An improved U-Net for nerve fibre segmentation in confocal corneal microscopy images

Zhou, Xinxin, Chen, Xinjian, Feng, Shuanglang, Shi, Fei

Xinxin Zhou, Xinjian Chen, Shuanglang Feng, Fei Shi, "An improved U-Net for nerve fibre segmentation in confocal corneal microscopy images," Proc. SPIE 11313, Medical Imaging 2020: Image Processing, 113131Z (10 March 2020); doi: 10.1117/12.2548257

SPIE.

Event: SPIE Medical Imaging, 2020, Houston, Texas, United States

An Improved U-Net for Nerve Fibre Segmentation in Confocal Corneal Microscopy Images

Xinxin Zhou[1], Xinjian Chen[1,2] , Shuanglang Feng[1] , Fei Shi[1,*]

[1]School of Electronics and Information Engineering, Soochow University, Suzhou, Jiangsu Province, 215006, China

[2] State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, 215123, China

ABSTRACT

Corneal confocal microscopy (CCM) is a new technique offering non-invasive and fast imaging useful for diagnosing and analyzing corneal diseases. The morphology of corneal nerve fibres can be clearly observed from CCM images. Segmentation and quantification of nerve fibres is important for analyzing corneal diseases such as diabetic peripheral neuropathy (DPN). In this paper, we propose an automated deep learning based method for corneal nerve fibre segmentation in CCM images. The main contributions of this paper are: (1)We add multi-scale split and concatenate (MSC) blocks to the decoding part of the four layer U-Net architecture. (2) A new loss function is applied that combining the Dice loss with the fibre length difference between the ground truth and the prediction. The method was tested on a dataset containing 90 CCM images from 4 normal eyes and 4 eyes with corneal diseases. The Dice coefficient of our approach can reach 87.96%, improves 1.6% compared with the baseline, and outperforms some existing deep networks for segmentation.

KEYWORDS: CCM imaging, nerve fibre segmentation, deep learning, MSC block, nerve fibre length difference

1. INTRODUCTION

There is a growing interest on the segmentation and analysis of corneal nerve images. These images, obtained in vivo, in conscious patients, through corneal confocal microscopy(CCM)[1], can document corneal nerve changes. Corneal injury and corneal diseases are often associated with changes in corneal nerve fibres. The quantification of nerve fibers offers information useful for analysis of many clinical cases, such as postoperative regeneration and repair of corneal injury, the effect of long-term contact lenses wearing, and different degrees of diabetic peripheral neuropathy (DPN) [2,3].

In the past, the analysis of corneal nerve parameters was usually based on a manual laborious task that is subjective and prone to errors[4-7]. At present, fully automatic methods of nerve fibre segmentation, requiring no specific expertise from the user, have been proposed. Ferreira *et al.* used adaptive histogram equalization to enhance image contrast, adopted wavelet transform filtering algorithm based on phase symmetry, and used a set of artificially selected seed points for neural reconstruction[8]. Scarpa *et al.* identified a set of seed points as the starting point of neural tracking, and used fuzzy c-means clustering to divide pixels into neural pixels and background pixels[9]. Dabbah *et al.* proposed an automatic neural analysis and classification system for corneal images under confocal microscopy based on multi-scale dual-model detection algorithm. In the classification stage, pixels are divided into neural or non-neural pixels based on random forest (RF) and neural network (NN)[10].

However，for some reasons, traditional methods cannot segment nerve fibres well. For example, corneal cells beside the nerve fibres may be a disturbance. Then for corneas with diseases, the abnormal areas in CCM images may reduce the segmentation accuracy. The segmentation of fine nerve fibres is also a big challenge. Therefore, in this paper, to deal with these problems, we consider applying a deep learning method with improved network architecture and loss function to obtain better segmentation performance.

*Corresponding author: E-mail: shifei@suda.edu.cn

Medical Imaging 2020: Image Processing, edited by Ivana Išgum, Bennett A. Landman,
Proc. of SPIE Vol. 11313, 113131Z · © 2020 SPIE · CCC
code: 1605-7422/20/$21 · doi: 10.1117/12.2548257

2. METHODS

2.1 Structure of the proposed deep network

U-Net was originally proposed to segment cells[11], and was proved to get quite good performance in lots of medical image segmentation tasks. We modify it to segment nerve fibres in corneal confocal microscopy images. Fig.1(a) shows the whole network and Fig.1(b) show the proposed multi-scale split and concatenate (MSC) block. As can be seen from Fig.1(a),we first reduce the original five-layer U-Net to a four-layer U-Net where the number of parameters is greatly reduced and the segmentation performance remains the same. Then in the decoding phase, we add the multi-scale MSC block after each up-sampling operation. The MSC block introduces various sizes of receptive fields in each network layer. Therefore the network can extract more feature information.

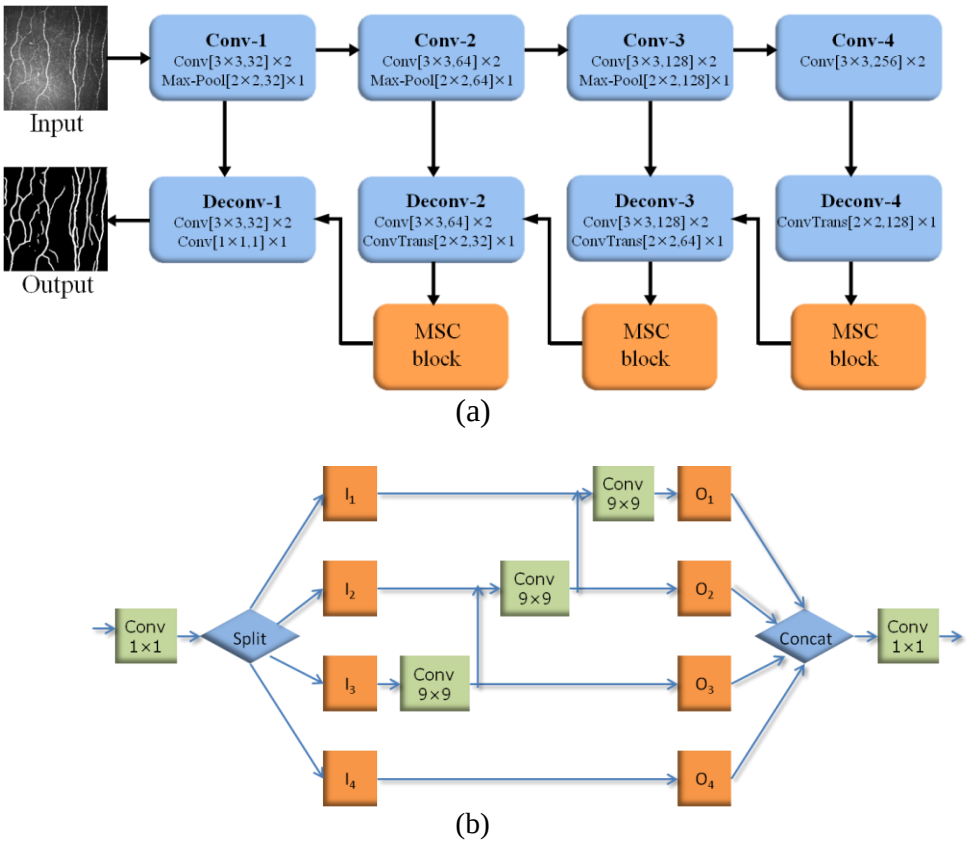


Fig.1. An overview of the improved U-net. (a) the network architecture. (b) the MSC block

2.2 The MSC block

The MSC block is inspired by the Res2Net model proposed for computer vision tasks[12], and is shown in Fig.1(b). After the 1×1 convolution, the feature maps are evenly split into 4 subsets of channels, denoted by I_1 to I_4. Except for I_4, each I_i goes through a corresponding 9×9 convolution, denoted by $K_i()$. We denote by O_i as the output of $K_i()$. The feature subset I_i is added with the output of $K_{i+1}()$, and then fed into $K_i()$.Thus, O_i can be written as:

$$O_i = \begin{cases} I_i & i = 4; \\ K_i(I_i + O_{i+1}) & 1 \le i < 4. \end{cases} \quad (1)$$

So the groups of feature maps go through 9×9 convolutions and concatenations in a hierarchical way. Finally, multi-scale feature maps O_1 to O_4 are obtained. To better fuse information at different scales, the four groups of feature maps are

concatenated and pass through a 1×1 convolution to get the output. Due to the combinatorial explosion effects, the output of the MSC block contains features with receptive fields of different sizes. As nerve fibres are linear structures with different width, the split and concatenation strategy can help to extract both thick and thin fibres. On the whole, the MSC block can help the U-Net to extract both global and local information, and thus enhances the ability of feature extraction.

2.3 Loss function

First, the Dice loss function[13] is applied to make the network more focused on the foreground, which occupies a small proportion of the image. However, with the Dice loss, thicker fibres contribute more, and the network can be biased toward accurate segmentation of thicker fibres and tends to ignore thinner and shorter ones. Therefore we consider incorporating the length difference into the loss function, so that the bias caused by thickness can be reduced.

We generate a binary segmentation map by applying a hard threshold 0.5 to the network output. Thinning operation is then applied to both the ground truth denoted as x and the segmentation map denoted as y. The total number of fibre pixels in the skeleton map are calculated and denoted as the total fibre length l_x and l_y. The mismatch ratio between them is defined as:

$$mr = \frac{|l_x - l_y|}{l_x} \tag{2}$$

Then the proposed loss function is defined as:

$$Loss = \alpha \times L_{Dice} \tag{3}$$

where

$$\alpha = 1 + mr \tag{4}$$

When the prediction map loses more nerve fibres, the index mr will be bigger. Then $Loss$ will be bigger and the network will adjust the parameters towards reducing the mr. Therefore this loss function can guide the model to pay more attention to the difference of nerve fibre length in the optimization process, and further solve the problem of class imbalance, which helps the detection of fine fibres.

3. RESULTS

3.1 Datasets

The dataset used in this paper includes 90 two-dimensional nerve fibre images obtained by a corneal confocal microscope with image size of 384×384, corresponding to 400μm×400μm. For these 90 images, 50 images were obtained from 4 normal eyes and the other 40 images were from 4 eyes with corneal diseases.

3.2 Implementation Details

Data augmentation including horizontal or vertical flipping, affine transformation and additive Gaussian noise was applied. In the training process, the stochastic gradient descent (SGD) algorithm with an initial learning rate of 0.01 and a momentum of 0.9 was used to optimize the network. The batchsize was 2, and the number of epochs was 80.

The training was done in an end-to-end way with the 90 nerve fibre images and corresponding ground truth segmentation maps. Four-fold cross validation was used in our experiments. In each fold, the training set and testing set are divided patient-wise.

We compare segmentation results of the proposed method (Baseline+MSC+αL_{Dice}) with those obtained by the original U-Net[2], the baseline (four-layer U-net architecture), and the state-of-the-art segmentation networks SegNet[14] and ERFNet[15]. We also compared with model variations where MSC block was used with only the Dice loss (Baseline+MSC), and where the new loss was applied but the MSC block was left out (Baseline+αL_{Dice}).

3.3 Metrics

The following five metrics: accuracy (*Acc*)[16], the Dice similarity coefficient (*DSC*), the area under the ROC curve(*AUC*), sensitivity(*Se*), and specificity (*Sp*) are calculated to quantitatively evaluate the performance of our method.

Some are defined as:

$$Acc = \frac{TP+TN}{TN+TP+FN+FP}\ ,\ DSC = \frac{2\times TP}{(TP+FP)+(TP+FN)}\ ,$$

$$Se = \frac{TP}{TP+FN}\ ,\qquad Sp = \frac{TN}{TN+FP}\ , \tag{5}$$

where *TP* denotes the true positive, *FP* denotes the false positive, *FN* denotes the false negative, *TN* denotes the true negative. The receiving operator characteristics (ROC) curve is computed with the true positive ratio (*Se*) versus the false positive ratio ($1-Sp$) with respect to a varying threshold, and the area under the ROC curve (AUC) is calculated for quality evaluation.

3.4 Results

Fig.2.shows the performance trend comparison(*Acc* and *DSC*) on the test set between baseline (four-layer U-net architecture) and our method(Baseline+MSC+αL_{Dice}) during the training process, it can be seen that for both methods, as the epoch increases, the performance improves and then tends to be stable. In addition, our method performs better than the baseline regarding both metrics. This trend in the training process proves the reliability of our method.

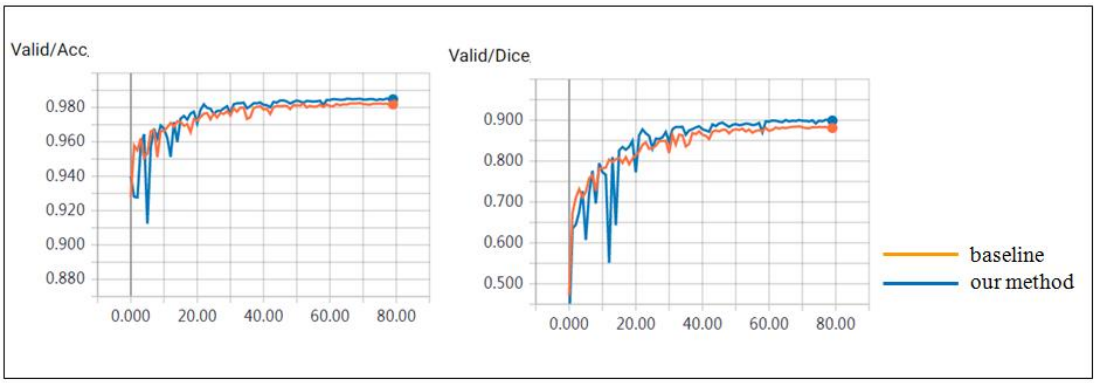


Fig.2. Validating Acc and DSC of baseline and our method.

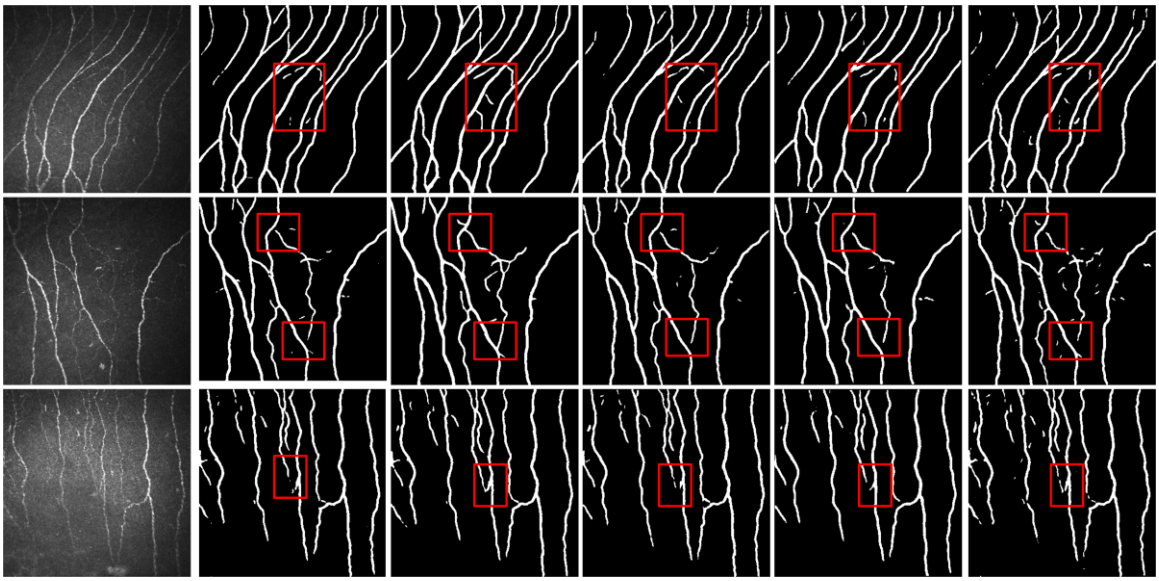
Table I shows the comparison of metric results on the normal, abnormal and total dataset. We can see that, compared with the baseline, the two proposed modifications can each obtain better performance, on both normal and abnormal data. The MSC block helps the network improve 0.6% on *DSC*, and the new loss function helps improve 1.2% on *DSC*. The combination of them obtains the best performance, with improvement of 1.62%. For the normal dataset, the proposed method achieves 0.8967, 0.9748 and 0.9347 for *DSC*, *Acc* and *AUC* respectively. For the abnormal dataset, the proposed method achieves 0.8581, 0.9834 and 0.9238 for *DSC*, *Acc* and *AUC* respectively. In terms of total dataset, our method achieves 0.8796, 0.9786 and 0.9299 for *DSC*, *Acc* and *AUC* respectively. Our approach also outperforms the original U-Net, ERFNet and SegNet.

Testing results of three representative CCM images are displayed in Fig.3. Fig.3.(a) shows the original nerve fibre images, where the first two images are from the normal dataset and the third is from the abnormal dataset. From fig.3.(b) to Fig.3.(f), it respectively shows the probability map generated by the baseline, our method, SegNet, U-Net and ERFNet . It can be seen that, compared with the baseline and other methods, our method can identify more thin and low-contrast nerve fibres on both normal and abnormal data.

4. CONCLUSIONS

In this paper, we propose a deep learning network for segmentation of corneal nerve fibres from CCM images, where the structure of U-net is improved by adding the MSC blocks, and the loss function is improved by combining the Dice loss and the fibre length difference. The MSC block introduces receptive fields of various sizes and helps extract more feature information. The loss function can guide the model to pay more attention to the difference of nerve fibre length and further solve the problem of class imbalance, which helps the detection of fine fibres. Experimental results

demonstrated the effectiveness of our method. The proposed method performs well for both normal and abnormal data. It provides more accurate information for the quantitative analysis of corneal nerve fibres.



(a) (b) (c) (d) (e) (f)

Fig.3. Experimental results in CCM images. (a) The original nerve fibre images (the first two images are from the normal dataset and the third is from the abnormal dataset). (b) The probability map generated by the baseline.(c) The probability map generated by our method. (d) The probability map generated by SegNet.(e) The probability map generated by U-Net. (f) The probability map generated by ERFNet.

Table I Comparison results on the normal, abnormal and total dataset.

Methods	DSC		Acc		AUC		Se		Sp	
	normal	total	normal	total	normal	total	normal	total	normal	total
	abnormal		abnormal		abnormal		abnormal		abnormal	
ERFNet	0.8689	0.8528	0.9680	0.9713	0.9191	0.9169	0.8523	0.8453	0.9845	0.9863
	0.8326		0.9795		0.9141		0.8366		0.9885	
SegNet	0.8857	0.8700	0.9721	0.9765	0.9267	0.9248	0.8647	0.8589	0.9874	0.9886
	0.8505		0.9820		0.9224		0.8517		0.9901	
U-Net	0.8820	0.8633	0.9713	0.9757	0.9260	0.9216	0.8642	0.8528	0.9866	0.9881
	0.8398		0.9811		0.9159		0.8385		0.9899	
Baseline	0.8821	0.8634	0.9714	0.9757	0.9262	0.9218	0.8646	0.8534	0.9865	0.9880
	0.8401		0.9811		0.9163		0.8369		0.9899	
Baseline+MSC	0.8884	0.8696	0.9729	0.9769	0.9304	0.9251	0.8726	0.8595	0.9871	0.9886
	0.8461		0.9819		0.9184		0.8433		0.9904	
Baseline+αL_{Dice}	0.8916	0.8754	0.9737	0.9779	0.9300	0.9254	0.8703	0.8587	0.9883	**0.9899**
	0.8553		0.9832		0.9196		0.8443		**0.9918**	
Baseline+MSC+αL_{Dice}	**0.8967**	**0.8796**	**0.9748**	**0.9786**	**0.9347**	**0.9299**	**0.8799**	**0.8679**	**0.9884**	0.9897
	0.8581		**0.9834**		**0.9238**		**0.8531**		0.9914	

5. ACKNOWLEDGEMENTS

This work was supported in part by the National Key R&D Program of China under Grant 2018YFA0701700, the National Basic Research Program of China (973 Program) under Grant 2014CB748600, and in part by the National Natural Science Foundation of China (NSFC) under Grant 61622114 61971298, 61771326.

6. REFERENCE

[1] R A Malik , P Kallinikos , C A Abbott, *et al*, "Corneal confocal microscopy: a non-invasive surrogate of nerve fibre damage and repair in diabetic patients," *Diabetologia*, vol. 46, no. 5, pp.683-688, 2003.

[2] X. Chen, J. Graham, M. Dabbah, *et al*, "An automatic tool for quantification of nerve fibres in corneal confocal microscopy images," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 4,pp.786-794, 2016.

[3] P. Hossain, A. Sachdev, R.A. Malik, "Early detection of diabetic peripheral neuropathy with corneal confocal microscopy,"*The Lancet*, vol. 366, no. 9494, pp.1340-1343,2005.

[4] D.V. Patel, C.N. McGhee, "Contemporary in vivo confocal microscopy of the living human cornea using white light and laser scanning techniques: a major review," *Clinical and Experimental Ophthalmology* ,vol.35, no.1, pp. 71 – 88,2007.

[5] C. Quattrini, M. Tavakoli, M. Jeziorska, *et al* "Surrogate markers of small fiber damage in human diabetic neuropathy," *Diabetes*, vol.56, no.8,pp. 2148 – 2154, 2007.

[6] D.V. Patel, C.N.J. McGhee, "Mapping of the normal human corneal sub-basal nerve plexus by in vivo laser scanning confocal microscopy,"*Investigative Ophthalmology and Visual Science* , vol.46, no.12, pp.4485 – 4488, 2005.

[7] P. Kallinikos, M. Berhanu, C. O ' Donnell, *et al*, "Corneal nerve tortuosity in diabetic patients with neuropathy,"*Investigative Ophthalmology and Visual Science* , vol. 45, no. 2, pp.418 – 422,2004.

[8] A. Ferreira , M. M. António, S. S. José, "A method for corneal nerves automatic segmentation and morphometric analysis," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 1, pp.53-60, 2012.

[9] F. Scarpa , E. Grisan , A. Ruggeri , "Automatic recognition of corneal nerve structures in images from confocal microscopy," *Investigative Ophthalmology and Visual Science*, vol. 45, no. 2, pp.4801-4807, 2008.

[10] M. a Dabbah, J. Graham, I. N. Petropoulos, *et al* "Automatic analysis of diabetic peripheral neuropathy using multi-scale quantitative morphology of nerve fibres in corneal confocal microscopy imaging," *Medical Image Analysis*, vol. 15, no. 5, pp. 738 – 747, 2011.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, pp. 234–241,2015.

[12] S. Gao, M. Cheng, K. Zhao, *et al*, "Res2Net: A new multi-scale backbone architecture," arXiv:1904.01169 [cs.CV], 2019.

[13] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *3D Vision (3DV), 2016 Fourth International Conference on.* IEEE, 2016,pp. 565–571.

[14] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," arXiv:1511.00561 [cs.CV], 2015.

[15] E. Romera, J. M. Alvarez, L. M. Bergasa, *et al*, "ERFNet: Efficient residual factorized convnet for real-time semantic segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol.19, no.1, pp.263-272, 2018.

[16] Z. Yan , X. Yang, K. T. Cheng , " Joint Segment-Level and Pixel-Wise Losses for Deep Learning Based Retinal Vessel Segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 15, no. 5, pp.1912-1923, 2018.