



Research article

Retinal vessel segmentation based on multi-scale feature and style transfer

Caixia Zheng^{1,2}, Huican Li², Yingying Ge¹, Yanlin He², Yugen Yi³, Meili Zhu¹, Hui Sun^{4,*} and Jun Kong^{2,*}

¹ Jilin Animation Institute, Changchun 130013, China

² College of Information Science and Technology, Northeast Normal University, Changchun 130117, China

³ School of Software, Jiangxi Normal University, Nanchang 330022, China

⁴ School of Science and Technology, Changchun Humanities and Sciences College, Changchun 130117, China

* **Correspondence:** Email: sunh333@nenu.edu.cn, kongjun@nenu.edu.cn.

Abstract: Retinal vessel segmentation is very important for diagnosing and treating certain eye diseases. Recently, many deep learning-based retinal vessel segmentation methods have been proposed; however, there are still many shortcomings (e.g., they cannot obtain satisfactory results when dealing with cross-domain data or segmenting small blood vessels). To alleviate these problems and avoid overly complex models, we propose a novel network based on a multi-scale feature and style transfer (MSFST-NET) for retinal vessel segmentation. Specifically, we first construct a lightweight segmentation module named MSF-Net, which introduces the selective kernel (SK) module to increase the multi-scale feature extraction ability of the model to achieve improved small blood vessel segmentation. Then, to alleviate the problem of model performance degradation when segmenting cross-domain datasets, we propose a style transfer module and a pseudo-label learning strategy. The style transfer module is used to reduce the style difference between the source domain image and the target domain image to improve the segmentation performance for the target domain image. The pseudo-label learning strategy is designed to be combined with the style transfer module to further boost the generalization ability of the model. Moreover, we trained and tested our proposed MSFST-NET in experiments on the DRIVE and CHASE_DB1 datasets. The experimental results demonstrate that MSFST-NET can effectively improve the generalization ability of the model on cross-domain datasets and achieve improved retinal vessel segmentation results than other state-of-the-art methods.

Keywords: retinal vessel segmentation; multi-scale feature; style transfer; pseudo-label learning; deep learning

1. Introduction

According to data from the World Vision Report [1] released by the World Health Organization in 2019, more than 418 million people suffer from eye diseases worldwide [2]. Eye diseases have seriously harmed human life and health. Therefore, it is necessary to prevent and screen for eye diseases early.

Information regarding the morphology of retinal vessels is essential for the treatment of many eye diseases and has been widely used in clinical diagnoses [3]. Hence, accurately segmenting retinal blood vessels is very important for doctors in diagnosing and treating eye diseases. However, the vessels in the optic cup region are significantly thicker than those in other regions. Additionally, the blood vessels in the fundus image are easily mixed with the background and difficult to distinguish, which leads to retinal vessel segmentation being very challenging.

Researchers have presented many retinal vessel segmentation approaches, which are broadly grouped into two classes: classical segmentation methods and segmentation methods based on deep learning. In the early years, many classical retinal vessel segmentation methods were proposed based on hand-crafted features (e.g., image gradient features and local texture features). For example, Soares et al. [4] employed the pixel intensity and two-dimensional Gabor wavelet transform responses taken at multiple scales as the feature vector of pixels to produce pixel-level vessel segmentation results. Orlando et al. [5] proposed a fully connected conditional random field method with an improved potential for feature extraction, which can perform fast and accurate blood vessel segmentation in retinal images. Nguyen et al. [6] employed a multiscale line detection framework to combine line detectors at varying scales to classify each pixel as either a retinal vessel or background in fundus images. The existing classical methods usually lack sufficiently discriminative information to accurately segment retinal vessels, and their performances are easily affected by the extracted hand-crafted features.

With the wide application of deep learning technology, retinal vessel segmentation approaches based on deep learning have attracted increasing attention. In 2015, Ronneberger et al. [7] proposed U-Net for medical image segmentation tasks. U-Net combines the encoder-decoder architecture and skip connections and is a well-known backbone network in the medical image segmentation community. Recently, researchers have presented a series of improved models based on U-Net [8]. For example, Wang et al. [9] proposed a dual-coded U-Net (DEU-Net), which improves the ability of the network to segment retinal blood vessels pixel-to-pixel. Zhang et al. [10] presented an attention guided network (AG-Net), which designs a filter named the attention guided filter to obtain improved segmentation results for fundus images. Wang et al. [3] developed a multi-scale integrated context network (MIC-Net) based on U-Net to fully fuse the multi-scale features from the encoder and the decoder to segment retinal vessels in fundus images. Wu et al. [11] designed Vesic-Net for vessel segmentation by embedding an initial residual convolution block into the U-type encoder-decoder structure. The residual convolution block is an effective strategy to resolve vanishing/exploding gradient problems and is widely used in various medical image analysis tasks; for example, Shin et al. [12] proposed a squeeze-and-excitation super-resolution residual network (SE-SRResNet) for transcranial focused

ultrasound simulation.

The retinal vessel segmentation approaches based on deep learning can continuously optimize the model according to the real labels of the training dataset so they can achieve good segmentation performances. However, when deep learning-based models trained well on one public dataset are directly applied to another dataset, the segmentation performance of the models will greatly decrease due to the domain shift problem [13]. Domain shifts often exist in different medical image datasets and are very common in practical applications. For example, different hospitals and institutions usually adopt different equipment to take fundus images, which causes a domain shift problem between the images acquired by different equipment. Even when adopting the same equipment, the appearance of the image usually varies with subjects, clinical operators and other factors thus the domain shift problem is still persistent. Figure 1 shows an example of a domain shift. In Figure 1, the style and size of images between two datasets (DRIVE and CHASE_DB1) are different, which means that there is a domain shift between these two databases. Since a domain shift will degrade the model performance, the development of a valid retinal vessel segmentation model that is robust to domain shifts is necessary.

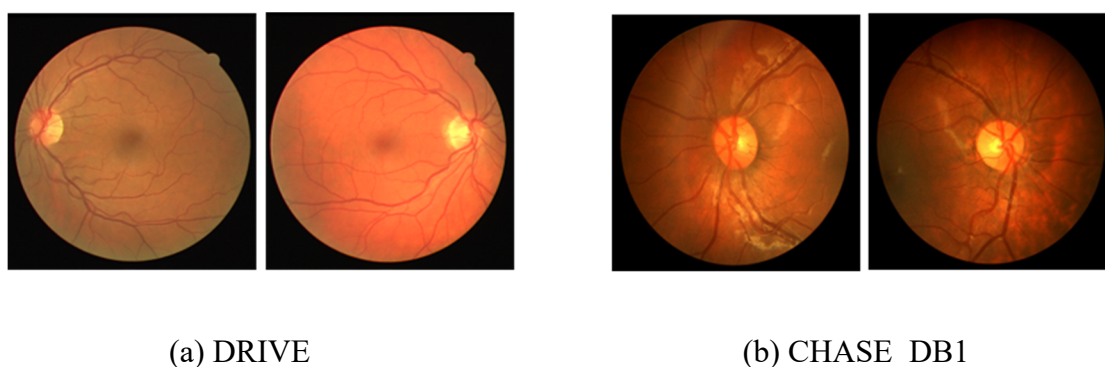


Figure 1. Comparison of images from different datasets.

In this paper, we propose a novel retinal vessel segmentation method based on multi-scale features and a style transfer (MSFST-NET), which can improve the segmentation results and is robust to domain shifts. The contributions of our work are as follows:

1) Small blood vessels in fundus images play a very important role in disease diagnoses. Ordinary segmentation models often use convolutions of the same size to extract features, which leads to poor small blood vessel segmentation. To solve this problem, we introduce the selective kernel (SK) into the segmentation model to learn more multi-scale feature information to improve the segmentation results of small vessels.

2) A domain shift is a common phenomenon in medical images. To alleviate the problem of model performance degradation caused by domain shifts, we propose a style transfer module to reduce the style difference between the source domain dataset and the target domain dataset.

3) To further enhance the robustness of the model, we additionally design a pseudo-label strategy to increase the model's generalization ability and reduce the effect of domain shifts.

4) To verify the effectiveness of our proposed method, we conducted a large number of experiments on the DRIVE and CHASE_DB1 datasets. The experimental results show that our proposed method is effective and superior to other advanced vessel segmentation methods.

2. Related work

2.1. Retinal vessel segmentation

Retinal vessel segmentation is used to localize blood vessels in fundus images. Early retinal vessel segmentation approaches are often designed by employing traditional shallow machine learning models, which include matching filtering and blood vessel tracking methods. Recently, with the development of deep learning, a large number of retinal vessel segmentation approaches have been presented by adopting the convolutional neural network (CNN), which achieves a superior performance compared to the approaches based on traditional shallow machine learning and has attracted increasing attention.

The fully convolutional network (FCN), which was proposed by Long et al. [14], is a classical semantic segmentation approach. The FCN extends the deep neural network from the classification of the whole image to the classification of each pixel in the image for the first time. To gain more accurate results in the area of medical image segmentation, a large number of improved methods based on FCN have been proposed. For instance, U-Net [7] is a deep convolutional network constructed by improving the original skip connection part of an FCN. Subsequently, many more advanced network models have been proposed based on U-Net (e.g., attention U-Net [15] and U-Net++) [16]. Attention U-Net introduces an attention gate mechanism into U-Net to reduce the redundant feature information brought on by the skip connection. U-Net++ introduces a jump path consisting of dense convolutional blocks and dense jump connections into the structure of U-Net. Meanwhile, U-Net++ adds a deep supervision mechanism to speed up the convergence of network training. In addition to the methods based on U-Net, some retinal vessel segmentation methods have been proposed based on generative adversarial networks [17–20]. For example, Yue et al. [20] proposed an improved generative adversarial network based on R2U-Net for retinal vessel segmentation. Generative adversarial network-based segmentation methods have difficulty training models; thus, they have not become widespread popular methods.

The abovementioned methods mainly focus on segmentation accuracy while usually ignoring segmentation efficiency. The efficiency of the model is of great significance to practical applications. To improve the segmentation model efficiency, Galdran et al. [21] proposed a new model called W-Net, which is a simple extension of the U-Net structure and can obtain an outstanding performance on retinal vessel segmentation. Hence, we employ W-Net as the backbone network of our proposed method and further boost its segmentation performance for small retinal vessels by introducing the multi-scale feature extraction module.

2.2. Cross-domain segmentation

In recent years, some researchers have found that when a deep learning model trained on a labeled dataset is directly applied to a different dataset, the segmentation performance of the model is greatly reduced [13]. This is because different fundus image datasets usually have the issue of domain shift. To alleviate this issue, many cross-domain segmentation methods have been presented to strengthen the domain transferability of models and mitigate cross-domain model performance decline. Cross-domain segmentation methods can be broadly categorized into two main classes: methods based on semi-supervised domain adaptation and methods based on unsupervised domain adaptation.

The methods based on semi-supervised domain adaptation [22–26] usually adopt fewer labeled

images and more unlabeled images in the target domain dataset to mitigate the negative impact of domain shift on model performance. For instance, Xia et al. [22] proposed an uncertainty-aware multi-view co-training (UMCT) method to achieve semi-supervised learning and domain adaptation, which can obtain an improved performance for cross-domain medical image segmentation. Chen et al. [23] proposed a dual-level domain mixing-based segmentation network that learns domain-invariant region-level and image-level features based on labeled samples from different domains and further enhances the model performance by using pseudo labels of unlabeled data. The performance of semi-supervised domain adaptation methods is limited by the number of labeled images; annotating these medical images is very difficult, which greatly reduces the scope of the practical application of these semi-supervised domain adaptation models.

The methods based on unsupervised domain adaptation aim to enhance model domain transferability and alleviate cross-domain performance degradation by only using unlabeled data in the target domain. In recent years, many unsupervised domain adaptation methods with a relatively good performance have been proposed [27–32]. Wang et al. [27] designed a cross-domain segmentation approach based on feature separation. This approach adopts a new unsupervised region adaptive strategy and the disentangled reconstruction neural network (DRNN) to reduce the impact of domain shift on the model's segmentation performance. Zuo et al. [28] developed an unsupervised domain adaptation method based on category-level adversarial self-ensembling, which aligns the source and target domain by constraining the descriptions. Xu et al. [29] designed self-ensembling attention networks to produce attention-aware features and used them to guide the model to compute the consistency loss in the target domain. Since the methods based on unsupervised domain adaptation do not require any labeled target domain data, they are more convenient for practical applications.

Inspired by these previous studies and to avoid models that are too complex, we design a simple yet effective method to address the issue of domain shift. Specifically, we combine Cycle-GAN [33], which is a popular algorithm to transfer images over different domains by cycle consistency losses, and a pseudo-label learning strategy to ensure that our segmentation method can still acquire good segmentation performance when dealing with images from cross-domains.

3. Proposed method

3.1. The overall framework of MSFST-NET

The overall structure of our MSFST-NET is shown in Figure 2. First, the labeled source domain images are input into our proposed segmentation model, named MSF-Net, to train and obtain the model with the optimal segmentation performance. Then, based on the style of the source domain images, style transformation is performed on the unlabeled target domain images by the proposed style transfer module. Finally, the target domain images after the style transformation are input into the trained MSF-Net to generate the prediction maps. Prediction maps are used as pseudo-labels that are fused with the labeled source domain images to further increase the segmentation accuracy and strengthen the generalization ability of the network. Overall, MSFST-NET includes the following three parts: the segmentation module MSF-Net, a style transfer module and a pseudo-label learning strategy.

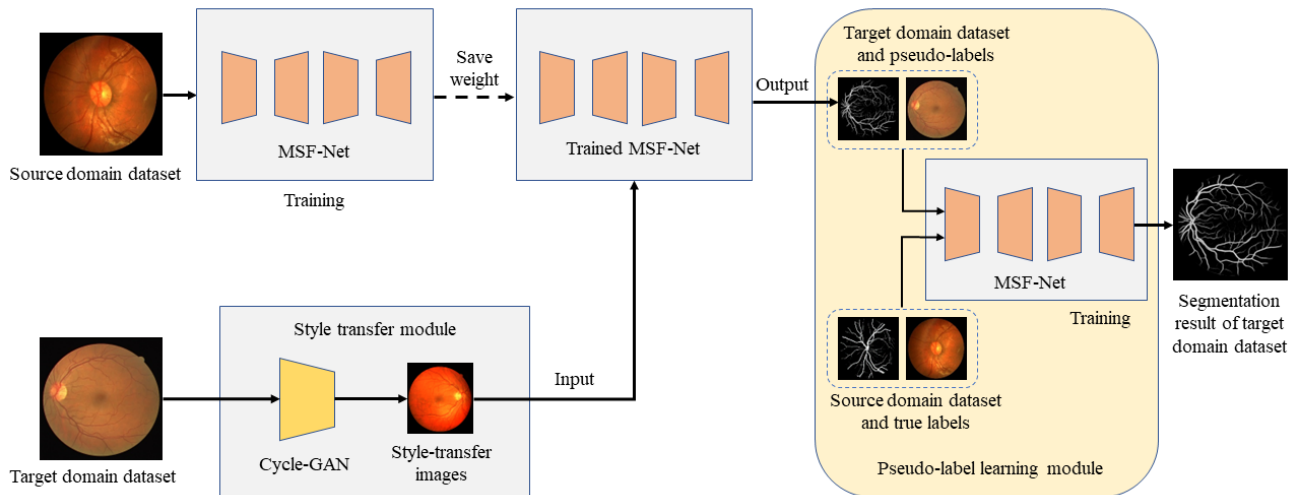


Figure 2. The overall structure of MSFST-NET.

Segmentation module MSF-Net: This part is the basic segmentation model of our proposed MSFST-NET. To make the model lightweight and have a good performance, we use two-cascade three-layer U-Net as the basic architecture of MSF-Net. In addition, considering that retinal vessels have different sizes and scales, it is difficult to simultaneously obtain the important feature information of blood vessels at each scale by only using convolution kernels of the same size. Therefore, we introduce SK into the encoding stage of each U-Net in MSF-Net, that is, we employ convolution kernels of different sizes to learn and adaptively fuse multi-scale features to reduce feature loss and further increase the accuracy of blood vessel segmentation.

Style transfer module: Since the issue of domain shift is common in the community of medical image segmentation, we construct a style transfer module to make the target domain images and the source domain images tend to assimilate in style, thereby reducing the style difference between them and easing the performance degradation when the model is directly applied to another dataset. In this module, the classical Cycle-GAN style transfer algorithm is adopted, and the model trained by the source domain dataset is employed to segment the target domain images after the style transfer to obtain improved segmentation results of the target domain images.

Pseudo-label learning strategy: To further strengthen the model performance when dealing with cross-domain datasets (i.e., there is a domain shift between two datasets), we further design a pseudo-label learning strategy. The pseudo-label of the target domain dataset generated by the style transfer module is utilized as the training data; then, the model is retrained by the target domain dataset and the source domain dataset, which can obtain a model with an improved segmentation performance and a stronger generalization ability.

3.2. MSF-Net

We adopt a deep learning model containing a two-cascade three-layer U-Net, which is called W-Net [21], as the backbone network of MSF-Net; this deep learning model is a lightweight model that has a good segmentation accuracy and few parameters. The original U-Net contains five layers of downsampling and upsampling. To make the model lightweight, W-Net reduces the five-layer network

structure to a three-layer network structure that has approximately 0.03 M parameters. To reduce the number of layers and parameters without losing too much model performance, W-Net concatenates two three-layer U-Nets, which has around 0.06 M parameters in total [21]. The segmentation results output by the first U-Net in W-Net are fused with the original images as the input of the second U-Net, that is, the blood vessel segmentation prediction map of the first U-Net is used as an attention map to weigh the original image and input it to the second U-Net so that the second U-Net pays more attention to the key parts of the image when segmenting blood vessels.

Although W-Net performs relatively well for retinal vessel segmentation, it cannot capture rich multi-scale image features; hence, the segmentation results of small blood vessels are not satisfactory. To solve this drawback, we embed the SK [34] into W-Net to construct a new network named MSF-Net to strengthen the multi-scale feature extraction ability of the model. Figure 3 shows the structure of our MSF-Net.

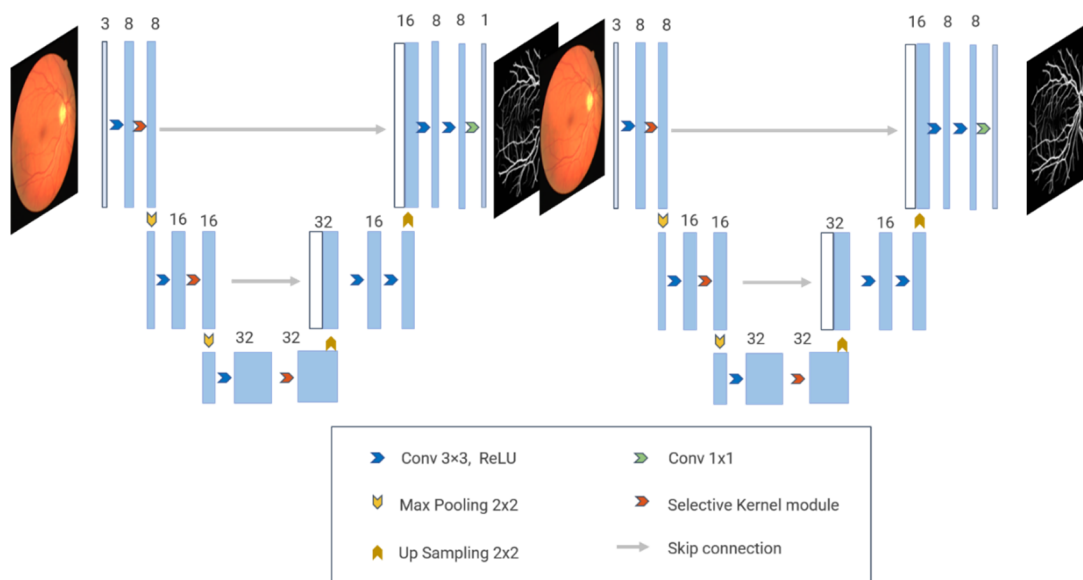


Figure 3. The MSF-Net architecture.

In MSF-Net, each encoder in the two cascaded U-Net networks contains three stages, and each stage contains a normal convolution operation and an SK module, along with a downsampling operation. Each decoder in U-Net consists of a convolution operation and an upsampling operation. In Figure 3, the blue rectangular block represents the feature map, and the number on each feature map denotes the number of channels. The blue arrow represents a 3×3 convolution, batch normalization (BN) and rectified linear unit (ReLU) activation function [35] (BN can accelerate network training and increase model accuracy [36], and ReLU can enhance nonlinear mapping learning and reduce the computational complexity of the network to prevent gradient disappearance). The red arrow represents the SK module, and BN and ReLU are added after SK. The yellow downward arrows represent the max pooling operation, which is employed to decrease the size of the feature map and extract features. The yellow up arrow represents the upsampling step used to gradually recover the details of the object and the corresponding spatial dimension. The gray arrow represents skip connections, which can provide richer feature information for the decoder.

The SK module [34] is a dynamic selective kernel that uses convolution kernels with different

receptive field sizes to extract multi-scale feature maps and adaptively selects the appropriate receptive fields for vessels of different sizes, thus obtaining rich multi-scale feature information for retinal vessel segmentation. Figure 4 shows the architecture of the SK module.

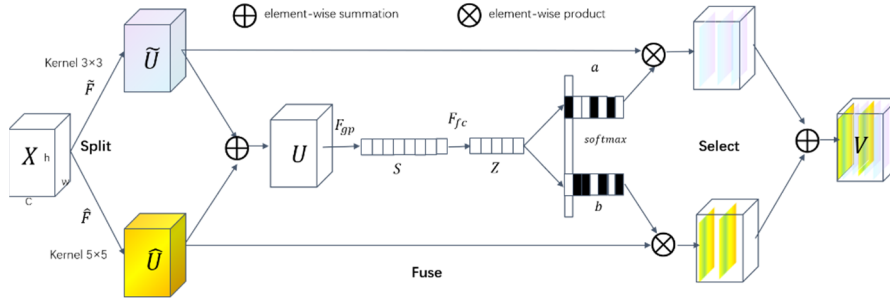


Figure 4. The SK module architecture.

The key to the SK module is that it employs convolution kernels with different receptive field sizes to adaptively extract and select multi-scale features, making the feature information learned by the network richer and more discriminative, which is conducive to improving the segmentation results of small blood vessels in the retinal image. SK consists of three operations: *split*, *fuse* and *select*.

The *split operation* extracts abundant multi-scale feature information and adopts convolution kernels with different receptive field sizes (3×3 and 5×5) to deal with the input features X to obtain the different scale feature maps \tilde{U} and \hat{U} , as shown in Figure 4.

The *fusion operation* first integrates \tilde{U} and \hat{U} by elementwise summation to obtain the fusion feature map U . Then, it utilizes the average pooling and fully connected layer to handle U to produce a compact feature map z . z will be utilized to learn the adaptive weights a_c and b_c for the different scale feature maps. This process can be described by the following formulas:

$$z = F_{fc}(s) = \delta_{relu}(BN(Ws)) \quad (1)$$

$$s_c = F_{gp}(U_c) = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w U_c(i, j) \quad (2)$$

where F_{fc} represents the fully connected layer, BN is the batch normalization, δ_{relu} represents the ReLU activation function, W is a learnable parameter and $W \in R^{d \times c}$, d represents the dimensions of the fully connected layer, F_{gp} denotes the global average pooling operation utilized to obtain the channel-wise statistics s , specifically, s_c represents the c -th element of s , U_c denotes the c -th channel of U , and h and w are the height and width of the feature map, respectively.

The *selection operation* adaptively selects the multi-scale features by the different attention weights. First, it employs the softmax operation to calculate the weights a_c and b_c of the feature maps with different scales by Eq (3). Then, the c -th channel V_c of the final feature map V is obtained by Eq (4):

$$a_c = \frac{e^{A_c z}}{e^{A_c z} + e^{B_c z}} \quad b_c = \frac{e^{B_c z}}{e^{A_c z} + e^{B_c z}} \quad (3)$$

$$V_c = a_c \cdot \tilde{U} + b_c \cdot \hat{U}, \quad a_c + b_c = 1 \quad (4)$$

where A_c represents the c -th row of A , B_c is the c -th row of B , A and B are learnable parameters, and $A, B \in R^{C \times d}$.

3.3. Style transfer module

Due to the differences in the shooting devices and the shooting angles, there are often stylistic differences between images from two different retinal vessel datasets, which is called a domain shift. Therefore, when a segmentation model trained on one fundus image dataset is directly applied to another fundus image dataset, the performance of the model will be greatly reduced. To address this problem, we develop a style transfer module to reduce the differences between the two datasets as much as possible to alleviate the domain shift problem. In the style transfer module, we adopt a classical model, termed Cycle-GAN [33], to transfer the style of the target domain images to that of source domain images. In medical image segmentation, interpretability is very important. Cycle-GAN can transform target domain images into source-like domain images, which is a more intuitive and highly interpretable approach. Furthermore, the transformed images can be directly utilized in our segmentation model without retraining. In contrast, other advanced domain generalization methods usually require aligning features from different domains in a high-dimensional feature space, which is less interpretable and creates greater training challenges. Hence, we selected the relatively simple yet effective Cycle-GAN as our style transfer module.

One advantage of Cycle-GAN is that it can be trained without paired datasets, and it can retain the details of the original image content as much as possible during style transfer. The Cycle-GAN structure is shown in Figure 5. I_t represents the target domain dataset and I_s represents the source domain dataset. There are two generators and two discriminators in the Cycle-GAN model: F is an image generator that converts the source domain style to the target domain style, and D_x is a discriminator that decides whether the generated image style is consistent with the target domain image style. G is an image generator that converts the target domain style to the source domain style and D_y is a discriminator that determines whether the style of the image generated by G is consistent with that of the source domain image.

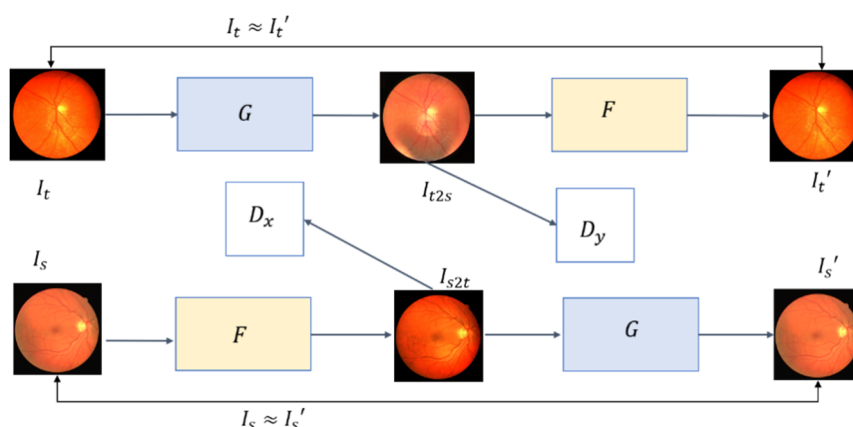


Figure 5. The architecture of Cycle-GAN.

Cycle-GAN mainly uses cycle consistency constraints to ensure that the content is consistent

between the output image of the generator and the original input image. As shown in Figure 5, in the process of cycle consistency, given two images, I_t and I_s , from different style domains, image I_t is first used to generate image I_{t2s} with image I_s style through the generator G , and then I_{t2s} is utilized to generate the image I_t' with the same style as I_t through the generator F . In this process, if the content and style of the image I_t and the image I_t' are consistent, it signifies that the image has gone through a cycle and the content and style of the image remain consistent before and after the cycle, which is the basic principle of cycle consistency.

Cycle consistency can ensure that the content of the original input image is preserved when the generator produces the style-transferred image so that the generated image can be returned to the input image by another generator during the second transformation. In our work, when performing a style transfer on fundus images, cycle consistency can preserve as much of the key information of blood vessels as possible.

3.4. Pseudo-label learning strategy

In the previous section, the style transfer module was proposed to ease the domain transfer problem between the different datasets to some extent. To further solve this problem and strengthen the generalization ability of our model, our work adopts a pseudo-label learning strategy. The pseudo-label learning strategy is a method based on the idea of semi-supervised learning, and its main purpose is to enhance the generalization ability of the model by employing unlabeled data.

The semi-supervised training process based on pseudo-label learning mainly contains the following four steps. First, the labeled source domain dataset is employed to train the segmentation model and save the optimal model parameters. Second, the trained model is utilized to estimate the labels of unlabeled target domain samples, and the estimated results are used as the pseudo-labels of these unlabeled target domain samples. Third, the labeled source domain samples and the target domain samples with pseudo-labels are employed as new training data to retrain the segmentation model, and the first to third steps are repeated until the model converges. Finally, the images in the target domain test dataset are predicted to obtain the final blood vessel segmentation results. The details of the pseudo-label learning strategy are shown in Algorithm 1.

Algorithm 1 Pseudo-label learning strategy

Input: Labeled source domain dataset $D_s(x_s, y_s)$, x_s represents source image and y_s is the corresponding label; Unlabeled target domain dataset $D_t(x_t)$, x_t represents the target image.

Initialization: The max number of iterations I_{max} ; Train MSF-Net using the samples from D_s and save the MSF-Net network f_1 which has the optimal parameters.

For $i = 1$ **to** I_{max} **do**

Predict the segmentation mask y_t' of each x_t by using f_i ;

Set y_t' as the pseudo-label of x_t and obtain a new target domain dataset $D_t'(x_t, y_t')$ with pseudo-labels;

$\tilde{D} \leftarrow D_s \cup D_t'$;

Train f_i again by \tilde{D} and obtain the network f_{i+1} with new parameters;

End for

Output: Final trained network $f_{I_{max}+1}$.

3.5. Loss function

Our proposed method adopts the binary cross-entropy (BCE) loss function [37]. Specifically, the formula of our final loss function f_{loss} is as follows:

$$f_{loss} = L_{BCE}(P_1, y) + L_{BCE}(P_2, y) \quad (5)$$

$$L_{BCE}(P_1, y) = -y \log P_1 - (1 - y) \log(1 - P_1) \quad (6)$$

$$L_{BCE}(P_2, y) = -y \log P_2 - (1 - y) \log(1 - P_2) \quad (7)$$

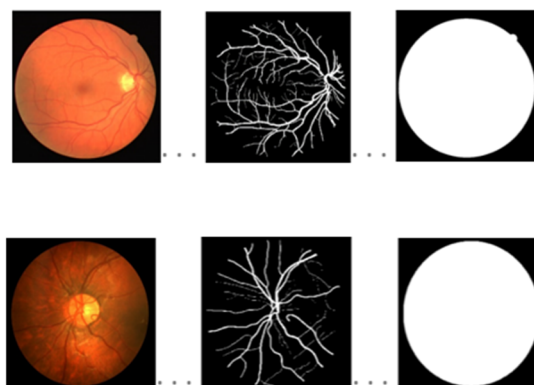
where x is the input image, y is the true label in the ground truth, P_1 is the prediction result of the first U-Net, P_2 is the prediction result of the second U-Net, and L_{BCE} represents the BCE loss, which is calculated based on the cross-entropy between the prediction result generated by the model and the true label in the ground truth.

4. Experiment results

Our proposed MSFST-NET is composed of MSF-Net, which is a style transfer module and a pseudo-label learning strategy. In this section, we first verify the effectiveness of MSF-Net for the retinal blood vessel segmentation task. Second, based on MSF-Net, the style transfer module and pseudo-label strategy are tested and evaluated to verify their effectiveness in solving the domain shift problem. Finally, the performance of our proposed MSFST-NET is evaluated by comparing it to other advanced methods and constructing an additional ablation experiment.

4.1. Datasets and evaluation criteria

The experimental datasets used in this paper are DRIVE [38] and CHASE_DB1 [39], which are frequently adopted in the fundus image segmentation community. Figure 6 shows the example images from these two datasets. In Figure 6, the images in the first row are from DRIVE, and the images in the second row are from CHASE_DB1.



(a) Images (b) Ground truth (c) Circular mask

Figure 6. The images from the DRIVE and CHASE_DB1 datasets.

The DRIVE dataset [38] contains a total of 40 color fundus images, where seven images contain diabetic lesions, and the remaining 33 images are normal. The resolution of each image is 584×565 pixels. The first 20 images in DRIVE are employed as the training data, and the last 20 images are adopted as the test data. Moreover, DRIVE provides a nearly circular mask with a diameter of approximately 540 pixels for each image, which is used to indicate the region for the model training.

The CHASE_DB1 dataset [39] consists of 28 color retinal images with a resolution of 999×960 pixels. The images in this dataset are taken from the left and right eyes of 14 school-aged children. CHASE_DB1 was evenly separated into two parts, each containing 14 images, and used as a training set and a test set. Additionally, this dataset provides a circular mask for each image.

To evaluate the performance of the proposed method, we adopt five evaluation criteria commonly used in retinal vascular segmentation, which include sensitivity (Se), accuracy (Acc), comprehensive evaluation index (F1-score), AUC (area under the curve) and Matthews correlation coefficient (MCC). The larger the value of these indicators, the better the prediction result. The AUC is calculated by using the implementation provided in the scikit-learn Python library. The specific formulas of the other evaluation criteria are as follows:

$$Se = \frac{TP}{TP+FN} \quad (8)$$

$$ACC = \frac{TP+TN}{TP+FP+TN+FN} \quad (9)$$

$$F1\text{-score} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (10)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (11)$$

where TP , TN , FP and FN are the number of true positive, true negative, false positive and false negative pixels, respectively.

4.2. Implementation details

In the experiment, our method is trained by the PyTorch 11.4 framework and implemented on a desktop PC with an Nvidia GeForce GTX 3060TI and 8 GB RAM.

1) Data enhancement

In the medical image processing domain, labeling data is difficult, so the medical image dataset is relatively small. To fully train the segmentation model, we adopt a variety of online data augmentation methods to preprocess the image data in the experiment, including horizontal flip, vertical flip, random 45-degree rotation transformation, and horizontal and vertical offset transformation.

2) Parameter setting

In the experiments, we utilize DRIVE as the target domain dataset and utilize CHASE_DB1 as the source domain dataset. In model training, the Adam optimization technology is used, and the number of iterations is set to 50 in each cycle, for a total of 20 rounds of training. The batch size is set to four. The learning rate is initially set to 0.01, and then cyclically decreased using the cosine law until the learning rate reaches 1×10^{-8} .

4.3. Verifying the performance of the MSF-Net module

Table 1. Comparison of MSF-Net and different segmentation methods on DRIVE.

Method	Year	AUC	Acc	Se	F1-score
U-Net [7]	2018	97.55%	95.31%	75.37%	81.42%
Yan et al. [42]	2018	97.50%	95.40%	76.30%	-
Kai Hu [43]	2018	97.59%	95.33%	77.72%	80.98%
BTS-UNet [44]	2019	98.06%	95.61%	78.91%	82.49%
CENe [45]	2019	97.79%	95.45%	83.09%	-
S-Unet [46]	2019	98.21%	95.67%	83.12%	-
IterNet [47]	2020	98.16%	95.73%	77.35%	82.05%
RVSeg-Net [48]	2020	98.17%	96.81%	81.07%	-
Cheng et al. [49]	2020	97.93%	95.59%	76.72%	-
Du et al. [50]	2021	97.80%	95.56%	78.14%	
Wu et al [51]	2021	98.16%	95.65%	78.69%	82.21%
Wu et al. [40]	2022	98.42%	96.86%	80.43%	81.79%
SegR-Net [41]	2023	-	-	82.06%	80.97%
MSF-Net	-	98.50%	96.05%	83.67%	84.46%

Table 2. Comparison of MSF-Net and different segmentation methods on CHASE_DB1.

Method	Year	AUC	Acc	Se	F1-score
U-Net [7]	2018	97.72%	95.78%	82.88%	77.83%
LadderNet [52]	2018	98.39%	95.33%	79.78%	80.31%
DEU-Net [9]	2019	98.12%	96.61%	80.37%	80.37%
AG-Net [10]	2019	97.79%	97.43%	81.86%	-
Lü et al. [53]	2020	97.82%	96.17%	81.35%	-
MSCNN-AM [54]	2020	98.38%	96.44%	81.32%	-
RVSeg-Net [48]	2020	98.33%	97.26%	80.69%	-
Cheng et al. [49]	2020	97.85%	94.88%	89.67%	-
Du et al. [50]	2021	97.84%	95.90%	81.95%	
Wu et al [51]	2021	98.46%	97.02%	79.42%	80.57%
Wu et al. [40]	2022	97.46%	97.46%	83.00%	80.05%
SegR-Net [41]	2023	-	-	83.29%	80.30%
MSF-Net	-	98.46%	96.21%	83.26%	81.44%

We compare our proposed MSF-Net to some methods with superior performances in recent years. Tables 1 and 2 show the segmentation results of MSF-Net and other methods on the DRIVE and CHASEDB1 datasets. It should be mentioned that in this experiment, there is no domain shift problem to consider, that is, the training and testing samples are from the same database. Here, we only evaluate the segmentation performance of MSF-Net, and we test the effectiveness of our proposed style transfer module, pseudo-label learning strategy and the proposed overall method MSFST-NET when dealing with the issue of domain shift in Sections 3.4 and 3.5.

From Table 1, we find that the AUC, Acc, Se and F1-score values of MSF-Net are 98.50%,

96.05%, 83.67% and 84.46%, respectively. In general, they are better than those of other methods. The Acc of MSF-Net is slightly lower than those of other methods; for example, the Acc of MSF-Net is slightly lower than that of Wu et al. [40], but the AUC, Se and F1-score of MSF-Net are higher than those of Wu et al. [40] by 0.08%, 3.24% and 2.67%, respectively. Hence, the overall performance of MSF-Net is superior to that of Wu et al. [40] on the DRIVE dataset. As shown in Table 2, the AUC, Acc, Se and F1-score of MSF-Net are 98.46%, 96.21%, 83.26% and 81.44%, respectively. MSF-Net gains an optimal performance in the two indicators of AUC and F1-score. The Acc of MSF-Net is not optimal; for example, the Acc is slightly lower than that of Wu et al. [40], but the AUC, Se and F1-score of MSF-Net are higher than those of Wu et al. [40] by 1.00%, 0.26% and 1.39, respectively. Moreover, this demonstrates that the overall performance of MSF-Net is superior to that of Wu et al. [40] on the CHASE_DB1 dataset. The Se of MSF-Net is slightly lower than that of SegR-Net [41] by 0.03% on the CHASE_DB1 dataset, but the F1-score of MSF-Net is higher than that of SegR-Net [41] by 1.14% on the CHASE_DB1 dataset, and both the Se and F1-score are higher than those of SegR-Net [41] by 1.61% and 3.49% on the DRIVE dataset, respectively. Therefore, by comprehensively comparing all the indicators, our proposed MSF-Net has the optimal segmentation precision for retinal blood vessels compared to other previously used methods.

To further demonstrate the performance of MSF-Net, we visualize and compare the segmentation results of MSF-Net and the classical U-Net. Figures 7 and 8 show the segmentation results of different approaches on DRIVE and CHASEDB1.

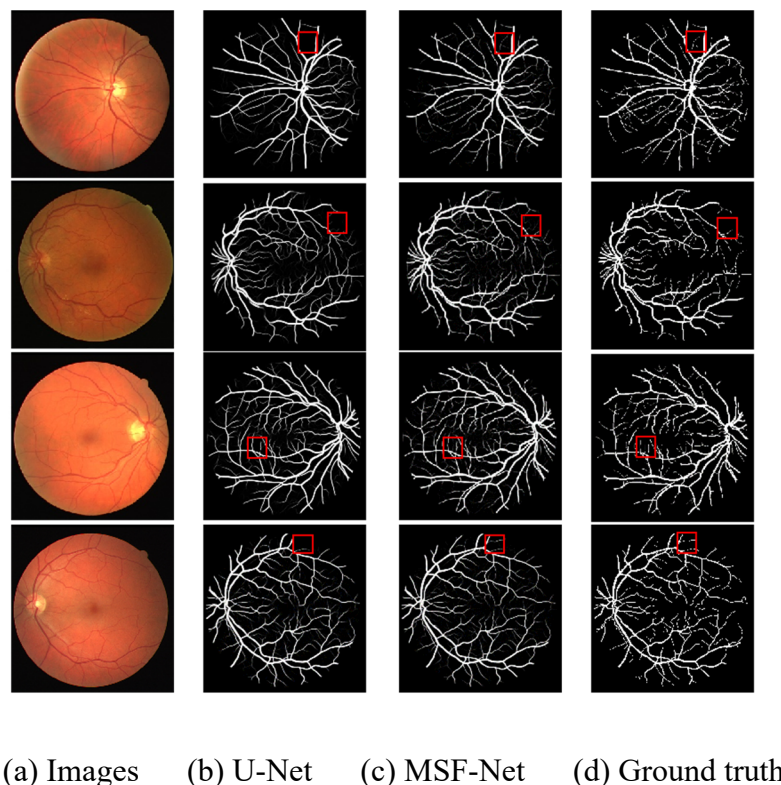


Figure 7. Segmentation results of the images from DRIVE.

From Figures 7 and 8, we find that compared to U-Net, our MSF-Net can obtain a result that is closer to the ground truth. For example, MSF-Net can segment very small blood vessels well, as shown

in the region marked by the red box in Figure 7. Moreover, U-Net misclassifies a large number of background pixels as blood vessels, while our MSF-Net can distinguish the background and blood vessels well, as shown in the region marked by the red box in Figure 8.

Based on the above results and analysis, we can conclude that compared to other segmentation approaches, the segmentation results of our MSF-Net are more consistent with the ground truth, that is, MSF-net achieves a better segmentation performance. This is because MSF-Net can extract more blood vessel feature information of different scales by improving the convolution operation based on SK, so it can effectively distinguish the background and blood vessels, especially for small blood vessels.

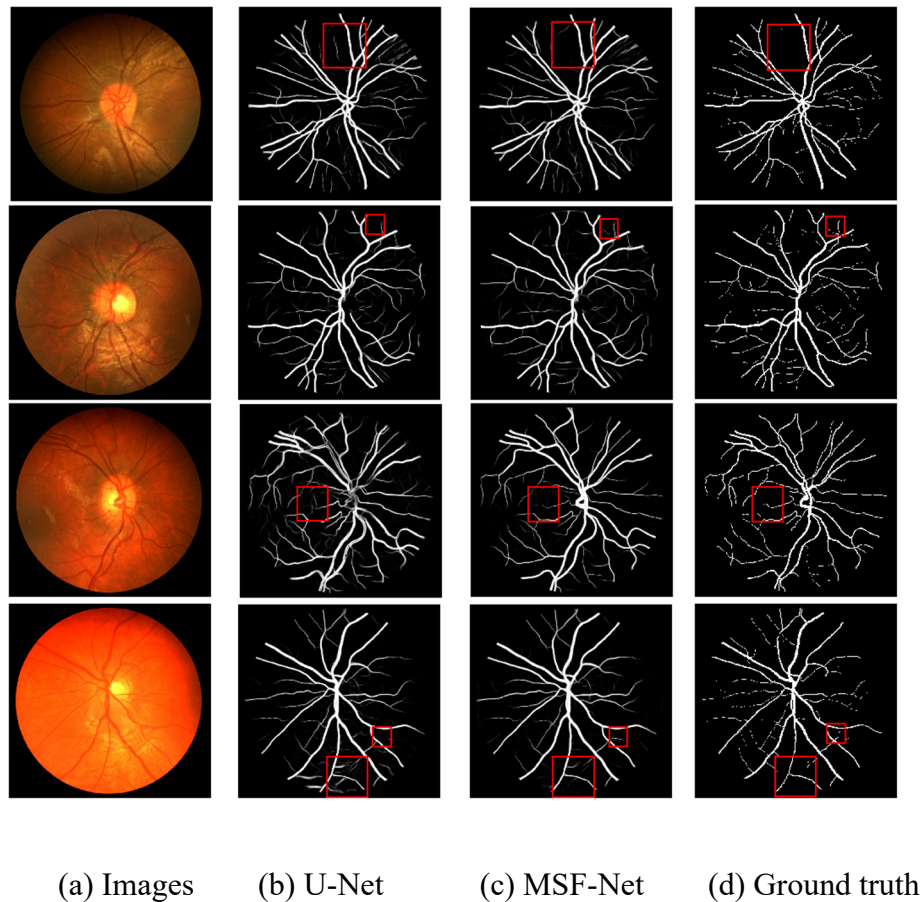


Figure 8. Segmentation results of the images from CHASE_DB1.

4.4. Verifying the performance of the style transfer module and pseudo-label learning strategy

In this experiment, we employ DRIVE as the target domain dataset and adopt CHASE_DB1 as the source domain dataset. After the style transfer, the DRIVE dataset is named DRIVE_style. The images obtained by the style transfer module are shown in Figure 9. In Figure 9, the images in the first row are from the source domain (CHASE_DB1), the images in the second row are from the target domain (DRIVE), and the images in the third row are the images obtained by transferring the style of the image in the target domain to the style of the source domain. From Figure 9, we can see that after the style transfer, the style of the images in the DRIVE dataset has generally been transformed into the style of the images in the CHASE_DB1 dataset.

To further verify the positive effect of style transfer on the retinal vessel segmentation task, the style-transferred images obtained by our style transfer module are input into the segmentation model MSF-Net to test the segmentation performance, and the experimental results are shown in Figure 10. In Figure 10, (a) is the original image in DRIVE, (b) is the segmentation result of the original image by employing the MSF-Net model trained on CHASE_DB1, (c) is the style-transferred image of the original image in DRIVE obtained by the style transfer module, (d) is the segmentation result of the image in (c) when using the MSF-Net model trained on CHASE_DB1, and (e) is the corresponding ground truth. Comparing the regions marked by the red rectangular box in Figure 10, it can be found that the segmentation results of the style-transferred images are significantly better than those of the original image. Therefore, this experiment verifies that our style transfer module can alleviate the degradation of model performance caused by the domain transfer.

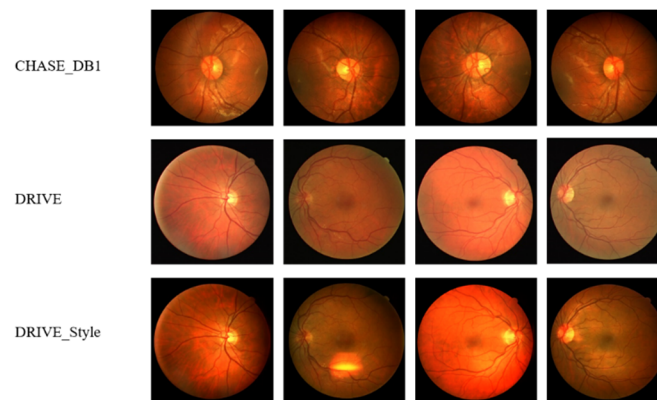


Figure 9. The style transfer results.

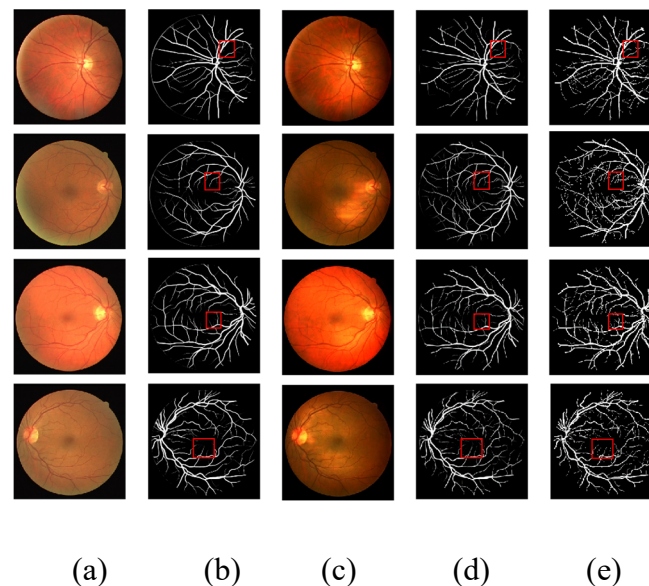


Figure 10. The style-transferred image segmentation results.

To further improve the generalization ability of our proposed approach, the pseudo-label learning method is adopted after the style transfer. To illustrate the validity of the pseudo-label learning strategy,

we compare the segmentation results obtained by the MSF-Net model with and without the pseudo-label learning strategy, as shown in Figure 11. In Figure 11, (a) is the original image in DRIVE, (b) is the segmentation result of the original image, (c) is the segmentation result of the style-transferred image obtained by the style transfer module, (d) is the segmentation result obtained after the style transfer and the pseudo-label learning strategies, and (e) is the ground truth. The abovementioned results are obtained by using MSF-Net trained on the training set of CHASE_DB1. Observing the red rectangular box in Figure 11, it can be found that with the gradual addition of the style transfer module and a pseudo-label learning strategy, the details of the blood vessels in the segmentation results become increasingly clear, which indicates that the use of the pseudo-label learning strategy based on style transfer can further mitigate the adverse effects of the domain shift on the segmentation performance of the model.

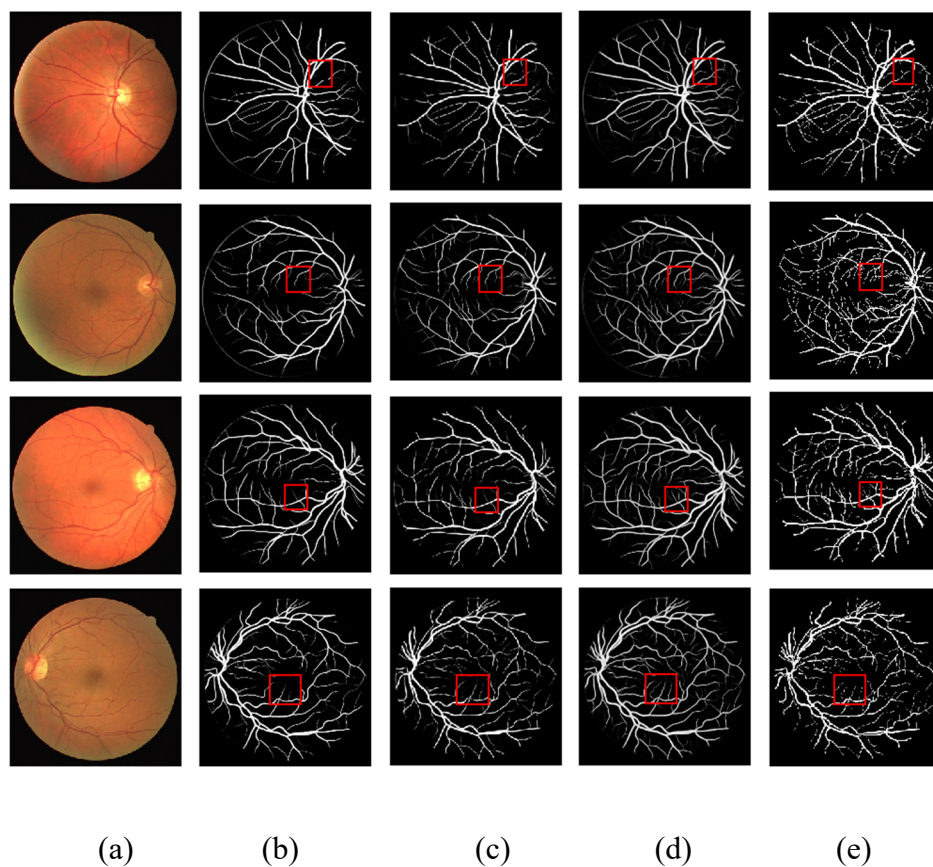


Figure 11. The segmentation results of the model without and with pseudo-label learning.

4.5. Comparison and analysis of the performance of MSFST-Net

In this section, we first test the performance of our proposed overall segmentation method MSFST-Net. Then, we perform an ablation experiment to quantitatively analyze the validity of each component in MSFST-Net. Finally, we test the effect of different parameters in the style transfer module on the segmentation performance.

4.5.1. Comparison of segmentation results

In this experiment, to verify the ability of the method to deal with cross-domain segmentation problems CHASE_DB1 is utilized as the source domain dataset, DRIVE is used as the target domain dataset, and we compare MSFST-Net with some cross-domain retinal vessel segmentation methods with superior performances in recent years. The adopted evaluation metrics include AUC, F1-Score, and MCC. The comparison results are listed in Table 3. From Table 3, we can find that MSFST-Net achieves 97.65%, 81.27% and 78.68% for the AUC, F1-Score and MCC, respectively, which is generally better than the other methods. The abovementioned results prove that our proposed MSFST-Net outperforms the compared methods.

It should be mentioned that although the evaluation metrics obtained by our proposed model are not significantly higher than those of other methods, our proposed model has fewer parameters than other models. Explicitly, most of the comparison models obtained good segmentation results by designing relatively complex network structures that have more parameters while we designed a relatively lightweight segmentation model, and our model can generally achieve an improved segmentation accuracy than that of these complex networks.

Table 3. Comparison of segmentation results of different methods.

Method	AUC	F1-Score	MCC
DANN [55]	97.21%	77.89%	76.02%
DRCN [56]	97.17%	79.17%	77.37%
DRNN [27]	96.80%	79.62%	77.92%
FD [57]	96.88%	79.13%	77.48%
KDFD [57]	97.13%	80.30%	78.57%
AMCD [58]	96.91%	78.60%	-
MSFST-Net	97.65%	81.27%	78.68%

4.5.2. Comparison of the number of parameters

Table 4 lists the number of parameters (Params) of the different methods and our proposed method. Fewer parameters indicate that the model is more lightweight. In practical applications, our proposed method trains Cycle-GAN offline before segmentation, so the main parameter number of our method in the training process is the parameter number of MSF-Net. From Table 4, we can find that our proposed method is the most lightweight. For example, the number of parameters of FD is 114.58 M, while the number of parameters of our proposed model is approximately 0.11 M. Although we propose a lightweight model, our model can generally achieve a better segmentation accuracy than other methods, as shown in Tables 1–3. In addition, for blood vessel segmentation, the average time to infer and save segmentation results per image is a total of 175 ms in our method, which can meet practical applications in medical scenarios.

Table 4. Comparison of the number of parameters.

Method	Params
U-Net [7]	35 M
SegR-Net [41]	0.64 M
FD [57]	114.58 M
Ours	0.11 M

4.5.3. Ablation experiment results

To objectively verify the performance of each module in MSFST-Net, we test the influence of each module in MSFST-Net on the segmentation performance when handling retinal vessel segmentation with a domain shift problem. The specific results are listed in Table 5.

Table 5. Ablation experiment results.

Method	AUC	Acc	F1-Score	MCC
Baseline	96.61%	94.89%	76.97%	74.65%
Baseline+Pseudo-label	96.63%	94.98%	78.20%	73.75%
Baseline+Cycle-GAN	96.99%	95.03%	77.79%	75.42%
Baseline+Cycle-GAN+Pseudo-label (MSFST-Net)	97.65%	95.45%	81.27%	78.68%

In Table 5, the baseline is the optimal MSF-Net model trained on the CHASE_DB1 dataset, and Baseline+Pseudo-label refers to the model that introduces the pseudo-label learning strategy into the baseline. Baseline+Cycle-GAN refers to the model that introduces the style transfer module into the baseline. Baseline+Cycle-GAN+Pseudo-label is the model that introduces both style transfer and pseudo-label learning into the baseline, that is, the final proposed method MSFST-Net. From Table 4, we can summarize the following points. First, after adding the pseudo-label learning strategy to the baseline, most of the evaluation indicator values are increased, which demonstrates that the pseudo-label learning strategy can significantly strengthen the segmentation performance. Second, after adding the style transfer module to the baseline model, most of the evaluation indicators are significantly improved, which means that the style transfer module can effectively reduce the style differences between the different datasets, thus further increasing the segmentation accuracy. Finally, after simultaneously introducing the style transfer module and pseudo-label learning strategy into the baseline model, the values of all evaluation indicators are significantly increased, which verifies that the introduction of the style transfer module and the pseudo-label learning strategy simultaneously have obvious roles in improving the segmentation performance.

4.5.4. Style transfer module parameter testing

The iteration times of the style transfer module in the training process have a great influence on the segmentation results. Therefore, we carry out different iterations when transferring the style of the image and compare the segmentation results obtained by the different iterations, as shown in Table 6.

In Table 6, the baseline is the MSF-Net model.

From Table 6, we can find the following:

1) Compared to the baseline, the values of AUC, Acc and F1-score slightly decrease when the iteration time is 200. The reason for this phenomenon is that Cycle-GAN cannot converge well under 200 iterations; thus, the quality of transferred images does not meet the demand, which further leads to the deterioration of the AUC, Acc and F1-score.

2) Baseline+Cycle-GAN performs better than the baseline when the number of iterations is 300 and 400, and Baseline+Cycle-GAN obtains optimal results in 400 iterations.

3) With an increase in the number of iterations (such as 600 and 800), the performance of Baseline+Cycle-GAN shows a slight decline again, which is due to the excessive number of iterations leading to overfitting of the network, which further leads to a decline in the segmentation performance.

Therefore, 400 is the optimal number of iterations and we select Cycle-GAN trained by 400 iterations as the style transfer module in our work.

Table 6. Comparison of segmentation results under the different iteration times.

Method	AUC	Acc	F1-score	Se
Baseline	96.61%	94.89%	76.97%	69.79%
Baseline+Cycle-GAN (200)	96.44%	94.75%	76.84%	70.34%
Baseline+Cycle-GAN (300)	96.78%	94.94%	77.36%	70.64%
Baseline+Cycle-GAN (400)	96.99%	95.03%	77.79%	71.27%
Baseline+Cycle-GAN (600)	96.62%	94.22%	76.94%	69.33%
Baseline+Cycle-GAN (800)	96.64%	94.77%	76.56%	69.85%

To more intuitively show the influence of different iteration times, we further visualize the style transfer results, as shown in Figure 12. From Figure 12, we can observe that the segmentation accuracy gradually improves after gradually increasing the number of training iterations. The preliminary style transfer results are obtained after 200 iterations; however, at this time, the image after the style transfer is missing many details, and the blood vessels are blurred. Conversely, when the number of iterations increases to 400, the blood vessels in the image gradually become clear.

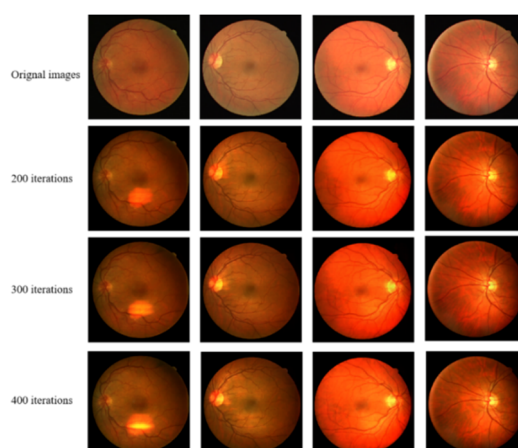


Figure 12. Results of style transfer with different numbers of iterations.

5. Conclusions

In this paper, a novel MSFST-NET was proposed for segmenting blood vessels, which first introduces SK to construct a new segmentation model MSF-Net to increase the model's segmentation ability for small blood vessels in fundus images. Then, to alleviate segmentation precision degradation caused by domain shift, we introduced a style transfer module and a pseudo-label learning strategy into MSFST-NET. The style transfer module is used to reduce the style difference between the different domain images to improve the segmentation performance. The pseudo-label learning strategy is combined with the style transfer method to further boost the generalization ability of the model. In the experiment, we adopted two datasets (DRIVE and CHASE_DB1) to test the performance of the proposed method. The experimental results verify that the proposed MSFST-NET has a good performance on the retinal vessel segmentation task, and style transfer and pseudo-label learning can strengthen the generalization ability of the model when dealing with cross-domain datasets that have the issue of domain shift.

In this paper, a limitation of our work is that it only focused on 2D medical images. In future work, we aim to explore and assess the performance of MSFST-NET in the context of 3D medical images. This extension will be a crucial step in evaluating the model's applicability in more complex medical imaging scenarios. Moreover, we plan to develop a network based on the large foundation model for medical image segmentation.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the Fund of the Jilin Provincial Science and Technology Department (No. 20220201157GX), the National Natural Science Foundation of China (No. 62272096), the Fund of Education Department of Jilin Province (No. JJKH20241729KJ), the Scientific Research Project of Jilin Animation Institute (Nos. KY22KZ01, KY21KZ04), the Outstanding Youth Project of Jiangxi Natural Science Foundation (No. 20212ACB212003), the Jiangxi Province Key Subject Academic and Technical Leader Funding Project (No. 20212BCJ23017) and the Qingpu District Industry University Research Cooperation Development Foundation (No. 202314).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. B. Swenor, V. Varadaraj, M. J. Lee, H. Whitson, P. Ramulu, World Health Report on vision: Aging implications for global vision and eye health, *Innovation Aging*, 4 (2020), 807–808. <https://doi.org/10.1093/geroni/igaa057.2933>

2. T. Li, W. Bo, C. Hu, H. Kang, H. Liu, K. Wang, et al., Applications of deep learning in fundus images: A review, *Med. Image Anal.*, **69** (2021), 101971. <https://doi.org/10.1016/j.media.2021.101971>
3. J. Wang, L. Zhou, Z. Yuan, H. Wang, C. Shi, MIC-Net: Multi-scale integrated context network for automatic retinal vessel segmentation in fundus image, *Math. Biosci. Eng.*, **20** (2023), 6912–6931. <https://doi.org/10.3934/mbe.2023298>
4. J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, M. J. Cree, Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification, *IEEE Trans. Med. Imaging*, **25** (2006), 1214–1222. <https://doi.org/10.1109/TMI.2006.879967>
5. J. I. Orlando, M. Blaschko, Learning fully-connected CRFs for blood vessel segmentation in retinal images, in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, Springer, (2014), 634–641.
6. U. Nguyen, A. Bhuiyan, L. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, *Pattern Recognit.*, **46** (2013), 703–715. <https://doi.org/10.1016/j.patcog.2012.08.009>
7. O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, Springer, (2015), 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
8. Y. He, H. Sun, Y. Yi, W. Chen, J. Kong, C. Zheng, Curv-Net: Curvilinear structure segmentation network based on selective kernel and multi-Bi-ConvLSTM, *Med. Phys.*, **49** (2022), 3144–3158. <https://doi.org/10.1002/mp.15546>
9. B. Wang, S. Qiu, H. He, Dual encoding u-net for retinal vessel segmentation, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019*, Springer, (2019), 84–92. https://doi.org/10.1007/978-3-030-32239-7_10
10. S. Zhang, H. Fu, Y. Yan, Y. Zhang, Q. Wu, M. Yang, et al., Attention guided network for retinal image segmentation, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019*, Springer, (2019), 797–805. https://doi.org/10.1007/978-3-030-32239-7_88
11. Y. Wu, Y. Xia, Y. Song, D. Zhang, D. Liu, C. Zhang, et al., Vessel-Net: Retinal vessel segmentation under multi-path supervision, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019*, Springer, (2019), 264–272. https://doi.org/10.1007/978-3-030-32239-7_30
12. M. Shin, Z. Peng, H. J. Kim, S. S. Yoo, K. Yoon, Multivariable-incorporating super-resolution residual network for transcranial focused ultrasound simulation, *Comput. Methods Programs Biomed.*, **237** (2023), 107591. <https://doi.org/10.1016/j.cmpb.2023.107591>
13. S. Suh, S. Cheon, W. Choi, Y. Chung, W. Cho, J. Paik, et al., Supervised segmentation with domain adaptation for small sampled orbital CT images, *J. Comput. Des. Eng.*, **9** (2022), 783–792. <https://doi.org/10.1093/jcde/qwac029>
14. J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2015), 3431–3440. <https://doi.org/10.1109/CVPR.2015.7298965>
15. O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, et al., Attention u-net: Learning where to look for the pancreas, *arXiv preprint*, (2018), arXiv:1804.03999. <https://doi.org/10.48550/arXiv.1804.03999>

16. Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation, in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, (2018), 3–11. https://doi.org/10.1007/978-3-030-00889-5_1
17. J. Son, S. J. Park, K. H. Jung, Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks, *J. Digital Imaging*, **32** (2019), 499–512. <https://doi.org/10.1007/s10278-018-0126-3>
18. H. Zhao, H. Li, S. Maurer-Stroh, Y. Guo, Q. Deng, L. Cheng, Supervised segmentation of un-annotated retinal fundus images by synthesis, *IEEE Trans. Med. Imaging*, **38** (2018), 46–56. <https://doi.org/10.1109/TMI.2018.2854886>
19. K. B. Park, S. H. Choi, J. Y. Lee, M-GAN: Retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks, *IEEE Access*, **8** (2020), 146308–146322. <https://doi.org/10.1109/ACCESS.2020.3015108>
20. C. Yue, M. Ye, P. Wang, D. Huang, X. Lu, SRV-GAN: A generative adversarial network for segmenting retinal vessels, *Math. Biosci. Eng.*, **19** (2022), 9948–9965. <https://doi.org/10.3934/mbe.2022464>
21. A. Galdran, A. Anjos, J. Dolz, H. Chakor, H. Lombaert, I. Ben Ayed, The little w-net that could: state-of-the-art retinal vessel segmentation with minimalistic models, *arXiv preprint*, (2020), arXiv:2009.01907. <https://doi.org/10.48550/arXiv.2009.01907>
22. Y. Xia, D. Yang, Z. Yu, F. Liu, J. Cai, L. Yu, et al., Uncertainty-aware multi-view co-training for semi-supervised medical image segmentation and domain adaptation, *Med. Image Anal.*, **65** (2020), 101766. <https://doi.org/10.1016/j.media.2020.101766>
23. S. Chen, X. Jia, J. He, Y. Shi, J. Liu, Semi-supervised domain adaptation based on dual-level domain mixing for semantic segmentation, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2021), 11018–11027. <https://doi.org/10.1109/CVPR46437.2021.01087>
24. Z. Wang, Y. Wei, R. Feris, J. Xiong, W. Hwu, T. S. Huang, et al., Alleviating semantic-level shift: A semi-supervised domain adaptation method for semantic segmentation, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, (2020), 936–937. <https://doi.org/10.1109/CVPRW50498.2020.00476>
25. X. Liu, F. Xing, N. Shusharina, R. Lim, C. Kuo, G. El Fakhri, et al., Act: Semi-supervised domain-adaptive medical image segmentation with asymmetric co-training, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (2022), 66–76. https://doi.org/10.1007/978-3-031-16443-9_7
26. Y. Chen, X. Ouyang, K. Zhu, G. Agam, Semi-supervised domain adaptation for semantic segmentation, in *2022 26th International Conference on Pattern Recognition (ICPR)*, (2022), 230–237. <https://doi.org/10.1109/ICPR56361.2022.9956524>
27. J. Wang, C. Zhong, C. Feng, J. Sun, Y. Yokota, Feature disentanglement for cross-domain retina vessel segmentation, in *2021 IEEE International Conference on Image Processing (ICIP)*, (2021), 26–30. <https://doi.org/10.1109/ICIP42928.2021.9506124>
28. Y. Zuo, H. Yao, C. Xu, Category-level adversarial self-ensembling for domain adaptation, in *2020 IEEE International Conference on Multimedia and Expo (ICME)*, (2020), 1–6. <https://doi.org/10.1109/ICME46284.2020.9102756>

29. Y. Xu, B. Du, L. Zhang, Q. Zhang, G. Wang, L. Zhang, Self-ensembling attention networks: Addressing domain shift for semantic segmentation, in *Proceedings of the AAAI Conference on Artificial Intelligence*, (2019), 5581–5588. <https://doi.org/10.1609/aaai.v33i01.33015581>
30. H. Lei, W. Liu, H. Xie, B. Zhao, G. Yue, B. Lei, Unsupervised domain adaptation based image synthesis and feature alignment for joint optic disc and cup segmentation, *IEEE J. Biomed. Health. Inf.*, **26** (2021), 90–102. <https://doi.org/10.1109/JBHI.2021.3085770>
31. S. Wang, L. Yu, X. Yang, C. W. Fu, P. A. Heng, Patch-based output space adversarial learning for joint optic disc and cup segmentation, *IEEE Trans. Med. Imaging*, **38** (2019), 2485–2495. <https://doi.org/10.1109/TMI.2019.2899910>
32. P. Liu, B. Kong, Z. Li, S. Zhang, R. Fang, CFEA: Collaborative feature ensembling adaptation for domain adaptation in unsupervised optic disc and cup segmentation, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019*, Springer, (2019), 521–529. https://doi.org/10.1007/978-3-030-32254-0_58
33. J. Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in *Proceedings of the IEEE International Conference on Computer Vision*, (2017), 2223–2232.
34. X. Li, W. Wang, X. Hu, J. Yang, Selective kernel networks, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2019), 510–519.
35. X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, (2011), 315–323.
36. S. Santurkar, D. Tsipras, A. Ilyas, A. Madry, How does batch normalization help optimization, *Adv. Neural Inf. Process. Syst.*, **31** (2018).
37. S. Jadon, A survey of loss functions for semantic segmentation, in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, (2020), 1–7. <https://doi.org/10.1109/CIBCB48159.2020.9277638>
38. J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, B. Van Ginneken, Ridge-based vessel segmentation in color images of the retina, *IEEE Trans. Med. Imaging*, **23** (2004), 501–509. <https://doi.org/10.1109/TMI.2004.825627>
39. M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, et al., An ensemble classification-based approach applied to retinal blood vessel segmentation, *IEEE Trans. Biomed. Eng.*, **59** (2012), 2538–2548. <https://doi.org/10.1109/TBME.2012.2205687>
40. J. Wu, Y. Liu, Y. Zhu, Z. Li, Atrous residual convolutional neural network based on U-Net for retinal vessel segmentation, *PIOS ONE*, **17** (2022), e0273318. <https://doi.org/10.1371/journal.pone.0273318>
41. J. Ryu, M. U. Rehman, I. F. Nizami, K. T. Chong, SegR-Net: A deep learning framework with multi-scale feature fusion for robust retinal vessel segmentation, *Comput. Biol. Med.*, **163** (2023), 107132. <https://doi.org/10.1016/j.combiomed.2023.107132>
42. Z. Yan, X. Yang, K. T. Cheng, Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation, *IEEE Trans. Biomed. Eng.*, **65** (2018), 1912–1923. <https://doi.org/10.1109/TBME.2018.2828137>
43. K. Hu, Z. Zhang, X. Niu, Y. Zhang, C. Cao, F. Xiao, et al., Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function, *Neurocomputing*, **309** (2018), 179–191. <https://doi.org/10.1016/j.neucom.2018.05.011>

44. S. Guo, K. Wang, H. Kang, Y. Zhang, Y. Gao, T. Li, BTS-DSN: Deeply supervised neural network with short connections for retinal vessel segmentation, *Int. J. Med. Inf.*, **126** (2019), 105–113. <https://doi.org/10.1016/j.ijmedinf.2019.03.015>
45. Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, et al., Ce-net: Context encoder network for 2d medical image segmentation, *IEEE Trans. Med. Imaging*, **38** (2019), 2281–2292. <https://doi.org/10.1109/TMI.2019.2903562>
46. J. Hu, H. Wang, S. Gao, M. Bao, T. Liu, Y. Wang, et al., S-unet: A bridge-style u-net framework with a saliency mechanism for retinal vessel segmentation, *IEEE Access*, **7** (2019), 174167–174177. <https://doi.org/10.1109/ACCESS.2019.2940476>
47. L. Li, M. Verma, Y. Nakashima, H. Nagahara, R. Kawasaki, Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks, in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, (2020), 3656–3665.
48. W. Wang, J. Zhong, H. Wu, Z. Wen, J. Qin, Rvseg-net: An efficient feature pyramid cascade network for retinal vessel segmentation, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference*, (2020), 796–805. https://doi.org/10.1007/978-3-030-59722-1_77
49. Y. Cheng, M. Ma, L. Zhang, C. Jin, L. Ma, Y. Zhou, Retinal blood vessel segmentation based on Densely Connected U-Net, *Math. Biosci. Eng.*, **17** (2020), 3088–3108. <https://doi.org/10.3934/mbe.2020175>
50. X. F. Du, J. S. Wang, W. Sun, UNet retinal blood vessel segmentation algorithm based on improved pyramid pooling method and attention mechanism, *Phys. Med. Biol.*, **66** (2021), 175013. <https://doi.org/10.1088/1361-6560/ac1c4c>
51. C. Z. Wu, J. Sun, J. Wang, L. F. Xu, S. Zhan, Encoding-decoding network with pyramid self-attention module for retinal vessel segmentation, *Int. J. Autom. Comput.*, **18** (2021), 973–980. <https://doi.org/10.1007/s11633-020-1277-0>
52. J. Zhuang, LadderNet: Multi-path networks based on U-Net for medical image segmentation, *arXiv preprint*, (2018), arXiv:1810.07810. <https://doi.org/10.48550/arXiv.1810.07810>
53. X. Lü, F. Shao, Y. Xiong, W. Yang, Retinal vessel segmentation method based on two-stream networks, *Acta Opt. Sin.*, **40** (2020), 0410002. <https://doi.org/10.3788/AOS202040.0410002>
54. Q. Fu, S. Li, X. Wang, MSCNN-AM: A multi-scale convolutional neural network with attention mechanisms for retinal vessel segmentation, *IEEE Access*, **8** (2020), 163926–163936. <https://doi.org/10.1109/ACCESS.2020.3022177>
55. Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, et al., Domain-adversarial training of neural networks, *J. Mach. Learn. Res.*, **17** (2016), 2096–2030.
56. M. Ghifary, W. B. Kleijn, M. Zhang, D. Balduzzi, W. Li, Deep reconstruction-classification networks for unsupervised domain adaptation, in *Computer Vision–ECCV 2016*, Springer, (2016), 597–613. https://doi.org/10.1007/978-3-319-46493-0_36
57. J. Wang, C. Zhong, C. Feng, Y. Zhang, J. Sun, Y. Yokota, Disentangled representation for cross-domain medical image segmentation, *IEEE Trans. Instrum. Meas.*, **72** (2022), 1–15. <https://doi.org/10.1109/TIM.2022.3221131>

-
58. J. Zhuang, Z. Chen, J. Zhang, D. Zhang, Z. Cai, Domain adaptation for retinal vessel segmentation using asymmetrical maximum classifier discrepancy, in *Proceedings of the ACM Turing Celebration Conference-China*, (2019), 1–6. <https://doi.org/10.1145/3321408.3322627>



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)