

Article

End-to-End Automatic Classification of Retinal Vessel Based on Generative Adversarial Networks with Improved U-Net

Jieni Zhang^{1,2}, Kun Yang^{1,3,*} , Zhufu Shen^{4,*}, Shengbo Sang^{1,3} , Zhongyun Yuan^{1,3,5}, Runfang Hao^{1,3}, Qi Zhang^{1,2} and Meiling Cai⁶

¹ Shanxi Key Laboratory of Micro Nano Sensor & Artificial Intelligence Perception, College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China; ineijgnahz@163.com (J.Z.); sunboa-sang@tyut.edu.cn (S.S.); yuanzhongyun@tyut.edu.cn (Z.Y.); txlhrf@163.com (R.H.); zhangqi0564@link.tyut.edu.cn (Q.Z.)

² Shanxi Institute of 6D Artificial Intelligence Biomedical Science, Taiyuan 030031, China

³ Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education, Taiyuan University of Technology, Taiyuan 030024, China

⁴ Department of Geriatrics, 920th Hospital of Joint Logistics Support Force, Teaching Hospital of Kunming Medical University, Kunming 650032, China

⁵ Speed Electronics Company Limited, Room 305, Tower A, Luohutouzikonggu Building, Qingshuihe 1st Road, Luohu, Shenzhen 518000, China

⁶ Image and Intelligence Laboratory, College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China; caimeiling0570@link.tyut.edu.cn

* Correspondence: yangkun01@tyut.edu.cn (K.Y.); shenzhufu@sohu.com (Z.S.)

Abstract: The retinal vessels in the human body are the only ones that can be observed directly by non-invasive imaging techniques. Retinal vessel morphology and structure are the important objects of concern for physicians in the early diagnosis and treatment of related diseases. The classification of retinal vessels has important guiding significance in the basic stage of diagnostic treatment. This paper proposes a novel method based on generative adversarial networks with improved U-Net, which can achieve synchronous automatic segmentation and classification of blood vessels by an end-to-end network. The proposed method avoids the dependency of the segmentation results in the multiple classification tasks. Moreover, the proposed method builds on an accurate classification of arteries and veins while also classifying arteriovenous crossings. The validity of the proposed method is evaluated on the RITE dataset: the accuracy of image comprehensive classification reaches 96.87%. The sensitivity and specificity of arteriovenous classification reach 91.78% and 97.25%. The results verify the effectiveness of the proposed method and show the competitive classification performance.

Keywords: retinal vessel classification; fundus image; GAN; U-Net; image segmentation



Citation: Zhang, J.; Yang, K.; Shen, Z.; Sang, S.; Yuan, Z.; Hao, R.; Zhang, Q.; Cai, M. End-to-End Automatic Classification of Retinal Vessel Based on Generative Adversarial Networks with Improved U-Net. *Diagnostics* **2023**, *13*, 1148. <https://doi.org/10.3390/diagnostics13061148>

Academic Editor: Jae-Ho Han

Received: 12 February 2023

Revised: 7 March 2023

Accepted: 13 March 2023

Published: 17 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the “World Vision Report” issued by WHO, at least 2.2 billion people in the world are visually impaired or blind, and nearly half of the visual impairments could be avoided through early prevention [1]. By 2020, the number of glaucoma patients reached 796 million worldwide, and age-related macular degeneration patients numbered 200 million people [2,3]. The number of patients with diabetic retinopathy is expected to reach 200 million by 2035 [4]. However, the noted irreversible blinding diseases can be prevented and treated in advance by an ophthalmologist’s examination of the ocular fundus [5]. In this relevant examination, the pattern and structure of retinal vessels are important clinical characterizations. In addition, retinal blood vessels in the human body are the only blood vessels that can be directly observed with non-invasive imaging technology. The morphology is also affected by various factors of cardiovascular disease, hypertension, arteriosclerosis, and other systemic diseases [6–8]. The changes in retinal vessels, including vascular caliber, branch morphology, and arteriolar-to-venular diameter ratio (AVR), can

be used as the diagnostic basis for vascular related diseases. For example, the method of observing the aspect ratio of retinal arterioles and the asymmetry of venous branches is an early monitoring method for Alzheimer's disease [9]. The risk value of coronary heart disease and other diseases is also associated with the ratio of arteriovenous diameters [10]. In addition, arterial stenosis in the cerebrovascular network is significantly correlated with retinal arteriolar diameter [11].

Therefore, it is of great significance to classify the retinal vessels accurately for the prevention and observation of many diseases. However, early retinal vessel segmentation and classification mainly rely on manual labeling by professional doctors [12], which requires a lot of time and effort. Moreover, due to the reliance on subjective criteria, the results of segmentation and classification can be different. Automatic segmentation and classification of retinal vessels can greatly reduce the workload of doctors and can also avoid the impact of different doctors' subjective factors on the classification results. With the development of computer vision technology, there are many methodologies for retinal vessel segmentation, including image filtering technology [13,14], machine learning algorithm for feature extraction [15], and neural network research [16]. However, related work that is dedicated to the classification of arteriovenous vessels is significantly less than that on vascular segmentation [17].

U-shaped convolution network (U-Net) [18] is widely used in the field of fundus blood vessel segmentation because of its excellent effect in the field of medical image segmentation. Specifically, the U-Net network is superior to convolutional neural network (CNN) and the fully convolutional networks (FCN) networks in the field of retinal vessel segmentation [19]. At the same time, generative adversarial network (GAN) has also been widely used in the field of fundus images and has been proven to be beneficial for various tasks [20]. In recent years, the GAN network has made good progress in the field of medical image segmentation. Some proposed GAN also obtained high accuracy for retinal vessel segmentation [19]. In addition, the atrous spatial pyramid pooling (ASPP) module enables the network to expand the receptive field and capture multi-scale contextual information without increasing the complexity of the algorithm parameters. It also reduces the loss of detailed information, which makes the vascular feature information better preserved and can enrich the microvascular information. In the task of blood vessel segmentation and classification, the number of pixels in the background of the fundus image is far greater than the number of blood vessels. In order to make the network pay more attention to the generation of blood vessel pixels, the attention module is also designed to be added to the network structure. At the same time, residual connections are introduced into the downsampling process of the U-Net of a generator network to alleviate the problem of gradient disappearance and increase the sensitivity of the generator network to weight changes, which makes the generator improve the vascular classification effect. In addition, low contrast is a major obstacle to the retinal image in optical imaging [21]. Before network classification, the preprocessing operations need to be conducted to enhance the contrast of the data.

To summarize, the contributions and novelty of the present study are highlighted as follows:

1. A model based on GAN and improved U-Net is proposed for the automatic end-to-end classification of fundus arteriovenous vessels. The introduction of ASPP and attention modules can also improve the classification capability of the model. The classification results of the proposed model are highly competitive.
2. A local contrast enhancement method was used to preprocess the input images. Through preprocessing, the problems of low overall brightness and poor contrast between blood vessels and background of the original fundus image data were solved. The effectiveness of this method was verified by ablation experiments.
3. The proposed method allows simultaneous classification of vessel crossings in fundus images in addition to the classification of arterioles, which is innovative in the study of fundus vascular classification.

The experimental materials used in this study are described in detail in Section 3. The design of the network model and the experimental process are described in Section 4. The results obtained are reported in Section 5. Section 6 summarizes this article.

2. Related Work

Although extensive research has been carried out in the field of retinal vessel segmentation, little attention has been paid to the field of automatic classification of retinal vessels [22]. Based on the existing research, the methods of retinal artery and vein classification can be divided into two categories: traditional machine learning based methods and depth learning based methods.

(A) Traditional Machine Learning Based Methods

Manual features. In the research of vessel classification based on traditional machine learning methods, it is usually necessary to manually extract features and then classify arteries and veins. It is often accompanied by some post-processing steps. Sathananthavathi et al. [23] extracted features manually according to the morphological structure of retinal vessels; the BAT evolutionary algorithm and the random forest classifier were used for main feature determination and classification, respectively; and, finally, the post-processing was used at the bifurcation of retinal vessels. Srinidhi et al. [24] and Xu et al. [25] also used manual features combined with random forest classifier to classify arteries and veins. Vázquez et al. [26] used the optic disc to divide the retinal vessels into many segments, and then the vessel segments were classified by color information, and the final classification result of whole blood vessel was determined by voting of the connected blood vessel segments.

Graph-based methods. Welikala et al. [27] avoided the use of hand-crafted features. The vascular network was first segmented from the retinal image. Bifurcations and crossover points were searched based on the retinal vascular skeleton, and vessel segments were segmented using the centerline. Finally, the vessel segments were fed into a convolutional neural network based on three convolutional and three fully connected layers to achieve the arteriovenous classification of retinal vessels. The classification rate with 47 features (the largest dimension tested) using OLPP in their own ORCADES dataset is only 90.56%, and the classification rate in the public dataset DRIVE is 86.7%. Zhao et al. [28] constructed the graph through image segmentation, skeletonization, and identification of significant nodes. They formalized the topology estimation and A/V classification into a pairwise clustering problem. The classification of blood vessels was effectively realized.

(B) Deep Learning Methods

Segmentation first and then classification. With the continuous development of deep learning, convolutional neural network has also been applied in the field of retinal vessel classification. Especially after U-Net [18] was proposed, it has performed well in the field of retinal vessel segmentation and classification, which leads to the realization of pixel-level segmentation and classification of fundus images. Li et al. [29] regarded arteriovenous classification as a three-classification task. First, the fundus image was preprocessed using the fuzzy removal technology, and then the image was classified by using the improved U-Net network. In order to improve the classification accuracy, the tracking algorithm was used as the post-processing method to further classify the blood vessels. Binh et al. [30] also regarded arteriovenous classification as a three-classification task, the improved U-Net model was used to classify retinal vessels, and the method of graph cutting was used for post-processing; the accuracy of their method is about 97.6%.

End-to-end classification. Morano et al. [31] decomposed the joint task into three segmentation problems: arteries segmentation, veins segmentation, and vessel segmentation. Their classification network consisted of the straightforward application of an FCNN with a custom loss. The accuracy of classifying retinal vessels reached 95.45%. Galdran et al. [32] used CNN as a task classification network. The previously segmented vascular tree did not need to be included in their classification method. Fully automatic classification of retinal

blood vessels was achieved. They also proposed a classification of uncertain blood vessels. For the benefit of retrieval, the investigation of comparison is concisely summarized in Table 1.

Table 1. Summary of contemporary retinal vessel classification methods.

Method	Category	Pros and Cons	Examples
Traditional	Manual features	The extracted vascular features can be explained, but feature processing requires more professional knowledge of fundus vascular medicine and images	[23–26]
	Graph-based methods	The interpretable features can be extracted automatically, and the vascular segments are classified. However, accurate segmentation at the pixel level is not achieved	[27,28]
Deep learning	Segmentation first and then classification	The accuracy of classification is good, but the vascular classification task is divided into multiple subtasks, and the errors in the vascular segmentation subtask will be directly transmitted to the vascular classification subtask	[18,29,30]
	End-to-end classification	Automatic blood vessel segmentation and pixel-level classification are implemented synchronously, but the amount of related research is insufficient	[31,32]

3. Experimental Materials and Preprocessing

3.1. Dataset

The Retinal Images vessel Tree Extraction (RITE) dataset [33] is used in this work, which is derived from the DRIVE dataset [34]. It has been widely used as evaluation criteria in research fields such as retinal vessel segmentation, vessel extraction, and vessel classification. As shown in Figure 1, the RITE dataset is composed of four parts: fundus image, mask image used to extract the region of interest, the vessel trees manually segmented, and the A/V reference standard. The A/V reference standard is generated by marking each vascular pixel. In Figure 1d, red represents artery (A), blue represents vein (V), green represents crossing parts of artery and vein (O), and white represents uncertain vessel (U), with a resolution of 565×584 pixels.

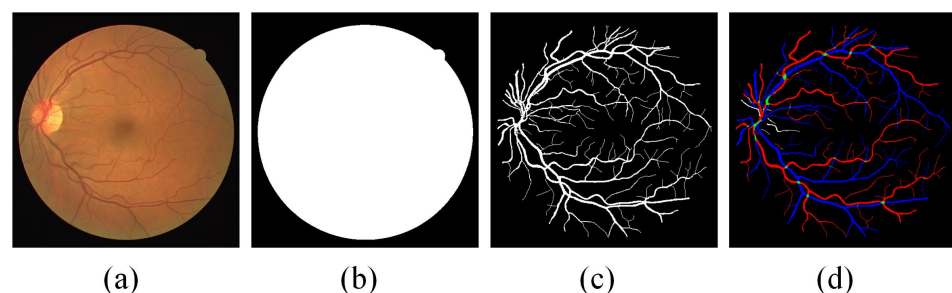


Figure 1. Examples of RITE dataset: (a) fundus images; (b) ROI mask; (c) ground truth of vessel segmentation; (d) ground truth of vessel classification.

3.2. Preprocessing

The collection process of image datasets inevitably has problems such as uneven lighting and noise. Similarly, the fundus image dataset is limited by imaging conditions such as low overall illumination and low contrast between blood vessels and background. These problems will have negative impacts on the further classification of fundus images. Therefore, this paper uses adaptive contrast enhancement (ACE) [35] to preprocess the original retinal fundus image.

The ACE algorithm divides the image into high frequency and low frequency. The low frequency part is obtained by smoothing, blurring, and other low-pass filtering methods. The high frequency part is directly obtained by subtracting the low frequency part from the original picture. In ACE, the high frequency part is amplified and added to the original low frequency part to obtain the enhanced image. The color constancy and brightness constancy of the enhanced image are improved, and the image contrast is changed. The details of ACE are as follows:

$$y(i, j) = m_x(i, j) + G(i, j)[x(i, j) - m_x(i, j)], \quad (1)$$

where $x(i, j)$ represents the pixel value corresponding to the image coordinate (x, j) before preprocessing; $m_x(i, j)$ represents the low-frequency part, $[x(i, j) - m_x(i, j)]$ represents the high-frequency part, and $G(i, j)$ represents the high-frequency amplification coefficient (contrast gain).

The low-frequency part represents the local average value of the area with the window size of $(2n + 1) \times (2n + 1)$ centered on the image coordinate (x, j) pixel. The specific formula is as follows:

$$m_x(x, j) = \frac{1}{(2n + 1)^2} \sum_{k=i-n}^{i+n} \sum_{l=j-n}^{j+n} x(k, l) \quad (2)$$

According to the relevant research on ACE [36], the high-frequency amplification coefficient $G(i, j)$ is defined as a variational constant, which is inversely proportional to the local mean square error, as shown in Equation (3).

$$G(x, j) = \alpha \frac{\delta}{\sigma_x(x, j)}, \quad (3)$$

where the value of δ is equal to the global mean square error of the image, which can reflect the dispersion of the image pixel value and the mean value. Constant α coefficient can linearly adjust the total amplification coefficient. The expression $\sigma_x(i, j)$ represents the local mean square deviation, which can reflect the contrast change of the gray value of each pixel in the local area of an image. The expression of σ is shown in Equation (4).

$$\sigma_x(x, j) = \sqrt{\frac{1}{(2n + 1)^2} \sum_{k=i-n}^{i+n} \sum_{l=j-n}^{j+n} [x(k, l) - m_x(x, j)]^2} \quad (4)$$

In addition, in order to avoid noise amplification or pixel value saturation caused by small local variance of the local part of image smoothing, the maximum value of $G(i, j)$ is limited. The setting of the maximum of $G(i, j)$ and the effects of all parameters is discussed in Section 5.5. The fundus images before and after preprocessing are shown in Figure 2. By comparing the original image, it can be seen that the contrast between the vascular pixels, especially the microvascular pixels and the background pixels, is significantly enhanced after preprocessing.

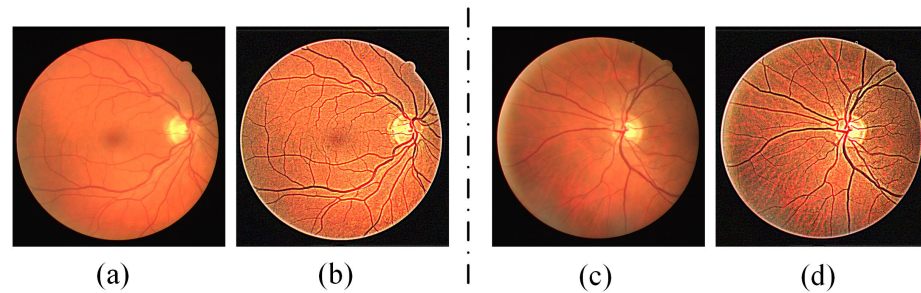


Figure 2. Two examples of fundus images before and after preprocessing: (a,c) original fundus image; (b,d) fundus image after preprocessing.

4. Methods

4.1. Method Architecture Overview

The overall flow chart of the proposed method is shown in Figure 3. In the proposed method, ACE is used to preprocess the original retinal fundus image before performing network training and testing. The proposed model is designed using GAN. GAN is based on the idea of dynamic adversarial and can be divided into the generator and the discriminator parts. In this paper, the modified U-Net network is used as the generator and the convolutional network is used as the discriminator.

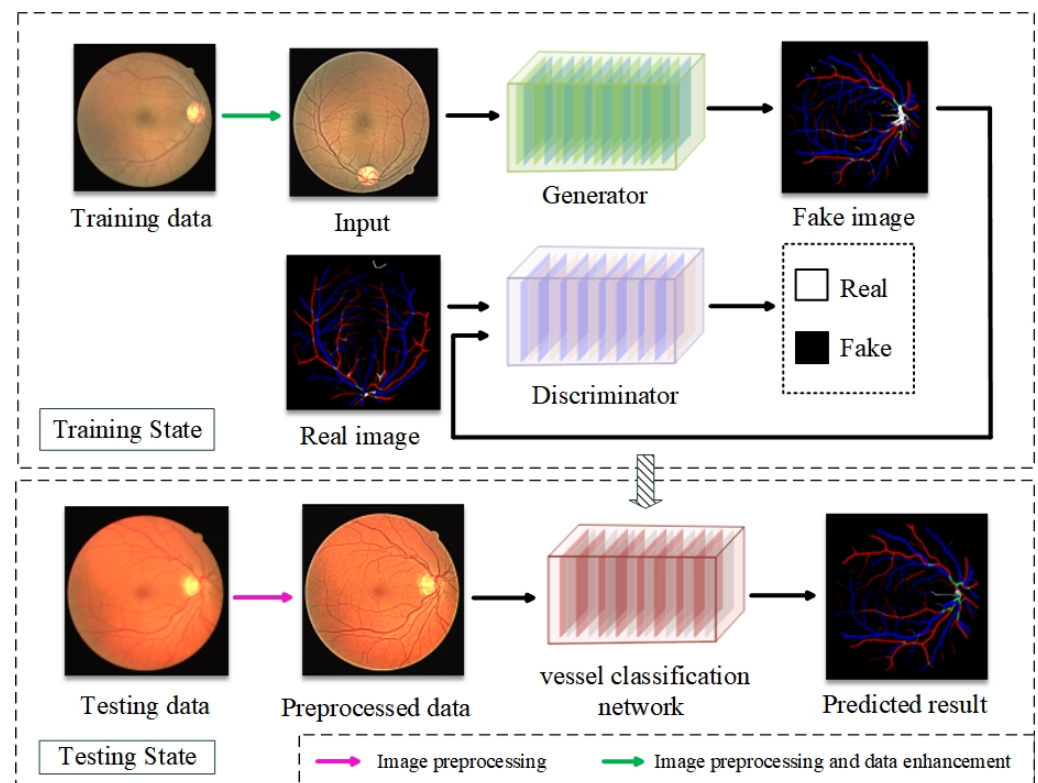


Figure 3. The overall flow chart of the proposed method.

The generator is used to generate the vessel classification prediction result from the input training fundus image data, which is labeled as Fake Image. The corresponding Real Image is the ground truth of vessel classification from trained fundus image data. The discriminator is used to discriminate the image source, and the discriminated results are marked as Real and Fake. Being marked as Real means the image comes from the real vascular classification data Real Image, and being marked as Fake means the image come from the vascular classification data Fake Image, which is generated by the generator. By repeatedly training the generator and discriminator, the Fake Image will be as close to

the Real Image as possible. When the discriminator cannot distinguish between real and false numbers, the required vessel classification network training is completed.

4.2. Network Structure

The design details of GAN are described in this section. The U-Net [18] network is used as the main design of the generator. The structure of the designed generator network is shown in Figure 4. The atrous spatial pyramid pooling (ASPP) [37] module is added to the downsampling process of the U-Net network. The attention module is introduced at the skip link of the U-Net network. To alleviate the problem of gradient disappearance, residual connections are introduced into the downsampling process of the U-Net network. The residual connections helps to increase the sensitivity of the generator network to weight changes, which makes the generator fully learn the distribution of retinal vascular pixels and improve the vascular classification effect.

The main role of the discriminator is to provide a descent gradient for the generator. Complex discriminator gradients will cause the gradient of the generator to disappear, which will not achieve effective adversarial training [38]. In this paper, ordinary convolutional networks are used as a discriminator, with the structure shown in Figure 5.

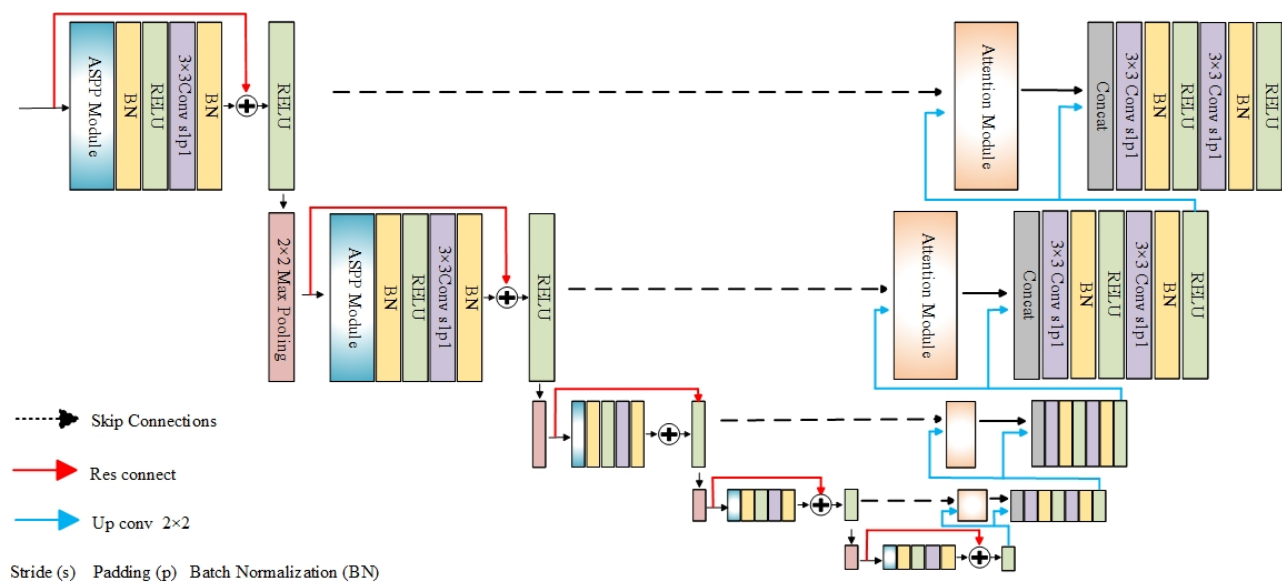


Figure 4. Improved U-Net network used as the generator (the ASPP Module is shown in Figure 6, and the Attention Module is shown in Figure 7).

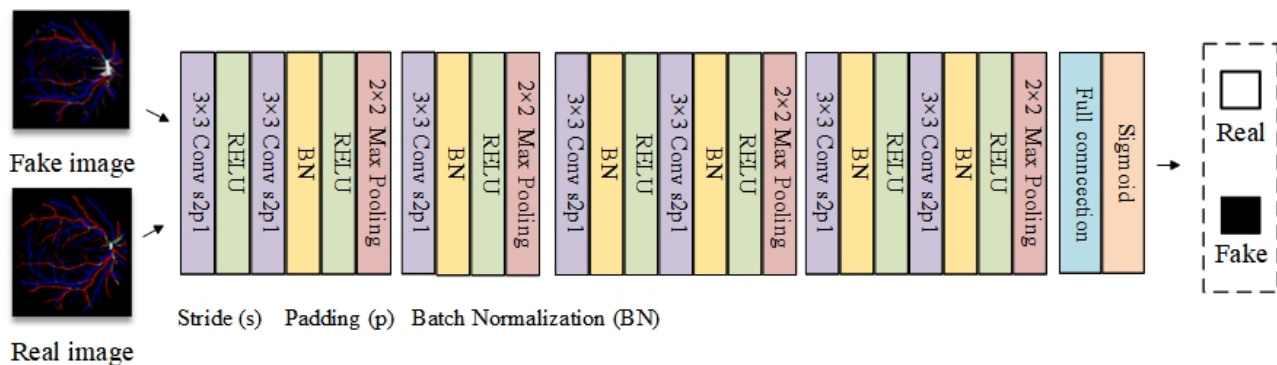


Figure 5. Discriminator network used.

4.2.1. ASPP Module

The original U-Net network uses downsampling to expand the receptive field while reducing the resolution, but the pooling layer of traditional convolution in U-Net will lose the details of retinal images, which leads to problems such as incomplete microvessel segmentation and susceptibility to breakage. In this paper, the ASPP module is added, which utilizes dilated/atrous convolution with multiple expansion rates to stack into a pyramidal structure instead of normal convolution. The structure of the ASPP module is shown in Figure 6. The structure enables the network to expand the receptive field and capture multi-scale contextual information without increasing the complexity of the algorithm parameters. It also reduces the loss of detail information, which makes the vascular feature information better preserved and can enrich the microvascular information.

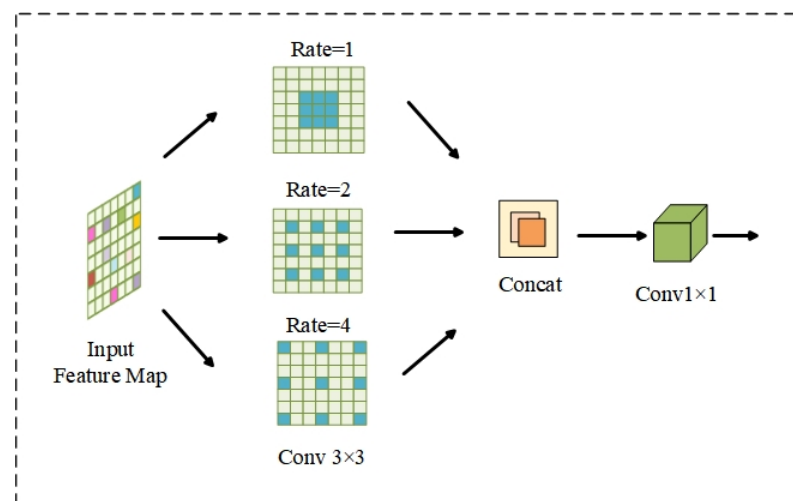


Figure 6. The structure of the ASPP module.

4.2.2. Attention Module

In the vascular classification task, the number of background pixels far exceeds the number of vascular pixels. The downsampled extracted feature maps are spliced directly with upsampling in the skip connection of traditional U-Net. Such a design produces a lot of redundant information and also leads to the deterioration of the extracted features. Inspired by Ashish Vaswani [39], the attention module is added to each skip to suppress excessive irrelevant information and make the model more concerned with the generation of vascular pixels. The details of the attention module are shown in Figure 7. In the figure, X is the feature map from downsampling and Y is the feature map from upsampling. After 1×1 convolution operation, X and Y are summed to highlight the features. After passing through ReLU and Sigmoid, the highlighted features are ranged between 0 to 1, which is the attention weight. It is assigned to the low-level feature after the attention weight is multiplied with X , and the attention allocation is achieved.

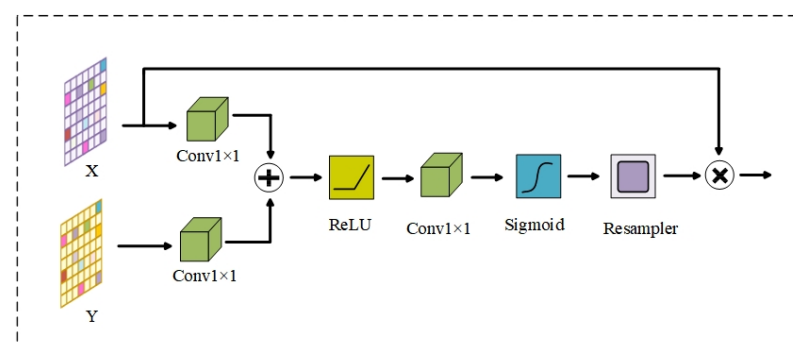


Figure 7. The structure of the attention module.

5. Experiments and Results

In this section, the experiments and results of the proposed method on the RITE dataset are detailed. Before going into the details of the implementation, the metrics to quantitatively evaluate the vascular classification results are described in Section 5.1. To ensure the validity of the proposed method, an ablation study is also described in Section 5.4. In addition, the effects of the parameters about the preprocessing are discussed in Section 5.5.

5.1. Evaluation Metrics

Considering the multiple cases of actual and predicted values in fundus images, a confusion matrix applicable to vascular classification studies is established, shown in Table 2. Because the proportion of arteriovenous crossing (A-V crossing) pixels and uncertain vessel pixels in fundus images is small, they are classified together as an uncertain vessel category.

Table 2. The confusion matrix for vascular classification of fundus images.

Actual Category	Forecast Classification Results			
	A (Artery)	V (Vein)	U (Uncertain Vessel)	B (Background)
A (Artery)	TA	FVa	FUa	FBa
V (Vein)	FAv	TV	FUv	FBv
U (Uncertain vessel)	FAu	FVu	TU	FBu
B (Background)	FAb	FVb	FUb	TB

According to the assessment methods widely used in the field of vascular classification [29,40–43], sensitivity (*Sens*), specificity (*Spec*), and accuracy (*Acc*) are used to assess the performance of the classification; the arteries are set as positive and the veins are set as negative. *Sens*, *Spec* and *Acc* are defined as follows:

$$Sens = \frac{TA}{TA + FVa} \quad (5)$$

$$Spec = \frac{TV}{TV + FAv} \quad (6)$$

$$Acc = \frac{TA + TV}{TA + FVa + TV + FAv} \quad (7)$$

In order to quantitatively analyze the accuracy of uncertain vessels (including A-V crossing) classification, the index *Acc_U* is also introduced. To reflect a more comprehensive performance, the overall accuracy *Acc_all* and the background segmentation accuracy *Acc_B* are added to evaluate the performance of the overall prediction of fundus images and the performance of the segmentation performance of non vascular pixels; *Acc_U*, *Acc_All*, and *Acc_B* are defined as follows:

$$Acc_U = \frac{TU}{TU + FVu + FAu} \quad (8)$$

$$Acc_{All} = \frac{TA + TV + TU + TB}{(TA + FAv + FAu + FAb + TV + FVa + FVu + FVb + TU + FUa + FUv + FUb + TB + FBa + FBv + FBu)} \quad (9)$$

$$Acc_B = \frac{TB}{TB + FAb + FUv + FVb} \quad (10)$$

5.2. Implementation Details

Twenty color fundus images of RITE are selected randomly as the training set, and the remaining images are used as the test set. To increase the training data, a series of data enhancement operations such as horizontal flipping and rotating the images was performed. The network implementation is based on the Python language and the Pytorch framework, and the model is trained on a PC with a core i9-12900k CPU (3.8 GHz) and NVIDIA RTX3090-24GGPC (32 GB of RAM). During training, the batch_size was set to 1, the learning rate was 0.0008, and the optimizer was Adam. The experiment was trained for a total of 855 iterations and took about 7 h.

5.3. Classification Results

The classification results of the proposed model on the RITE retinal images are presented in Figure 8. The validity of the proposed method is demonstrated by comparing the network prediction results of fundus images with the ground truth of vessel classification from the medical professional. In addition to the crude arteriovenous vessels, microvessels can also be more effectively and accurately classified. The details of the microvessel classification are magnified in Figure 8d. As described in Section 5.1, the results of the data quantified by the six evaluation metrics are shown in Table 3.

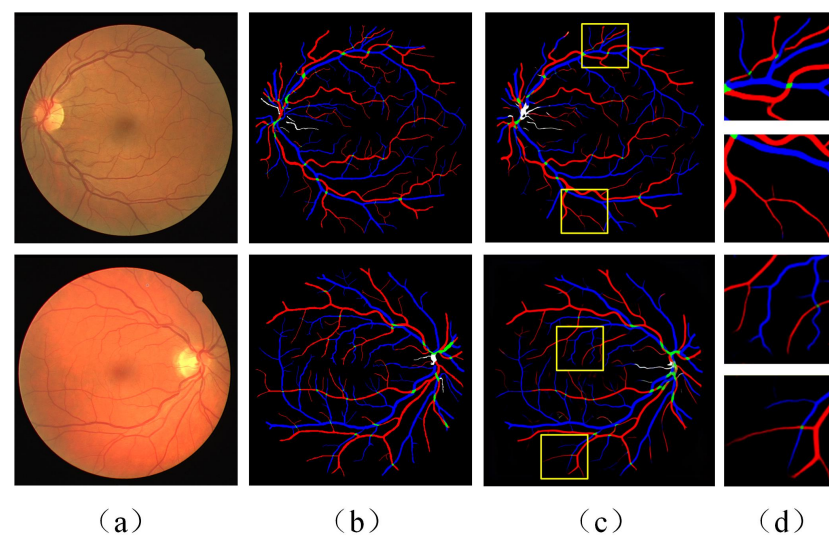


Figure 8. Example of vessel classification results of RITE fundus image dataset: (a) the original retinal fundus image; (b) ground truth of vessel classification; (c) vessel classification results of proposed method; (d) detail amplification of microvascular region.

Table 3. Vessel classification assessment results.

Formula	Index	Result	Comments
(1)	<i>Sens</i>	91.78%	The ability to classify arteries
(2)	<i>Spec</i>	97.25%	The ability to classify veins
(3)	<i>Acc</i>	96.86%	The ability to comprehensively classify the arteries and veins
(4)	<i>Acc_U</i>	63.23%	The ability to classify uncertain vessels (including A-V crossing)
(5)	<i>Acc_{All}</i>	96.87%	The ability to accurately classify the whole fundus image
(6)	<i>Acc_B</i>	98.88%	The ability to classify non-vascular regions

In Table 4, the experimental result is compared with the results of related studies based on the three important evaluation statistics of *Sens*, *Spec*, and *Acc*. As can be seen from Table 4, the proposed method outperforms most models in all evaluation metrics and achieves the highest scores on *Acc* and *Sens*. It shows the competitive performance of the proposed method compared to state-of-the-art methods. The evaluation metric *Sens* can be used to indicate arterial classification performance. The value of *Sens* of the proposed

classification method increased to 91.78%, which proves the advantages of the proposed method in arterial classification performance. Among the methods, Morano et al. [31] designed an FCNN classification network with custom losses. Their method has a greater classification performance for veins. The value of *Spec* was increased to 98.67% in their method. However, the value of *Sens* only reached 78.07%. The balance between arterial classification and venous classification is the worst compared with other methods. The value of *Sens* also decreased by 13.71% compared with the proposed method.

Table 4. Comparison of the assessment results for the related advanced arteriovenous classification.

Methods	Acc	Sens	Spec	Datasets	A-V Crossing	Years	End to End
Xu et al. [44]	92.30%	90.00%	90.00%	RITE	No	2017	No
Dashtbozorg et al. [25]	87.40%	73.00%	84.00%	DRIVE	No	2014	No
Kang et al. [41]	90.81%	88.63%	92.97%	RITE	No	2020	No
Huang et al. [45]	72.00%	70.90%	73.80%	DRIVE	No	2018	No
Galdran et al. [32]	89.00%	89.00%	90.00%	RITE	Yes	2019	Yes
Morano et al. [31]	96.05%	78.07%	98.67%	RITE	No	2021	Yes
This work	96.86%	91.78%	97.25%	RITE	Yes	2023	Yes

The best results are highlighted in bold.

In the whole fundus image, the pixel of A-V crossings is a small proportion. As such, the corresponding classification training samples that can be obtained are few. Therefore the accurate classification of the A-V crossings is a greater challenge. In most current studies of fundus image classification [31,32,43–45], the identification of the A-V crossings region is not considered. However, A-V crossings are an area of concern for physicians. For example, the retinal crossover sign is one of the common alterations of retinal vessels in the fundus of hypertensive eyes [46]. The phenomenon of compression at the A-V crossings regions is the retinal arteriovenous crossing sign. When the A-V crossings are compressed, they are called retinal arteriovenous crossings signs. Therefore, A-V crossings also need to be considered in the classification studies of retinal vessels in the fundus images. The classification results of the A-V crossings are shown in Figure 9. The results show that A-V crossings far from the optic disc regions would enable a more accurate classification. However, due to the dense and complex distribution of blood vessels near the optic disc, the classification in these regions is not ideal.

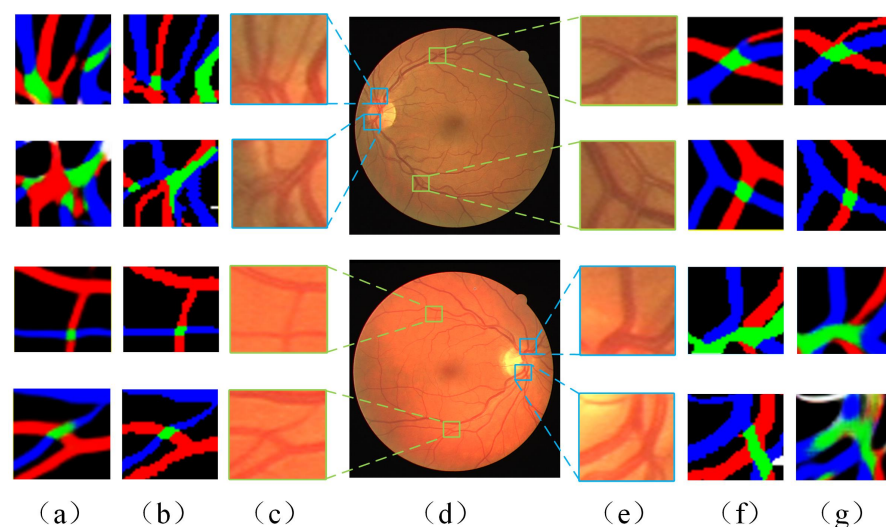


Figure 9. Details of A-V crossings classification results (the areas bounded by the blue rectangle is the examples near the optic disc, the areas bounded by the green rectangle is the examples far away from the optic disc). (a,g) A-V crossing classification results; (b,f) ground truth of A-V crossing; (c,e) fundal images of A-V crossing locally; (d) fundus image.

5.4. Ablation Study

In order to verify the effectiveness of the proposed preprocessing and network improvement modules, six sets of ablation studies are conducted on the RITE dataset. The results of the ablation studies are shown in Table 5. The baseline model is the backbone network, where the original U-Net network is used as the generator network. Compared to baseline, adding residual connections *RS* improved *Acc* by 3.64% and *Sens* by 7.23%. After the addition of the ASPP module (*ASPPM*), *Acc* was improved by 3.10%, and *Sens/Spec* were improved, respectively, by 15.74%/3.02%. The addition of the attention module (*AttM*) increased *Acc/Sens/Spec* by 4.56%/14.66%/4.58%. Combining the residual connections, the ASPP module, results in an overall improvement for *Acc/Sens/Spec/Acc_U* by 4.91%/16.59%/5.54%/7.78%, which verified the effectiveness in the improved network. Moreover, with the preprocessing method (*Pre*), the results of *Acc*, *Sens*, and *Acc_U* reach the optimum. In summary, the proposed method is verified to have potential in the vascular classification task.

Table 5. Ablation study results of baseline and improved network.

Methods					Vessel Classification			
Baseline	<i>RS</i>	<i>ASPPM</i>	<i>AttM</i>	<i>Pre</i>	<i>Acc</i>	<i>Sens</i>	<i>Spec</i>	<i>Acc_U</i>
Yes	No	No	No	No	91.13%	72.67%	92.33%	52.08%
Yes	Yes	No	No	No	94.77%	79.90%	96.61%	54.31%
Yes	No	Yes	No	No	94.23%	88.41%	95.35%	58.98%
Yes	No	No	Yes	No	95.69%	87.33%	96.91%	57.11%
Yes	Yes	Yes	Yes	No	96.04%	89.26%	97.87%	59.86%
Yes	Yes	Yes	Yes	Yes	96.86%	91.78%	97.25%	61.23%

5.5. Discussion

In Equation (1), $G(x, j)$ represents the high-frequency amplification coefficient (contrast gain). When it is a fixed constant, the high-frequency part is amplified in the same proportion. An overexposure phenomenon occurs in over-enhanced areas such as the edge of the fundus image and the center of the optic disc, as shown in Figure 10a. To avoid this phenomenon, the high-frequency amplification coefficient $G(x, j)$ is defined as a variational constant in this work, which is inversely proportional to $\delta_x(i, j)$, as shown in Equation (3); $\delta_x(i, j)$ represents the local mean square deviation, and its value is directly related to n . Figure 10b shows an example of a result where n is equal to 50; the overexposure phenomenon at the edge of the fundus image and the center of the optic disc is significantly improved compared with Figure 10a. Due to the addition of $\delta_x(i, j)$, the high frequency amplification factor $G(x, j)$ becomes a variable of spatial adaptation. In places where the image changes violently, the high-frequency amplification coefficient decreases accordingly, which can avoid the Ringing effect.

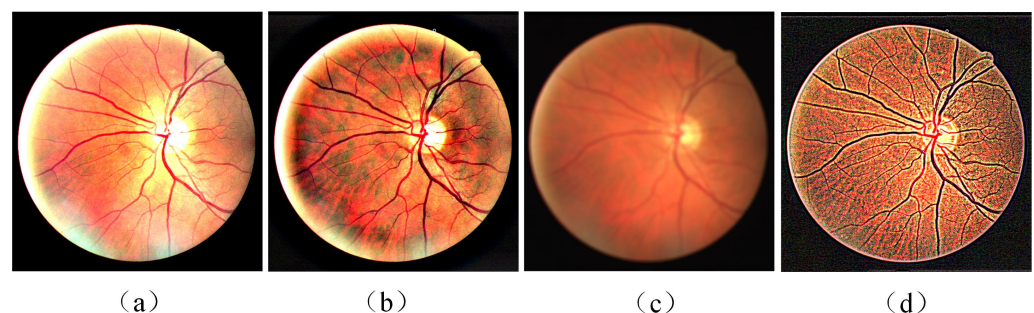


Figure 10. Examples of fundus images before and after preprocessing: (a) example of the overexposure phenomenon; (b) example of n equal to 50; (c) example of α equal to 0.01; (d) example of excessive noise.

The function of α is to linearly adjust the value of $G(x, j)$ to control the enhancement effect of the preprocessing. For example, the effect of α equal to 0.01 is shown in Figure 10c. In addition, since $G(x, j)$ is inversely proportional to the local mean squared deviation, the local mean squared deviation may be small in areas where the image is smooth. If the local mean squared deviation is too small, and the $G(x, j)$ will become large, which will lead to noise amplification or pixel saturation (i.e., pixel value exceeds 255). For example, this phenomenon can be observed easily, as shown in Figure 10d. When the noise is too large, the tiny vascular regions are greatly disturbed. Therefore, the maximum value of $G(x, j)$ is limited to obtain better results. After comparing different preprocessing effects, this paper limits the maximum value of $G(x, j)$ to 5.

As a further detailed supplement, the parameter sensitivity analysis is shown in Figure 11. According to the analysis, the results show a slightly increasing trend when n is less than 6, and a significantly decreasing trend when n is greater than 6. When α is less than 0.9, it has a significant impact on the results, and when α is more than 0.9, it has a more subtle impact on the results. When the maximum setting is less than or equal to 5, for Acc_U has a slight improvement, and other results are slightly affected. When the maximum setting is greater than 5, all results basically show a downward trend.

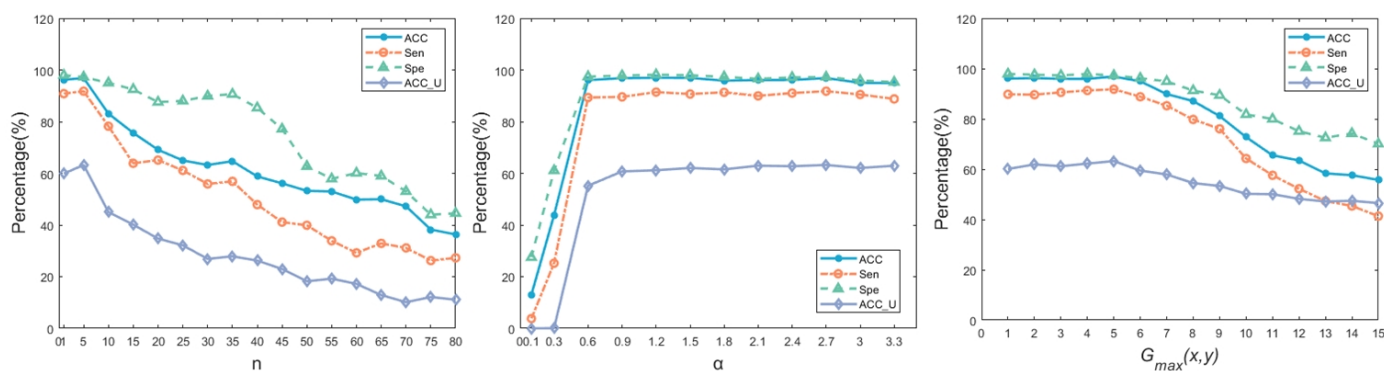


Figure 11. The parameter sensitivity analysis.

5.6. Limitations

In fundus images, the pixel and background of blood vessels are not balanced, which is also a common problem in medical images. In particular, the classification of A-V crossings are involved in this work. Its content and background are seriously unbalanced, and vascular crossing needs further attention by the network. At the same time, the classification accuracy is greatly affected by the preprocessing effect, but the data preprocessing needs to rely on manual adjustment. In the experimental process, we found that when the number of epochs increases to a higher level, the classification performance for arterial vein and vascular crossing rapidly decline. In future work, better preprocessing methods and solutions to imbalance problems can improve the accuracy of classification of vascular crossing and achieve more stable performance.

6. Conclusions

In this paper, a method based on the combination of the GAN and U-Net networks is proposed for vessel classification of fundus images. This method preprocesses the input images by ACE to solve problems such as unclear data of original fundus images. In addition, the confusion matrix established in this study for vessel classification has reference value for comprehensive analysis and evaluation of prediction results. In the design of GAN, the original U-Net network is improved by introducing the residual connection, the null pyramid module, and the attention module. The proposed network structure not only achieves the end-to-end prediction for arterial and venous vessels, but also predicts the arterial–venous crosspoint region. Due to the complexity of the optic

disc region and the small percentage of vascular cross-pixels, there is still much room for improvement in the accuracy of vascular cross-point recognition. Compared with the existing end-to-end methods, this study has improved the accuracy of vessel classification, which is important for retinal vessel classification to 96.86%. Experimental results show that the proposed method can effectively realize the automatic classification of arteriovenous and A-V crossings. This can be applied to the basic stage of screening and diagnosis of retinal vascular related diseases to improve the diagnostic efficiency of doctors, which has important significance for clinical practice. The proposed model design can also provide important reference for similar semantic segmentation tasks. In addition, the distribution of retinal vessels, arteriovenous intersection, and other characteristics have individual uniqueness, which have potential application value in biological recognition and other fields. In addition, some reliable future research can be continued based on the advantages and limitations of our research methods. For example, research directions to improve the accuracy of A-V crossings are needed. In addition, the fundus image enhancement method based on depth learning can be tried to solve the disadvantages of data optical imaging and improve the classification accuracy. Automated measurement of arteriovenous caliber ratios can be achieved based on arteriovenous classification, which has important implications for the diagnosis of diabetes.

Author Contributions: Methodology, J.Z. and K.Y.; Validation, K.Y., J.Z. and Z.Y.; Writing, J.Z. and K.Y.; Supervision, Z.S., R.H. and S.S.; Visualization, Q.Z. and M.C.; Funding acquisition, Z.Y., S.S., Z.S. and K.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 62003232, No. 52275568, and No. 62031022), the Shanxi Provincial Key Medical Scientific Research Project (2020XM06), the Shanxi Provincial Basic Research Project (No. 202103021221006), and the Natural Science Foundation of Yunnan Province (Grant No. 202201AY070001-287).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. *World Report on Vision*; World Health Organization: Geneva, Switzerland, 2019.
2. Quigley, H.A.; Broman, A.T. The number of people with glaucoma worldwide in 2010 and 2020. *Br. J. Ophthalmol.* **2006**, *90*, 262–267. [[CrossRef](#)] [[PubMed](#)]
3. Wong, W.L.; Su, X.; Li, X.; Cheung, C.M.G.; Klein, R.; Cheng, C.Y.; Wong, T.Y. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis. *Lancet Glob. Health* **2014**, *2*, e106–e116. [[CrossRef](#)]
4. Guariguata, L.; Whiting, D.R.; Hambleton, I.; Beagley, J.; Linnenkamp, U.; Shaw, J.E. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res. Clin. Pract.* **2014**, *103*, 137–149. [[CrossRef](#)] [[PubMed](#)]
5. Dong, L.; Liang, W. Clinical observation of fundus photography screening for primary glaucoma in community patients with diabetes mellitus type 2. *China Pract. Med.* **2020**, *15*, 2. (In Chinese)
6. David, S.A.; Mahesh, C.; Kumar, V.D.; Polat, K.; Alhudhaif, A.; Nour, M. Retinal Blood Vessels and Optic Disc Segmentation Using U-Net. *Math. Probl. Eng.* **2022**, *2022*, 8030954. [[CrossRef](#)]
7. Madala, B.S.; Adabifirouzjaei, F.; Lando, L.; Yarmohammadi, A.; Long, C.P.; Mas, C.; Goldbaum, M.H.; Sarraf, D.; Demaria, A.N.; Envelope, M.F.; et al. Retinal Ischemic Perivascular Lesions, a Biomarker of Cardiovascular Disease. *Ophthalmol. Retin.* **2022**, *6*, 865–867. [[CrossRef](#)]
8. Zhao, X. Comparison of different methods in early screening for diabetic retinopathy. *China Contin. Med. Educ.* **2020**, *12*, 152–154. (In Chinese)
9. Cheung, C.Y.; Mok, V.; Foster, P.J.; Trucco, E.; Wong, T.Y. Retinal imaging in Alzheimer’s disease. *J. Neurol. Neurosurg. Psychiatry* **2021**, *12*, 459–509. [[CrossRef](#)]
10. Wang, N.; Liang, C. Relationship of Gensini score with retinal vessel diameter and arteriovenous ratio in senile CHD. *Open Life Sci.* **2021**, *16*, 737–745. [[CrossRef](#)]
11. Rhee, E.J.; Chung, P.W.; Wong, T.Y.; Song, S.J. Relationship of retinal vascular caliber variation with intracranial arterial stenosis. *Microvasc. Res.* **2016**, *108*, 64–68. [[CrossRef](#)]
12. Ciecholewski, M.; Kassjański, M. Computational Methods for Liver Vessel Segmentation in Medical Imaging: A Review. *Sensors* **2021**, *21*, 2027. [[CrossRef](#)]

13. Cao, X.; Xue, L.; Lin, J.; Yu, L. Retinal vessel segmentation method based on matched filtering and automatic threshold. *J. Electron. Meas. Instrum.* **2017**, *1*, 51–57.
14. Dash, S.; Verma, S.; Kavita; Jhanjhi, N.Z.; Masud, M.; Baz, M. Curvelet Transform Based on Edge Preserving Filter for Retinal Blood Vessel Segmentation. *Comput. Mater. Contin.* **2022**, *71*, 2459–2476. [\[CrossRef\]](#)
15. Palanivel, D.A.; Natarajan, S.; Gopalakrishnan, S. Retinal vessel segmentation using multifractal characterization. *Appl. Soft Comput.* **2020**, *94*, 106439. [\[CrossRef\]](#)
16. Chen, D.; Yang, W.; Wang, L.; Tan, S.; Lin, J.; Bu, W. PCAT-UNet: UNet-like network fused convolution and transformer for retinal vessel segmentation. *PLoS ONE* **2022**, *17*, e0262689. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Mookiah, M.R.K.; Hogg, S.; MacGillivray, T.J.; Prathiba, V.; Pradeepa, R.; Mohan, V.; Anjana, R.M.; Doney, A.S.; Palmer, C.N.; Trucco, E. A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification. *Med. Image Anal.* **2021**, *68*, 101905. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
19. Chen, C.; Chuah, J.H.; Ali, R.; Wang, Y. Retinal vessel segmentation using deep learning: A review. *IEEE Access* **2021**, *9*, 111985–112004. [\[CrossRef\]](#)
20. You, A.; Kim, J.K.; Ryu, I.H.; Yoo, T.K. Application of generative adversarial networks (GAN) for ophthalmology image domains: A survey. *Eye Vis.* **2022**, *9*, 1–19. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Moccia, S.; De Momi, E.; El Hadji, S.; Mattos, L.S. Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics. *Comput. Methods Programs Biomed.* **2018**, *158*, 71–91. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Abdulsahib, A.A.; Mahmoud, M.A.; Mohammed, M.A.; Rasheed, H.H.; Mostafa, S.A.; Maashi, M.S. Comprehensive review of retinal blood vessel segmentation and classification techniques: Intelligent solutions for green computing in medical images, current challenges, open issues, and knowledge gaps in fundus medical images. *Netw. Model. Anal. Health Inform. Bioinform.* **2021**, *10*, 1–32. [\[CrossRef\]](#)
23. Sathananthavathi, V.; Indumathi, G. BAT optimization based Retinal artery vein classification. *Soft Comput.* **2021**, *25*, 2821–2835. [\[CrossRef\]](#)
24. Srinidhi, C.L.; Aparna, P.; Rajan, J. Automated method for retinal artery/vein separation via graph search metaheuristic approach. *IEEE Trans. Image Process.* **2019**, *28*, 2705–2718. [\[CrossRef\]](#)
25. Dashtbozorg, B.; Mendonca, A.M.; Campilho, A. An Automatic Graph-Based Approach for Artery/Vein Classification in Retinal Images. *IEEE Trans. Image Process.* **2014**, *23*, 1073–1083. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Vázquez, S.; Cancela, B.; Barreira, N.; Penedo, M.G.; Rodríguez-Blanco, M.; Pena Seijo, M.; de Tuero, G.; Barceló, M.A.; Saez, M. Improving retinal artery and vein classification by means of a minimal path approach. *Mach. Vis. Appl.* **2013**, *24*, 919–930. [\[CrossRef\]](#)
27. Welikala, R.A.; Foster, P.J.; Whincup, P.H.; Rudnicka, A.R.; Owen, C.G.; Strachan, D.P.; Barman, S.A. Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort. *Comput. Biol. Med.* **2017**, *90*, 23–32. [\[CrossRef\]](#)
28. Zhao, Y.; Xie, J.; Zhang, H.; Zheng, Y.; Zhao, Y.; Qi, H.; Zhao, Y.; Su, P.; Liu, J.; Liu, Y. Retinal Vascular Network Topology Reconstruction and Artery/Vein Classification via Dominant Set Clustering. *IEEE Trans. Med. Imaging* **2020**, *39*, 341–356. [\[CrossRef\]](#)
29. Li, P.; Deng, Q.; Li, H. The Arteriovenous Classification in Retinal Images by U-net and Tracking Algorithm. In Proceedings of the 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), Beijing, China, 10–12 July 2020; pp. 182–187.
30. Binh, N.T.; Hien, N.M.; Tin, D.T. Improving U-Net architecture and graph cuts optimization to classify arterioles and venules in retina fundus images. *J. Intell. Fuzzy Syst.* **2022**, *42*, 4015–4026. [\[CrossRef\]](#)
31. Morano, J.; Hervella, Á.S.; Novo, J.; Rouco, J. Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images. *Artif. Intell. Med.* **2021**, *118*, 102116. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Galdran, A.; Meyer, M.; Costa, P.; Mendonça, Campilho, A. Uncertainty-aware artery/vein classification on retinal images. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 556–560.
33. Hu, Q.; Abramoff, M.D.; Garvin, M.K. Automated construction of arterial and venous trees in retinal images. *J. Med. Imaging* **2015**, *2*, 044001. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Staal, J.; Abramoff, M.D.; Niemeijer, M.; Viergever, M.A.; Van Ginneken, B. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* **2004**, *23*, 501–509. [\[CrossRef\]](#)
35. Narendra, P.M.; Fitch, R.C. Real-time adaptive contrast enhancement. *IEEE Trans. Pattern Anal. Mach. Intell.* **1981**, *PAMI-3*, 655–661. [\[CrossRef\]](#)
36. Imageshop. Available online: <https://www.cnblogs.com/Imageshop/p/3324282.html> (accessed on 23 December 2021).
37. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the International Conference on Machine Learning, PMLR, Sydney, Australia, 6–11 August 2017; pp. 214–223.

39. Mou, L.; Zhao, Y.; Chen, L.; Cheng, J.; Gu, Z.; Hao, H.; Qi, H.; Zheng, Y.; Frangi, A.; Liu, J. CS-Net: Channel and spatial attention network for curvilinear structure segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 721–730.
40. Xu, X.; Wang, R.; Lv, P.; Gao, B.; Li, C.; Tian, Z.; Tan, T.; Xu, F. Simultaneous arteriole and venule segmentation with domain-specific loss function on a new public database. *Biomed. Opt. Express* **2018**, *9*, 3153–3166. [[CrossRef](#)]
41. Kang, H.; Gao, Y.; Guo, S.; Xu, X.; Li, T.; Wang, K. AVNet: A retinal artery/vein classification network with category-attention weighted fusion. *Comput. Methods Programs Biomed.* **2020**, *195*, 105629. [[CrossRef](#)] [[PubMed](#)]
42. Ma, W.; Yu, S.; Ma, K.; Wang, J.; Ding, X.; Zheng, Y. Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 769–778.
43. Estrada, R.; Allingham, M.J.; Mettu, P.S.; Cousins, S.W.; Tomasi, C.; Farsiu, S. Retinal artery-vein classification via topology estimation. *IEEE Trans. Med Imaging* **2015**, *34*, 2518–2534. [[CrossRef](#)]
44. Xu, X.; Ding, W.; Abramoff, M.D.; Cao, R. An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image. *Comput. Methods Programs Biomed.* **2017**, *141*, 3–9. [[CrossRef](#)] [[PubMed](#)]
45. Huang, F.; Dashtbozorg, B.; Tan, T.; ter Haar Romeny, B.M. Retinal artery/vein classification using genetic-search feature selection. *Comput. Methods Programs Biomed.* **2018**, *161*, 197–207. [[CrossRef](#)] [[PubMed](#)]
46. Ilakkiya, M.; Devi, S.S. Detection of Arteriovenous Nicking in Retinal Fundus Images for Screening Hypertensive Retinopathy. 2020. Available online: https://www.researchgate.net/publication/343263226_Detection_of_Arteriovenous_Nicking_in_Retinal_Fundus_Images_For_Screening_Hypertensive_Retinopathy (accessed on 12 March 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.