable guidance to the paper.

References

1. Tchinda B S, Tchiotso D, Noubom M, et al. Retinal blood vessels segmentation using classical edge detection filters and the neural network[J]. *Informatics in Medicine Unlocked*, 2021, 23(3):100521.
2. Mohammedhasan M, Uguz H. A New Deeply Convolutional Neural Network Architecture for Retinal Blood Vessels Segmentation[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2020.
3. Francia G A, Pedraza C, Aceves M, et al. "Chaining a U-Net With a Residual U-Net for Retinal Blood Vessels Segmentation," *IEEE Access*, vol. 8, pp. 38493-38500., (2020).
4. Liang LM, Sheng XQ, Lan ZM, et al. U-Shaped Retinal Vessel Segmentation Algorithm Based on Adaptive Scale Information[J]. *Acta Optica Sinica*, 2019, 39(8):0810004.
5. Wang X, Jiang X. Retinal vessel segmentation by a divide-and-conquer funnel-structured classification framework[J]. *Signal processing*, 2019, 165(Dec.):104-114.
6. Yan Z, Yang X, Cheng K T. A Three-Stage Deep Learning Model for Accurate Retinal Vessel Segmentation[J]. *Biomedical and Health Informatics, IEEE Journal of*, 2019, 23(4):1427-1436.
7. Odstrcilik J, Kolar R, Budai A, et al. Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database[J]. *IET Image Processing*, 2013, 7(4):373-383.
8. D. A. Dharmawan, B. P. Ng and N. Borijindaragoon. Design of Optimal Adaptive Filters for Two-Dimensional Filamentary Structures Segmentation. *IEEE Signal Processing Letters*, vol. 26, no. 10, pp. 1511-1515, Oct. 2019, doi: 10.1109/LSP.2019.2938631.
9. Rodrigues LC, Marengoni M, et al. Segmentation of optic disc and blood vessels in retinal images using wavelets, mathematical morphology and Hessian-based multi-scale filtering[J]. *Biomedical signal processing and control*, 2017: 36, 39-49.
10. Hugo A R, Gabriel A, Ivan C A, et al. Blood vessel segmentation in retinal fundus images using Gabor filters, fractional derivatives, and Expectation Maximization[J]. *Applied Mathematics and Computation*, 2018, 339:568-587.

11. Moosavi A, Figueiredo N, Prasanna P, et al. Imaging features of vessels and leakage patterns predict extended interval aflibercept dosing using ultra-widefield angiography in retinal vascular disease: findings from the PERMEATE study[J]. *IEEE Transactions on Biomedical Engineering*, 2020, 68(6): 1777-1786.
12. Bhuvana, J., Mirnalinee, T. T., Bharathi, B., Sneha, I.: Efficient Generative transfer learning framework for the detection of COVID-19. *Computer Science and Information Systems*, Vol. 19, No. 3, 1241-1259. (2022), <https://doi.org/10.2298/CSIS220207033B>
13. Chen C, Chuah J H, Ali R, et al. Retinal vessel segmentation using deep learning: a review[J]. *IEEE Access*, 2021, 9: 111985-112004.
14. Arhami M, Desiani A, Yahdin S, et al. Contrast enhancement for improved blood vessels retinal segmentation using top-hat transformation and otsu thresholding[J]. *International Journal of Advances in Intelligent Informatics*, 2022, 8(2): 210-223.
15. Lee C. LSTM-CRF models for named entity recognition[J]. *IEICE transactions on information and systems*, 2017, 100(4): 882-887.
16. J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai and L. Wang, "SuperPCA: A Superpixelwise PCA Approach for Unsupervised Feature Extraction of Hyperspectral Imagery," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4581-4593, Aug. 2018, doi: 10.1109/TGRS.2018.2828029.
17. Mookiah M R K, Hogg S, MacGillivray T J, et al. A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification[J]. *Medical Image Analysis*, 2021, 68: 101905.
18. Dong H, Zhang T, Zhang T, et al. Supervised learning-based retinal vascular segmentation by m-unet full convolutional neural network[J]. *Signal, Image and Video Processing*, 2022, 16(7): 1755-1761.
19. Aguirre-Ramos H, Avina-Cervantes J G, Cruz-Aceves I, et al. Blood vessel segmentation in retinal fundus images using Gabor filters, fractional derivatives, and Expectation Maximization[J]. *Applied Mathematics and Computation*, 2018, 339: 568-587.
20. Fu Q, Celenk M, Wu A. An improved algorithm based on CLAHE for ultrasonic well logging image enhancement[J]. *Cluster Computing*, 2019, 22(5): 12609-12618.
21. Qingwu Shi, Shoulin Yin, Kun Wang, et al.?Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation.?Evolving Systems?(2021). <https://doi.org/10.1007/s12530-021-09392-3>
22. Lee M, Lee J G, Kim N, et al. Hybrid Airway Segmentation Using Multi-Scale Tubular Structure Filters and Texture Analysis on 3D Chest CT Scans[J]. *Journal of Digital Imaging*, 2018, 32(5):1-14.
23. Shoulin Yin, Hang Li, Desheng Liu and Shahid Karim. Active Contour Modal Based on Density-oriented BIRCH Clustering Method for Medical Image Segmentation [J]. *Multimedia Tools and Applications*. Vol. 79,?pp. 31049-31068, 2020.
24. Yang H, Wu X T, He B G, et al. Image fusion based on multiscale guided filters[J]. *Guangdianzi Jiguang/Journal of Optoelectronics Laser*, 2015, 26(1):170-176.
25. Ding B, Wen G, Ma C, et al. Target recognition in synthetic aperture radar images using binary morphological operations[J]. *Journal of Applied Remote Sensing*, 2016, 10(4):046006.
26. Gai D, Shen X, Chen H, et al. Medical image fusion using the PCNN based on IQPSO in NSST domain[J]. *IET Image Processing*, 2020(5).
27. Zhang Z, Xi X, Luo X, et al. Multimodal image fusion based on global-regional-local rule in NSST domain[J]. *Multimedia Tools and Applications*, 2020(5):1-27.
28. R Yang, Du B, Duan P, et al. Electromagnetic Induction Heating and Image Fusion of Silicon Photovoltaic Cell Electrothermography and Electroluminescence[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(7):4413-4422.
29. Jing Yu, Hang Li, Shoulin Yin. Dynamic Gesture Recognition Based on Deep Learning in Human-to-Computer Interfaces [J]. *Journal of Applied Science and Engineering*, vol. 23, no. 1, pp.31-38, 2020.

30. Wang Z, Wang L, Elimelech M. Viability of Harvesting Salinity Gradient (Blue) Energy by Nanopore-Based Osmotic Power Generation[J]. Engineering, 2021.
31. Premjith B, Soman K P. Deep learning approach for the morphological synthesis in malayalam and tamil at the character level[J]. Transactions on Asian and Low-Resource Language Information Processing, 2021, 20(6): 1-17.
32. Y. Yuan, Z. Xu and G. Lu, "SPEDCCNN: Spatial Pyramid-Oriented Encoder-Decoder Cascade Convolution Neural Network for Crop Disease Leaf Segmentation," in IEEE Access, vol. 9, pp. 14849-14866, 2021, doi: 10.1109/ACCESS.2021.3052769.
33. Atli B, Gedik O S. Sine-Net: A fully convolutional deep learning architecture for retinal blood vessel segmentation[J]. Engineering Science and Technology an International Journal, 2021, 24(2): 271-283.
34. Upadhyay K, Agrawal M, Vashist P. Unsupervised multiscale retinal blood vessel segmentation using fundus images[J]. IET Image Processing, 2020, 14(11).
35. K. -B. Park, S. H. Choi and J. Y. Lee, "M-GAN: Retinal Blood Vessel Segmentation by Balancing Losses Through Stacked Deep Fully Convolutional Networks," in IEEE Access, vol. 8, pp. 146308-146322, 2020, doi: 10.1109/ACCESS.2020.3015108.
36. Rezaee K, Haddadnia J, Ashkan T. Optimized clinical segmentation of retinal blood vessels by using combination of adaptive filtering, fuzzy entropy and skeletonization[J]. Applied Soft Computing, vol. 52, pp. 937-951, 2017. doi: 10.1016/j.asoc.2016.09.033.

Feng Lijuan is with School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology.

Fan Zhang is with the Zhengzhou University of Technology, Zhengzhou 450000, China. Research direction: Information security, image processing and cloud computing.

Received: November 20, 2022; Accepted: April 12, 2023.