Optic Disc Detection via Deep Learning in Fundus Images

Peiyuan Xu[1], Cheng Wan[1(✉)], Jun Cheng[2], Di Niu[1,3], and Jiang Liu[3]

[1] Nanjing University of Aeronautics and Astronautics, Nanjing 210000, China
wanch@nuaa.edu.cn
[2] Institute for Infocomm Research, A*STAR, Singapore, Singapore
[3] Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo, China

Abstract. In order to realize the localization of optic disc (OD) effectively, a new end-to-end approach based on CNN was proposed in this paper. CNN is a revolutionary network structure which has shown its power in fields of computer vision like classification, object detection and segmentation. We intend to make use of CNN in the study of fundus images. Firstly, we use a basic CNN on which specialized layers are trained to find the pixels probably in OD region. Then we sort out candidate pixels furtherly via threshold. By calculating the center of gravity of these pixels, the location of OD is finally determined. The method has been tested on three databases including ORIGA, MESSIDOR and STARE. In totally 1240 images to be tested, the OD of 1193 are successfully located with the rate of 96.2%. Besides the accuracy, the time cost is another advantage. It takes only 0.93 s to test one image on average in STARE and 0.51 s in MESSIDOR.

Keywords: Optic disc localization · Convolution neural networks · Retinal fundus image

1 Introduction

The optic disc (OD) is one of the main physiological structures of the retina, from which optic nerve and blood vessels stretch to the surrounding areas. In fundus camera pictures, these blocks are reflected as a round bright yellow area, where few but thick blood vessels also exist. Researchers pay attention to the automatic detection of the OD for the reason that diagnosis of some ocular fundus lesions are based on its correct detection [1, 2]. The main application of this approach is to pre-process the retinal images for further studies such as the segmentation of optic disc or the detection of Macular region. Till now, to detect OD, there have occurred several analysis techniques, which can be divided into two categories: early methods mainly take the characteristics of the OD like brightness, contrast, shape etc. for example, the method in [3] locate the OD by searching for the center of rectangle region in which the amplitude of variation of gray level is highest. In [4, 5], Hough transform, which is convenient for specific shape object detection, is used to detected the OD region. These early methods can simply and timely get the results because the OD is usually brighter than other areas in fundus camera pictures and occurs as a regular ellipse. However, considering that the retinal images are not always high quality and several diseases may lead to the change of OD [19, 20], these

M.J. Cardoso et al. (Eds.): FIFI/OMIA 2017, LNCS 10554, pp. 134–141, 2017.
DOI: 10.1007/978-3-319-67561-9_15

methods might fail. The other methods can be described as the methods based on vascular feature detection. The OD, where the main blood vessels converge in the retinal, can be easily detected when vascular features are known. In [6], two parabolas are used to describe the blood vessels in the left and right direction of OD. The center of OD is just located in the public vertex of two parabolas. In [7], the OD region is determined via calculating the confluence of blood vessels. Though the methods above could get relatively high detection accuracy, they should make sure that the detection of vessels is exact, which in low quality or abnormal images can be hard to realize. Besides, the OD detection algorithms based on vascular feature detection are usually complex and time-costing.

Except for two main methods mentioned above, there are several methods considering multiple features in OD region. The description of multi-source information about OD characteristics is undoubtedly helpful to improve the accuracy and robustness of OD detection. However, simple threshold or models could not make full use of these information. To overcome this drawback, supervised learning method, which requires great effort to design such a template that can integrate multiple information, are used. Supervised learning method increases the complexity of the OD detection algorithm, which is not suitable for real time applications.

Overall, the speed and accuracy of OD detection are hard to balance. The methods with appearance characteristics, which are fast, could not get a high accuracy in abnormal fundus images. In contrast, the methods based on vascular feature detection could get a relatively high accuracy in no matter normal or abnormal fundus images, with the cost of high algorithmic complexity and long locating time. Therefore, how to improve the accuracy and reduce the complexity of the algorithm as well is worth in further research. In this paper, we propose an end-to-end method for automatic localization of OD based on the Convolution Neural Network (CNN). The OD features are learned through combination of features from different layers of CNN. After this, each pixel of image will get a probability to judge whether it belongs to OD region. In order to avoid interference, only points with high probability are picked as candidate points. Then, we calculate the center of these points and finally the location of OD is determined. The rest of this paper is organized as follows. In Sect. 2 we show the details of our proposed method. Section 3 shows the experimental results and discussions. In Sect. 4, the conclusions are mentioned here.

2 Proposed Method

2.1 Preprocessing

All images in training set are firstly subtracted the mean value of each color channel. Then, these images are resized into a fixed size of 400×600, which makes the training faster and reduce the computing complexity. Considering the condition of GPU, this size can be set as other values. There's no need to pre-process the images in testing set for the reason that pre-process will not lead to obvious changes in testing results. However, restricted to hardware condition, images of ORIGA and MESSIDOR are resized into the same size as training set. The images of STARE are kept as the same.

2.2 Feature Representation by CNN

We leverage a successful deep learning architectures called VGG network, which originally used for the classification of natural images. VGG network is similar to AlexNet, which is a classic CNN. They both consist of 5 blocks. Between each block, there is pooling layer. Each block of VGG network contains several convolution layers with 3×3 size filters, differs from AlexNet which only has one convolution layer with 7×7 size filter. This can be seen as imposing a regularization on the 7×7 filters, forcing them to have a decomposition through the 3×3 filters [8]. because of the reduce of filter size, VGG has much more channels than AlexNet, which is considered to lead to higher accuracy.

Considering that the CNN are used as feature detector, we remove the fully connected layers of VGG network. The architecture remained are mainly consisted by convolutional layers with Rectified Linear Unit (ReLU) activations and max pooling layers. These layers have been pre-trained on millions of images. Features detected by deep blocks are rougher than shallow blocks duo to the decrease in size. We decide to combine features from 3 deeper blocks rather than any one of them to get better results. These features are then forwarded into deconvolution layers and crop layers to be the same size as original images. The flow chart of our method is shown in Fig. 1.

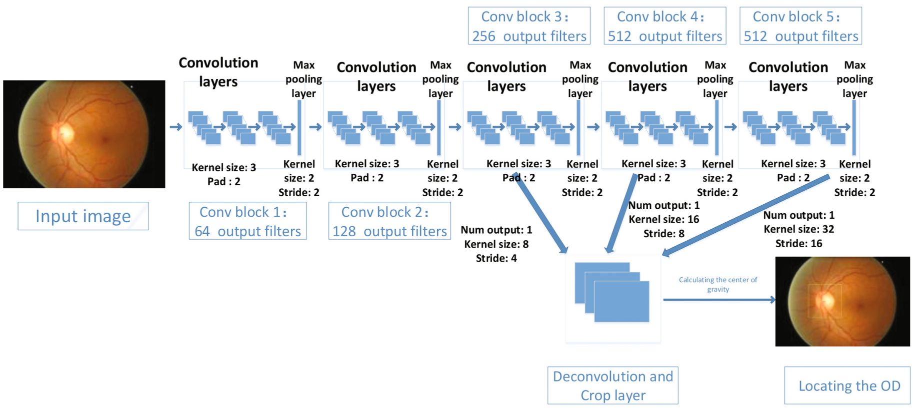


Fig. 1. The proposed model architecture

The train starts by loading the weights of VGG network. Then, input images are put into the designed network. We extract the features of last three pooling layers, which are then forwarded into corresponding deconvolution layers. the last layers, combined by these deconvolution layers, output the probabilities of each pixel of the whole image. The class of each pixel can be determined by setting a threshold. To avoid interference, we set a high threshold $T = 0.9$, which means the network judges this pixel as OD with the probability of 0.9. Finally, by calculating center of gravity of the pixels above, we realize the localization of OD. In this condition, even there are several pixels are missclassified, the center will still in OD region. The center of gravity can be calculated as following:

$$X_c = \frac{\sum P(x_i)}{\sum P_i} \qquad Y_c = \frac{\sum P(y_i)}{\sum P_i} \tag{1}$$

where P_i represents the value, x_i, y_i are the location of pixel.

To train the network, the cross entropy loss is adopted to update the weights through descent back-propagation. Cross entropy loss is used to measure the similarity between two probability distributions. It can be defined as following:

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)}\log(h_\theta(x^{(i)})) + (1 - y^{(i)})\log(1 - h_\theta(x^{(i)})) \tag{2}$$

where, input images are $X^{(i)} = (1, x_1^{(i)}, x_2^{(i)}, \cdots, x_p^{(i)})^T$ and $y^{(i)}$ are predicted labels, θ denotes the parameters of CNN. Here, $\theta = (\theta_0, \theta_1, \theta_2, \cdots, \theta_p)^T$. The hypothesis function is defined as $h_\theta(x^{(i)}) = \frac{1}{1 + e^{-\theta^T x^{(i)}}}$.

At training time, we fine-tune the entire network with 880 images (305 from ORIGA and 575 from MESSIDOR, which will be introduced later) for 150000 iterations. The input images are operated via CNN with batches of size 1. We use a small learning rate ($lr = 10^{-9}$, which will decrease as training time increases). Stochastic gradient descent is used to minimize error function with momentum = 0.9.

At testing time, the OD detector is realized on 2.8 GHz Intel Xeon E5-1603 v4 CPU with GTX 1080 using python. The last layer of network outputs the probability of each pixel of testing image. Then we set a threshold T = 0.9, which removes most of noise and ensure that pixels remained are mostly in OD region. Finally, we get the location of OD via calculating the center of these pixels. Considering that some researchers like [9, 10] proposed that in different fundus images, the diameter of OD is about 1/5 to 1/8 of the ROI region. So, based on the center point determined, we draw a Rectangular box with length of about 1/4 of ROI region (here we set length = 100 in ORIGA and MESSIDOR. In STARE, we set as 150).

3 Experiments and Results

3.1 Databases

The approach has been tested on the following three datasets. ORIGA [11], MESSIDOR [12] and STARE [13]. MESSIDOR and STARE are available publically, among which MESSIDOR totally includes 1200 images with three sizes: 1440 × 960, 2200 × 1488, 2304 × 1536. Differs from MESSIDOR which contains relatively high quality images with few lesion, STARE includes many images with lesion and low quality. A large number of OD detection methods have been tested and compared on this image dataset. ORIGA is a database with 650 images. All images in ORIGA with the size of 3072 × 2048 are captured via a high resolution retinal fundus camera and well selected, which ensure the quality.

3.2 Results and Discussions

Table 1 shows the accuracy of our method in different datasets. Of all the 1240 fundus images, 1193 images can find the right OD location and the success rate is 96.2%. In addition, only the rectangle box contains the whole OD region, which comes from label image, can be seen as success. In Table 2 we compare different methods in MESSIDOR. Experimental results show that our method could achieve the accuracy of 99.43% with 0.51 s per image, which is better than other methods. In Table 3, it can be seen that the method based on vascular characteristics like [17] could achieve a high accuracy with the cost of long processing time. The methods making use of appearance information like [5] are not competitive in accuracy comparing with other methods for that the information they use are not stable during different images. Considering both accuracy and processing speed, our method or method in [18] are more practical. The results of other methods mentioned above are original ones in their papers.

Table 1. The accuracy of our method in different datasets

Datasets	Images	Abnormality	Accuracy
ORIGA	314	–	100%
MESSIDOR	526	54.5%	99.43%
STARE	400	91%	89%
Total	1240	–	96.2%

Table 2. The accuracy of different methods in MESSIDOR database

Methods	Accuracy	Running time
Yu [14]	99%	6.6 s
Alghamdi [15]	99.20%	–
Zubair et al. [16]	98.65%	3.5 s
Our method	99.43%	0.51 s

Table 3. The accuracy of different methods in STARE database

Methods	Accuracy	Running time
Foracchia [17]	97.5%	120 s
ZHANG [18]	91.4%	13.2 s
Haar [5]	67.9%	–
Alghamdi [15]	86.71%	–
Our method	89%	0.93 s

Our method could get a high accuracy except STARE. This is because STARE is a dataset containing many images with serious lesions. In STARE, many optic discs of fundus images are damaged and some optic discs of fundus images can only be partially observed. Besides, images with the bright yellow lesions which is similar to the OD in appearance are also contained. Of all images failed to detected by our method, images with low contrast are the most common situation.

Restricted to the length of the article, only some of the representative fundus images were selected, which can be seen in Fig. 2. In these samples, all kinds of conditions, in which OD are hard to be detected correctly, are included. (a), (b), (e) and (f) are images with regions produced by lesions. (d), (j), (l) are images caused by bleeding in small or large scales. (i) and (l) are images in which OD were covered by lesions. (k) is the classic image that OD region locates in the edge of ROI in fundus image. (c), (g) and (h) are those with low contrast. This shows the power of our method in detecting OD region from abnormality images.

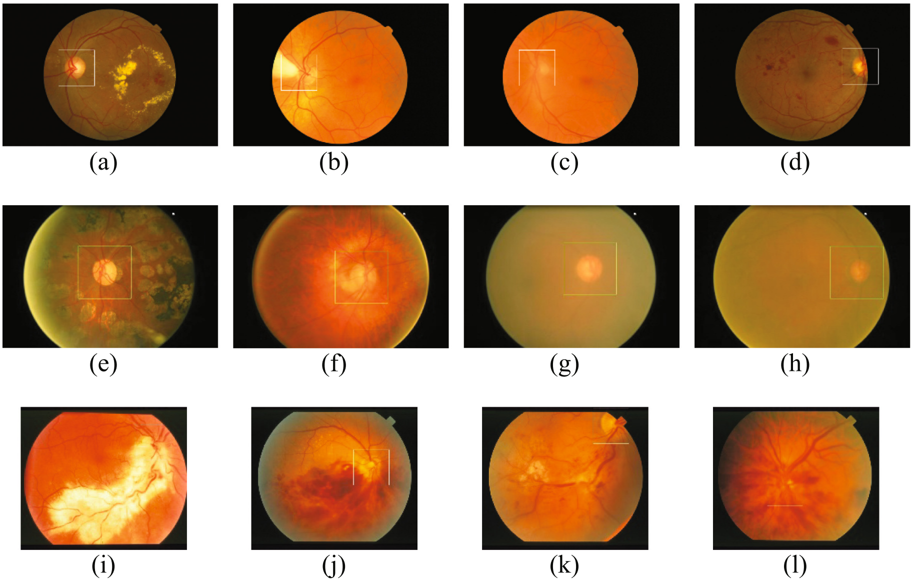
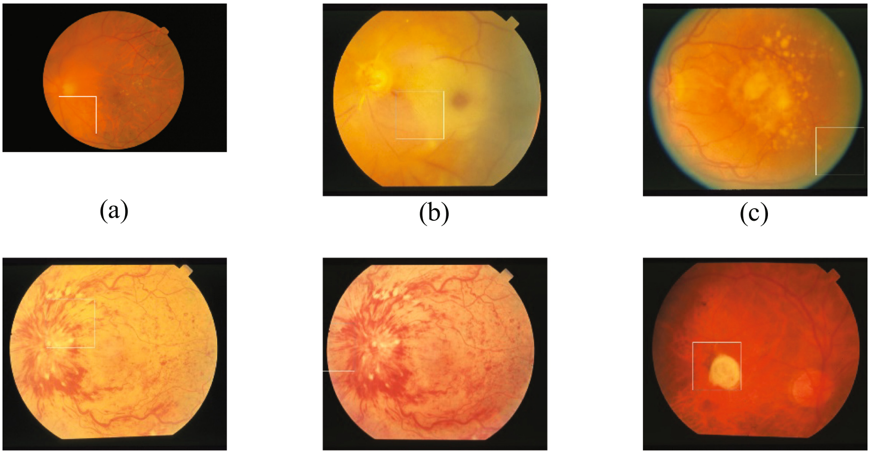


Fig. 2. Several testing results hard to be detected.

In Fig. 3, we show some images detected incorrectly. From the samples, (a)–(c) are because of low contrast, which accounts for the most of error detected images. in (d)–(e), OD region, covered by lesion, is hard to identify. Though detecting is failed, the region given by our method is surrounding the OD. We can conjecture that our method has learned some information of vessels. In (f), our method judges the lesion region as OD which are quite similar in appearance, from which we could infer that the main features CNN learns are mainly about the appearance of OD. Further research is needed to make the network learns more about vessels.



(a) (b) (c)

(d) (e) (f)

Fig. 3. Several testing results detecting incorrectly.

4 Conclusion

In this paper, we proposed a new approach for OD detection. Current experiment results show that the method has good robustness. Unlike those complex and time-costing methods which need for vascular feature detection, our method is fast and accurate in both normal and abnormal images.

In the future, we are going to evaluate our method on more database in order to obtain more objective and comprehensive test results. What's more, model will be improved by exploiting other algorithms to overcome the drawback of our method. We also want to add the method of pre-training and image enhancement before the whole architecture to achieve a better result.

References

1. Zhang, Z., Lee, B.H., Liu, J., Wong, D.W.K., et al.: Optic disc region of interest localization in fundus image for Glaucoma detection in ARGALI. In: Industrial Electronics & Applications, pp. 1686–1689 (2010)
2. Cheng, J., Liu, J., Xu, Y., Yin, F., Wong, D.W.K., Tan, N.M., Tao, D., Cheng, C.Y., Aung, T., Wong, T.Y.: Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. IEEE Trans. on Med. Imaging **32**(6), 1019–1032 (2013)
3. Boyce, J.F., Cook, H.L., et al.: Automated location of the optic disk, fovea, and retinal blood vessels from digital color fundus images. Br. J. Ophthalmol. **83**(8), 902–910 (1999)
4. Barrett, S.F., Naess, E., Molvik, T.: Employing the hough transform to locate the optic disk. Biomed. Sci. Instrum. **37**(1), 81–86 (2001)
5. Haar, F.T.: Automatic localization of the optic disc in digital color images of the human retina. Utrecht University, Utrecht (2005)

6. Osareh, A.: Automated identification of diabetic retinal exudates and the optic disc. University of Bristol, Bristol (2004)
7. Lalonde, M., Beaulieu, M., Gagnon, L.: Fast and robust optic disk detection using pyramidal decomposition and Hausdorff based template matching. IEEE Trans. Med. Imaging **20**(11), 1193–1200 (2001)
8. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: Computer Science (2014)
9. Li, H., Chutatape, O.: Automatic location of optic disk in retinal images. In: International Conference on Image Processing, vol. 2, pp. 837–840 (2001)
10. Klein, R., Klein, B., Moss, S., Davis, M., et al.: The Wisconsin epidemiologic study of diabetic retinopathy II. Arch. Ophthalmol. **102**(4), 520–526 (1984)
11. Zhang, Z., Yin, F., Liu, J., Wong, D.W.K., Tan, N.M., Lee, B.H., Cheng, J., Wong, T.Y.: Origa-light: an online retinal fundus image database for glaucoma analysis and research. In: International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3065–3068 (2010)
12. Decencière, E., et al.: Feedback on a publicly distributed database: the Messidor database. Image Anal. Stereology **33**(3), 231–234 (2014)
13. Hoover, A.: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans. Med. Imaging **19**(3), 203–210 (2000)
14. Yu, H., Barriga, E.S., Agurto, C., et al.: Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets. IEEE Trans. Inf. Technol. Biomed. **16**(4), 644–657 (2012)
15. Alghamdi, H., Tang, H., Waheeb, S., Peto, T.: Automatic optic disc abnormality detection in fundus images: a deep learning approach. In: OMIA 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, Iowa Research Online, pp. 17–24 (2016)
16. Zubair, M., Yamin, A., Khan, S.A.: Automated detection of optic disc for the analysis of retina using color fundus image. In: IEEE International Conference on Imaging Systems and Techniques, Beijing (2013). doi:10.1109/IST.2013.6729698
17. Foracchia, M., Grisan, E., Ruggeri, A.: Detection of optic disc in retinal images by means of a geometrical model of vessel structure. IEEE Trans. Med. Imaging **23**(10), 1189–1195 (2004)
18. Zhang, D.B., Yi, Y., Zhao, Y.Y.: Projection based optic disc detection method for retinal fundus images. Chin. J. Biomed. Eng. **32**(4), 477–483 (2013)
19. Anastasi, M., Lodato, G., Cillino, S.: VECPs and optic disc damage in diabetes. Doc. Ophthalmol. Adv. Ophthalmol. **66**(4), 331–336 (1987)
20. Artes, P.H., Chauhan, B.C.: Longitudinal changes in the visual field and optic disc in glaucoma. Prog. Retinal Eye Re. **24**(3), 333 (2005)