

Retinal Blood Vessel Segmentation Based on Multi-Scale Deep Learning

Ming Li

Information Science and Technology College,
Dalian Maritime University
Dalian, 116026, China
Email: limingdlmu@outlook.com

Qingbo Yin, Mingyu Lu

Information Science and Technology College,
Dalian Maritime University
Dalian, 116026, China,
Email: {qingbo, lumingyu}@dlmu.edu.com

Abstract—Fundus images are one of the main methods for diagnosing eye diseases in modern medicine. The vascular segmentation of fundus images is an essential step in quantitative disease analysis. Based on the previous studies, we found that the category imbalance is one of the main reasons that restrict the improvement of segmentation accuracy. This paper presents a new method for supervised retinal vessel segmentation that can effectively solve the above problems. In recent years, it is a popular method that using deep learning to solve retinal vessel segmentation. We have improved the loss function for deep learning in order to better handle category imbalances. By using a multi-scale convolutional neural network structure and label processing approach, our results have reached the most advanced level. Our approach is a meaningful attempt to improve blood vessel segmentation and further improve the diagnostic level of eye diseases.

I. INTRODUCTION

RETINAL fundus images have been widely used for diagnosis, screening and treatment of cardiovascular and ophthalmologic diseases[1], including age-related macular degeneration(AMD), diabetic retinopathy(DR), glaucoma, hypertension, arteriosclerosis and choroidal neovascularization, among which AMD and DR have been considered as two leading causes of blindness[2]. Vessel segmentation is a basic step for the quantitative analysis of retinal fundus images[3]. The segmented vascular tree can be used to extract the morphological attributes of blood vessels, such as length, width, branching and angles.

Moreover, the vascular tree has been adopted in multimodal retinal image registration [4]and retinal mosaic [5]as the most stable feature in the images. In [6], the vascular tree is also used for biometric identification due to its uniqueness. Manual segmentation of the vascular tree in retinal images is a tedious task that requires experience and skill. In the development of a computer-assisted diagnostic system for ophthalmic disorders, automatic segmentation of retinal vessels has been accepted as a vital and challenging step. The size, shape and intensity level of retinal vessels can vary hugely in different local areas. The width of a vessel often ranges from 1 to 20 pixels, depending on both the anatomical width of the vessel and the image resolution. The existence of vessel crossing, branching and centerline reflex makes it difficult to segment the vessels accurately using artificially designed features. Pathologies in

the form of lesions and exudates can further complicate the automatic segmentation. In the past decades, several methods have been proposed for the segmentation of vessels in retinal images, and they can be divided into two categories: unsupervised and supervised methods.

Both classic one-stage object detection methods, like boosted detectors [5]and DPMs(Deformable Parts Model) , and more recent methods, like SSD(Single Shot Multi-Box Detector) , face a large problem of class imbalance during training. These detectors evaluate huge candidate locations per image but only a few locations contain objects. This imbalance causes two problems: (1) training is inefficient as most locations are easy negatives that contribute no useful learning signal; (2) Simultaneously, the negatives can overwhelm training and lead to degenerate models. A common solution is to perform some form of hard negative mining [6] that samples hard examples during training or more complex sampling/reweighting schemes[7]. In contrast, we show that our proposed focal loss naturally handles the class imbalance faced by a one-stage detector and allows us to efficiently train on all examples without sampling and without easy negatives overwhelming the loss and computed gradients.

This paper presents a segmentation method that is suitable for class imbalance. This paper proposes a multi-scale convolutional neural network and improves the traditional loss function, and improves the class labels. Our method outweighs the most advanced methods reported in terms of sensitivity, specificity and accuracy. (1) The proposed method solves the problem of class imbalance in the traditional segmentation method, so that deep learning can better handle the task of image segmentation of the fundus. (2) A series of procedures proposed in the article can be used not only in segmentation tasks but also in various task types such as packet classification detection, and have a wide range of versatility.

II. THE PROPOSED METHOD

A. Motivation

We carefully combed the work of related work and found that most of the previous methods used to perform the two-class task were not evenly sampled. We have found that class imbalances lead to submerged samples, which is usually not

what we want. We explore a solution to the problem from three parts: loss function, network structure and category labels.

The others have done much work in designing robust loss functions (e.g., Huber loss) that reduce the contribution of outliers by down-weighting the loss of examples with large errors (hard examples). So our focal loss functions is designed to address class imbalance by down-weighting inliers (easy examples) such that their contribution to the total loss is small even if their number is large, rather than outliers. In other words, the focal loss performs the opposite role of a robust loss: it just trains on a sparse set of hard examples.

The multi-scale network structure is mainly composed of two deep convolutional network stacks, according to Jarrett . Thus a good multi-level network model to achieve effective target recognition is an important part in our work. The network model in this paper is shown in Figure 1. However, the traditional multi-scale method scales the image segments to different scales and cannot express the boundary of the target region accurately. Thus the method in this paper is based on the selection of each pixel in the image as the center. We can see in Figure 1 two different scale image segments generate input of two deep convolutional networks. The low-level feature extraction network structure first extracts low-level feature information of large-scale image segments, and then uses local refined network structure to capture local region feature information from small-scale image segments. To compare with the traditional multi-scale method, the difference is not that multiple scales separately extract features of different size images. In this paper, the low-level features are combined with the image features extracted from the first layer of the local refinement network. After subsequent network operations, a dense and complete feature vector is obtained, which greatly improves the accuracy of image pixel category prediction.

B. Loss function improvement

The usual method in the objective function of network optimization is the cross-entropy loss function, as followed:

$$E = - \sum_x p(x) \log(q(x)) \quad (1)$$

Where $p(x)$ is the true distribution of the sample and $q(x)$ is the estimated probability obtained through training. When the cross-entropy loss function is used for a two-category task, its form is:

$$E = \begin{cases} -\log(p), & \text{if } y = 1 \\ -\log(1 - p), & \text{otherwise} \end{cases} \quad (2)$$

The terrestrial truth category is specified in $y \in \{\pm 1\}$ above and $p \in [0, 1]$ is the estimated probability of the model for the category of label $y = 1$ For symbol convenience, we define p :

$$p = \begin{cases} p, & \text{if } y = 1 \\ 1 - p, & \text{otherwise} \end{cases} \quad (3)$$

However, in the category imbalance problem, sometimes the difference between positive and negative sample ratios is very different. Taking a fundus image as an example, blood

vessel pixels in a single image are only one-fifth that of non-vascular pixels. This imbalance leads to two kinds of situations: (1) Excessive samples of negative examples cause the information of positive examples to be difficult to be effectively learned and even concealed. (2) Simple negative factors lead to training shifts and make the model degenerate. Therefore, we propose an improvement to the cross-entropy loss function that can solve this kind imbalance problem. More crucially, this improvement can make the loss function able to calculate the difficulty in judging every sample, then we can give higher weight to those samples that are more difficult to distinguish. We call the new loss function LCE (Le Cross Entropy), LCE is defined as:

$$E = -\alpha * \cos(\beta * p) * \log(p) \quad (4)$$

$$\alpha = \frac{\alpha^*}{\beta} \quad (5)$$

Among them, the value of β is $\frac{\pi}{2}$. The reason for the value of β is that the range of p is $(0, 1)$, so that $\beta * p$ can have the same mapping range as p . The weight α is the balance coefficient, usually $\alpha \in [0, 1]$. This means that the real balance coefficient is actually α^* , but for the sake of clearly, we uses α to describe it. When using LCE as a loss function, we noticed that it has the following two characteristics: (1) When an example is misclassified and p is small, the loss will not be affected. When p is larger, then α goes to zero, and the loss of well-classified examples is reduced. (2) The balance factor α effectively adjusts the weight of the instance. With the change of α , LCE can adapt to different degrees of class imbalance. In our experiment, the best results are obtained when α is taken as 0.25. About the first feature, we take a specific example to illustrate. In the case of $\alpha = 0.2$, and $p = 0.9$, the results of LCE is 33.3 times lower than CE, and the value of loss function when $p \simeq 0.2$ is 50 times lower. From the data shown in the above examples, we can find that correcting misclassification examples is necessary.

C. The Proposed Architecture

In this study, we designed a multi-scale convolutional neural network structure to segment blood vessels from fundus images. In the figure 1, we will show the details of this network in this work. The network consists of two consecutive convolution structures with input sizes of $13 * 13$ and $17 * 17$ respectively. The convolution kernel we have chosen on each convolutional layer is $5 * 5$ in size and each time we move kernel one pixel. Networks of different scales have a similar convolutional layer structure. The first convolutional layer has 64 feature plots and the second convolutional layer has 128 feature plots, and the third convolutional layer has 256 feature plots. After each convolutional layer, we use a rectifying linear unit (ReLU) excitation as an activation function. As shown in this article[8], using ReLU as an activation function for the convolutional layer can speed up the training of the network. After the features are extracted from the convolution layer, we add the feature maps and connect them with the fully

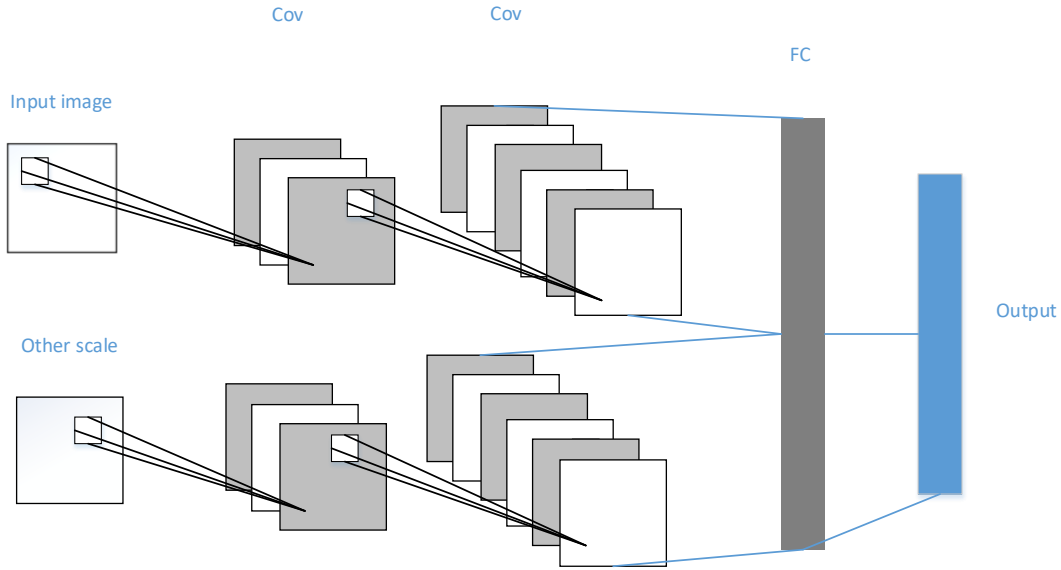


Fig. 1. The Structure of Multi-Scale Convolutional Neural Network

connected layer. It should be noted that the full-connected layer is not only one layer. Only one layer is drawn on the graph for the sake of simplicity of the view. In fact, there are three full-connected layers. We used dropout learn more powerful skills and reduce overfitting. The technique sets the output of each neuron to zero with a probability of 0.5. Finally, we use the softmax function to classify and use our improved cross-entropy function as an energy function.

D. Label Processing

In deep learning, the handling of sample tags is a part that is easily overlooked, especially in binary tasks. The tag y_- is simply designed as $y_- \in \{\pm 1\}$ or $y_- \in (0, 1)$. However, in the traditional machine method, the processing of the label has been proved to be an effective method to improve the accuracy of the classifier, and this method does not have any additional requirements for the computing power. Therefore, we try to introduce sparse variables into the label design of deep convolutional neural networks. The new tag y^* is defined as:

$$y^* = \alpha * y_- + \varepsilon \quad (6)$$

Where α is called the proportional coefficient, and its size depends on the proportion of the positive and negative samples in the training set. This coefficient was specially designed to solve the imbalance problem. ε is called a sparse variable. It is to bias the same kind of sample. We usually convert segmentation tasks into classification tasks, and convolutional neural networks as a classifier, its role can be simplified to find the optimal classification plane between different types of samples. By adding sparse variables to the category labels, increasing the distance between the classes' classification

TABLE I
PERFORMANCE OF ε -DRAGGING ON DATA POINT IN TWO CLASSES

	class	y	y after ε -Dragging	
x1	1	[1,0]	$[1 + \varepsilon_{11}, -\varepsilon_{12}]$	$\varepsilon_{11}, \varepsilon_{12} > 0$
x2	1	[1,0]	$[1 + \varepsilon_{21}, -\varepsilon_{22}]$	$\varepsilon_{21}, \varepsilon_{22} > 0$
x3	2	[0,1]	$[-\varepsilon_{31}, 1 + \varepsilon_{32}]$	$\varepsilon_{31}, \varepsilon_{32} > 0$
x4	2	[0,1]	$[-\varepsilon_{41}, 1 + \varepsilon_{42}]$	$\varepsilon_{41}, \varepsilon_{42} > 0$

planes has been used in traditional machine learning. We refer to this approach to deep convolutional neural networks.

Table 1 further explains our method, which reports the four data points in the two categories. Their class label vector is listed in the third column. Now, if we group together the first element of the class labeling vector, we get "1,1,0,0". In this way, a binary class partition can be obtained in which the first two is divided into one class, and the latter two classified data points are classified into another class. After label processing is performed, their image will change from "1,1,0,0" to " $1 + \varepsilon_{11}, 1 + \varepsilon_{12}, \varepsilon_{13}, \varepsilon_{14}$ " since all are non-negative, this processing can help expand the distance between classes after data point mapping.

E. Model training

Our method does not require image preprocessing including image enhancement, which greatly simplifies the difficulty of segmentation tasks and improves the versatility of the method. According to the mask image, the fundus image part of the original image is centered on each pixel, and we can get the classification of the preset scale, and after we delete the sample of the edge part sample beyond the mask range, remain about 150 samples. Millions of sample drawings. One million of

them are training sets and 500,000 are test sets. However, our method does not use all the training set samples for training. We used a special selection method to obtain samples that was only one-tenth of the original training set. This greatly speeds up the training time without losing too much classification accuracy.

We trained and tested the network on an Intel core computer and implemented it using Anconda+TensorFlow. The processor we used was the Intel Xeon(R) CPU E5-2680V3. The training on these lasted 18 hours. Between the limits of our experimental equipment, we chose a batch size of 64. During training, the weights were updated by stochastic gradient descent algorithm with a momentum of 0.9 and a weight decay of . The biases in convolutional layers and fully-connected layer were initialized to 1. The number of epochs was tuned on a validation set consisting of patches from one randomly selected subject in the training set. The learning rate was set to initially.

III. RESULT

A. Implementation Details

We have evaluated the nature of our proposed method on a very popular DRIVE dataset, which consists of 40 retinal images. The dataset is divided into two subsets, training sets and test sets, each set contains 20 images. The image is 565 x 584 pixels, 8 bits per color channel. Only the images of training set can be used for the training of the network. First, we clip the training image into 23x23 tiles and mark each tile by the label image. Then add a slack variable to the tag based on the method above and use the mask contained in the DRIVE database to identify the FOV. All pixels of each patch should be in the FOV area, and a total of 100000 patches are obtained for 20 training images. But not all small pieces will be trained. In fact, we only used 50,000 small pieces, which accounted for only about 2

Test the performance of the segmentation algorithm in the test set. In 20 images of the test set, four images are pathological and the other are normal images. During the testing phase, we also use the first expert's tag as a basic fact. The images in all test sets are used for testing to evaluate the performance of the algorithm.

B. Evaluation Criterion

In vessel segmentation, there are two class labels: vessel and non-vessel. By comparing the segmentation results with the manual ground truth, we obtain four measures: the vessel pixels that are predicted as vessels are denoted as true positives (TP), the vessel pixels that are predicted as non-vessels are denoted as false negatives (FN), the non-vessel pixels that are predicted as non-vessels are denoted as true negatives (TN), and the non-vessel pixels that are predicted as vessels are denoted as false positives (FP).

Usually we use three criteria to compare the performance of the proposed method with other state-of-the-art methods: sensitivity (Se), specificity (Sp) and accuracy (Acc). Evaluation indicators are only calculated for pixels within the FOV.

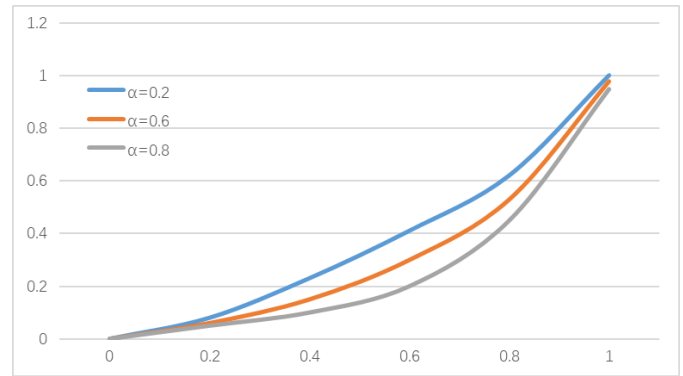


Fig. 2. Results of different segmentation methods

These metrics are defined as

$$Se = \frac{TP}{TP + FN} \quad (7)$$

$$Sp = \frac{TN}{TN + FP} \quad (8)$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \quad (9)$$

Because true positive score (Se) and false positive score (Sp) are sensitive to the number of sample categories. That is, when the number of samples of the binary classification problem is not balanced, these indicators cannot accurately reflect the performance of the classifier, so we also use the performance of the area under the ROC curve (AUC) evaluation method. The AUC is equal to 1 when the classifier can perfectly classify the sample. In addition, we have additionally introduced the concept of interclass accuracy.

This indicator better reflects the performance of the classifier when dealing with unbalanced categories.

C. Performance of the proposed method

Figure 2 is a comparison of our segmentation results with the traditional CNN segmentation results. Among them are (a) the original image, (b) the ground truth, (c) the segmentation result of the traditional CNN, and (d) the segmentation result of the proposed method. It can be clearly seen that our segmentation method is more delicate, and the segmentation accuracy of blood vessel details is higher and closer to the truth on the ground. The reason why our performance better is our method considers the effect of category imbalance on the segmentation of blood vessels. After modify the loss function and use multi-scale convolutional neural networks, we obviously reduce the impact of this problem on the results. In this method, the importance of the loss function is highlighted. Table III shows the comparison between the traditional CE and our proposed LCE. When using our structure at the same time, the accuracy of LCE is significantly higher than that of CE. Figure 3 shows the influence of different parameters ϵ on the results. And through the experiments, we can get the best results when $\epsilon=0.2$. The proposed method is evaluated

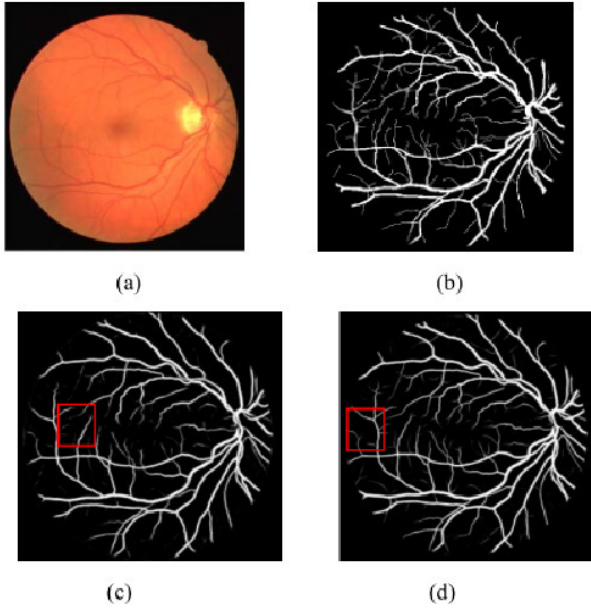
Fig. 3. The Effect of Different ϵ_A Values on Loss in LCE Functions

TABLE II

COMPARISON OF CORRECT RATIOS OBTAINED BY DIFFERENT SAMPLE RATIOS ON DRIVE DATA SETS

Num	1:1	1:2	1:3	1:4
Acc	0.917	0.932	0.944	0.956

on DRIVE and STARE database images with available real ground images. The performance results are shown in Tables II and III. The performance index is calculated based on the first human observer. The accuracy, sensitivity, specificity, and AUC of the DRIVE database were 0.8347, 0.9796, 0.951, and 0.9792, respectively. The accuracy of the segmentation results of the STARE database is 0.956; the sensitivity, specificity, and AUC are 0.94471, 0.99432, and 0.988388, respectively. Figure 4 shows the performance of ROC curves on the DRIVE and STARE datasets. The average AUC of the ROC curves on the two datasets is 0.9792, 0.9743. As we know, the closer the AUC value is to 1, the better the performance of the classifier. So our retinal vessel segmentation results are excellent.

Then, we compare the proposed method with several advanced retinal vessel segmentation methods. In general, supervised learning methods have better classification accuracy than unsupervised learning methods. The methods in the reference document achieve better results than other methods because of

TABLE III

COMPARISON OF CORRECT RATE OF CE AND LCE ON DIFFERENT DATA SETS

	DRIVE	STARE
CE	0.937	0.932
LCE	0.951	0.956

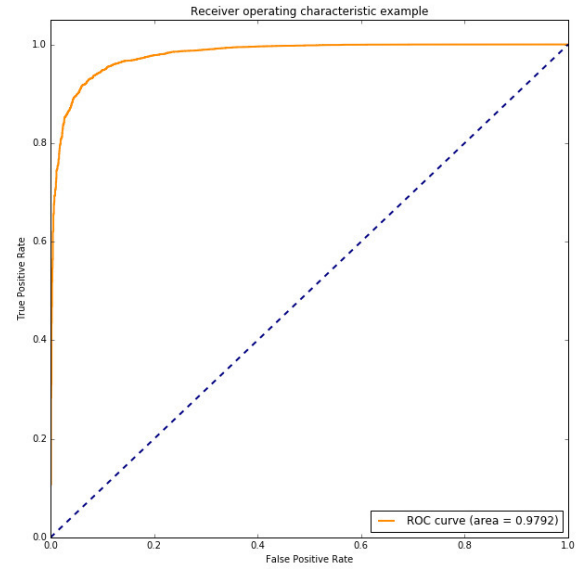


Fig. 4. ROC curve of proposed method

the use of ensemble learning methods. In a single classifier, our method has better average accuracy than other methods.

IV. DISCUSSIONS

Table II shows the accuracy of the classifier at different sample rates. On the DRIVE data set, we use the same network structure and evaluation indicators, different positive and negative sample ratios. From the above figure, we can clearly see that as the proportion of the sample much closer to the true proportion of the data set, the segmentation accuracy of our network becomes more accurate. Although it is not pursuit the unbalanced proportion on purpose, when randomly selecting small blocks from the sample set, it is actually obtained an imbalanced sample. This method can quickly classify the network, but it has no practical meaning. Because the network learning is the distribution state of the sample set itself, rather than the real characteristics of the sample.

After realizing that the above approach is not rigorous, we consider how to solve this problem. The modification of the loss function and network structure is the method that first enters our mind. Table III shows the performance of different loss functions on two data sets. After we used the improved loss function, the performance of the network has improved significantly. It is worth noting that our proposed improvement of the loss function not only has a broad prospect in this network but also in the broader field of deep learning. We are conducting experiments in this area and will soon have results. According to the definition of LCE, ϵ_A is a very important hyper parameter, which determines the performance of this loss function on a specific problem. Figure 3 shows the impact of different ϵ_A on network performance when we use LCE.

TABLE IV
PERFORMANCE COMPARISON OF VESSEL SEGMENTATION METHODS ON DRIVE IMAGES

No	Methods	Se	Sp	Acc	Auc
1	Fraz[1]	0.7302	0.9472	0.9422	N.A
2	Fraz[9]	0.7406	0.9807	0.9480	0.9747
3	Soares[10]	0.7283	0.9788	0.9466	0.9616
4	George[3]	0.7655	0.9704	0.9442	0.9614
5	Nicola[11]	0.7731	0.9724	0.9467	0.9588
6	Aslani[12]	0.7545	0.9801	0.9513	0.9682
7	Maji[13]	N.A	N.A	0.9470	0.9283
8	Lahiri[14]	0.7500	0.9800	0.9480	0.9500
9	Martina[15]	0.7276	0.9785	0.9466	0.9749
10	Avijit[16]	0.7691	0.9801	0.9533	0.9744
11	Fu[17]	0.7294	N.A	0.9470	N.A
12	Proposed method	0.8347	0.9796	0.9510	0.9792

TABLE V
PERFORMANCE COMPARISON OF VESSEL SEGMENTATION METHODS ON STARE IMAGES

No	Methods	Se	Sp	Acc	Auc
1	Hoover[18]	0.6747	0.9384	0.9348	N.A
2	Jiang[19]	N.A	N.A	0.9009	N.A
3	Mendonca[20]	0.6996	0.9730	0.9440	N.A
4	Lam[21]	N.A	N.A	0.9567	0.9739
5	You[22]	0.7260	0.9756	0.9479	N.A
6	Marin[15]	0.6944	0.9819	0.9526	0.9769
7	Fraz[1]	0.7548	9763	0.9534	0.9768
8	Proposed method	0.8231	0.9782	0.9560	0.9743

After our experiments, we found that when $\epsilon\hat{A}=0.2$, the best effect was obtained.

After using the method, we mentioned above, we have achieved very good results on the two data sets. The accuracy, sensitivity, specificity, and AUC of the DRIVE database were 0.8347, 0.9796, 0.951, and 0.9792, respectively. The accuracy of the segmentation results of the STARE database was 0.956; the sensitivity, specificity, and AUC were 0.94471, 0.99432, and 0.988388, respectively. In particular, except that our experiment is based on the results of a balanced sample size, other experimental results can be obtained when the sample is not balanced.

For blood vessel segmentation tasks, we are more likely to get blood vessel pixels than non-vascular pixels because blood vessel pixels are very rare and their value is much higher than non-vascular pixels. Therefore, in Table IV and Table V, our results showed a significantly higher specificity than other results. That means, in the case of a balanced sample, we only lost a little bit of accuracy, but we improved our specificity significantly. This is what we are happy to get. Our experiment has embarked on a new direction for the next study, which is not to regard accuracy as the first criterion, but rather to focus on specificity.

V. CONCLUSION

By comparing the differences of the experimental results, we found that there are unbalanced samples in the fundus

image segmentation task, and we hope to improve the segmentation accuracy of blood vessels. Further we propose three solutions. By using these three methods together, we can get a better result. The deep neural network can learn hierarchically preprocessed images from it. It has a great potential in medical image processing and can help doctors easily diagnose accurately. In this paper, firstly, we use the slack variable method to increase the distance among different categories, thereby improving the performance of the classifier. Secondly, we propose a multi-scale convolutional neural network to extract the difference in information among different views, so that we can make accurate judgments. Finally, we solve the problem of unbalanced quantity among different types of samples by modifying the loss function. Our proposed method has performed well on two common data sets.

As we said above, we are conducting more tests and improvements on LCE so that it can perform well when dealing with unbalanced tasks for example target detection. The other limitation of our method is that it requires more training time than the previous method. Obviously, multi-scale networks have more parameters and the introduction of slack variables, which slows down the training speed of the network. Although the loss of time seems unavoidable, it can be acceptable to improve the accuracy. In the future, we hope to continue to improve the network structure so that it can be trained and tested more quickly.

REFERENCES

- [1] M. M. Fraz, A. Basit, and S. A. Barman, "Application of morphological bit planes in retinal blood vessel extraction." *Journal of Digital Imaging*, vol. 26, no. 2, pp. 274–286, 2013.
- [2] M. D. Abrf'd'moff, J. C. Folk, D. P. Han, J. D. Walker, D. F. Williams, S. R. Russell, P. Massin, B. Cochener, P. Gain, and L. Tang, "Automated analysis of retinal images for detection of referable diabetic retinopathy," *Jama Ophthalmology*, vol. 131, no. 3, p. 351, 2013.
- [3] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable cosfire filters for vessel delineation with application to retinal images," *Medical image analysis*, vol. 19, no. 1, pp. 46–57, 2015.
- [4] F. Zana and J.-C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and hough transform," *IEEE transactions on Medical Imaging*, vol. 18, no. 5, pp. 419–428, 1999.
- [5] P. Dollár, Z. Tu, P. Perona, and S. Belongie, "Integral channel features," in *British Machine Vision Conference, BMVC 2009, London, UK, September 7-10, 2009. Proceedings*, 2009.
- [6] P. F. Felzenszwalb, R. B. Girshick, and D. McAllester, "Cascade object detection with deformable part models," in *Computer vision and pattern recognition (CVPR), 2010 IEEE conference on*. IEEE, 2010, pp. 2241–2248.
- [7] S. R. Buló, G. Neuhof, and P. Kotschieder, "Loss maxpooling for semantic image segmentation," *CVPR*, July, vol. 7, 2017.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [9] M. 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 Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.
- [10] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE Transactions on medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.
- [11] N. Strisciuglio, G. Azzopardi, M. Vento, and N. Petkov, "Supervised vessel delineation in retinal fundus images with the automatic selection of b-cosfire filters," *Machine Vision and Applications*, vol. 27, no. 8, pp. 1137–1149, 2016.
- [12] S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomedical Signal Processing and Control*, vol. 30, pp. 1–12, 2016.
- [13] D. Maji, A. Santara, P. Mitra, and D. Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," *arXiv preprint arXiv:1603.04833*, 2016.
- [14] A. Lahiri, A. G. Roy, D. Sheet, and P. K. Biswas, "Deep neural ensemble for retinal vessel segmentation in fundus images towards achieving label-free angiography," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 1340–1343.
- [15] M. Melinščak, P. Prentašić, and S. Lončarić, "Retinal vessel segmentation using deep neural networks," in *VISAPP 2015 (10th International Conference on Computer Vision Theory and Applications)*, 2015.
- [16] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*. IEEE, 2017, pp. 248–251.
- [17] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 698–701.
- [18] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, pp. 203–210, 2000.
- [19] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, pp. 131–137, 2003.
- [20] A. M. Mendonca and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE transactions on medical imaging*, vol. 25, no. 9, pp. 1200–1213, 2006.
- [21] B. S. Lam, Y. Gao, and A. W.-C. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling," *IEEE Transactions on Medical Imaging*, vol. 29, no. 7, pp. 1369–1381, 2010.
- [22] X. You, Q. Peng, Y. Yuan, Y.-m. Cheung, and J. Lei, "Segmentation of retinal blood vessels using the radial projection and semi-supervised approach," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2314–2324, 2011.