Attention residual convolution neural network based on U-net (AttentionResU-Net) for retina vessel segmentation

Shun Zhao*, Tao Liu*, Bowen Liu, Kun Ruan
Sichuan Agricultural University, College of Science Ya’an, China

*Corresponding author e-mail: liutao@sicau.edu.cn, *m13209488592@163.com

Abstract. Extraction of retinal vessels from fundus images is one of the basic steps for diagnosis of diabetic retinopathy. Although some scholars have proposed several segmentation methods, this segmentation is still challenging due to differences in retinal vascular network and image quality. At present, the main challenges of retinal vascular segmentation are noise (due to uneven light) and capillaries. Due to the high complexity of retinal vascular characteristic information, the existing algorithms exist some problems such as microvascular segmentation and optic disc segmentation. We propose a network segmentation model based on fusion residual block, Attention mechanism and U-net. Firstly, the image was enhanced by limiting contrast histogram equalization. Secondly, Gamma correction is used to improve image brightness information and reduce artifact interference. Finally, AttentionResU-net segmentation model is used for segmentation. This algorithm was tested on DRIVE dataset, and its AUC reached 97.93%.

1. Introduction (Heading 1)
Human retina is a photosensitive tissue with abundant vascular information, and it is the only one with non-invasive and non-traumatic visualization [1], by analyzing the number, Angle, branch and curvature of retinal blood vessels [2], doctors can be assisted in diagnosing patients’ diseases. Therefore, the automatic analysis of retinal vascular system has become a hot topic in the field of medical imaging.

Most existing methods rely on manually marking the differences between vascular and non-vascular pixels. For example, the fundus segmentation method based on multi-scale matching filtering [3], in this method, retinal blood vessels with approximately Gaussian intensity distribution are used to enhance blood vessels before threshing. Although most of the tiny blood vessels are enhanced, there are still some problems such as insufficient segmentation at the intersection of blood vessels and missegmentation of lesions. Soares [4] using 2d-gabor filtering of different scales is an effective alternative to training classifier to realize pixel detection of blood vessels, but this method still has the problem of microvascular segmentation and fracture. Other features include ridge feature [5], moment invariant feature [6], local phase feature [7], COSFIRE filtering [8], etc. Most of the features of blood vessels can be extracted well, but the anti-noise property is poor, which is easy to cause micro vessel segmentation and fracture. Existing retinal micro vessels usually have low contrast, and the difference in strength from the wide vessels is often greater than the change in background. In addition, there are also the effects of optic disc, macula, pathology and artifact. Although the above methods have achieved good segmentation results, However, when these artificially selected features are used to solve these two problems, they are still...
unable to solve the trend of vascular changes and the invariance of vascular information in a relatively robust way, resulting in problems such as insufficient segmentation of micro vessels, and incorrect segmentation of lesions and optic discs.

In recent years, the feature learning method based on deep learning has been widely used in the retinal vein segmentation of fundus. This method is different from the artificial feature extraction method, and it needs to select a better classifier to complete the final vascular segmentation. Deep learning convolutional neural network (CNN) combines feature extraction with classifier and has better generalization ability and robustness [9]. Orlando [10] successfully applies the intensive conditional random field (CRF) model and CNN orientation to retinal vascular segmentation. In this model, remote links are established within the image, so the problem of "shrinkage deviation" is well solved, but the phenomenon of lesion missegmentation still exists. Zhou [11] used CNN to extract vascular features and a set of filters to promote micro vessels. Finally, intensive CRF was used for vascular segmentation, which solved the problem of insufficient segmentation of micro vessels. However, some micro vessels were still broken and easily linked.

2. METHODS

In view of the shortcomings and experience of the above algorithms, this paper proposes a segmentation model of fusion residual block [12] and Attention mechanism based on U-net [13]. Residuals block can help us train deeper, learning networks; meanwhile, we used an attention-gating model for medical imaging. The model can automatically focus on target structures of different shapes and sizes, and the trained model can suppress the irrelevant areas in the input image, and highlight the significant features that are useful for specific tasks, and has the characteristics of low overhead, improved model sensitivity and prediction accuracy. AttentionResU-Net was used to segment the processed image, which solved the problems of insufficient segmentation of small blood vessels and wrong segmentation of optic disc and lesion. The overall flow chart of this paper is shown in figure 1.

Figure 1. Overall flow chart of the algorithm.
2.1. Image Preprocessing

2.1.1. Gray Level Conversion: There are many methods for grayscale conversion, some of which are based on a color channel, while others conduct grayscale conversion by multiplying the three channels by a coefficient respectively. In order to standardize the intensity value, the sum of the three coefficients must be 1. \((c_r + c_g + c_b = 1)\) The formula is as follows:

\[
I = c_r * R + c_g * G + c_b * B,
\]

Because the contrast of green channel is relatively high, in our paper, the three color channel coefficients are respectively \(c_r=0.3\), \(c_g=0.6\), \(c_b=0.1\).

2.1.2. Normalization: We use z-score standardization, which can convert two or more groups of data into unit-free z-score, so as to unify the data standard. The formula is as follows:

\[
z = \frac{x - \mu}{\sigma}
\]

\(\mu\) The mean of the sample data, \(\sigma\) the standard deviation of the sample data. After standardization, the optimal parameters can be obtained more easily to accelerate the convergence.

2.1.3. CLAHE: In CLAHE, contrast limits are mandatory for every small area. CLAHE is trying to limit the size of the histogram by clipping it with predefined thresholds before calculating the CDF. The slope of the CDF is limited, and therefore the slope of the transformation function. The value of Histogram reduction depends on the distribution of histogram, and therefore also depends on the value of domain size. In general, it is not good to simply ignore the parts that are outside the clipping limit of the histogram. Instead, the clipped parts should be distributed evenly to the rest of the histogram. After equalization, we use bilinear interpolation to remove defects in the frame of the block.

2.1.4. Gamma adjustment: Gamma adjustment increases the representation accuracy of darker values, while decreases the representation accuracy of brighter values. According to different pixel characteristics of blood vessels and background, gamma adjustment is carried out on different areas of retinal images respectively, so as to improve the pictures quality and suppress the influence of uneven illumination and other factors.

2.2. AttentionResU-net model

the existing U-Net network is prone to fracture of microvascular segmentation and complexity of retinal fundus image features, Attention mechanism and residual block are introduced to suppress the irrelevant areas in the image, and meanwhile to highlight the prominent features, the repeated utilization rate of network information is increased, so that more vascular information can be segmented in the u-shaped network segmentation model, so as to achieve the optimal segmentation of retinal image training with fewer samples.
Figure 2. $x^l$ is used to scale attention coefficient $\alpha$. The spatial region is selected by analysing the activation and context information provided by the gated signal collected at large scale. Finally, three linear interpolation is used to complete the grid resampling of the attention coefficient.

This combination brings us three benefits: 1) the residuals unit makes our network easier to train. 2) The residual element and the leaping connections between shallow and deep layers of the network facilitate the dissemination of information, so we can design neural networks with fewer parameters and stronger learning ability.

Figure 3. Architecture of the AttentionResU-Net

This model solves the interference of the existing lesions and artifacts in the retinal images of the fundus, so as to retain more information about the characteristics of blood vessels and obtain more tiny
blood vessels. Moreover, the phenomenon of overfitting is less likely to occur in retinal images with fewer training sets, it improves the segmentation accuracy of test set and has strong robustness. The algorithm network module in this paper is shown in figure 3. In this paper, we utilize a 9-level architecture of deep AttentionResU-Net for blood vessels segmentation.

3. Experimental Setup and Results
In this paper, the manually segmented fundus vascular images of the first expert are selected as training labels on the dataset DRIVE, and the manually segmented fundus images of the second expert are adopted as the reference standard for the final segmentation results. The platform of this experiment was PyCharm, Using Keras and its Tensor Flow port, Inter(R) CoreTM i7-8750H CPU@2.20GHZ, 16GB of memory, Nvidia GeForce GTX 2080GPU, and 64bit Win10.

3.1. Database Summary
The DRIVE dataset contains 40 three-channel retinal images, 20 samples for training and 20 samples for testing. The original image size is 565*584. We used the sliding window with a step size of 5 to capture 48*48 patches. The input sample is shown in Fig.4.

Figure 4. Input patches in the left and outputs patches in the right

3.2. Quantitative Analysis Approaches
The following three indicators are used as bright criteria in this paper for model performance evaluation.

\[
AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}
\]

\[
SE = \frac{TP}{TP + FN} \tag{4}
\]

\[
SP = \frac{TN}{TN + FP} \tag{5}
\]

TP, which is correctly divided into positive examples; FP, the number of positive examples wrongly classified; FN, the number of negative examples wrongly classified; TN, the number of negative examples correctly divided.

The ROC (Receiver Operating Characteristic) curve uses "true example rate" (TPR) as the Y-axis and "false positive example rate" (FPR) as the X-axis.
The AUC value is the area under the ROC curve. The better the classifier performance, the greater the AUC value.

\[
TPR = \frac{TP}{TP + FN} \tag{6}
\]

\[
FPR = \frac{FP}{TN + FP} \tag{7}
\]

Figure 5. Retinal image segmentation results

Figure 6. ROC curve on DRIVE dataset

3.3. Result
We used the gold standard as the base fact, using 20 images from the DRIVE test dataset, considering only pixels belonging to the FOV, using the mask included in the DRIVE database to identify FOV. To
improve performance, the vascular probability of each pixel point is obtained by averaging multiple predictions. Sliding padding of 5 pixels in both height and width is used to extract multiple continuous overlaps from each test image. Then, for each pixel, the vascular probability can be obtained by averaging the probability balls of patches covering all pixels. The precise segmentation results achieved with the proposed AttentionResU-Net model are shown in Fig. 5. Figs. 6 show the AUC of the ROC curve when using the DRIVE dataset. The results of quantitative analysis are shown in Table 1. As can be seen from the table, in all cases, the proposed AttentionResU-Net model has a good performance in terms of AUC and accuracy. The ROC of the highest AUCs of the AttentionResU-Net model on the drive dataset.

| Methods          | SE   | SP   | AC   | AUC  |
|------------------|------|------|------|------|
| Staal et al[5]   | -    | -    | 0.9460 | -    |
| Soares et al[4]  | -    | -    | 0.9460 | -    |
| Marín et al [6]  | 0.7060 | 0.9800 | 0.9450 | 0.8430 |
| Orlando and Blaschko[10] | 0.7850 | 0.9670 | - | 0.8760 |
| Chen[14]         | 0.7252 | 0.9798 | 0.9474 | 0.9648 |
| Azzopardi[8]     | 0.7655 | 0.9704 | 0.9442 | 0.9614 |
| Roychowdhury[15] | 0.7250 | 0.9830 | 0.9520 | 0.9620 |
| Liskowsk[16]     | 0.7763 | 0.9768 | 0.9495 | 0.9720 |
| Qiaoliang Li[17] | 0.7569 | 0.9816 | 0.9527 | 0.9738 |
| U-Net            | 0.7537 | 0.9820 | 0.9531 | 0.9755 |
| Residual U-Net   | 0.7726 | 0.9820 | 0.9553 | 0.9779 |
| AttentionResU-Net| 0.7823 | 0.9815 | 0.9562 | 0.9793 |

4. Summary and prospect
Retinal retinal vascular segmentation is an important process in the medical field from manual analysis to automatic diagnosis. In this paper, the residual convolutional neural network and attention mechanism are used to extend the U-net architecture, and the proposed model is called attentionresu-net the scroll down window on the left of the MS Word Formatting toolbar. The model was tested in the DRIVE dataset. The false positive phenomenon of insufficient segmentation of micro vessels, and the fact that lesions and optic discs are easily divided into blood vessels by mistake is solved. Meanwhile, compared with U-net and Residual U-net, it has better segmentation performance with the same network parameters. In addition, the model not only guarantees good performance in the training process, but also guarantees good performance in the test stage, however, the segmentation rupture of blood vessels can still occur in images with lesions. In the future, we will continue to explore how to extract the lesion and repair the vascular rupture at the lesion.

References
[1] Zheng, Yalin, et al. "A comprehensive texture segmentation framework for segmentation of capillary non-perfusion regions in fundus fluorescein angiograms." PloS one 9.4 (2014): e93624. L. C. Perelman, J. Paradis, and E. Barrett, Mayfield Handbook of Technical and Scientific Writing, Mayfield, Mountain View, California (1997).
[2] Fraz, Muhammad Moazam, et al. "Blood vessel segmentation methodologies in retinal images—a survey." Computer methods and programs in biomedicine 108.1 (2012): 407 - 433.
[3] Al-Rawi, M., Qutaishat, M., Arrar, M. J. C. i. B. & Medicine. An improved matched filter for blood vessel detection of digital retinal images. 37, 262 - 267 (2007). K. Elissa, “Title of paper if known,” unpublished.
[4] Soares, J. V., Leandro, J. J., Cesar, R. M., Jelinek, H. F. & Cree, M. J. J. I. T. o. m. I. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. 25, 1214-1222
[5] Staal, J., Abràmoff, M. D., Niemeijer, M., Viergever, M. A. & Van Ginneken, B. J. I. t. o. m. i. Ridge-based vessel segmentation in color images of the retina. 23, 501 - 509 (2004).

[6] Marín, D., Aquino, A., Gegúndez-Arias, M. E. & Bravo, J. M. J. I. t. o. m. i. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. 30, 146 - 158 (2010).

[7] Läthén, G., Jonasson, J. & Borgia, M. J. P. R. L. Blood vessel segmentation using multi-scale quadrature filtering. 31, 762 - 767 (2010).

[8] Azzopardi, G., Strisciuglio, N., Vento, M. & Petkov, N. J. M. i. a. Trainable COSFIRE filters for vessel delineation with application to retinal images. 19, 46 - 57 (2015).

[9] Melinščak, M., Prentašić, P. & Lončarić, S. in 10th International Conference on Computer Vision Theory and Applications (VISAPP 2015).

[10] Orlando, J. I., Prokofyeva, E. & Blaschko, M. B. J. I. t. o. B. E. A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. 64, 16 - 27 (2016).

[11] Zhou, L. et al. Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement. 148, 13 - 25 (2017).

[12] He, K., Zhang, X., Ren, S. & Sun, J. in Proceedings of the IEEE conference on computer vision and pattern recognition. 770 - 778.

[13] Ronneberger, O., Fischer, P. & Brox, T. in International Conference on Medical image computing and computer-assisted intervention. 234 - 241 (Springer).

[14] Cheng, E. et al. Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features. 25, 1779 - 1792 (2014).

[15] Roychowdhury, S., Koozekanani, D. D., Parhi, K. K. J. I. j. o. b. & informatics, h. Blood vessel segmentation of fundus images by major vessel extraction and subimage classification. 19, 1118 - 1128 (2014).

[16] Liskowski, P. & Krawiec, K. J. I. t. o. m. i. Segmenting retinal blood vessels with deep neural networks. 35, 2369 - 2380 (2016).

[17] Li, Q. et al. A cross-modality learning approach for vessel segmentation in retinal images. 35, 109 - 118 (2015).