The Application of Deep Learning in Vessel Extraction with Fundus Image

Zuman Yang, Jiaxing Wu, Kagyu Jiang

Computer science and technology, Liaoning University NO.66, Chongshan Road, HuangGu District, Shenyang, P.R. China
Email: 906622326@qq.com

Abstract: Fundus image is a very prevalent Biomedical data, so it is very important to choose a method to extract blood vessel from fundus images. Extracting vessel from the fundus image can be used for personal identity verification and assist doctor to diagnose diseases [1] such as diabetes. However, traditional extracted methods such as manually annotation area waste of time and energy. There is some another method [2] based on the image recognition, extracting the feature with the filter and segment with mathematical morphology method. But in the fact of the application, this method has low stability and accuracy. Traditional machine learning methods such as SVM [3] are also used for vessel extraction, but this method can only apply to dealing with one-dimensional data while the images showed up as matrix form. So, we have to stretch a matrix line by line or column by column into one-dimensional vectors [4-6]. However, by doing so, we will be face that some important information can’t be reflected because the position relation of rows and rows in the original matrix will be lost. Deep learning can process image data naturally and extract features automatically with convolution, and it is also popular for its good learning ability and high accuracy. U-net [7] is a convolutional network based on deep learning and widely used in biological image extraction, which constructs a U-shaped network from input image and output image. In this paper, building a U-Net convolutional network, which takes fundus image as input and vessel image as output, to extract vessels in fundus image. We established a training set with 20 fundus images and a test set with 20 fundus images. Finally, the accuracy rate of our U-net model on the training set can reach 99 percent, while the accuracy rate on the test set was stable above 99.5 percent.

1.Introduction
Fundus image has a wide range of applications, and vessel extraction is the most important information in fundus image. It is very important to extract the blood vessel from fundus image effectively and accurately. Vessel extraction with fundus image has been widely used in daily life such as the identity recognition. Fundus blood vessels are different from person to person, so it can be used for identification just like the fingerprint. Besides, fundus vessels are the feature inside the body and it is difficult to change by operations, which ensure the accuracy of identification. Furthermore, fundus vessel structure is the most important observable structure in the fundus image, which makes it useful in medical diagnosis. Some eye diseases can change fundus vessel structure, so eye doctors can use fundus vessels images to diagnose diseases. For example, high blood pressure and cerebral hemorrhage can cause fundus hemorrhage at the time of onset, and arteriosclerosis can cause the fundus vascular wall thickness. More accurate and clear fundus blood vessel images will make the diagnosis of these diseases easier than before.
There are several methods [8-10] that have been used to extract vessels from fundus images. One kind of approach is manually annotation, where doctors mark the position of the blood vessel picture by picture. There is no doubt that this method is very accurate, but it also wastes the time of experienced doctors. The increasing number of people with eye diseases also generates a lot of fundus images, which increases the workload of doctors and makes manually annotation inapplicable. Another type of method is based on the image processing and mathematical morphology, which divides images into different regions with some vision algorithms and classifies these regions into vessels and non-vessels according to their attributes. However, when the quality of the fundus image is not ideal, the recognition accuracy of this technique will be greatly reduced.

Traditional machine learning methods are also used to extract vessels from fundus images. The features of fundus images are extracted and used to train traditional machine learning classifiers such as KNN, SVM and random forest, and these classifiers are used to predict whether it is a blood vessel point by point.

An unignored disadvantage of this method is that it cannot deal with images directly and it has to extract features from images. However, it is difficult to construct features and to choose which features are effective for distinguishing blood vessels. Another disadvantage is that it can only predict vessels pixels by pixels but cannot directly output images, which increases the amount of calculation and lowers the prediction accuracy.

With the development of computer technology, deep learning is becoming more and more popular, especially in the area of image recognition. Deep learning has great advantages compared with previous traditional technologies. Firstly, different from traditional machine learning methods, deep learning can directly process two-dimensional image data. In other words, the image can be directly used as the input of convolutional neural network, while traditional machine learning methods can only take one-dimensional feature vector as input, which avoids the data lost caused by the feature extraction stage. Secondly, deep learning can extract the features of images automatically with convolutions on different scales. After the convolution stage, pooling layers are set up to process image data at different resolution, which can extract the features in different resolutions so that the image information more complete. Deep learning can also avoid the problem of over-fitting by dropout in training procedure and reduce the computation amount significantly by weight sharing of convolution kernels. Generally speaking, deep learning is suitable for image recognition with its high accuracy and learning ability. U-net is a convolutional network based on deep learning and widely used in biological image extraction and recognition.

In this paper, we construct a U-shaped network which takes fundus image as input and outputs vessel image directly. We don’t need to extract features from fundus image, since the convolution and pooling layers in our network can extract features. We can also input fundus image and get the vessel image directly from our U-net. But in real applications, we can’t directly input the whole fundus image, because the compete image has a lot of different features so that reduce the prediction accuracy of the model. We randomly generate a lot of patches from the image and take patches as input. With this method, the sample size is increased, and the blood vessel characteristics are reflected more accurately, which can improve the recognition accuracy. The rest part of this article is organized as follows. In section 2, we briefly introduce our training and testing data, and the preprocess for original fundus images. The procedure for patches generation from original images is also presented in this section. In section 3, we construct our U-net model for blood vessel extraction and introduce its structure in detail. We explain how to get predicted vessel images from predicted patches in section 4. The forecast results and accuracies are showed in the section Finally, we make a simple conclusion for the paper.

2. Data description

2.1. DRIVE database
The image data is from the DRIVE database. Images in this database were obtained from a diabetic retinopathy screening program in The Netherlands. The screening population consisted of 400 diabetic subjects between 25-90 years of age. In this paper we selected 20 images as train set and 20 images as test set. Both training and testing data consist of three parts, part of original fundus images (Figure 1), part of hand-marked vessel (Figure 2), and one part of hand-marked eyeball boundary (Figure 3). Examples of DERIV database is show in Figure 1-3. These images all have the same size 584x565. This data is publicly available at http://www.isi.uu.nl/Research/Databases/DRIVE/.

The extracted photo reflects a clearer vascular structure that will help the doctor diagnose. Therefore, the application of deep learning in fundus image recognition has practical significance.

2.2. Original image preprocessing

In order to enhance the information of the image, we preprocess the input original fundus images with three main steps, graying, normalization and equalization. First, we convert RGB images to grayscale images by weighted average with weights 0.299, 0.587, 0.114 for red, green, blue channel respectively [11-12]. Then we take z-score normalization for each image and perform min-max normalization among all images to improve the convergence speed of the model. To improve contrast in images, we also use the CLAHE (Contrast Limited Adaptive Histogram Equalization) to perform histogram equalization on images. Finally, we perform histogram stretching on the image.

2.3. Patch

Obviously, 20 fundus images are not enough to train a deep learning network. Moreover, directly inputting the whole fundus image will cause the model to be over-fitted because it does not accurately reflect the growth of blood vessels. In order to solve these problems, we randomly extract a lot of patches from the pictures. We randomly extract 48x48 patches from the pre-processed training set image, and draw 10000 patches per image, which generates a total of 200,000 patches. These patches solve the problem of insufficient training set, and reflect the growth of blood vessels more accurately. The position of the original image and the position of the real blood vessel image are consistent. We present 50 extracted patches in Figure 4 and Figure 5.
3. U-Net for Vessel Extraction

U-net is a deep convolutional neural network proposed by Ronneberger and widely used in biomedical image segmentation. The traditional U-net convolutional network is shown in Figure 6. This structure contains a downsampling path and an upsampling path. The full convolution method is used in the downsampling path and the feature channel is reduced by half each time downsampling. Upsampling and downsampling are symmetrically present, forming a U-shaped structure. The network does not contain any link structure, only the valid part of each convolution.

We also constructed a U-net network for the extraction vessel structure from fundus images. This network structure is shown in Figure 7.
Figure 7: U-net model in this paper

Layer 1:
The input 48x48 patches are processed with two convolution layers with kernel size 3x3 and a dropout layer with dropout rate is 0.2, and we get 32 48x48 images as the output of this layer.

Layer 2:
The input 32 48x48 patches are processed with two convolution layers with kernel size 3x3, a maxpooling layer with pool size 2x2 and a dropout layer with dropout layer rate is 0.2. We get 48 24x24 images as output of this layer.

Layer 3:
The input 48 24x24 patches are processed with two convolution layers with kernel size 3x3, and then padded with zeros, a dropout layer with dropout rate is 0.2 and a maxpooling layer with pool size 2x2. We conduct upsampling with size 2x2 for each image and get 128 24x24 images as output of this layer.

Layer 4:
The input 128 24x24 images from layer3 and 64 24x24 images from layer 2 are processed with a concatenated layer, the patches from two layers concatenated into the same set. Two convolution layers with kernel size 3x3, a dropout layer with dropout rate is 0.2 and a upsampling layer with size 2x2 for each image. Output 64 48x48 images from this layer.

Layer 5:
The input 64 48x48 patches from layer 4 and 32 48x48 patches from layer 1 are processed with a concatenated layer, two convolution layers with kernel size 3x3 and a dropout layer with dropout rate is 0.2. We get 32 48x48 images as the output of this layer.

Layer 6:
The input 32 48x48 patches are processed with a convolution layer with kernel 1x1. The image with size 48x48 is reshaped to a vector with length 2304 in reshape layer and exchange length and width vectors in permute layer. The Stochastic gradient descent (SGD) algorithm are employed to minimize categorical cross-entropy loss. We set epoch number to be 20 and each batch with 32 patches. The weights with minimal loss are saved as our final best weights.
4. Prediction and Results
Since our U-net is designed for 48x48 patches, we need to segment the original images (584x565) into 48x48 patches for prediction. We stride a 48x48 window on the original image with step length 5 pixels to obtain corresponding patches. 11445 patches are generated in total for each image, and then used to predict vessel locations with our U-net. The predictions for each pixel are averaged and considered as the final prediction.

The confusion matrix is presented in Table and the accuracy is 95.2%.

| Prediction | Real | True Rate |
|------------|------|-----------|
| F          | 3908311 | 0.9868241 |
| T          | 165428  | 0.7136185 |
| TPR         | 0.9593916 | 0.88763447 |

The receiver operating characteristic (ROC) curve and precision recall (PR) curve are plotted in Figure 8. The area under ROC curve and PR curve are 0.9749 and 0.8980, respectively.

![ROC curve](image1.png)

![Precision - Recall curve](image2.png)

Figure 8: ROC and PR AUC
We see that our U-net performs well in Vessel Extraction for retinal Images. All original retinal images, true vessels and predicted vessels with U-net are summarized in Figure 9.

![Figure 9: Extracted result](image)

5. Conclusions

Nowadays, people use their eyes more frequently, which to cause the incidence rate of eye diseases to increase year by year. Vascular extraction of fundus images can help doctors diagnose these diseases. In this article, we have chosen a U-net network that is widely used in biomedical image segmentation to extract fundus images. In 20 test sets, the ROC AUC reached 0.97, and the PR AUC reached 0.89. It can be said that we have effectively reduced the difficulty of blood vessel extraction from fundus images.

References

[1] Shang Xingyu, Zhang dong-bo. Blood Extraction Approach for Retinopathy Image[J], Lecture Notes in Electrical Engineering, 2012(135):145-150.
[2] B.Cote,W.Goldbaum,P.Kube,and M.Nelson. Classification of Blood vessel in ocular fundus images[M]. Computer Science and Engineering Dept, Univ of California, San Diego, Tech Rep.1994
[3] Hang Xiao, Gu Shuo. Artificial Intelligence of Diabetic Retinopathy Image Recognition Used in the Real Word[G], Technology Intelligence Engineering,2018
[4] E.Ricci,R.Perfetti.Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification[J].IEEE Transactions on Medical Imaging, 2007,26(10):1357-1365.
[5] A.Hoover and M.Goldbaum, Locationg the optic never in a retinal image using the fuzzy convergence of the blood vessel[J],IEEE Trans.Med.Img,2003,22(8):951-958.
[6] X.Jiang,D.Mojon.Adaptive Local Thresholding by Verification-Based Multithreshold Probing with Application to Vessel Detection in Retinal Images [J].IEEE Trans. On Pattern Analysis and Machine Intelligence, 2003,25(1):131-137.
[7] O.Ronneberger,P.Fischer and T.Brox. U-Net: Convolution Network for Biomedical Image Segmentation.2012
[8] M.Kuivalainen Retinal image analysis using machine vision[D].Master’s thesism,Lappeenranta University of technology,2005
[9] Benson S Y L,Hong Y.A novel vessel segmentation algorithm for pathological retinal images based in the divergence of vector fields[J],IEEE Transactions on Medical Imaging, 2008 27 (02):237-246.
[10] Young Kyung Par.Ratinex method based in adaptive smoothing for illumination invariant face recognition[J]. Signal Processing.2008,88:1929-1945
[11] M.Foracchia,E.Grisan.Luminosity and contrast normalization in retinal images [J].Medical Image Analysis,2005,3(9):179-190.