Mathematical Biosciences and Engineering

MBE, 19(10): 9948–9965.
DOI: 10.3934/mbe.2022464
Received: 18 March 2022
Revised: 23 June 2022
Accepted: 27 June 2022
Published: 12 July 2022

http://www.aimspress.com/journal/MBE

Research article

SRV-GAN: A generative adversarial network for segmenting retinal vessels

Chen Yue[1,2], Mingquan Ye[1,2,*], Peipei Wang[1,2], Daobin Huang[1,2] and Xiaojie Lu[1,2]

[1] School of Medical Information, Wannan Medical College, Wuhu 241002, China
[2] Research Center of Health Big Data Mining and Applications, Wannan Medical College, Wuhu 241002, China

* **Correspondence:** Email: ymq@wnmc.edu.cn; Tel: +8605533932972.

Abstract: In the field of ophthalmology, retinal diseases are often accompanied by complications, and effective segmentation of retinal blood vessels is an important condition for judging retinal diseases. Therefore, this paper proposes a segmentation model for retinal blood vessel segmentation. Generative adversarial networks (GANs) have been used for image semantic segmentation and show good performance. So, this paper proposes an improved GAN. Based on R2U-Net, the generator adds an attention mechanism, channel and spatial attention, which can reduce the loss of information and extract more effective features. We use dense connection modules in the discriminator. The dense connection module has the characteristics of alleviating gradient disappearance and realizing feature reuse. After a certain amount of iterative training, the generated prediction map and label map can be distinguished. Based on the loss function in the traditional GAN, we introduce the mean squared error. By using this loss, we ensure that the synthetic images contain more realistic blood vessel structures. The values of area under the curve (AUC) in the retinal blood vessel pixel segmentation of the three public data sets DRIVE, CHASE-DB1 and STARE of the proposed method are 0.9869, 0.9894 and 0.9885, respectively. The indicators of this experiment have improved compared to previous methods.

Keywords: deep learning; retinal image segmentation; generative adversarial networks; attention; loss functions

1. Introduction

Retinal blood vessels are continuous and have dendritic structures. The branches start from the optic disc, and the width of the blood vessel decreases as it moves away from the optic disc. At the same time, the optic disc is the confluence of the main blood vessels. Vascular caliber is important for assessing cardiovascular disease risk [1]. The diameter, size and morphology of retinal blood vessels are closely related to diabetes, macular disease, glaucoma and other diseases, and there will be hard exudate and other pathological features in the diseased retina. Therefore, the challenges faced by retinal vessel segmentation technology include low capillary and background contrast, mis-segmentation of optic disc boundaries and interference from pathological spots. In the past, doctors determined the morphology of retinal blood vessels by manual segmentation, but this method is time-consuming and laborious, and the efficiency is very low. Therefore, it is of great significance to segment the morphology of retinal blood vessels by computer vision. Over the years, many scholars worldwide have studied automatic retinal segmentation algorithms. At present, there are supervised and unsupervised methods to segment retinal vessels according to whether labeled or unlabeled data are required. Common unsupervised retinal vessel segmentation methods include conventional matched filtering, image morphology processing, vessel tracking, threshold segmentation, region growth, active contour-based methods and graph-based methods.

Based on the centerline of blood vessels, Mendoca et al. [2] adopted an iterative region growing method in the segmentation, which combined images generated by morphological filters and achieved good results. Based on probability tracking, Yin et al. [3] used a Bayesian method to detect blood vessel edge points and achieved good segmentation accuracy on three publicly available retinal datasets. Ye et al. [4] proposed a three-dimensional multi-scale enhancement filter. This method uses three-dimensional Hessian matrix eigenvalues, which can improve the saliency of tiny blood vessels and increase the speed of calculation. Lazar et al. [5] proposed a new region growth method that defines the pixel response as a vector, and the nearest neighbor classifier is used to filter the seed points. In order to overcome the false response of the optic disc boundary, a symmetrically constrained multi-scale filtering technique was also proposed. Neto et al. [6] proposed a coarse-to-fine retinal vessel segmentation method, which uses spatial correlation, probability and statistics data, curvature analysis, morphological reconstruction and adaptive local thresholds to improve segmentation accuracy on multiple datasets. A method proposed by Nguyen et al. [7] is to obtain line detectors of different scales by changing the length of a basic line detector and linearly combining the line responses of different scales. This method is efficient and scalable.

The supervised method is to first extract the features of retinal blood vessels, then train the classifier with manually labeled images and finally use the trained classifier to segment retinal blood vessels. Feature extraction methods include discrete wavelet change, Gaussian filtering, vascular filtering, etc., and the classifier usually adopts a support vector machine, artificial neural networks, or the k-nearest neighbor algorithm. Staal et al. [8] proposed a supervised model based on ridge lines for automatic segmentation of retinal vessels. The method uses the sequential forward selection algorithm to get the best eigenvalue of the pixel on the ridge, and it then uses the k-nearest neighbor algorithm to classify each pixel. Ricci et al. [9] proposed a retinal vascular segmentation method combining line operations with a support vector machine algorithm. In this method, two orthogonal detectors are combined with the gray values of pixels to extract feature images, and then the support vector machine algorithm is used to complete the classification of pixels. Wilfred et al. [10] used the structure of an

artificial neural network with multiple hidden layers to segment the retinal vascular structure and experiments show that the accuracy of this method on the DRIVE data set is good.

2. Related works

In recent years, the deep learning method has been widely applied in various fields. Due to its good segmentation effects and high computational efficiency, more and more researchers have used the deep learning method to segment retinal vessels. Scholars have proposed AlexNet [11], VGG [12], GoogLeNet [13], Residual Net [14], DenseNet [15] and other models. In terms of semantic segmentation, the fully convolutional neural network proposed by Long et al. has achieved better performance than other convolutional neural networks. Since 2015, the U-Net model proposed by Ronneberger et al. [16] is a kind of fully convolutional neural network. The U-Net model is composed of an encoder and decoder as well as skip connections, and it has the abilities of pixel location and feature extraction. Based on the U-Net model, many derivative methods have been produced. Zhou et al. [17] proposed UNet++ for lung nodule segmentation, colon polyp segmentation, cell nucleus segmentation and liver segmentation. The core idea is to use dense connection modules in the skip connection part to fuse the semantic gap. Jin et al. [18] introduced deformable convolution in retinal image segmentation. The model combines the advantages of deformable units and U-shaped networks. Oktay et al. [19] proposed attention U-Net. Its core idea is to introduce an additive attention gate (AG) at the skip connection, which can suppress the feature response of irrelevant background areas. Ding et al. [20] proposed a multi-channel neural network that can accurately segment the end of blood vessels and achieved good results on retinal datasets. Sun et al. [21] proposed two new data enhancement modules, channel random gamma correction and channel random blood vessel enhancement, so that the model can recognize more features globally and locally. The dense connection network proposed by Li et al. [22] extracted retinal vascular information through dense connection blocks, which could alleviate the gradient disappearance in the feature extraction process. Alom et al. [23] proposed a cyclic residual network based on the U-Net model, in which residual connections can alleviate gradient disappearance and train deep network information. Combined with a recurrent neural network (RCNN), it can accumulate features and achieve high performance in segmentation tasks, and it performed well on retinal data sets.

To produce a composite image that more closely matches the real data, Goodfellow et al. [27] proposed GAN, which generally consists of a generator and discriminator. These two models compete against each other in the process of training and learning and finally reach a game equilibrium. The purpose of a discriminator is to determine whether the input data comes from real data or from a generator. The purpose of a generator is to learn the characteristics of a sample to produce data that confuses the judgment of the discriminator. In the traditional GAN, the input of the generator is a random noise signal. Through the training, finally the generator can output a high quality sample; however, because the random noise of input is uncontrollable, the sample type obtained by the generator is difficult to control. This is not suitable for an accurate-to-pixel task. Therefore, the conditional GAN proposed by Mirza and Osindro [28] adds conditional information to both the generator and the discriminator to guide the training of the model, realizing the controllable generated content. FCN and U-Net are mature models, which are widely used in various computer vision tasks with remarkable effects. However, it can be found that the detail features of images are often ignored in previous studies, while GAN can improve the image synthesis performance of the convolution

model [29]. Radford et al. [30] proposed a deep convolutional generative adversarial network. By combining a deep convolutional network with GAN, GAN can be more stable in the training process and accelerate its training, which performs well in various fields of medical image processing. Since then, deep convolutional GANs have been widely used. For example, Pix2Pix proposed by Isola et al. [31] is a general framework for image translation based on conditional GAN, which realizes the generalization of model structure and loss function, and it has achieved remarkable results on many image translation data sets. Isola et al. designed a generator structure similar to U-Net and a convolutional discriminator structure, PatchGAN, which inputs local image blocks into the discriminator and achieves superior performance on various data sets. The convolutional GAN proposed by Yang et al. [32], which combines short and dense connections, can detect more tiny vessels and is superior to many methods previously proposed in terms of sensitivity and specificity. The GAN proposed by Son et al. [33] was used in retinal vascular segmentation and optic disc segmentation. The results showed that the indexes in the vascular segmentation task were significantly improved, while optic disc segmentation was not. In the GAN proposed by Dong et al. [34], U-Net was used as the generator, and a fully convolutional neural network was used as the discriminator to perform segmentation experiments on multiple thoracic organs, which proved the reliability and feasibility of GAN for medical images. Zhang et al. [35] proposed an improved dense GAN combined with U-Net and proposed a multi-layer attention mechanism for lung CT image segmentation, which improved the segmentation accuracy compared with other methods.

Although there are extensive studies on GAN in medical image processing in the field of radiology, there are few applications of GAN in retinal image processing in the field of ophthalmology. In previous studies on retinal vascular segmentation, it can be found that conditional GAN performs better than U-Net and other convolutional models [33]. GAN can deliver important performance in the absence of large tagged datasets and data shortages [36,37]. Therefore, this paper proposes a conditional GAN model based on deep convolution for retinal vascular image segmentation, in which a controllable variable is used as an additional input of the generator and discriminator, so as to control the output types of the generator. In our network, this controllable variable is the original fundus image. This setting ensures that the samples generated by the generator in our network are controllable, and the input image pair of the discriminator ensures the mapping between the original fundus image and the vascular segmentation image. Figure 1 shows the overall network structure, where x represents the original fundus image, G represents the generator, G(x) is the segmentation image generated by the generator, D represents the discriminator, and y is the manually labeled image. In the network we designed, the input of the generator is the original fundus image, and the output is the probability map of the same size as the fundus image we input. Obviously, the value range on the probability graph is 0–1, where the value corresponding to each pixel point represents the probability value of a blood vessel. The input of the discriminator is an image pair, namely, the original image and vascular diagram. The task of the discriminator is to distinguish whether the vascular diagram in the image pair is artificially annotated or generated by the generator.

The main work of this paper includes the following: This generator combines a residual unit and cyclic unit, and it uses the R2U-Net model, which can accumulate characteristic information and alleviate gradient disappearance. In the last layer of convolution, spatial attention and channel attention are used to extract global information features and reduce the interference of redundant information. In previous studies, densely connected networks have realized high accuracy in the classification task, so we use dense connection modules in the discriminator. They can alleviate the gradient dissipation

problem in the process of training, and due to the large number of features reused, a large number of features can be generated using a small amount of convolution kernels. According to previous experience, we usually mix the loss function of the GAN with the traditional loss function to get good results. However, we cannot guarantee that the generated blood vessel map and the original fundus image correspond pixel by pixel, that is, the generated blood vessel map is not very close to our labeled results, so we also need a loss function that makes the results of the generator correspond to the labeled map. Kamran et al. [38] proposed using the mean square error loss function to generate a probability segmentation graph that is closer to ground truth through RV-GAN. Therefore, we adopted this loss in this experiment.

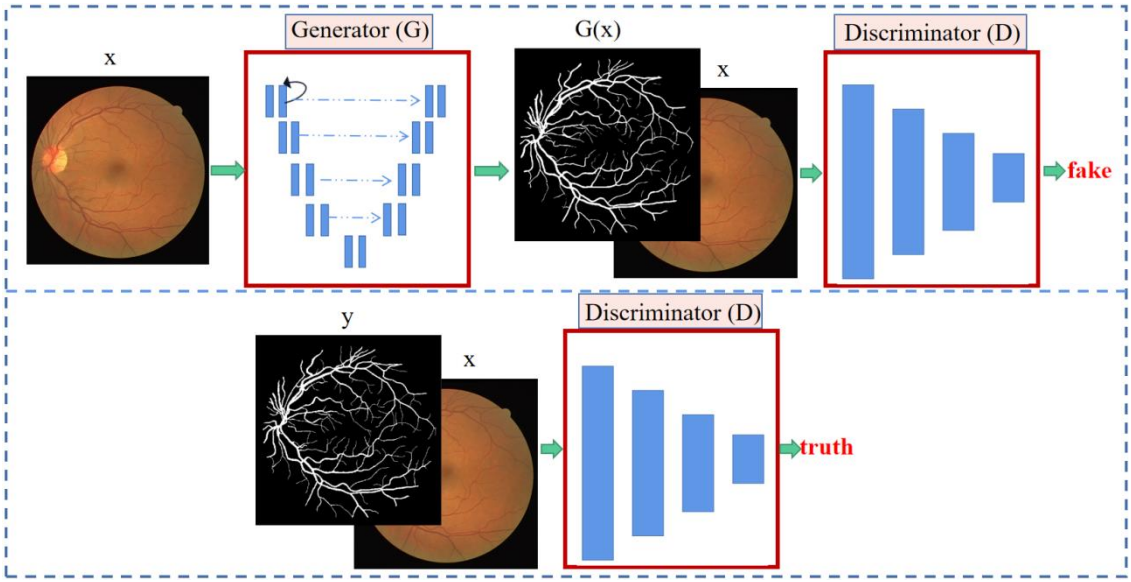


Figure 1. The overall structure of SRV-GAN.

3. Methods

3.1. Generator

RCNN and its variants have shown superior performance in target recognition tasks using different benchmarks [39,40]. Alom et al. proposed R2U-Net and applied it to medical image segmentation, which has excellent segmentation performance in a variety of data sets. Therefore, we take R2U-Net as the basic model of our generator. As shown in Figure 2, R2U-Net uses recurrent residual blocks instead of the traditional conv+relu layer in the encoding and decoding process, which can effectively increase the network depth; using feature summation at different time steps to obtain more expressive features helps to extract lower levels. In skip connections, instead of cutting in the original U-Net, a cascading operation is used. The loop structure deepens the network level, and the residual structure avoids the problem of gradient disappearance as the depth increases. The advantages of U-Net, residual network and RCNN are combined.

In previous studies, we can see that attention mechanism has been widely used in image segmentation tasks. Liu et al. [24] proposed a residual network model fused with an attention mechanism, which highlights shallow details in channel and spatial dimensions through a reverse

attention mechanism, thereby effectively fusing deep local features and shallow global information, and high segmentation indexes were obtained on three retinal data sets. The multi-scale fusion network proposed by Yang et al. [25] introduces both channel attention and spatial attention, which adapts the weights through the channel attention module to improve the segmentation performance, and it captures long-range feature dependencies through the position attention module. It is superior to other methods on publicly available retinal data sets. Based on R2U-Net, this generator introduces a channel attention module and position attention module to reduce the redundancy of information in the physical signs and extract more effective information. As shown in Figures 3 and 4, we adopt the dual-attention module proposed by Fu et al. [26] to establish the interdependence relationship in the channel and spatial dimensions, respectively.

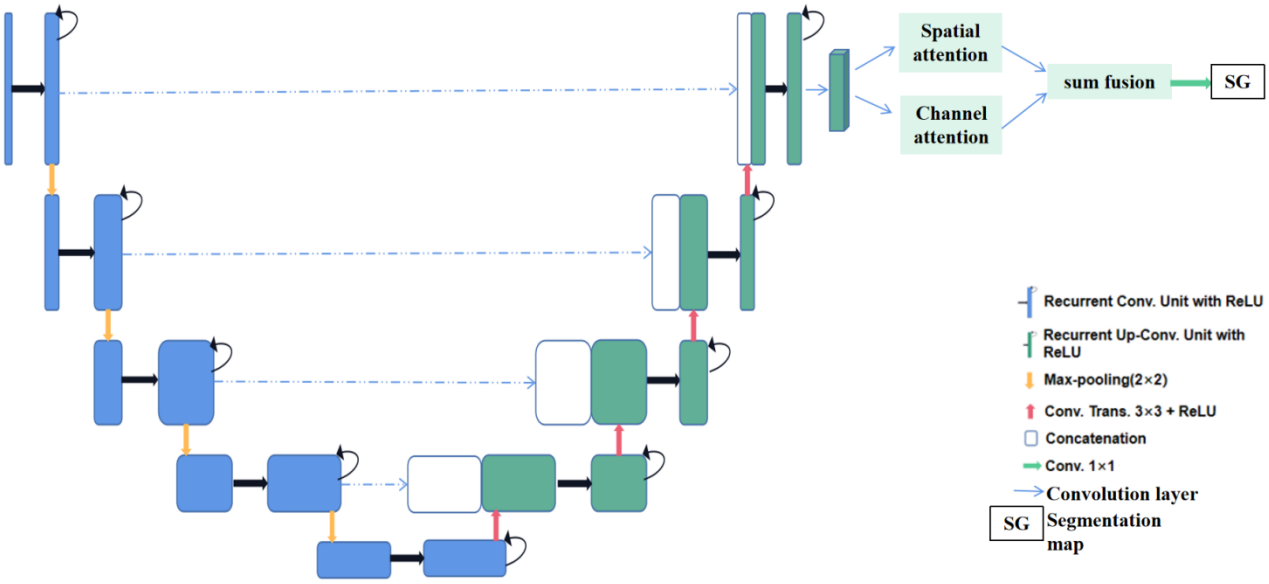


Figure 2. The structure of the generator.

Feature A $\in$ $R^{C\times H\times W}$, and after feature A is subjected to the convolution operation, features B, C, D are obtained, {B, C} $\in$ $R^{C\times H\times W}$. B is subjected to reshape and transpose operations, and C is subjected to reshape operation. N = H × W, where N refers to pixels, so $R^{C\times N}$ is obtained. The new features obtained by B and C are matrix multiplied and then go through softmax to obtain S, S $\in$ $R^{N\times N}$, as shown in formula (1).

$$s_{ji}=\frac{\exp\left(B_i\cdot C_j\right)}{\sum_{i=1}^{N}\exp\left(B_i\cdot C_j\right)} \tag{1}$$

Feature D $\in$ $R^{C\times H\times W}$. Reshape feature D to get $R^{C\times N}$, and then matrix multiply with S. Then, reshape them into $R^{C\times H\times W}$, and finally element-wise sum it with feature A to get the final feature E, which is multiplied by the scale parameter α. α learns weight from 0, E $\in$ $R^{C\times H\times W}$, as shown in formula (2).

$$E_j=\alpha\sum_{i=1}^{N}\left(s_{ji}D_i\right)+A_j \tag{2}$$

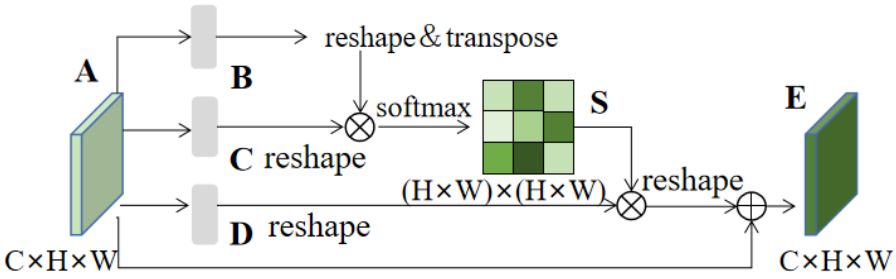


Figure 3. Position attention module.

First, the reshaped feature A and the reshaped and transpose A are matrix multiplied and then go through softmax to obtain X, $X \in R^{C \times C}$. Then, X and the reshaped A are matrix multiplied, and get $R^{C \times H \times W}$ after reshaping. Then, element-wise sum it with feature A to get the final feature E, which is multiplied by scale parameter β. β learns weight from 0, $E \in R^{C \times H \times W}$, as shown in formulas (3) and (4).

$$x_{ji} = \frac{\exp\left(A_i \cdot A_j\right)}{\sum_{i=1}^{C} \exp\left(A_i \cdot A_j\right)} \tag{3}$$

$$E_j = \beta \sum_{i=1}^{C} \left(x_{ji} A_i\right) + A_j \tag{4}$$

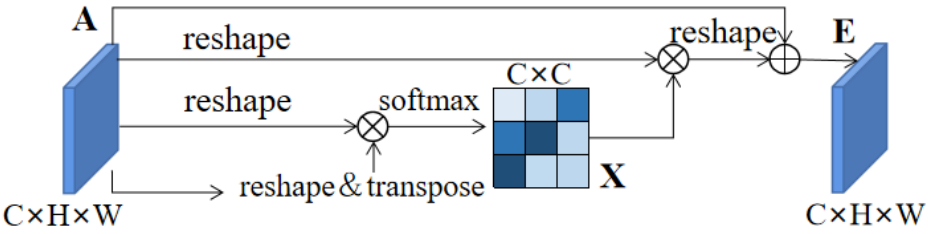


Figure 4. Channel attention module.

3.2. Discriminator

If the discriminator is slow to respond, the resulting images will converge, and the patterns will start to collapse. Conversely, when the discriminator performs well, the gradient of the generator's loss function vanishes, and learning is slow. Earlier, we mentioned that R2U-Net has a high segmentation accuracy in the retina datasets. Therefore, we should explore an excellent discriminator, compared to the deep residual network (ResNet), to achieve the same accuracy. The number of parameters to be learned by DenseNet is much lower than that of the ResNet, so the learning efficiency is higher. Compared with the ordinary convolutional neural network, the special structure of DenseNet can not only reduce the gradient disappearance problem faced by the deep network but also strengthen the "understanding ability" of the network due to the repeated use of feature graphs. Therefore, this paper

adopts the dense connection module in the discriminator. As shown in Figure 5, the discriminator is composed of two convolutional layers and four densely connected blocks. The structure of the convolutional layer is Conv3 × 3-BN-ReLU. The small convolution kernel can ensure the visual perception domain and reduce the parameters, and densely connected modules are composed of BN-ReLU-Conv3 × 3. They improve the back propagation of the gradient, make the network easier to train and can achieve feature reuse, improve efficiency, reduce the amount of parameters and calculation costs and achieve superior performance. Through a series of operations, the sample features can be extracted, and finally the generated samples and the output real samples can be judged by sigmoid, so as to distinguish the ground truth and the generated retinal blood vessel segmentation images.

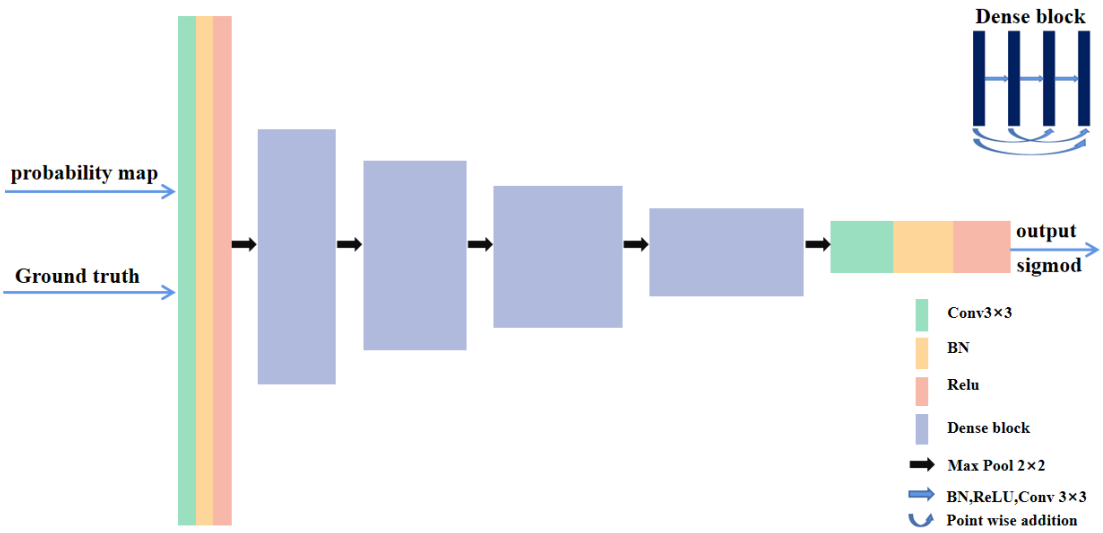


Figure 5. The structure of the discriminator.

On the basis of the residual network, Huang et al. proposed a densely connected network. The densely connected network is similar to the residual network. The input of the latter layer is related to the previous layer. Unlike the residual network, the input of each layer is the output of all previous layers, and the output of each layer is also the input of all subsequent layers. Because of this special structure, the densely connected network has the advantage of improving the effect when the number of parameters is reduced, and the densely connected network effectively alleviates the phenomenon of gradient disappearance and reduces the loss of feature information. Its specific formula is shown in formula (5).

$$X_i = H_i\left(\left[X_0, X_{i-1}\right]\right) \tag{5}$$

In the formula, X_0, X_1, X_2...X_{i-1} refers to the feature maps fused by layer 0, layer 1, layer 2, layer i-1, and H_i is a composite function composed of BN-ReLU-Conv3 × 3.

3.3. Objective function

The loss function that removes the noise vector in the original GAN is Eq (6).

$$L(G,D) = E_{x,y\sim P_{data(x,y)}}[\log D(x,y)] + E_{x\sim P_{data(x)}}[\log(1 - D(x,G(x)))] \quad (6)$$

$$L_{rec}(G) = E_{x,y}||G(X) - y||^2 \quad (7)$$

x represents the input image, and y represents ground truth. However, the above formula is not our objective function. Based on the above formula, we introduce the reconstruction loss (mean squared error) in the generator to confirm the difference between the output images of the generated network and the real images, as shown in Eq (7). By using this loss, we ensure that the synthesized images contain more realistic blood vessels.

Combining the loss function of formula (6) with this function, the loss functions of the generator and discriminator are obtained as formulas (8) and (9), respectively.

$$G^* = \arg\min_G L_{gen}(G,D) = \arg\min_G \max_D L(G,D) + \lambda L_{rec}(G) \quad (8)$$

$$D^* = \arg\min_D L_{dis}(G,D) = \arg\min_G \max_D L(G,D) \quad (9)$$

In the formula, $L_{gen}(G, D)$ is the generator loss, $L_{dis}(G, D)$ is the discriminator loss, λ is a hyperparameter and is set to 10, and G and D are two opposing training processes. First, fix G to train D, and then fix D to train G, and so on. In the end, the capabilities of both sub-networks can be improved. When the sample image generated by G is judged by D to be an artificially annotated image, the training ends. At this time, after the newly input image passes through G, the image generated by G can be used as the correct segmentation image.

4. Results and discussion

To complete the implementation, we used the Pytorch and Tensorflow frameworks on a single GPU machine with 16GB of RAM and an NIVIDIA GEFORCE GTX-1650 SUPER. We tested them on three retinal image datasets: DRIVE [8], CHASE_DB1 [41] and STARE [42]. The optimizer uses the Adam optimizer with learning rate $\alpha = 0.0002$, $\beta_1 = 0.5$, $\beta_2 = 0.999$, batch_size = 24, and the number of iterations of the three data sets is 100. It took 24–48 hours to train.

4.1. Database and preprocessing methods

Table 1 shows the pixel sizes of the three data sets and the numbers of training and test images.

Table 1. Database.

	DRIVE	STARE	CHASE_DB1
Pixel size	565 × 584	700 × 605	999 × 960
Training\test	20\20	15\5	21\7

Preprocessing: In the fundus image, because the green channel has high contrast and clear blood vessels, the green channel of the fundus retinal image is selected for processing in the experiment.

Then, the image is preprocessed by histogram equalization, normalization and gamma transformation. Considering that the number of training sets is too small, the model is prone to overfitting, resulting in poor classification performance. Therefore, data expansion is required. The expansion methods used in this paper are rotation, mirroring, translation and so on.

4.2. Experimental indicators and chart analysis

In order to evaluate this experiment objectively, we used 5 evaluation indicators for analysis: sensitivity (SE), specificity (SP), accuracy (AC), F1-score and AUC. The formulas are as follows.

$$SE = \frac{P_{TP}}{P_{TP} + P_{FN}} \tag{10}$$

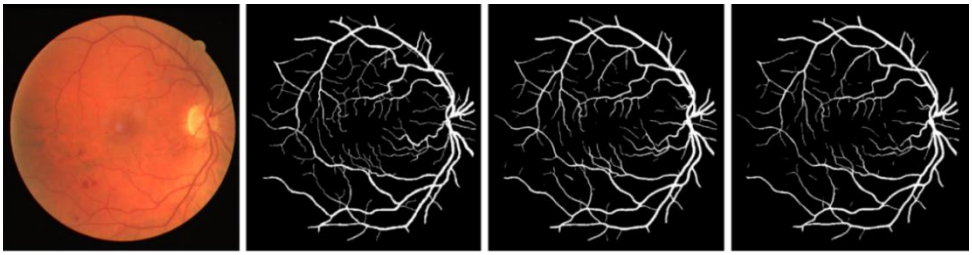
$$SP = \frac{P_{TN}}{P_{TN} + P_{FP}} \tag{11}$$

$$AC = \frac{P_{TP} + P_{TN}}{P_{TP} + P_{FP} + P_{FN} + P_{TN}} \tag{12}$$

$$PR = \frac{P_{TP}}{P_{TP} + P_{FP}} \tag{13}$$

$$F1 = 2 \times \frac{PR \times SE}{PR + SE} \tag{14}$$

In order to explore whether the dual attention mechanism improves the segmentation performance in this experiment, we conducted comparative experiments. Figures 6–8 are the segmentation comparison images of the SRV-GAN model with and without the dual attention mechanism on the DRIVE, CHASE-DB1 and STARE data sets. The first column is the original image, the second column is ground truth, the third column is the segmentation graph with the dual attention mechanism, and the fourth column is the segmentation graph without attention. From the figures, we can see that the segmentation graph with the dual attention mechanism is closer to ground truth and has higher classification accuracy, which indicates that after the addition of the attention mechanism, the segmentation performance of the generator is indeed improved. The attention mechanism can pay attention to the important parts of feature information and suppress the interference of invalid information. Tables 2–4 are the index data in the experiment. It can be seen from the data in the tables that various segmentation metrics of the model have been improved when the attention mechanism is added, especially the F1-score and sensitivity on the STARE dataset, which further proves that adding an attention mechanism is very necessary.

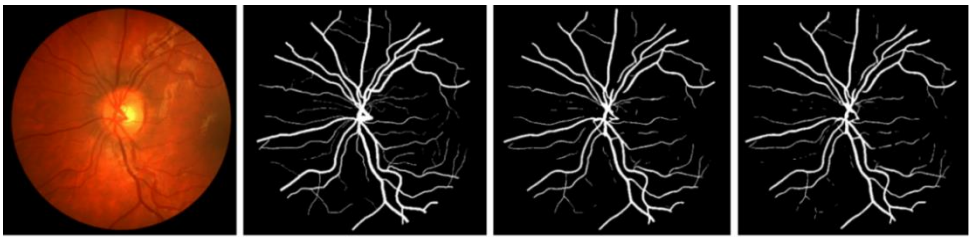


(a) (b) (c) (d)

Figure 6. Retinal vessel segmentation map on DRIVE. (a) Fundus. (b) Ground truth. (c) SRV-GAN. (d) SRV-GAN without dual attention.

Table 2. Indicators of SRV-GAN on DRIVE.

Method	Dual attention	F1	SE	SP	AC	AUC
SRV-GAN	×	0.8322	0.8214	0.9814	0.9669	0.9825
SRV-GAN	√	0.8452	0.8337	0.9850	0.9702	0.9869

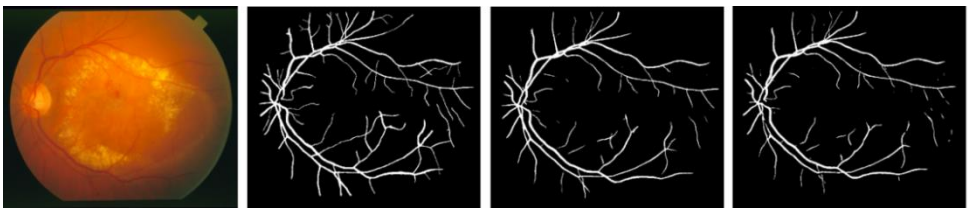


(a) (b) (c) (d)

Figure 7. Retinal vessel segmentation map on CHASE-DB1. (a) Fundus. (b) Ground truth. (c) SRV-GAN. (d) SRV-GAN without dual attention.

Table 3. Indicators of SRV-GAN on CHASE-DB1.

Method	Dual attention	F1	SE	SP	AC	AUC
SRV-GAN	×	0.8059	0.7998	0.9815	0.9652	0.9889
SRV-GAN	√	0.8201	0.8132	0.9837	0.9673	0.9894



(a) (b) (c) (d)

Figure 8. Retinal vessel segmentation map on STARE. (a) Fundus. (b) Ground truth. (c) SRV-GAN. (d) SRV-GAN without dual attention.

Table 4. Indicators of SRV-GAN on STARE.

Method	Dual attention	F1	SE	SP	AC	AUC
SRV-GAN	×	0.7787	0.7938	0.9812	0.9707	0.9881
SRV-GAN	√	0.8102	0.8344	0.9884	0.9712	0.9885

In this article, we applied our proposed model, U-Net, LadderNe and IterNet models to three publicly available retinal data sets, as shown in Figure 9. The first column shows fundus images, the second column shows the segmentation results of U-Net model, the third column shows the segmentation results of LadderNet, the fourth column shows the segmentation results of IterNet, the fifth column shows the segmentation results of SRV-GAN, and the sixth column shows ground truth. According to the experimental results, it can be seen from Figures 10–12 that SRV-GAN has the highest AUC value, indicating that compared with these models, SRV-GAN achieves the best segmentation accuracy on retinal images. In addition, we list the results of various experimental indicators of other models on the same data sets in Table 5, including sensitivity (SE), specificity (SP), accuracy (AC), F1-score (F1) and AUC. It can be seen that our model is superior to U-Net derived architecture and recent models in AUC-ROC of DRIVE, SA-UNet [43] performs better in CHASE-DB1, and R2U-Net performs best in the STARE data set. However, the whole, most of the indicators of the method proposed in this paper are not inferior to or even better than the original methods. Sensitivity and AUC-ROC are representative of segmentation performance, so we have to work harder to improve these two indicators.

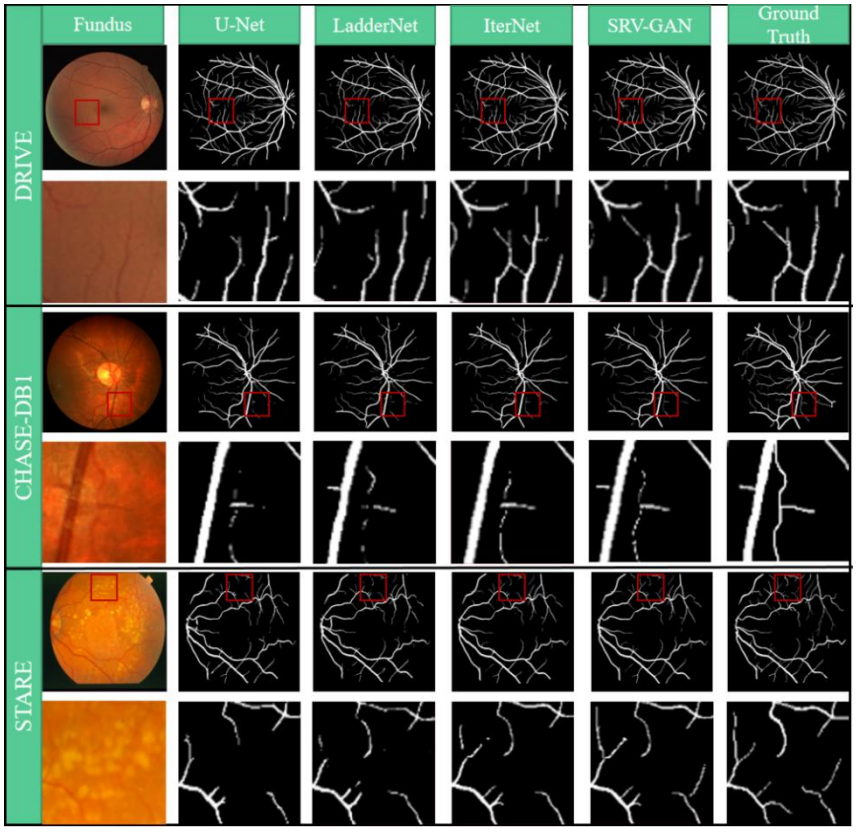


Figure 9. Segmentation results of SRV-GAN and other models.

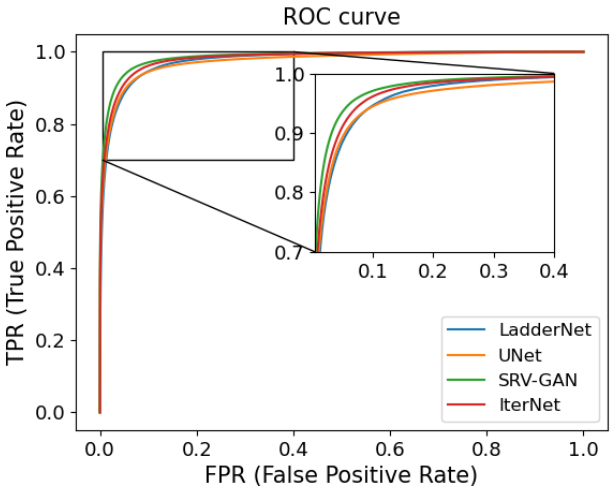


Figure 10. The ROC curves of SRV-GAN and other models on DRIVE.

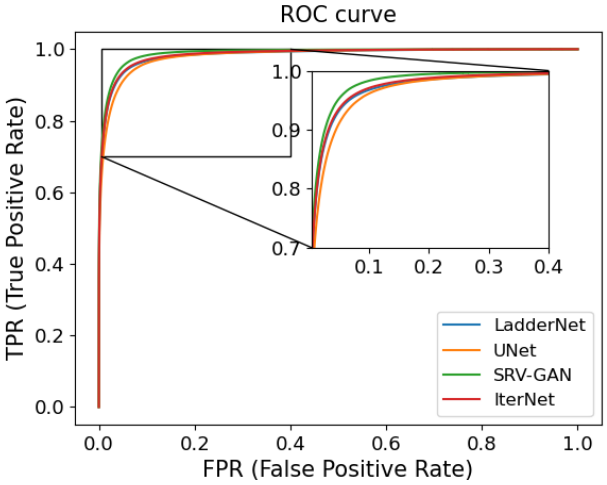


Figure 11. The ROC curves of SRV-GAN and other models on CHASE-DB1.

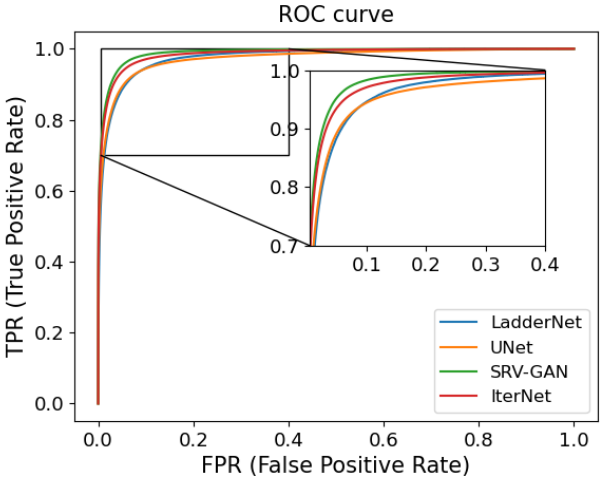


Figure 12. The ROC curves of SRV-GAN and other models on STARE.

Table 5. Performance comparison on DRIVE, CHASE-DB1 and STARE.

Dataset	Method	Year	F1	SE	SP	AC	AUC
DRIVE	U-Net [18]	2018	0.8174	0.7822	0.9808	0.9555	0.9752
	R2U-Net [23]	2018	0.8171	0.7792	0.9813	0.9556	0.9784
	LadderNet [44]	2018	0.8202	0.7856	0.9810	0.9561	0.9793
	IterNet [45]	2019	0.8205	0.7735	0.9838	0.9573	0.9816
	FAU-Net [47]	2020	0.8320	–	–	0.9698	0.9853
	SA-UNet [43]	2020	0.8263	0.8212	0.9840	0.9698	0.9864
	MRU-Net [46]	2020	0.8444	0.8618	–	0.9611	0.9837
	SRV-GAN		0.8452	0.8337	0.9850	0.9702	0.9869
CHASE-DB1	U-Net [18]	2018	0.7993	0.7841	0.9823	0.9643	0.9812
	R2U-Net [23]	2018	0.7928	0.7756	0.9820	0.9634	0.9815
	LadderNet [44]	2018	0.8031	0.7978	0.9818	0.9656	0.9839
	IterNet [45]	2019	0.8073	0.7970	0.9823	0.9655	0.9851
	SA-UNet [43]	2020	0.8153	0.8573	0.9835	0.9755	0.9905
	SRV-GAN		0.8201	0.8132	0.9837	0.9673	0.9894
STARE	U-Net [18]	2018	0.7595	0.6681	0.9915	0.9639	0.9710
	R2U-Net [23]	2018	0.8475	0.8298	0.9862	0.9712	0.9914
	IterNet [45]	2019	0.8146	0.7715	0.9886	0.9701	0.9881
	SUD-GAN [32]	2020	–	0.8334	0.9897	0.9663	0.9734
	MRU-Net [46]	2020	0.8143	0.7887	–	0.9662	0.9856
	SRV-GAN		0.8102	0.8344	0.9884	0.9712	0.9885

5. Conclusions

In this article, we propose an improved GAN for retinal image segmentation, and we achieved good segmentation results in three publicly available datasets. The experimental results show that, compared with U-Net, LadderNet and IterNet models, the SRV-GAN model proposed in this paper shows better performance in segmentation tasks. We found that GAN has not yet been used in clinical trials, so the performance of external data sets independent of training sets cannot be guaranteed. Whether GAN technology can improve the performance of machine learning in clinical diagnosis needs further research. In the future, we will explore more accurate and stable methods of adversarial training, so that they can be put into clinical trials more quickly.

Acknowledgments

This work was funded and supported in part by the National Natural Science Foundation of China, under Grant 61672386; the Anhui Provincial Natural Science Foundation of China, under Grant 1708085MF142; and the Key Research and Development Plan of Anhui Province, China, under Grant 2022a05020011.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. C. Y. Cheung, D. Xu, C. Y. Cheng, C. Sabanayagam, T. Y. Wong, A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre, *Nat. Biomed. Eng.*, **5** (2021), 498–508. https://doi.org/10.1038/s41551-020-00626-4
2. A. M. Mendonca, A. Campilho, Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction, *Ieee. T. Med. Imaging.*, **25** (2006), 1200–1213. https://doi.org/10.1109/tmi.2006.879955
3. Y. Yin, M. Adel, S. Bourennane, Automatic segmentation and measurement of vasculature in retinal fundus images using probabilistic formulation, *Comput. Math. Methods. Med.*, **2013** (2013), 260410. https://doi.org/10.1155/2013/260410
4. D. H. Ye, D. Kwon, I. D. Yun, S. U. Lee, Fast multiscale vessel enhancement filtering, in *Proceedings of SPIE - The International Society for Optical Engineering*, **6914** (2008), 691423. https://doi.org/10.1117/12.770038
5. I. Lázár, A. Hajdu, Segmentation of retinal vessels by means of directional response vector similarity and region growing, *Comput. Biol. Med.*, **66** (2015), 209–221. https://doi.org/10.1016/j.compbiomed.2015.09.008
6. L. C. Neto, G. Ramalho, J. Neto, R. Veras, F. Medeiros, An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images, *Expert. Syst. Appl*, **78** (2017), 182–192. https://doi.org/10.1016/j.eswa.2017.02.015
7. U. Nguyen, A. Bhuiyan, L. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, *Pattern. Recogn.*, **46** (2013), 703–715. https://doi.org/10.1016/j.patcog.2012.08.009
8. J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, *IEEE. T. Med. Imaging.,* **23** (2004), 501–509. https://doi.org/10.1109/TMI.2004.825627
9. E. Ricci, R. Perfetti, Retinal blood vessel segmentation using line operators and support vector classification. *IEEE. T. Med. Imaging.*, **26** (2007), 1357–1365. https://doi.org/10.1109/TMI.2007.898551
10. S. Franklin, S. Rajan, Computerized screening of diabetic retinopathy employing blood vessel segmentation in retinal images, *Biocybern. Biomed. Eng.*, **34** (2014), 117–124. https://doi.org/10.1016/j.bbe.2014.01.004
11. A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM*, **60** (2017), 84–90. https://doi.org/10.1145/3065386
12. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *Comput. Sci*, 2014. https://doi.org/10.48550/arXiv.1409.1556
13. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, A. Rabinovich, Going deeper with convolutions, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2015), 1–9. https://doi.org/10.1109/CVPR.2015.7298594

14. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), 770–778. https://doi.org/10.1109/CVPR.2016.90
15. G. Huang, Z. Liu, V. Laurens, K. Weinberger, Densely connected convolutional networks, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), 2261–2269. https://doi.org/10.1109/CVPR.2017.243
16. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *International Conference on Medical image computing and computer-assisted intervention*, (2015), 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
17. Z. Zhou, M. Siddiquee, N. Tajbakhsh, J. Liang, UNet++: A nested U-Net architecture for medical image segmentation, *Lecture Notes in Computer Science*, Springer, Cham, **11045** (2018). https://doi.org/10.1007/978-3-030-00889-5_1
18. Q. Jin, Z. Meng, T. Pham, Q. Chen, L. Wei, R. Su, DUNet: A deformable network for retinal vessel segmentation, *Know.-Based Syst.*, **178** (2019), 149–162. https://doi.org/10.1016/j.knosys.2019.04.025
19. O. Oktay, J. Schlemper, L. Folgoc, M. Lee, M. Heinrich, K. Misawa, et al., Attention U-Net: Learning where to look for the pancreas, 2018. https://doi.org/10.48550/arXiv.1804.03999
20. J. Ding, Z. Zhang, J. Tang, F. Guo, A multichannel deep neural network for retina vessel segmentation via a fusion mechanism, *Front. Bioeng. Biotechnol.*, **9** (2021), 663. https://doi.org/10.3389/fbioe.2021.697915
21. X. Sun, X. Cao, Y. Yang, L. Wang, Y. Xu, Robust retinal vessel segmentation from a data augmentation perspective, *Ophthalmic Medical Image Analysis, Lecture Notes in Computer Science*, Springer, Cham, **12970** (2021), 189–198. https://doi.org/10.1007/978-3-030-87000-3_20
22. Z. Li, M. Jia, X. Yang, M. Xu, Blood vessel segmentation of retinal image based on Dense-U-Net Network, *Micromachines*, **12** (2021), 1478. https://doi.org/10.3390/mi12121478
23. M. Alom, M. Hasan, C. Yakopcic, T. Taha, V. K. Asari, Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation, preprint, arXiv: 1802.06955.
24. W. Liu, Y. Jiang, J. Zhang, Z. Ma, RFARN: Retinal vessel segmentation based on reverse fusion attention residual network, *PLoS ONE*, **16** (2021). https://doi.org/10.1371/journal.pone.0257256
25. Q. Yang, B. Ma, H. Cui, J. Ma, AMF-NET: Attention-aware multi-scale fusion network for retinal vessel segmentation, in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (2021), 3277–3280. https://doi.org/10.1109/EMBC46164.2021.9630756
26. J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, et al., Dual attention network for scene segmentation, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, (2019), 3141–3149. https://doi.org/10.1109/CVPR.2019.00326
27. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative adversarial nets, in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, **27** (2014), 2672–2680. https://doi.org/10.48550/arXiv.1406.2661
28. M. Mirza, S. Osindero, Conditional generative adversarial nets, *Comput. Therm. Sci.*, (2014), 2672–2680. https://doi.org/10.48550/arXiv.1411.1784
29. B. Lei, Z. Xia, F. Jiang, X. Jiang, S. Wang, Skin lesion segmentation via generative adversarial networks with dual discriminators, *Med. Image. Anal.*, **64** (2020), 101716, https://doi.org/10.1016/j.media.2020.101716

30. A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, preprint, arXiv: 1511.06434.
31. P. Isola, JY. Zhu, T. Zhou, AA. Efros, Image-to-Image translation with conditional adversarial networks, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2017), 5967–5976. https://doi.org/10.1109/CVPR.2017.632
32. T. Yang, T. Wu, L. Li, C. Zhu, SUD-GAN: Deep convolution generative adversarial network combined with short connection and dense block for retinal vessel segmentation, *J. Digit. Imaging.*, **33** (2020), 946–957. https://doi.org/10.1007/s10278-020-00339-9.
33. J. Son, S. Park, K. Jung, Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks, *J. Digit. Imaging.*, **32** (2019), 499–512. https://doi.org/10.1007/s10278-018-0126-3
34. X. Dong, Y. Lei, T. Wang, M. Thomas, L. Tang, W. J. Curran, et al., Automatic multiorgan segmentation in thorax CT images using U-Net-GAN, *Med. Phys.*, **46** (2019), 2157–2168. https://doi.org/10.1002/mp.13458
35. J. Zhang, L. Yu, D. Chen, W. Pan, C. Shi, Y. Niu, et al., Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT images, *Biomed. Signal. Process. Control.*, **69** (2021), 102901. https://doi.org/10.1016/j.bspc.2021.102901
36. A. You, J. Kim, I. Ryu, T. Yoo, Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey, *Eye. Vis. (Lond)*, **9** (2022), 1–19, https://doi.org/10.1186/s40662-022-00277-3
37. V. Bellemo, P. Burlina, Y. Liu, T. Wong, D. Ting, Generative adversarial networks (GANs) for retinal fundus image synthesis, in *Computer Vision – ACCV 2018 Workshops, Lecture Notes in Computer Science*, Springer, Cham, **11367** (2018), 289–302. https://doi.org/10.1007/978-3-030-21074-8_24
38. S. Kamran, K. Hossain, A. Tavakkoli, S. Zuckerbrod, K. Sanders, S. Baker, RV-GAN: Segmenting retinal vascular structure in fundus photographs using a novel multi-scale generative adversarial network, in *MICCAI 2021: Medical Image Computing and Computer Assisted Intervention, Lecture Notes in Computer Science*, Springer, **12908** (2021), 34–44. https://doi.org/10.1007/978-3-030-87237-3_4
39. M. Alom, M. Hasan, C. Yakopcic, T. Taha, Inception recurrent convolutional neural network for object recognition, preprint, arXiv: 1704.07709.
40. M. Alom, M. Hasan, C. Yakopcic, T. Taha, V. Asari, Improved inception-residual convolutional neural network for object recognition, preprint, arXiv: 1712.09888.
41. C. Owen, A. Rudnicka, R. Mullen, S. Barman, D. Monekosso, P. Whincup, et al., Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (CAIAR) program, *Invest. Ophth. Vis. Sci.*, **50** (2009), 2004–2010. https://doi.org/10.1167/iovs.08-3018
42. A. D. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, *IEEE. T. Med. Imaging.,* **19** (2000), 203–210. https://doi.org/10.1109/42.845178
43. C. Guo, M. Szemenyei, Y. Yi, W. Wang, B. Chen, C. Fan, SA-UNet: Spatial attention U-Net for retinal vessel segmentation, in *2020 25th International Conference on Pattern Recognition (ICPR)*, (2021), 1236–1242. https://doi.org/10.48550/arXiv.2004.03696

44. J. Zhuang, LadderNet: Multi-path networks based on U-Net for medical image segmentation, preprint, arXiv: 1810.07810
45. L. Li, M. Verma, Y. Nakashima, H. Nagahara, R. Kawasaki, Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks, in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, (2020), 3645–3654. https://doi.org/10.1109/WACV45572.2020.9093621
46. H. Ding, X. Cui, L. Chen, K. Zhao, MRU-Net: A U-shaped network for retinal vessel segmentation, *Appl. Sci.*, **10** (2020), 6823. https://doi.org/10.3390/app10196823
47. D. Huang, L. Yin, H. Guo, W. Tang, T. Wan, FAU-Net: Fixup initialization channel attention neural network for complex blood vessel segmentation, *Appl. Sci.*, **10** (2020), 6280. https://doi.org/10.3390/app10186280

AIMS Press