Journal of Imaging Science and Technology® 66(4): 040408-1–040408-8, 2022.

Improved U-Net Fundus Image Segmentation Algorithm Integrating Effective Channel Attention

Junhua Liang[†], **Lihua Ding**[†], **Xuming Tong, and Zhisheng Zhao**
College of Information Science and Engineering, Hebei North University, Zhangjiakou, China
E-mail: zhaozhsh@hebeinu.edu.cn

Jie Li
Department of Pathology, China-Japan Friendship Hospital, Beijing, China

Junqiang Liang
BeiJing HealthOLight Technology Co., Ltd, Beijing, China

Beibei Dong and Yanhong Yuan
College of Information Science and Engineering, Hebei North University, Zhangjiakou, China

Abstract. *Fundus blood vessel segmentation is important to obtain the early diagnosis of ophthalmic- related diseases. A great number of approaches have been published, yet micro-vessel segmentation is still not able to deliver the desired results. In this paper, an improved retinal segmentation algorithm incorporating an effective channel attention (ECA) module is presented. Firstly, the ECA module is imported into the downsampling stage of a U-shape neural network (U-Net) to capture the cross-channel interaction information. Secondly, a dilated convolutional module is added to expand the receptive field of the retina, so that more micro-vessel features can be extracted. Experiments were performed on two publicly available datasets, namely DRIVE and CHASE_DB1. Finally, the improved U-Net was used to validate the results. The proposed method achieves high accuracy in terms of the dice coefficient, mean pixel accuracy (mPA) metric and the mean intersection over union (mIoU) metric. The advantages of the algorithm include low complexity and having to use fewer parameters.*
[DOI: 10.2352/J.ImagingSci.Technol.2022.66.4.040408]

1. INTRODUCTION

Retinal vascular images play an important role in attempting to diagnose fundus diseases [1]. Consequently, these images are used in computer-assisted diagnoses and artificial intelligence in the medical field [2–4]. Fundus image segmentation assists with primary filtering of patients with retinopathy, hypertension, diabetes, and other related diseases [5]. However, various problems still exist, including diverse morphology of fundus blood vessels, the existence of bleeding point exudates, insufficient fundus image resolution, and low micro-vascular endothelium. Thus, the topic of improving the accuracy of fundus image segmentation has attracted a great number of scholars. Retinal blood vessel segmentation methods include tracking detection based on blood vessel orientation, segmentation techniques based on mathematical morphology methods and matched filtering, segmentation techniques using deformation models, and machine learning [6–9]. With deep learning rapidly advancing in the medical field, retinal blood vessel segmentation has also achieved remarkable results [10–13]. The residual network model combined with the DenseNet can assist in learning more robust morphological structure information. However, the excessive use of DenseNet may take up too much computing resources and lead to an overly complex network [14]. A fundus image segmentation based on an improved U-Net model has been proposed algorithms [15] to replace the convolutional layer serial connection mode of the U-Net model with the superposition method of residual mapping. This replacement improves the efficiency of feature use by reducing the dimension and strengthening the information fusion of high-dimensional features and low-dimensional features. A study by Li [16] introduced the attention mechanism in the decoding structure, and took advantage of this combination to decouple the features and map them to the low-dimension space to restore and extract small blood vessels of the fundus retinal image. Detailed features will improve the accuracy of segmentation, but the process of feature dimension reduction will negatively impact the channel attention prediction and destroy the direct correspondence between module channels and their weights. Some scholars, from the perspective of the receptive field, integrate the hollow convolutional network and DenseNet into the U-Net structure. This increases the overall perception of the network and improves the network's ability to repeatedly use features while providing play to the dense connection ratio between layers [17]. Traditional networks have the advantage of producing fewer output dimensions and avoiding to learn redundant target feature information. However, all aforementioned methods involve

[†]These authors contributed equally to this work.

Received Feb. 24, 2022; accepted for publication Apr. 2, 2022; published online May 6, 2022. Associate Editor: Ping Wang.
1062-3701/2022/66(4)/040408/8/$25.00

the reduction of dimension, less channel information in the network structure, and lack of interaction between local cross-channel information.

To resolve the problem of the complexity of the structure of fundus images and meet the need to improve the accuracy of fine segmentation, this study integrates the Efficient Channel Attention (ECA) module based on the traditional U-Net, and proposes an improved U-Net. The improved U-Net has two main features. Firstly, the ECA module is imported in the down sampling stage so that feature mapping can be quickly and effectively realized through one-dimensional convolution. Once the ECA module is added, it can avoid dimension reduction, appropriately capture cross-channel interactive information, improve the performance of the network, and make the features of the retinal blood vessels more apparent. Secondly, the U-Net contraction path is capable of extracting the features of the region of interest. However, it is likely to lose the detailed feature information of the image with frequent contraction. Moreover, it is difficult to restore the loss in the expansion path. To extract more detailed vessel features, a dilated convolution module is added to the network in the contraction and the expansion path to expand the receptive field of the convolution operation without providing additional network parameters. The accuracy of segmentation was improved and more features of microvascular vessels were retained by combining the ECA with dilated convolution, which was verified on public datasets.

2. MATERIALS AND METHODS

2.1 *Materials*

For this study, two established open-access retinal image datasets were available, namely the Digital Retinal Images for Vessel Extraction (DRIVE) and Child Heart and Health Study in England Databases (CHASE_DB1). These databases were used in the experiment to verify the effectiveness of the algorithm. The DRIVE dataset was established and published in 2004 by Niemeijer et al. [18] for the screening of diabetic retinopathy in the Netherlands. The original images were from 453 subjects aged 31–86 years, including seven pathologic images of exudate, hemorrhage, and pigment epithelial cells. The entire dataset consists of 40 colored fundus images and their corresponding labeled images. Each colored fundus image is 565×584 pixels in size, and each image contains labeled results manually segmented by two expert groups.

The CHASE_DB1 dataset includes 28 retinal images of the left and right eyes of 14 school-aged multi-ethnic children with an image resolution of 999×960 pixels. Each image also contains the labeled results of two experts' manual segmentations [19]. The images from this dataset have an uneven background light, low vascular contrast, a wide artery and a bright vascular reflection stripe in the center.

In this study, the improved U-Net (after adding an efficient channel module, named ECA-Unet), can optimize the network and gain accuracy from inter-channel information. Python 3.6 was used as the development integrated environment. The specific experimental environment configuration is shown in Table I. To solve the problem of an insufficient sample size, the available fundus images were enhanced. This was achieved by random up and down rotation, left and right rotation, noise addition and other operations to prevent overfitting. The learning rate of experimental iteration training was set at 0.001. The loss function was cross-entropy, with the epoch and batch_size set as 200 and 4, respectively. To fit the data more easily, the Kaiming method was adopted for weight initialization [20].

Table I. Experimental environment configuration.

Experimental Environment	Configuration
The operating system	Ubuntu 18
Linux Graphics card	RTX 2080 Ti
Processor	Interl© CoreTM i7-5500H CPU (16GB)
Learning framework	Pytorch
IDE	Pycharm

2.2 *Methods*

The overall framework of the algorithm based on U-Net [21] is shown in Figure 1. The improved ECA-Unet is composed of a contraction path (left) and an expansion path (right). The left and right sides each contain four mutually symmetric structural blocks. The blue bar box in each structural block corresponds to the multi-channel feature map, which represents 3×3 convolution. The pink, red, orange, and green bar boxes correspond to 3×3 dilated convolution, downsampling operations, upsampling operations, and deconvolution, respectively.

The left-hand contraction path follows the typical architecture of the convolutional network and consists of two 3×3 convolutions. The second convolution is a dilated convolution with an expansion rate of 2. After integrating the ECA module with the Rectified Linear Unit (ReLU) as the activation function, a 2×2 maximum pooling operation was used in step 2 for the subsampling, as shown in Figure 2.

Each step in the expansion path of the right half includes the upsampling operation of the feature map, the 2×2 convolution operation and the dilated convolution with an expansion rate of 2. The contraction path feature mapping doubles the number of feature channels and vice versa for the expanding path. At the last level of output, the mapping class of eigenvectors for each component uses a 1×1 convolution, as shown in Figure 3.

2.2.1 *Efficient Channel Attention Module*

Traditional attention modules [22–24] have been used to improve the segmentation accuracy at the expense of model computational complexity or to control the complexity of the model through dimension reduction. However, increasing the complexity will inevitably lead to a greater computational burden. Dimension reduction will adversely affect channel attention prediction, and it is inefficient and unnecessary to

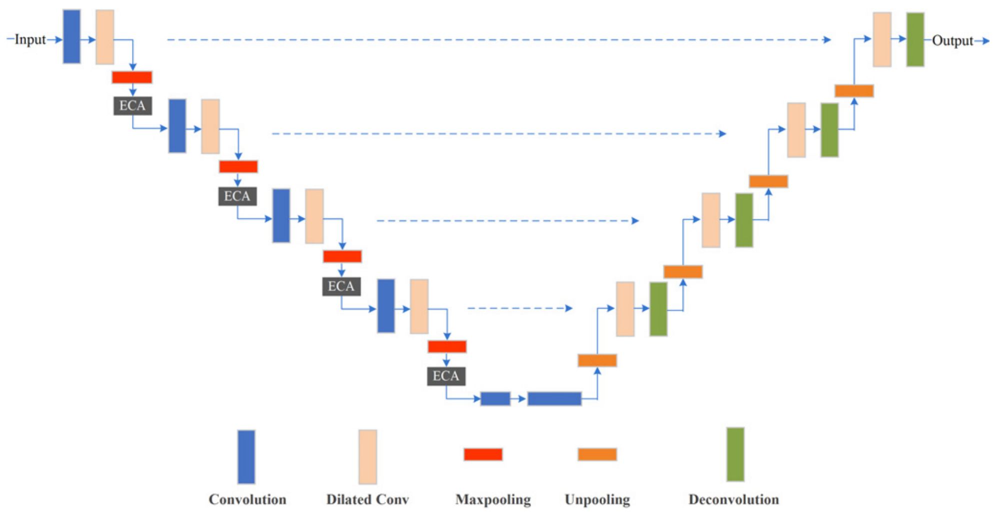

Figure 1. Improved ECA-Unet structural framework.

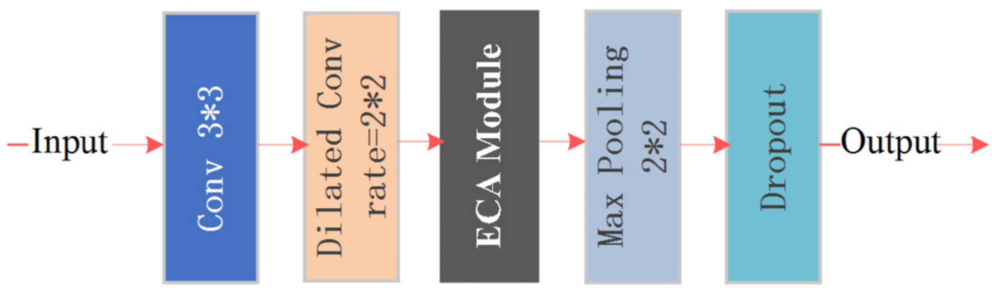

Figure 2. The contraction path structure diagram.

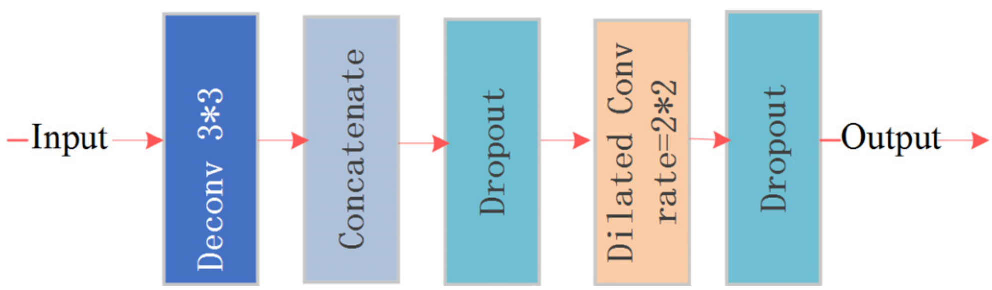

Figure 3. The expansion path structure diagram.

capture the dependencies between all channels. For example, the traditional Squeeze-and-Excitation (SE) module [22] is mainly composed a Global Average Pooling layer (GAP), two Fully Connected layers (FC) and a Sigmoid function. The first FC layer is mainly used for dimension reduction to control the complexity of the model. To solve the contradiction between model performance and complexity, an ECA can use only a few parameters in the calculation process, and bring about significant performance gains [25]. The ECA modules are shown in Figure 4, $\psi \epsilon R^{W*H*C}$ where W, H, C represent the width, height and channel dimension (number of filters), respectively. The variable k is the size of the convolution, representing the coverage of local cross-channel interaction. The variable ψ is the result processed by the ECA, and is the product of elements.

As shown in Fig. 4, and compared to the traditional SE module, the ECA removes the FC layer and learns directly through a 1D convolution that can share weight on the features after GAP. This change reduces the number of parameters of the ECA module from $k \times C$ to C in our framework. This effectively reduces the algorithm complexity while avoiding dimension reduction.

2.2.2 *Dilated Convolution Module*

The receptive field size of the convolutional kernel is determined by its size. To achieve the purpose of fewer parameters, the convolution with a smaller size is generally

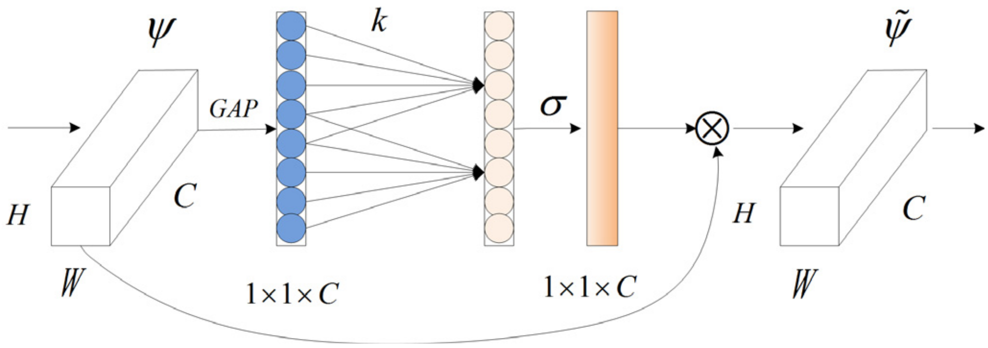

Figure 4. ECA module structure diagram.

chosen, and the range of the receptive field of the feature mapping is increased by subsampling or adding the void convolution [26]. However, the sampling operation includes processing a lot of detailed feature information. Repeated subsampling will lose the resolution of the feature map, resulting in the loss of information, which cannot be recovered by the upsampling operation. Therefore, the dilated convolution was added into the inter-frame convolution. Additionally, the number of downsampling layers was reduced so that the network could have a larger receptive field when extracting intra-frame features. For the dilated convolution of two-dimensional images, the relationship between input x and output y is as follows:

$$y = \sum_{m} x[p + r * m]w[m], \tag{1}$$

where, p represents each pixel point on the feature map, w is the void convolution kernel, m is the size of the dilated convolution kernel, and the distance between adjacent elements in the convolution kernel is represented by the expansion rate r. Following the expansion, the dilated convolution kernel is

$$m' = m + (m - 1)(r - 1). \tag{2}$$

As shown in Figure 5, the convolution kernel is selected for 3×3 dilated convolution, which is still a 3×3 convolution kernel at a dilated rate $r = 1$, When the dilated rate $r = 2$, it is a dilated convolution kernel of $r = 5$. When the dilated rate is $r = 4$, it is a 9×9 dilated convolution kernel. Thus, as the convolution kernel r increases, the interval of adjacent elements in the convolution kernel also increases by a corresponding multiple, which expands the receptive field without significantly increasing the number of parameters. For example, the 3×3 convolution kernel with a dilated rate of 2 has the same receptive field of 5×5 as the convolution kernel, but only uses nine parameters.

In this proved algorithm, the dilated convolution is applied to the convolution layer near the bottom of U-Net network. The appropriate receiver field is obtained by adjusting the expansion rate to reduce the number of downsampling layers and avoid the irreversible loss of detailed features caused by too many downsampling operations.

3. RESULTS

3.1 *Evaluation Indicators*

The Dice coefficient (Dice Score), Mean Pixel Accuracy (mPA) metric and Mean Intersection Over Union (mIoU) metric are the quantitative evaluation indexes for this study, which are also three commonly used indicators in the field of image segmentation [27, 28].

The Dice score is a measurement function to measure the similarity of pixel sets between two samples. The formula is as follows:

$$Dice\ score = \frac{2|X \cap Y|}{|X| + |Y|}, \tag{3}$$

where X is the real value, Y is the predicted value. The Dice score is limited in the range of [0,1]. The mPA can be calculated as the proportion of the correct classification pixels of each category to all pixels of this category, from which the mean value is then obtained:

$$mPA = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij}}. \tag{4}$$

The mIoU is the intersection and union ratio of the mean values, which is used to calculate the intersection and union ratio of the real value and the predicted value. After this calculation, the IoU of each kind is accumulated. The average is then determined to obtain the global evaluation:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}}, \tag{5}$$

where, i is the true value, j is the predicted value. p_{ii} represents that i is predicted to be i. p_{ij} represents that i is predicted to be j. p_{ji} represents that j is predicted to be i.

3.2 *Experimental Results*

The two datasets, DRIVE and CHASE_DB1, were applied to the experiments to prove the feasibility of the proposed method. Figure 6 shows the changing process of the segmentation effect of the algorithm on two different datasets

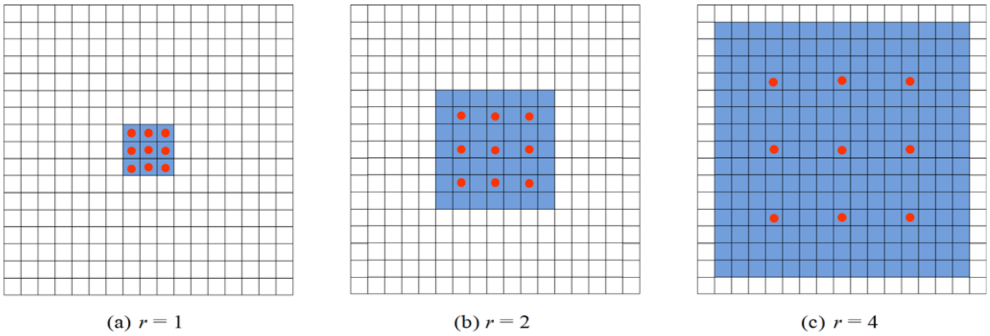

Figure 5. Schematic diagram of dilated convolution with size 3 × 3.

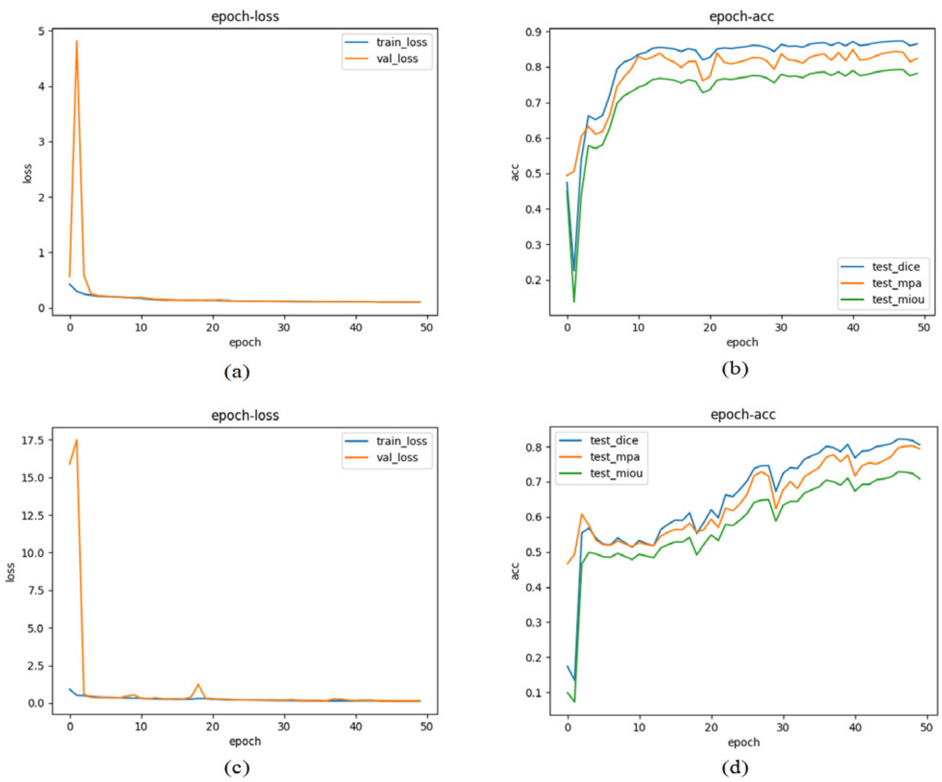

Figure 6. The segmentation effect on two datasets using the improved algorithm. (a) Loss changes along with the increase of epoch in the DRIVE dataset. (b) Three indicators changed along with the increased epoch in the DRIVE dataset. (c) Loss changes along with epoch in CHASE_DB1 dataset. (d) Three indexes changed along with an increased epoch in the CHASE_DB1 dataset.

with the increase of epoch. Fig. 6(a) and (b) show the training and testing process of 50 iterations on the DRIVE dataset, respectively. The pixel error rate was minimized through cross-validation. The loss value decreased gradually and converged to a fixed value. In Fig. 6(b), with the increase of epoch, the three segmentation index values (Dice score, mPA, mIoU) of the test set gradually increased and became stable, indicating that the algorithm design is feasible and effective. The experimental process of the CHASE_DB1 dataset also verifies the feasibility of the algorithm, as shown in Fig. 6(c) and (d).

To further verify the superiority of the network model, the average segmentation coefficient of the trained model was taken after a single sheet test on the test set, the segmentation results obtained are consistent with the verification results in the training process. Contrary to the traditional U-Net, the effectiveness of this method was proven using experiments. To show the superiority of the proposed algorithm intuitively, the segmentation tests were carried out on vessels in two different datasets, and a comparison experiment of the segmentation results between the traditional U-Net and the ECA-Unet was given. The verified effect of adding the ECA model onto the segmentation results. As shown in Figure 7, the ECA-Unet was able to segment smaller blood vessels in the images selected from the DRIVE dataset compared with expert markers. In the images selected from the CHASE_DB1

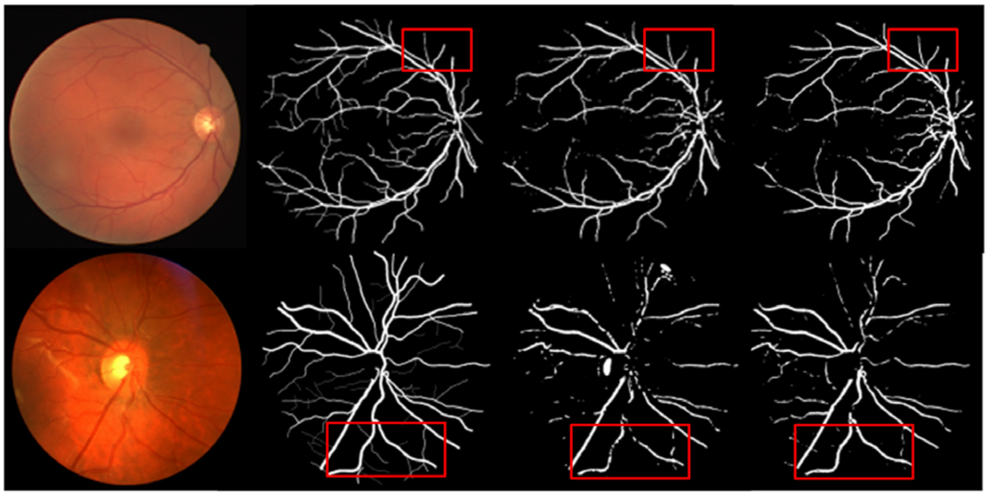

Figure 7. Comparison of the segmentation effect on the test pictures using ECA-Net and U-Net algorithms. Among them, Col. 1: the two original images from the data set; Col. 2: the same images labeled by the first expert; Col. 3: the same images using U-Net algorithm; Col. 4: the same images using ECA-Net algorithm. Lin. 1: the image randomly extracted from the DRIVE data set; Lin. 2: the image randomly extracted from the CHASE_DB1 dataset.

Table II. Comparison of segmentation results using different approaches in the DRIVE dataset.

Evaluation Index	Methods				
	U-Net	Ref. [29]	Ref. [30]	CBAM-Unet	ECA-Unet
Dice score	0.808	0.813	0.829	0.867	0.892
mPA	0.768	0.755	0.798	0.848	0.897
mIoU	0.715	0.720	0.731	0.793	0.818

Table III. Comparison of segmentation results using different approaches in CHASE_DB1 dataset.

Evaluation Index	Methods				
	U-Net	Ref. [29]	Ref. [30]	CBAM-Unet	ECA-Unet
Dice score	0.786	0.806	0.817	0.851	0.883
mPA	0.724	0.765	0.762	0.816	0.857
mIoU	0.690	0.714	0.725	0.765	0.807

dataset, the ECA-Unet was clearer than the coarse blood vessels extracted from the traditional U-Net.

Tables II and III provide the three indexes of Dice score, mPA and mIoU corresponding to vascular segmentation on the DRIVE and CHASE_DB1 datasets, respectively. Based on the gold standard of retinal vascular segmentation in each fundus image by the second ophthalmologist, the ECA-Unet index results were statistically analyzed by comparing it to the traditional U-Net [21], reference [29, 30], and the CBAM-Unet [31]. In Table II, the segmentation effect of the algorithm-added modules is significantly higher than that of the classic U-Net, indicating the necessity of adding modules to improve the segmentation performance. Compared to the attention mechanism CBAM of the convolutional module, the segmentation effect of the ECA module is significantly improved.

Since the CHASE_DB1 dataset had more interference than the DRIVE dataset with uneven background light and low contrast of blood vessels, the value of the three indexes were reduced. However, Table III shows that all indicators of the ECA-Unet are improved compared to those of the traditional U-Net. Additionally, the introduction of the ECA module is more effective than the attention modules from previous studies [14–16].

4. DISCUSSION

Fig. 6 implies that the loss of the proposed segmentation algorithm was convergent with the increase of epoch in two public fundus datasets, DRIVE and CHASE_DB1. This indicates that the U-shaped network designed in this study is effective and feasible. Fig. 7 illustrated that the ECA-Unet can obtain more imperceptible blood vessel characteristics compared to the traditional U-Net. The Dice score, mPA and mIoU metrices, three commonly used quantitative evaluation indexes in the field of image segmentation, were applied in the contrast experiment. The results showed that the ECA-Unet improved the retinal blood vessel segmentation's accuracy for both the DRIVE and CHASE_DB1 dataset. This approach achieved an average Dice score of 0.892 and 0.883, an mPA of 0.897 and 0.857, and an mIoU of 0.818 and 0.807, respectively for the DRIVE and CHASE_DB1 dataset. The results show that this approach is highly effective for retinal images and can improve the efficiency and effectiveness of retinal vessel segmentation. This is mainly due to the ability of an ECA module, imported during the downsampling stage. This approach can capture the local cross-channel information so that feature mapping can be quickly and effectively realized through one-dimensional convolution. Dilated convolution in the process of feature mapping application expands the receptive field of retinal blood vessels, which balances the algorithm complexity and segmentation accuracy. It should be noted that the traditional U-Net, reference [29, 30], CBAM-Unet [31] and the ECA-Unet tends to produce false vessel detection around the optic disk and pathological regions such as dark and bright lesions. This decreases the overall accuracy in the CHASE_DB1 dataset compared to that of in DRIVE dataset. In the future work, we may further verify the robustness of the algorithm against a complex background.

5. CONCLUSION

The aim of this study was to minimize the difficulty of the segmentation of small blood vessels in the colored fundus images. An improved U-shape neural network method was proposed in this paper. Firstly, the ECA module was introduced into the network, which strengthens the cross-channel information of the feature map and greatly reduces the number of network parameters. The U-shaped symmetric structure makes the network easier to optimize. The combination of these two modules further improves the performance of the network. Secondly, the dilated convolution was introduced into the encoder and decoder structure network to accurately extract the blood vessel image while preserving more small blood vessels. Compared to other traditional methods, the algorithm in this paper achieved high accuracy in terms of the Dice score, mPA and mIoU. In addition to greatly reducing network parameters and improving computational efficiency, the proposed algorithm also retained more intact small vessels and shows improved segmentation performance.

ACKNOWLEDGMENT

The authors would like to thank the students of Medical Big Data Research Team for their technical support.

Author Contributions

J. H. Liang, Z. S. Zhao and J. Q. Liang conceived the idea and designed the experiments. J. H. Liang and L. H. Ding contributed equally to this work. X. M. Tong, J. Li, B. B. Dong and Y. H. Yuan conducted the experiments. All authors contributed equally to the writing of the manuscript.

Fundings

Major Foundation Program of Hebei Education Department [Grant number ZD2018241], Youth Foundation Program of Hebei Education Department [Grant number QN2018155], the Funding Project of Science & Technology Research and Development of Hebei North University [Grant number XJ2021002], the Fundamental Research Funds for Provincial Universities in Hebei in 2020 [Grant number JYT2020029] and the Foundation of the Population Health Informatization in Hebei Province Technology Innovation Center.

REFERENCES

[1] D. A. Lotliker, "Survey on diabetic retinopathy detection from retinal images," Int. J. Res. Appl. Sci. Eng. Technol. **10**, 237–241 (2020).

[2] A. T. Soomro, J. Gao, M. T. Khan, F. M. A. Hani, A. U. M. Khan, and M. Paul, "Computerised approaches for the detection of diabetic retinopathy using retinal fundus images: a survey," Pattern Anal. Appl. **20**, 927–961 (2017).

[3] K. Mittal and M. A. V. Rajam, "Computerized retinal image analysis-a survey," Multimedia Tools Appl. **79**, 22389.0–22421.0 (2020).

[4] V. Bellemo, G. Lim, T. H. Rim, G. S. Tan, C. Y. Cheung, S. Sadda, M. G. He, A. Tufail, M. L. Lee, W. Hsu, and D. S. W. Ting, "Artificial intelligence screening for diabetic retinopathy: the real-world emerging application," Curr. Diabetes Rep. **19**, 1–12 (2019).

[5] U. T. Nguyen, A. Bhuiyan, L. A. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," Pattern Recognit. **46**, 703–715 (2013).

[6] J. Dash and N. Bhoi, "A survey on blood vessel detection methodologies in retinal images," *2015 Int'l. Conf. on Computational Intelligence and Networks* (IEEE, Piscataway, NJ, 2015), pp. 166–171.

[7] M. Bansal and N. Singh, "Retinal vessel segmentation techniques: a review," *Lecture Notes in Engineering and Computer Science* (2017), pp. 442–447.

[8] S. Guo, Q. Yang, B. Gao, X. Cai, Y. Zhao, N. Xiao, Y. He, and C. Zhang, "An improved threshold method based on histogram entropy for the blood vessel segmentation," *2017 IEEE Int'l. Conf. on Cyborg and Bionic Systems (CBS)* (IEEE, Piscataway, NJ, 2017), pp. 221–225.

[9] J. Almotiri, K. Elleithy, and A. Elleithy, "Retinal vessels segmentation techniques and algorithms: a survey," Appl. Sci. **8**, 155–155 (2018).

[10] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher, "Deep learning-enabled medical computer vision," NPJ Digit. Med. **4**, 1–9 (2021).

[11] A. Oliveira, S. Pereira, and A. C. Silva, "Retinal vessel segmentation based on fully convolutional neural networks," Expert Syst. Appl. **112**, 229–242 (2018).

[12] Z. Yan, X. Yang, and K.-T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation," IEEE Trans. Biomed. Eng. **65**, 1912–1923 (2018).

[13] C. Wu, Y. Zou, and Y. Liu, "Preliminary study on deep-learning for retinal vessels segmentation," *2020 15th Int'l. Conf. on Computer Science Education (ICCSE)* (IEEE, Piscataway, NJ, 2020), pp. 687–692.

[14] C. Wu, B. Yi, Y. Zhang, S. Huang, and Y. Feng, "Retinal vessel image segmentation based on improved convolutional neural network," Acta Opt. Sinaca **38**, 125–131 (2018).

[15] H. Gao, T. Qiu, Y. Chou, M. Zhou, and X. Zhang, "Blood vessel segmentation of fundus images based on improved U network," Chin. J. Biomed. Eng. **38**, 1–8 (2019).

[16] D. Li and Z. Zhang, "Improved u-net segmentation algorithm for the retinal blood vessel images," Acta Opt. Sin. **40**, 64–72 (2020).

[17] L. Liang, X. Sheng, Z. Lan, G. Yang, and X. Chen, "U-shaped retinal vessel segmentation algorithm based on adaptive scale information," Acta Opt. Sin. **39**, 0810004–0810004 (2019).

[18] M. Niemeijer, J. Staal, B. Ginneken, M. Loog, and M. Abràmoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," Proc. SPIE **5370** (2004).

[19] M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," IEEE Trans. Biomed. Eng. **59**, 2538–2548 (2012).

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," *Proc. IEEE Int'l. Conf. on Computer Vision (ICCV)* (IEEE, Piscataway, NJ, 2015), pp. 1026–1034.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," *Int'l. Conf. on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2015), pp. 234–241.

[22] J. Hu, L. Shen, G. Sun, and S. Albanie, "Squeeze-and-Excitation Networks," IEEE Trans. Pattern Anal. Mach. Intell. 99 (2017).

[23] S. Woo, J. Park, J. Y. Lee, and S. I. Kweon, "Cbam: convolutional block attention module," *Proc. European Conf. on Computer Vision (ECCV)* (Springer, Cham, 2018), pp. 3–19.

[24] Y. Chen, Y. Kalantidis, J. Li, S. Yan, and J. Feng, "A^2-nets: double attention networks," *32nd Conf. on Neural Information Processing Systems (NeurIPS)* (Curran Associates, Inc., Red Hook, NY, 2018), pp. 1–10.

[25] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: efficient channel attention for deep convolutional neural networks," *2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Piscataway, NJ, 2020), pp. 11 531–11 539.

[26] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," Preprint arXiv:1511.07122 (2015).

[27] X. Shen, X. Wang, and Y. Yao, "Spatial frequency divided attention network for ultrasound image segmentation," J. Comput. Appl. **41**, 1828–1835 (2021).

[28] M. O. Berezsky and Y. O. Pitsun, "Evaluation methods of image segmentation quality," Radio Electron. Comput. Sci. Control **1**, 119–128 (2018).

[29] X. Gao, Y. Cai, C. Qiu, and Y. Cui, "Retinal blood vessel segmentation based on the gaussian matched filter and u-net," *2017 10th Int'l. Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (IEEE, Piscataway, NJ, 2017), pp. 1–5.

[30] J. Son, S. J. Park, and K.-H. Jung, "Retinal vessel segmentation in fundoscopic images with generative adversarial networks," Preprint arXiv:1706.09318 (2017).

[31] J. Wang, W. Jin, P. Tang, and J. Hu, "Retinal vessel segmentation based on visual attention enhancement cbam-u-net model," J. Comput. Appl. **37**, 321–323 (2020).