Deep Classification and Segmentation Model for Vessel Extraction in Retinal Images

Yicheng Wu[1], Yong Xia[1,2(✉)], and Yanning Zhang[1]

[1] National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China
`yxia@nwpu.edu.cn`
[2] Centre for Multidisciplinary Convergence Computing (CMCC), School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China

Abstract. The shape of retinal blood vessels is critical in the early diagnosis of diabetes and diabetic retinopathy. Segmentation of retinal vessels, particularly the capillaries, remains a significant challenge. To address this challenge, in this paper, we adopt the "divide-and-conque" strategy, and thus propose a deep neural network-based classification and segmentation (CAS) model to extract blood vessels in color retinal images. We first use the network in network (NIN) to divide the retinal patches extracted from preprocessed fundus retinal images into wide-vessel, middle-vessel and capillary patches. Then we train three U-Nets to segment three classes of vessels, respectively. Finally, this algorithm has been evaluated on the digital retinal images for vessel extraction (DRIVE) database against seven existing algorithms and achieved the highest AUC of 97.93% and top three accuracy, sensitivity and specificity. Our comparison results indicate that the proposed algorithm is able to segment blood vessels in retinal images with better performance.

Keywords: Retinal vessels segmentation · Deep learning Classification and segmentation

1 Introduction

Diabetic retinopathy (DR) is one of the most serious and common complications, the leading cause of visual impairment in many countries. Early diagnosis of diabetes and DR, in which vessels segmentation in color images of the retina plays a pivotal role, is critical for best patient care.

A number of automated retinal vessels segmentation algorithms have been published in the literature, which can be roughly categorized into three groups. First, vessels segmentation algorithms can be guided by the prior domain knowledge. Staal et al. [1] introduced the vessels location and others heuristics to the

vessels segmentation process. Second, Lupascu et al. [2] jointly employed filters with different scales and directions to extract 41-dimensional visual features and applied the AdaBoosted decision trees to those features for vessels segmentation. Next, recent years have witnessed the success of deep learning techniques in retinal vessels segmentation. Liskowski et al. [3] trained a deep network with the augmented blood vessels data at variable scales to facilitate segmentation. Li et al. [4] adopted an auto-encoder to initialize the neural network for vessel segmentation and avoided the preprocessing of retinal images. Ronneberger et al. [5] proposed a fully convolutional network called U-Net for retinal vessels segmentation. Fu et al. [6] applied deep neural network and fully connected conditional random field (FC-CRF) to improve the performance. Despite their success, deep learning-based algorithms still suffer from mis-segmentation, particularly in capillary regions.

Comparing to major arteries and veins, capillaries have smaller diameter, extremely lower contrast and very different appearance in retinal images. Therefore, we suggest that applying different methods to segment retinal vessels with different widths. In this paper, we adopt the "divide and conque" strategy and propose a deep neural network-based classification and segmentation (CAS) model to extract blood vessels in color retinal images. We first extract patches from preprocessed retinal images, and then classify these patches into three categories: wide-vessel, middle-vessel and capillary patches. Next, we construct three U-Nets to segment three categories of patches. Finally, we use those segmented patches to reconstruct the completed retinal vessels. We have evaluated the proposed algorithm against seven existing retinal vessels segmentation algorithms on the benchmark digital retinal images vessel extraction (DRIVE) database [1].

2   Dataset

The DRIVE database used for this study comes from a diabetic retinopathy screening program initiated in Netherlands. It consists of 20 training and 20 testing fundus retinal color images of size 584×565. These images were taken by optical camera from 400 diabetic subjects, whose ages are 25–90 years. Among them, 33 images do not have any pathological manifestations and the rest have very small signs of diabetes. Each original image is equipped with the ground truth from two experts' manual segmentation and the corresponding mask.

3   Method

The proposed CAS model consists of retinal images preprocessing, retinal patches extraction, classification and segmentation and retinal vessels reconstruction. A diagram that summarizes this algorithm is shown in Fig. 1.

3.1   Retinal Image Preprocessing

To avoid the impact of hue and saturation, we calculate the intensity at each pixel, and thus convert each color retinal image into a gray-level image. Then,

we apply the contrast limited adaptive histogram equalization (CLAHE) algorithm [7] and gamma adjusting algorithm to each gray-level image to improve its contrast and suppress its noise.
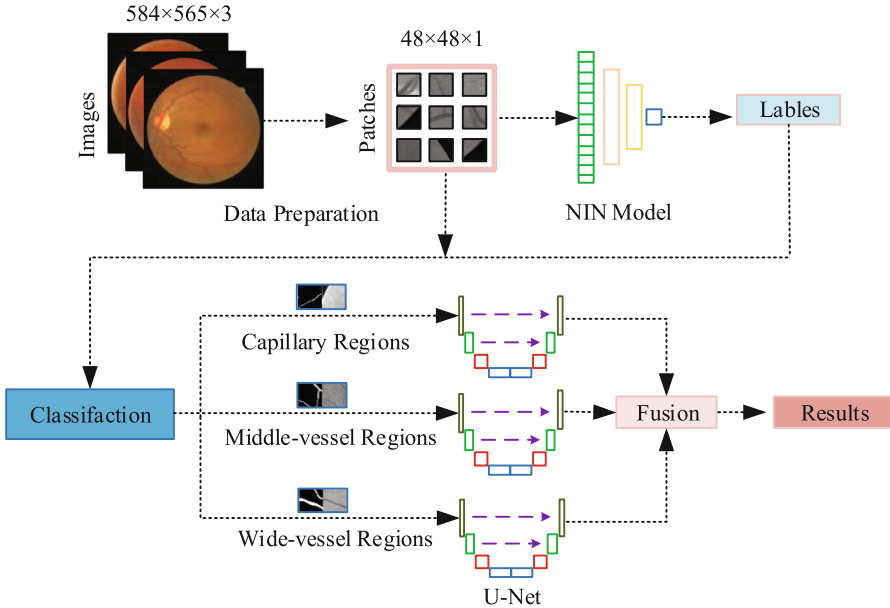

Fig. 1. Diagram of the proposed CAS model (color figure online)

3.2 Training a Patch Classification Network

A fundus retinal image contains a dark macular area, a bright optic disc region, high contrast major vessels and low contrast capillaries (see Fig. 2). To address the difficulties caused by such variety, we extract partly overlapped patches in each image, classify them into three groups, and design a deep neural network to segment each group of patches.

We randomly extract 9,500 patches of size 48×48 in each training image, and thus have totally 190,000 patches for training. In each training patch, we firstly calculate the distance between each vessel pixel and the background pixel nearest to it. The maximum distance along each line that is perpendicular to the direction of the vessel is defined as the radius of the vessel at that point. A vessel segment with a radius larger than T_1 is defined as wide vessel, a vessel segment with a radius smaller than T_2 is defined as capillary, and other vessel segments are defined as middle vessels. For this study, the threshold T_1 and T_2 are empirically set to five and three, respectively. Three types of retinal vessels in two training images are illustrated in Fig. 3.

Accordingly, each training patch can be assigned to one of three classes based on the majority type of vessels within it. Then, we use those annotated patches
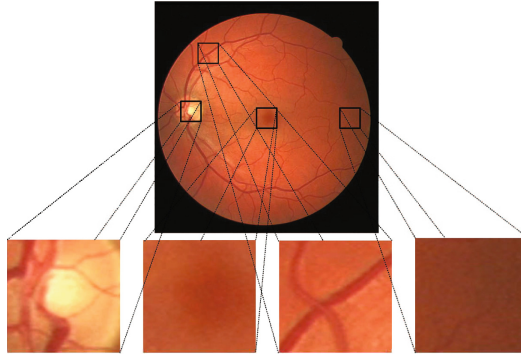
Fig. 2. A fundus retinal image: the optic disc region (1st column), macular area (2nd column), high contrast patch (3rd column) and low contrast patch (4th column) (color figure online)
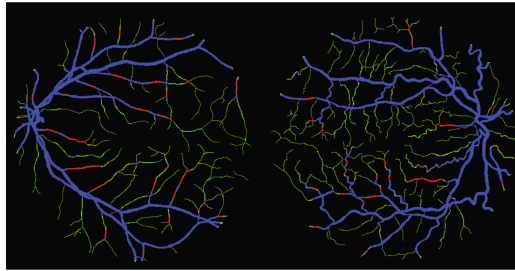

Fig. 3. Three types of vessels, including (blue) the wide-vessel, (red) middle-vessel, and (green) capillary, in two training images (Color figure online)

to train a patch classifier. Since the number of training patches is too small to train a very deep neural network, we use a network in network (NIN) model [8] as the patch classifier, which contains nine convolutional layers, including two layers with 5×5 kernels, one layer with 3×3 kernels and six layers with 1×1 kernels (see Fig. 4). To train this NIN, we fix the maximum iteration number to 120, choose the min-batch stochastic gradient decent with a batch size of 100, use the learning rate from 0.01 to 0.0004, and set the weight decay as 0.0005.

3.3   Training Patch Segmentation Networks

We use each class of patches to train a U-Net, which predicts each pixel in a patch belonging to either blood vessels or background. The U-Net consists of an input layer that accepts 48×48 patches, a Softmax output layer and five blocks in the middle (see Fig. 5). Block-1 has two convolutional layers, each consisting of 32 kernels of size 3×3, and a 2×2 max-pooling layer. Block-2 is identical to Block-1 except for using 64 kernels in each convolution layer. In Block-3, the number of kernels in each convolutional layer increases to 128, and there is an
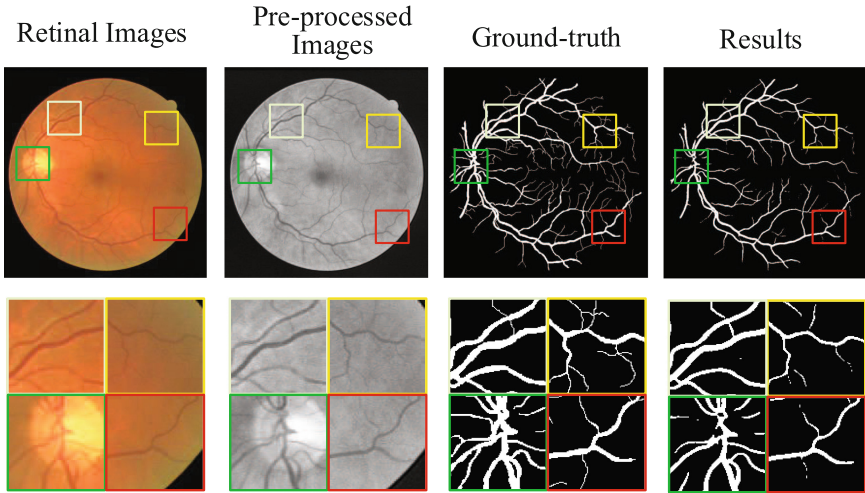
extra 2×2 up-sampling layer. Block-4 accepts the combined output of Block-3 and Block-2. The rest part is identical to Block-2 except that the last layer is a 2×2 up-sampling layer. Block-5, that accepts the combined output of Block-4 and Block-1, consists of two convolutional layers with 32 kernels of size 3×3 and a 1×1 convolution layer. Moreover, there is a dropout layer with a dropout rate of 20% between any two adjacent identical convolutional layers.

Fig. 4. Architecture of the NIN [8] model used in our algorithm

Fig. 5. Architecture of the U-Net [5] used in our algorithm

To train each U-Net, we fix the maximum iteration number to 150 and set the batch size as 32. Considering the small dataset, we first use all training patches to pre-train a U-Net, and then use three classes of patches to fine-tune this pre-trained U-Net, respectively.

Fig. 6. An example test image (1st column), pre-processed image (2nd column), ground-truth (3rd column) and segmentation result (4th column) (color figure online)

3.4   Testing: Retinal Vessels Segmentation

Applying the proposed CAS model to retinal vessels segmentation consists of four steps. First, partly overlapped patches of size 48×48 are extracted in each retinal image, which has been preprocessed in the same way, with a stride of 5 along both horizontal and vertical directions. Second, each patch is inputted to the trained NIN model for a class label. Third, wide-vessel patches and capillary patches are segmented by the U-Net fine-tuned on the corresponding class of training patches, whereas each middle-vessel patch is segmented by three U-Nets and labelled by binarizing the average of three probabilistic outputs with the threshold 0.5. Fourth, since patches are heavily overlapped, each pixel may appear in multiple patches. Hence, we use the majority voting to determine the class label of each pixel in the final segmentation result.

4   Results

Figure 6 shows an example test image, its preprocessed version, segmentation result and ground-truth. To highlight the accurate segmentation obtained by our algorithm, we randomly selected four areas in the image, enlarged them, and displayed them in the bottom row of Fig. 6. It shows that our algorithm is able to detect most of retinal vessels, including low contrast capillaries.

Table 1 gives the average accuracy, specificity, sensitivity and area under curve (AUC) of our algorithm and seven existing retinal vessels segmentation algorithms. It reveals that our algorithm achieved the highest AUC and the top three performance when measured in terms of accuracy, specificity and sensitivity.

Table 1. Performance of seven retinal vessels segmentation algorithms on DRIVE

Algorithm	AUC (%)	Accuracy (%)	Specificity (%)	Sensitivity (%)
Lupascu et al. [2]	95.61	95.97	98.74	67.28
Fraz et al. [9]	97.47	94.80	98.07	74.06
Fu et al. [6]	94.70	N.A	N.A	72.94
Lahiri et al. [10]	N.A	95.33	N.A	N.A
Liskowski et al. [3]	97.90	95.35	98.07	78.11
Fu et al. [11]	N.A	95.23	N.A	76.03
Dasgupta et al. [12]	97.44	95.33	98.01	76.91
Orlando et al. [13]	95.07	N.A	96.84	78.97
CAS model	***97.93***	***95.62***	***98.30***	***77.21***

5   Discussions

5.1   Computational Complexity

Due to the use of deep neural networks, the proposed CAS model has a very high computation complexity. It takes more than 32 h to perform training (Intel Xeon CPU, NVIDIA Titan Xp GPU, 128 GB Memory, 120 GB SSD and Keras 1.1.0). However, applying this algorithm to segmentation is relatively fast, as it takes less than 20 s to segment a 584×565 color retinal image on average.

Table 2. Performance of our algorithm when using one to four patch classes

# Patch classes	1 (%)	2 (%)	3 (%)	4 (%)
AUC	97.90	97.86	***97.93***	95.99
Accuracy	95.58	95.48	***95.62***	93.28
Specificity	98.03	97.71	***98.30***	94.82
Sensitivity	78.79	80.14	*77.21*	**82.78**

5.2   Number of Patch Classes

To deal with the variety of retinal patches, we divided them into three classes and employed different U-Nets to segment them. We also tried to categorize those patches into one to four classes, and displayed the corresponding performance in Table 2. It shows that (a) using three classes of patches performed best, (b) increasing the number of patch classes to four resulted in worst performance, and (c) separating patches into two or four classes performed worse even than not separating patches at all.

6   Conclusion

In this paper, we propose the CAS model to extract blood vessels in color images of the retina. The intra- and inter-image variety can be largely addressed by separating retinal patches into wide-vessel, middle-vessel and capillary patches and by using three U-Nets to segment them, respectively. Our results on the DRIVE database indicates that the proposed algorithm is able to segment retinal vessels more accurately than seven existing segmentation algorithms.

References

1. Staal, J.J., Abramoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE TMI **23**(4), 501–509 (2004)
2. Lupascu, C.A., Tegolo, D., Trucco, E.: FABC: retinal vessel segmentation using AdaBoost. IEEE TITB **14**(5), 1267–1274 (2010)
3. Liskowski, P., Krawiec, K.: Segmenting retinal blood vessels with deep neural network. IEEE TMI **35**(11), 2369–2380 (2016)
4. Li, Q., Feng, B., Xie, L.P., Liang, P., Zhang, H., Wang, T.: A cross-modality learning approach for vessel segmentation in retinal images. IEEE TMI **35**(1), 109–118 (2016)
5. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
6. Fu, H., Xu, Y., Wong, D.W.K., Liu, J.: Retinal vessel segmentation via deep learning net-work and fully-connected conditional random fields. In: ISBI, pp. 698–701. IEEE, Prague (2016)
7. Setiawan, A.W., Mengko, T.R., Santoso, O.S., Suksmono, A.B.: Color retinal image enhancement using CLAHE. In: ICISS, pp. 1–3. IEEE, Jakarta (2013)
8. Lin, M., Chen, Q., Yan, S.: Network in network. In: ICLR. IEEE, Banff (2014)
9. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G.: An ensemble classification-based approach applied to retinal blood vessel segmentation. IEEE TBE **59**(9), 2538–2548 (2012)
10. Lahiri, A., Roy, A.G., Sheet, D., Biswas, P.K.: Deep neural ensemble for retinal vessel segmentation in fundus images towards achieving label-free angiography. In: EMBC, pp. 1340–1343. IEEE, Orlando (2016)
11. Fu, H., Xu, Y., Lin, S., Kee Wong, D.W., Liu, J.: DeepVessel: retinal vessel segmentation via deep learning and conditional random field. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 132–139. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_16

12. Dasgupta, A., Singh, S.: A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation. In: ISBI, pp. 248–251. IEEE, Melbourne (2017)
13. Orlando, J., Prokofyeva, E., Blaschko, M.: A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. IEEE TBE **64**(1), 16–27 (2017)