Novel Retinal Vessel segmentation method based on U-net and FPN

Xinyuan Zuo[1*]

1. School of Computer Science, University of Nottingham, Nottingham, UK
*corresponding author
E-mail: ezyxz@nottingham.ac.uk

***Abstract*—The structure and appearance of the blood vessel network in the retinal fundus image is an important basis for the diagnosis of fundus diseases and cardiovascular diseases. Fast and accurate retinal image segmentation has important clinical significance for predicting information and discovering abnormal information. This paper uses a novel convolutional neural network based on the improvement of the U-net network and the FPN network framework. This method uses skip-connection and convolutional layers of different sizes to generate residual blocks of different scales which obtain accurate acquisition of position information, then generates feature maps through feature fusion, and finally obtains semantic segmentation maps through a decoder layer. Many experiments have been performed on two public datasets. CHASE_DB1 and DRIVE have achieved subdivision accuracy of 97.20% and 96.48%, respectively. The performance results show that the automatic retinal vessel segmentation algorithm proposed in this paper is better than the latest in many verification indicators. The algorithm proposed by the algorithm can not only detect tiny blood vessels, but also capture large-scale high-level semantic ship features.**

Keywords-Retinal vessel segmentation; Convolutional Neural Network; Residual network; U net; Feature Pyramid Networks

I. INTRODUCTION

Retinal blood vessels are the only deep-seated microvessels in human tissues that can be observed in a non-surgical state [1]. Its morphology and structure are the same as those of people's diseases such as glaucoma, diabetes, hypertension, arteriosclerosis, and other fundus diseases. According to changes in their diameter, degree of curvature, and color, they can effectively help doctors in the judgment of ophthalmology, internal medicine and other diseases [2]. The distribution of blood vessels in fundus images is complicated and there are many small blood vessels. Traditional manual segmentation method is an effective method for clinical diagnosis, but it is a huge amount of manual labor for doctors. In addition, the results of manual segmentation by different doctors are very different. It is difficult to accurately analyze and judge the retinal image only by the doctor's manual segmentation. Automatic detection and precise analysis of retinal blood vessels can help doctors improve the efficiency of screening for diseases like diabetic retinopathy. Machine-assisted diagnosis of retinal images can be comparable to manual segmentation, at the same time it is faster and more reliable. There have been many proposals for medical image segmentation technology. Due to the influence of lighting conditions and camera equipment, the image is inevitably mixed with noise, and there are more tiny blood vessels at the end, which makes machine segmentation obstacles.

Retinal blood vessel segmentation methods based on image segmentation technology can be roughly divided into supervised and unsupervised methods.

Unsupervised methods utilize prior knowledge about vascular structure and rely on rule-based schemes.

Initially in 1989, S Chaudhuri et al. proposed to use a matched filter [3], which mainly used the grayscale distribution of the intersection of blood vessels into a Gaussian shape, and the grayscale of the background was basically the same. According to the matched filtering principle of signal processing, a (Gaussian) filters whose transfer function is consistent with the distribution of blood vessels is selected to obtain the output of the maximum signal-to-noise ratio, which can enhance the blood vessels. In 2002, Marks et al. [4] proposed a threshold modeling under the stochastic resonance method by adding uniformly distributed random noise to the image to obtain the optimal value of the output signal under the threshold, and the resulting image is more in line with human vision habit. In 2014, Azzopardi et al. [5] proposed a powerful method that uses several filters based on difference of Gaussian. They aggregate those filter responses by using weighted harmonic averages. The weights are carefully designed to maximize the system's response to the bar pattern. One year later, Zhao et al. [6] proposed a new infinite active contour model, which uses the mixed region information of the image to achieve blood vessel segmentation.

Generally, unsupervised methods can generalize the entire image well because they are not learned from samples that are supervised objects. However, due to their diverse appearance, it may be difficult to contain both normal and abnormal blood vessels.

Supervised learning can be divided into traditional machine learning and deep learning. The supervised machine learning model requires a classifier training process, the training label needs to manually label the blood vessel and non-vessel pixels, and the output result of the model is constantly close to the label pixels.

Niemeijer et al. [7] proposed a supervised retinal blood vessel segmentation method, which is based on the feature vector constructed using a multi-scale Gaussian filter, and the k-NN classifier is used to identify blood vessels and non-vascular pixels.

In the literature [8], a method based on fully connected conditional random fields and using structured output support vector machine (SVM) to automatically learn parameters is proposed to segment retinal blood vessels.

Compared with the traditional segmentation model based on gradient algorithm, the stochastic resonance response

model has unique advantages in image enhancement but lacks the ability of self-adaptation and learning and has no obvious advantages in image segmentation efficiency. On the contrary, a nonlinear artificial neural network built by imitating the structural characteristics of human brain neurons has strong adaptive and self-learning capabilities.

As early as the end of the 20th century, Yann LcCunl et al. [9] proposed a convolutional neural network and applied it to handwritten digit recognition tasks and achieved amazing results. In 2012, Hinton et al. [10] used a convolutional neural network structure to design AlexNet and divided about one million images into 1,000 categories with a fairly high accuracy and efficiency. In 2014, Karen Simonyan and Andrew Zisserman proposed the VGG convolutional neural network model [11], whose name comes from the abbreviation of the author's Visual Geometry Group at Oxford University. Among them, the most widely used is VGG16, which consists of 13 convolutional layers and 3 fully connected layers. The network structure is unconventional, easy to modify, and it has a good adaptability to other data sets, which greatly promotes the development of CNN. Jonathan Long [12] proposed the Fully Convolutional Networks (FCN) model. In his model, all the fully connected layers of the model are replaced with the volume base layer for the first time. Deconvolution and Bilinear upsampling are proposed to realize the upsampling layer to restore the image to its original size and combine different depths. The skip structure of the layer results ensures robustness and accuracy, which makes it possible to start semantic segmentation using neural networks.

In 2015, Olaf Ronneberger first proposed the U-net model network [13]. The author adds an upsampling stage based on the FCN model and adds many feature channels. Therefore, the more original image texture information can be spread in the high-resolution layer. The U-net network uses features to be spliced together in the channel size to form thicker features, while the addition of corresponding points used in FCN fusion will not form thicker features, which makes the shallow network retain obvious content information. Then in 2017, Tsung-Yi Lin et al. [14] proposed the Feature Pyramid Network (FPN) model. They proposed the feature fusion of Residual blocks to generate feature maps and then decode, which maximizes the acquisition of low-level target information and high-level semantic information. The method in this article will also follow this idea.

II. METHOD

In this paper, our method mainly uses the FPN framework to build feature pyramids and impresses on the skip connection and encoder-decoder ideas in U-net to complete a novel convolutional neural network model for retinal vessel segmentation task.

In the past few years, many neural network structures just use a single high-level feature map for object detection, prediction and other tasks. For instances, the earliest proposed VGG [11] model consists of only 13 convolutional layers and 3 fully connected layers to predict; the faster R-CNN [15] uses four times the down-sampled convolutional layer Conv4 to perform subsequent object classification and bounding box regression. Nevertheless, the low-level feature semantic information is relatively small, but the pixel information is richer and the target position is more accurate. It is easily lost during multiple down sampling. The FPN framework makes each layer of feature maps of different scales have strong semantic information, which is able to merge low-resolution feature map with Strong semantic and high-resolution feature map with weak semantic but rich spatial information.

Basic FPN network can be divided into two stages: Bottom-up pathway and Top-down pathway. In Bottom-up pathway, it firstly generates residual blocks which record semantic information at different scales as {C2, C3, C4, C5} by conv2, conv3, conv4, and conv5. Because of large memory footprint and poor semantic info in C1, the FPN author abandoned it. Second stage is feature fusion. It uses ResNet [16] generating feature maps {P2, P3, P4, P5} which reflect multi-dimensional information, by concating same stage residual block after conv1*1 and upper stage residual block after double upsampling.

FPN feature fusion, this process can be expressed by mathematical formulas.

$$P_i, P_{i+1}, P_{i+2}, \ldots P_{i+n}, = \int (C_i, C_{i+1}, C_{i+2}, \ldots C_{i+n}) \quad (1)$$

Finally, based on these feature maps, it can be combined with other network models to perform tasks such as target detection and segmentation. Figure 1 shows the image task framework based on FPN networks.

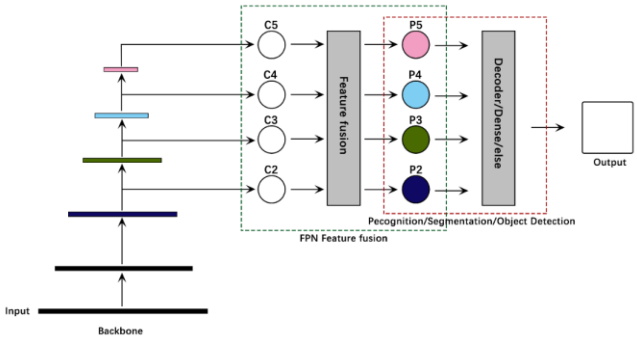

Figure 1. whole FPN frame

Our method innovates residual blocks and feature maps generation way and decoder to a segmentation map of the same size as the input image.

A. *Residual blocks Generation*

In our method, the residual blocks layer generation is improved and the idea of ResNets [16] is incorporated. For C0, our model takes input as single image and obtained C0 residual block after 64 conv3*3 and max pooling3*3. For C1, skip-connection is applied, which combines the module after con1*1, con3*3 and con1*1 and the module con1*1. Upper residual blocks generation is a simplified version of the Inception [17] module, which uses different filter sizes to capture multiple concepts in different ranges at the same time, so that it can obtain more information for the neural network to choose from. The Inception structure uses

multiple convolution kernels of different scales to enhance the generalization and structure expression capabilities of the network, and adds more nonlinearity to the network model, which greatly improves the learning feature ability of the convolutional neural network. For {C2, C3, C4, C5}, out model changes the first con3*3 to con1*1, stride as 2 to produce semantic information at different scales. Figure2 shows residual blocks generation process.

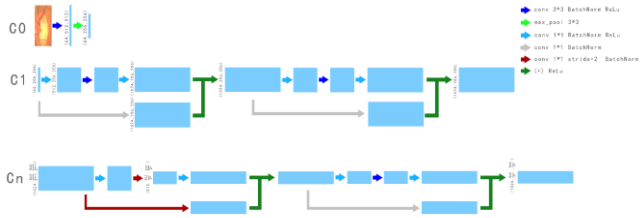

Figure 2. Residual blocks

B. *Feature maps generation – Feature fusion*

Our method uses top-down one-way fusion of residual blocks to generate feature maps. The generation of feature maps of different scales are combined with the feature maps of the lower layer, which carries out 4 times of upsampling and uses skip connection in the same stage. The residual block is first subjected to con1 * 1, then upsampleing is cascaded with the upper residual block, finally con3 * 3 is performed to obtain the corresponding function map. The existing commonly used upsampling methods include bilinear interpolation, nearest interpolation and deconvolution. Since the network structure in this article is complex enough to capture the underlying semantic information, and deconvolution will bring more calculations, the nearest interpolation method is adopted.

Now, the high-level features have been enhanced. The feature map used in each layer of prediction combines features of different resolutions and different semantic strengths, which can complete the detection and measurement of objects with corresponding resolutions, ensuring that each layer has appropriate resolution and Strong semantic features. Figure3 represents feature fusion process.

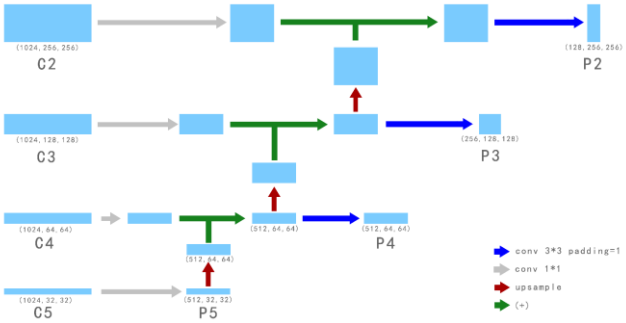

Figure 3. Feature maps

C. *Decoder*

Now, we get different scale feature maps {P2, P3, P4, P5}.

The semantics of the retinal blood vessel image is relatively simple, the structure is relatively fixed. Therefore, the image stability does not require a lot of processing noise or segmentation of different types of objects, while the accuracy of small blood vessel segmentation requires excessive attention. To achieve this, our method directly uses C0 as P1 rather than processing C0 with skip-connection to generate P1 to make the first feature map retain high position information without retaining too much semantic information. U-net neural network has been extremely outstanding in medical image segmentation once it was proposed. Medical images are images of fixed organs instead of full-body images. The semantic information is not particularly rich, so both high-level semantic information and low-level features are very important. U-net network has a great advantage in the skip connection and symmetric network structure of the same stage. Our method uses the idea of the decoder part of the U-net network. After upsampling, the high-level feature map is superimposed with the low-level feature map，then conv3*3 is performed twice. Finally, segmentation map is generated by using con1*1 to reduce feature map thickness. Figure4 represents decoder process.

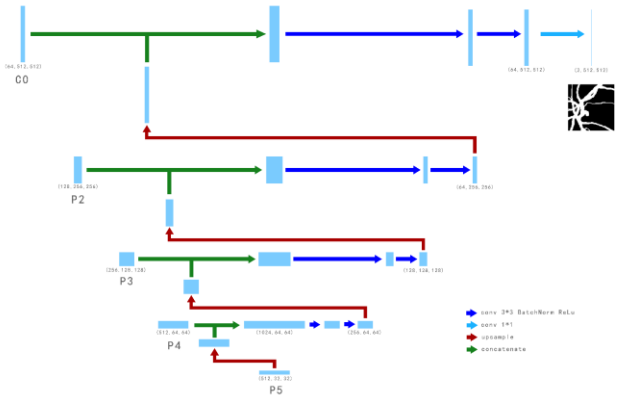

Figure 4. Decoder

III. DATASETS AND IMAGE PROCESSING

The segmentation performance of the method proposed in this paper is verified on CHASE_DB1[18] and DRIVE [19].

The DRIVE data set includes 40 color fundus images of the retina. These images come from a diabetic retinopathy screening project in the Netherlands. Each image has 584×565 pixels. In this paper, the first 20 images of 40 images are used for training, and the remaining 20 images are used for testing. A binary field of view mask and GT image are provided for each image in the data set. CHASE_DB1 has 28 retinal color images, and the pixel size of each image is 999×960. In this paper, the first 20 images are used for training, and the remaining 8 images are used for testing.

The original picture is a tiff format picture with a size of 999*960, a total of 20 pictures. In order to facilitate the training of this experiment, the 20 pictures were randomly sliced into 512*512 and a total of 4000 pictures were used as the training set. During the batch training, 6 pictures will be randomly taken out and randomly horizontally flipped each time, then rotated around the center point from -15° to 15° and randomly scaled between 0.75 and 1.5 to obtain more possible training sets. Finally, the normalization is carried out by each channel is divided by 255, then means for each channel is subtracted, the standard deviations of each channel is divided. The data is mapped to a specific distribution with mean value (0.315, 0.319, 0.470) and standard deviation (0.144, 0.151, 0.211), so that the grayscale and exposure of the picture has a uniform standard, which eliminates the overexposure to a certain extent the influence of poor quality or noise on the model. In CNN, the standard language also makes the RGB component pixel value distribution more similar, and the gradient descent method will be more stable and fast during training. It can also effectively solve the problem of gradient explosion and gradient disappearance. Figure 5 shows that the outline of the picture is more obvious after normalization.

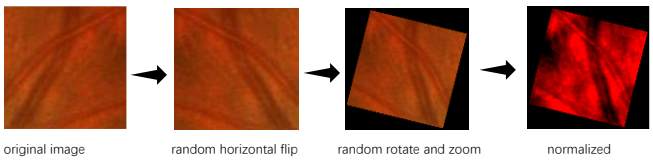

Figure 5. Image processing renderings

IV. TRAIN

This image segmentation task can be regarded as a two-class classification task, predicting the value of each pixel to be 1 or 0, so the loss function is adopted as the cross entropy.

$$loss = -\frac{1}{n}\sum_{i=0}^{n}[\, y_i \ln \hat{y}_i \; - \; (1-y_i)\ln(1-\hat{y}_i\,)] \tag{2}$$

Where: n represents the total number of pixels; y_i represents the corresponding value of the i-th pixel label, $\hat{y}_i$ is the predicted value of the pixel.

The cross-entropy loss function has the property that when the error is large, the weight is updated quickly; when the error is small, the weight is updated slowly. In this paper, the Adam optimization method is used to optimize the parameters, and the learning rate is dynamic learning rate.

$$\alpha_i \; = \; \alpha_{i-1}\,(0.99^{(i-1)\,|10}) \tag{3}$$

Where: α_i represents learning rate of the i-th.

The model development integrated environment in this paper is PyCharm, based on the Pytorch framework. The operating platform is Linux, the GPU is GeForce RTX 2080Ti, the memory is 32GB, the number of experimental iteration training is 200, the batch size is 6, and the training time is about 20 hours.

Whole project code and dataset is available on github:

https://github.com/ezyxz/Retinal_Vessel_Segmentation_Project

V. PREDICT AND RESULTS

The horizontal and vertical pixels of the original image are enlarged by 5 times, and then divided into multiple 256*256 images which are first standardized. After that, these images are input into the model, then the outputs are pieced together to form a complete retinal segmentation image. Finally, it is resized to the original image size using the ANTIALIAS filter.

The characteristics of retinal images are complex and vary from person to person, and it is easily affected by external conditions and the disease itself. Figure 6 shows some comparison images of blood vessel segmentation in some local areas, including blood vessels at the intersection, small blood vessels under low contrast, and diseased areas.

Figure 6. Detail segmentation comparison

Figure 7 shows the real image, the horizontal comparison of the manual annotation and the predicted value of the method in this paper. From the effect of the segmentation map, the segmentation result obtained by the algorithm proposed in this paper is basically the same as the standard image of the expert segmentation, especially in It has a good effect on the segmentation of small blood vessels.

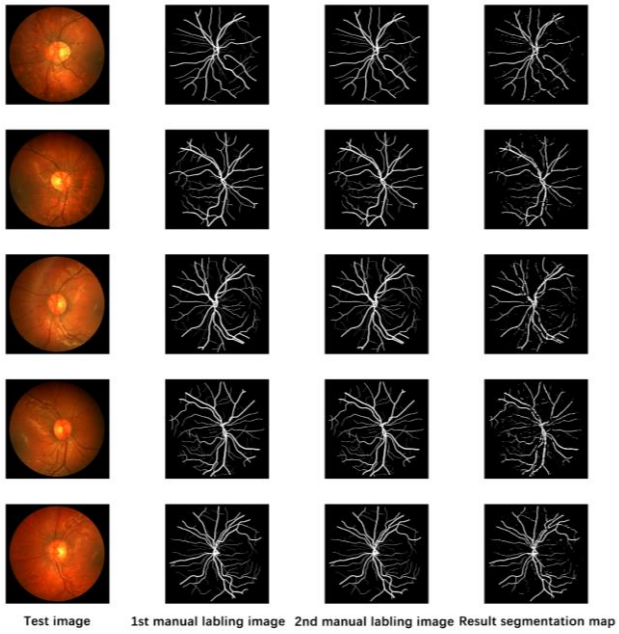

Figure 7. Full image comparison on CHASE_DB1

This article uses IoU, accuracy, sensitivity, and specificity to evaluate the segmentation effect. IoU (intersection over union) which calculates the ratio of the intersection and union of the predicted boundary and the true boundary.

Sensitivity refers to the proportion of samples that are actual positive that are judged to be positive. The closer the index is to 1, the higher the sensitivity of the model to small blood vessels, and it is easy to segment small blood vessels.

Specificity refers to the proportion of samples that are actual negative. The closer the index is to 1, the model will not additionally segment blood vessels.

Accuracy, also known as efficiency, is expressed as the percentage of the total number of true positives and true negatives in the number of subjects tested. The closer the index is to 1, the closer the model segmentation effect is to the real artificial segmentation level, and the better the segmentation effect.

$$\text{Sensitivity} = \frac{P_{TP}}{P_{TP}+P_{FN}} \tag{4}$$

$$\text{Specificity} = \frac{P_{TN}}{P_{TN}+P_{FN}} \tag{5}$$

$$\text{Accuracy} = \frac{P_{TP}+P_{TN}}{P_{TP}+P_{FP}+P_{Fn}+P_{TN}} \tag{6}$$

Where: True Positive (TP) means the number of pixels that correctly classify the blood vessel; True Negative (TN) means the number of pixels that correctly classify the background; False Positive (FP) means that the background is misclassified Number of pixels; False Negative (FN) means the number of pixels that misclassify blood vessels.

After prediction, the average IoU, accuracy, sensitivity, and specificity of the 8 images test set in CHASE_DB1 are 0.8192,0.9720, 0.7354, and 0.765. The average Accuracy, Sensitivity, and Specificity in the 20 test sets of DRIVE are 0.8059, 0.9720, 0.7354, and 0.765. The two tables below show how this method compares with other deep learning methods on two datasets.

TABLE I. COMPARISON OF THE METHOD PROPOSED ON CHASE_DB1

Method	Year	Accuracy	Sensitivity	Specificity
Multilayer NN[19]	2014	0.9530	0. 6867	0. 9824
U – Net[20]	2015	0.9500	0. 7758	0. 9755
DeepVessel[21]	2016	0.9523	0. 7063	N A
Deep FCN[22]	2017	0.9521	0. 7779	0. 9780
LadderNet[23]	2018	0.9663	0. 7856	0. 9810
Residual U – Net[24]	2018	0.9553	0.7726	0.9820
Dense U – Net[25]	2019	0.9511	**0. 7986**	0. 9736
DUNet [26]	2019	0.9566	0. 7963	0. 9800
Our Method	2021	**0.9720**	0.7354	**0.9961**

TABLE II. COMPARISON OF THE METHOD PROPOSED ON DRIVE

Method	Year	Accuracy	Sensitivity	Specificity
YIN[27]	2012	0.9267	0.6522	0.9710
ZHAO[28]	2015	0.9540	0.7420	0.9820
COSFIRE filters [5]	2015	0.9442	0.7655	0.9704
Song[29]	2017	0.9499	0.7501	0.9795
LadderNet[23]	2018	0.9561	**0.7856**	0.9810
YAN[30]	2018	0.9542	0.7653	0.9818
WANG[31]	2019	0.9541	0.7648	0.9817
Our Method	2021	**0.9648**	0.7454	**0.9858**

Compared with the benchmark U-Net, all indicators have achieved better performance, and there is a big gap. Compared with other proposed algorithms, it has excellent performance in accuracy and specificity.

VI. CONCLUSION

In this article, in view of the problem of segmenting blood vessel boundaries and low-contrast areas in retinal blood vessel segmentation, we propose a novel network which consists of residual blocks generation, feature fusion and decoding mechanism to make the network better segment the blood vessel boundaries, while also improving the network positioning and classification capabilities. In residual blocks generation, Inception structure is used to obtain more features. Top-down one-way fusion is used in feature fusion and concatenating with lower layer feature map is used in decoder layer, which achieves accurate segmentation.

REFERENCES

[1] Kirbas C. A review of vessel extraction techniques and algorithms[J]. Acm Computing Surverys 2004, 36(2):81-121

[2] Nayebifar B,Moghaddam H A. A novel method retinal vessel tracking using particle filters[J].Computers in Biology and Medicine, 2013,43(5): 541-548.

[3] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M. and Goldbaum, M., 1989. Detection of blood vessels in retinal images using two-dimensional matched filters.

[4] Marks R hompson BEI-Sharkawi,Stochastic resonance thresholc etector Image visualization and explanation.IEEE International Symposium on Circuits and Systems, 2002,4: 521-523

[5] Azzopardi, G., Strisciuglio, N., Vento, M., Petkov, N.: Trainable cosfire filters for vessel delineation with application to retinal images. Med. Image Anal.19(1), 46–57 (2015)

[6] Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retinal Images

[7] Niemeijer M, Staal J, van Ginneken B, Loog M, Abramoff MD, editors. Comparative study of retinal vessel segmentation methods on a new publicly available database. SPIE medical imaging. 2004;5370:648–656.

[8] Orlando J I, Prokofyeva E, Blaschko M B . A discriminatively trained fully connected conditional random field model for blood vessel segmentation fundus images [J] . IEEE transactionsonBiomedical Engineering, 2017, 64 (1):16-27

[9] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W. and Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), pp.541-551.

[10] Krizhevsky, A., Sutskever, I. and E. Hinton, G.,ImageNet classification with deep convolutional neural networks | Communications of the ACM. [online] Dl.acm.org.

[11] Simonyan, K. and Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition. [online] arXiv.org.

[12] Long, J., Shelhamer, E. and Darrell, Fully Convolutional Networks for Semantic Segmentation. [online] Ieeexplore.ieee.org.

[13] Ronneberger, O., Fischer, P. and Brox U-Net: Convolutional Networks for Biomedical Image Segmentation

[14] Lin, T., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S., 2017. Feature Pyramid Networks for Object Detection. [online] arXiv.org. Available at: <https://arxiv.org/abs/1612.03144v2> [Accessed 9 March 2021].

[15] Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. [online] arXiv.org.

[16] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016.

[17] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going Deeper with Convolutions. [online] arXiv.org. Available at: <https://arxiv.org/abs/1409.4842>

[18] Owen C G, Rudnicka A R, Mullen R, et al. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (CAIAR) program[J].Investigative Ophthalmology & Visual Science, 2009, 50(5): 2004-2010.

[19] Mo J, Zhang L. Multi-level deep supervised networks for retinal vessel segmentation[J]. International Journal of Computer Assisted Radiology and Surgery, 2017;12 (12) :2181－2193.

[20] Ronneberger O,Fischer P,Brox T.U-Net: convolutional networks for biomedical image segmentation [c]International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015:234－241.

[21] Fu H Z, Xu Y W, Lin S,et al. Deep vessel: retinal vessel segmentation via deep learning and conditional random field [c]International Conference on Medical Image Computing and Computer-Assisted Intervention － MICCAI 2016. Athens: Springer,2016: 132－139.

[22] Wilfred F S, Edward R S. Computerized screening of diabetic retinopathy employing blood vessel segmentation in retinal images ［J］. Biocybernetics and Biomedical Engineering,2014, 34(2) : 117－124.

[23] Zhuang J. LadderNet: multi-path networks based on U-Net for medical image segmentation［EB /OL］. (2019－08－28)[2019－10－10]. https: / /arxiv.org /abs/1810.07810.

[24] Alom M Z, Hasan M, Yakopcic C, et al. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation[J]. ArXiv Preprint, 2018, arXiv:1802.06955.

[25] Wang C, Zhao Z Y,Ren QQ,et al. Dense U-Net based on patch-based learning for retinal vessel segmentation［J］. Entropy,2019,21(2) : 168－182.

[26] Jin Q, Meng Z, Pham T D, et al. DUNet: a deformable network for retinal vessel segmentation[J] . Knowledge-Based Systems,2019,178: 149－162.

[27] YIN Y, ADEL M, BOUR ENNANES. Retinal vessel segmentation using a probabilistic tracking method [J]. Pattern Recognition, 2012,45(4):1235-1244

[28] ZHAO, Y., Rada, L. and Chen, K., 2015. Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retinal Images. [online] Ieeexplore.ieee.org. 34(9):1797-1807.

[29] Song J, Lee B. Development of automatic retinal vessel segmentation method in fundus images via convolutional neural networks [C] // 2017 Annual International Conference of the IEEE Engineering in Medicine and Biology, 2017: 681-684.

[30] Yan ZQ, Yang X, Cheng KT. Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation [J]. IEEE Transactions on Biomedical Engineering, 2018, 65(9): 1912-1923.

[31] Wang, X., Jiang, X. and Ren, J., 2018. Blood vessel segmentation from fundus image by a cascade classification framework. .[J] Pattern Recognition,88:331-341