Feature extraction and classificatio of glaucoma *

Shuai Lu[1,2], Shiyuan Ke[1], and Yi Yang[1]

[1] Mathematics,Beijing University of Chemical Technology, Beijing, China
[2] Technical department,Beijing Shanggong Medical Technology Co.,ltd,Beijing ,China
1292226727@qq.com

Abstract. Glaucoma is an ophthalmic disease which can cause blindness by optic nerve damage.Early diagnosis of glaucoma is the key to preventing loss of vision.Cup to disc ratio is an important evidence for the diagnosis of glaucoma.Thus,the automatic segmentation of optic disc and optic cup and diagnosis of glaucoma are basic task.In our paper,we use convolutional neural networks to automatically segment images and automatically classify them.We turn all the fundus images to grayscale as input images,which can improve the generalization ability of the algorithm.For the segmentation of optic disc and optic cup,the U-Net with less filters can reduce training times.For the classification of glaucoma,we extract features and classify by Xception network.

Keywords: Deep learning· images segmentation· U-Net · Xception · glaucoma diagnosis.

1 Introduction

Glaucoma is a progressive optic neuropathy characterized by structural changes in the optic nerve head, which the World Health Organization ranks as the second most common blind eye disease in the world[1].Early diagnosis of glaucoma is extremely difficult, and computer-aided diagnosis is an efficient and powerful tool.At present, computer-aided diagnosis of glaucoma mainly focuses on classification and image segmentation.

The paper contains two tasks:(i)segmentation of Optic Disc and Cup;(ii) classification of clinical glaucoma.In the past few years, Deep Learning has shown very good performance in solving many problems such as visual recognition, speech recognition, and natural language processing.Therefore,our method is mainly composed of convolutional neural network.All images are processed with preprogressing,such as image enhancements.

The train set comprises 400 fundus images taken using Zeiss Visucam 500 with resolution of 2124×2056 pixels including 360 normal eyes and 40 glaucomatous eyes.But the test set comprises 400 fundus images taken using Canon CR-2 camera with resolution of 1634×1634pixels.There is a big difference between the two data sets.All input images are grayscale to eliminate color differences in the images datasets.

* Supported by organization x.

2 Segmentation of Optic Disc and Cup

2.1 Methods

For the optic disc and optic cup segmentation task, our proposed method comprises tree main steps:(i)Pre-processing:image preprocessing includes changing grayscale,image enhancement,and cropping;(ii) Neural network:U-Net[2] architecture is adapted for medical image segmentation;(iii):Post-processing:to obtain the final segmentation result by obtaining the largest connected region.Moreover, ellipse fitting just doing for optic disc segmentation.

Pre-processing firstly,we padding the images of train set to 2124×2124,then resize the images to 1634 1634.For the optic disc segmentation,we cut the images to 817×817,and resize to 256×256 for the optic cup segmentationlike the fig.2. After segmenting the optic disc, we find the largest rectangle containing the disc, and extend 100 sizes outward, then scale to 128×128 as the input image scale. We tried to eliminate dataset differences and increase the number of dataset by

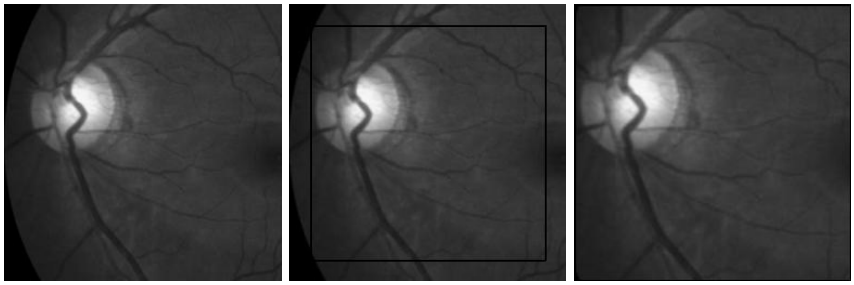

Fig. 1. Process of cutting pictures

preprocessing. The steps of preprocessing are as follows:
Step1: We turn all images into grayscale,like the fig.1 and use power law transformation to change image brightness. With four different brightness, each picture becomes 4,which is increased from 400 to 1600.
Step2: Randomly flip the image up and down with a probability of 50%;
Step3: Randomly flip the image left and right with a probability of 50%;
Step4: Randomly enlarge the image with a probability of 50%, and cut the image with size 256*256.
For the segmentation of optic cup, preprocessing operations are used for all images, but we only change the brightness of the test set images by power law transformation.

Neural Network Our approach is primarily based on deep learning techniques U-Net[2].the architecture presented in the paper is depicted in fig4. Like the

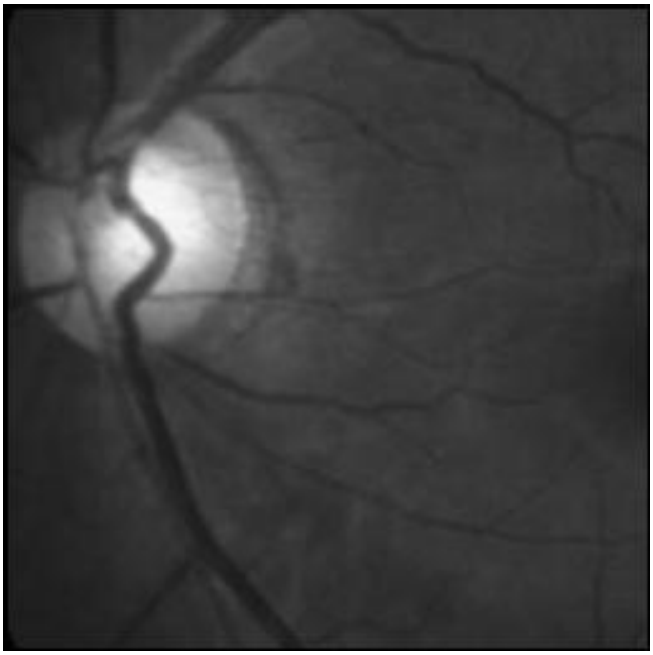

Fig. 2. grayscale pictures

original U-Net, it consists of contracting path (left side) and an expansive path (right side). Network is not fully connected, only convolution and down sampling . Input image is firstly passed through a convolutional layer with filters of 3×3 pixels spatial resolution; number of filters in a layer is shown in the figure above a blue rectangle representing layers output. Contracting path structurally repeats a typical architecture of convolutional part of the classification network, can capture context information. Expansive path can positioning precisely. Information is merged from layers of contracting path of appropriate resolution and layers of expansive path of lower resolution. Compared to original U-Net, the presented architecture has less filters in all convolutional layers and does not possess an increasing number of filters for decreasing resolution, so our architecture need less parameters and training time.
As a loss function,we use $l\,(A, B)$:

$$l\,(A, B) = -\log d\,(A, B) \tag{1}$$

where:

$$d\,(A, B) = \frac{2\sum_{i,j} a_{ij} b_{ij}}{\sum_{i,j} a_{ij}^2 + \sum_{i,j} b_{ij}^2} \tag{2}$$

where $A = (a_{ij})_{i=1}^{n}{}_{j=1}^{n}$ is predicted output map and $B = (b_{ij})_{i=1}^{n}{}_{j=1}^{n}$ is a correct

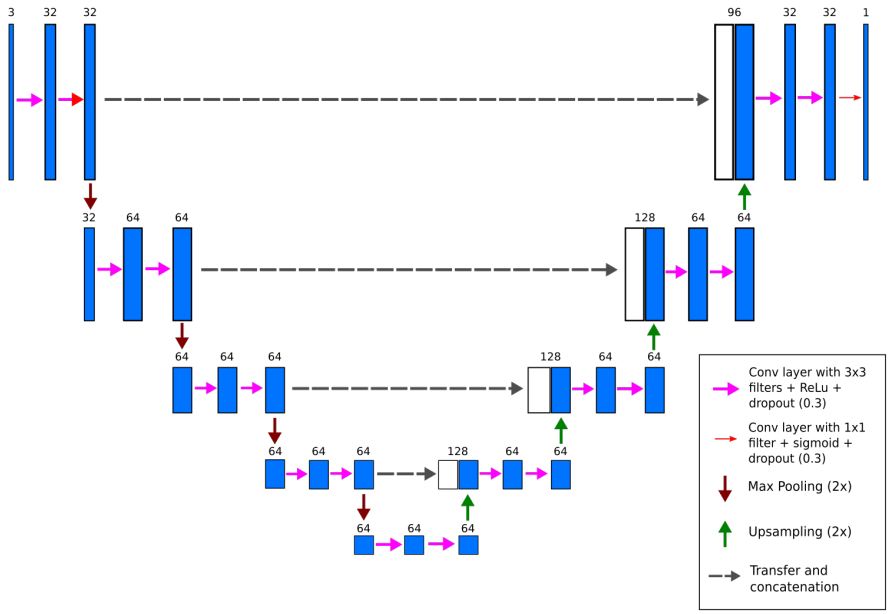

Fig. 3. U-Net architecture

binary output map.$d\,(A, B)$ is an extension of Dice coefficient :

$$Dice\,(A, B) = \frac{2\,|A \cap B|}{|A + B|} \tag{3}$$

We use a decay learning rate r_n for the optic disc and optic cup, like the formulation:

$$r_n = r_{n-1}\left(1 + \frac{0.001}{epoch}\right) \tag{4}$$

Where $epoch = 1000$.Learning rate decays as the number of iterations increases. Momentum is set to 0.95. Resolution of input images was set to 256×256 for optic disc and to 128×128 for optic cup segmentation.

Post-processing Since the output map contains extra connected areas,thus the misclassified class label shoud be corrected by. The object of maximum area is retained by considering it as the optic disc/optic cup region and classifying any other small objects as background class label.After outputting the predicted binary images, we use the ellipse fitting just for optic disc.

2.2 Experiment

We use a 4 fold cross validation scheme with 360 training images and 40 test images in each fold. In addition,each fold is composed of 90 normal eyes and 10

glaucomatous eyes.The performance of the proposed method for segmenting the optic disc and optic cup when compared with the ground truth was evaluated using Dice coefficient,like the equation (2),the Mean Absolute Deviation(MAD) of cup to disc ratio,which is defined as follows:

$$MAD - CDR = \frac{1}{n}\sum_{i=1}^{n}\left|c_i - c_i^{'}\right| \tag{5}$$

where $C = (c_i)_{n\times 1}$ is presented the predicted value of cup to disc ratio and $C^{'} = \left(c_i^{'}\right)_{n\times 1}$ is presented the real value of cup to disc ratio.The cup to disc ratio is the key to diagnosis glaucoma.This reflects the actual prediction error.

3 Classification

3.1 Methods

Pre-progressing our method is mainly composed of convolutional neural network for the classification of normal eyes and glaucomatous eyes.Same as above,we padding the image of train set to size 2124*2124, then resize the image to size 512×512.The input images is 512×512.The process of pretreatment is the same as above.We turn all the images to grayscale.The smaller the scale, the higher the training efficiency.

Nerual Netwook The neural network which is applied to extract the feature from the fundus images and classify the normal eyes and glaucoma is Xception[3]. The architecture is presented in fig.4. The Xception architecture has 36 convolutional layers forming the feature extraction base of the network. The basic idea is the channel separation convolution. Each convolution kernel in the depthwise convolution convolution only corresponds to one channel of input.
The activation function is ReLu. Momentum is set to 0.95.The learning rate is like equation (3). Learning rate decays as the number of iterations increases.

3.2 Experiment

As same above,we use 4-fold cross validation for the classification experiment.Each fold contacts 90 normal eyes and 10 glaucomatous eyes.The performance of classification is evaluate by sensitivity,AUC,which is defined as follows:

$$Sensitivity = \frac{t_p}{t_p + f_n} \tag{6}$$

where t_p ,t_n,f_p,f_n refer to true positivity,true negativity,false positivity,false negativity,respectively.AUC is the area value of the ROC curve.The ROC space defines the false positive rate (FPR or $1 - specificity$)as the X axis and the true positive rate (TPR or sensitivity) as the Y axis.Where:

$$FPR = \frac{f_p}{f_p + t_n} \tag{7}$$

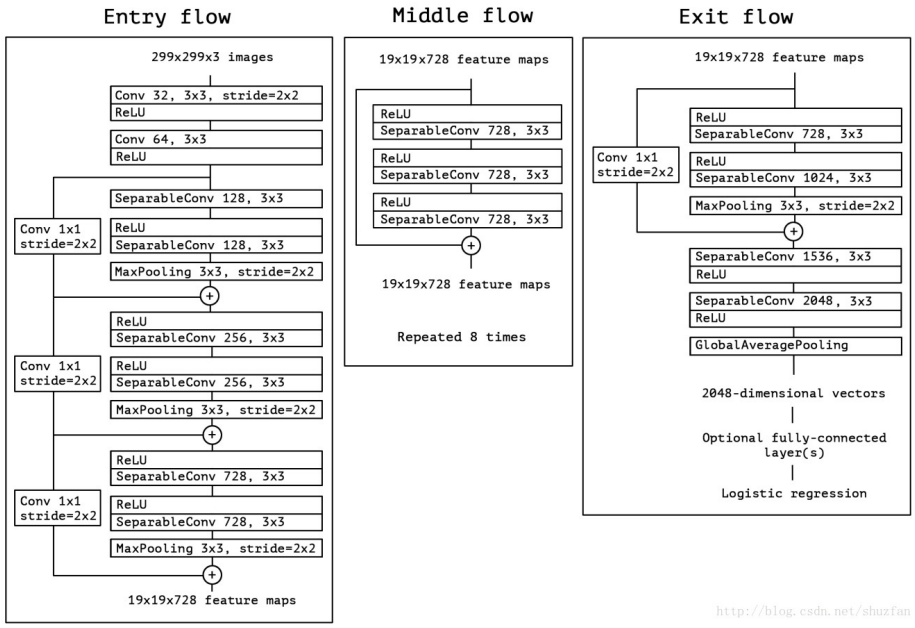

Fig. 4. the architecture of Xception

4 Conclusion

Compared with the original image, the grayscale image we use can preserve the information while eliminating the difference in image color and improving the generalization ability of the algorithm.We increase the data set by flipping the image, changing the brightness parameters, etc.In addition, we crop the images and reduce the network structure in order to improve training efficiency.
We use cross-validation to experiment on the training set to expect good performance on the test set.

References

1. Klein B E, Klein R, Sponsel W E, et al. Prevalence of glaucoma. The Beaver Dam Eye Study. Ophthalmology, **1992**, 99(10):1499-1504.
2. Sevastopolsky A. Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network. Pattern Recognition & Image Analysis.**2017**,27(3):618-624.
3. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. **2016**:1800-1807.