Fovea Localization in Fundus Photographs by Faster R-CNN with Physiological Prior

Jun Wu[1], Jiapei Wang[1], Jie Xu[2(✉)], Yiting Wang[1], Kaiwei Wang[1], Zongjiang Shang[1], Dayong Ding[3], Xirong Li[4,5], Gang Yang[5], Xuemin Jin[6], Yanting Wang[6], Fangfang Dai[6], and Jianping Fan[7]

[1] School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China

[2] Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology and Visual Science Key Lab, Beijing, China
fionahsu920@foxmail.com

[3] Vistel AI Lab, Visionary Intelligence Ltd., Beijing, China

[4] School of Information Science and Technology, University of Science and Technology of China, Hefei, China

[5] School of Information, Renmin University of China, Beijing, China

[6] Henan Provincial Peoples' Hospital, Zhengzhou, China

[7] University of North Carolina-Charlotte (UNCC), Charlotte, NC, USA

Abstract. A macular fovea is a physiological structure of the human retina, which is an essential optical center. The distance between the lesion area and the foveal center determines the severity degree of visual impacts. Therefore, accurate fovea localization is the basis of the computer-aided ophthalmic diagnosis and vision screening. A simple but effective fovea localization algorithm based on the Faster R-CNN and physiological structure prior is presented. First, a fovea localization model and an optic disc localization model are trained separately. Then, for each fundus photograph, both candidate areas of the fovea and the location of the optic disc are predicted using two pre-trained models. Next, prior knowledge of the physiological adjacent relationship between a fovea and an optic disc is applied to eliminate unreasonable candidate bounding boxes. Finally, the ultimate bounding box of the fovea is determined by the best candidate. Experiments were conducted on a private dataset with 5,203 fundus photographs and the public Messidor dataset including 1200 fundus photographs. The accuracy of the foveal location in the offset scale of 1/2 optic disc diameter on the Messidor is 99.58%, which is 0.71% higher than the state-of-the-art (98.87%).

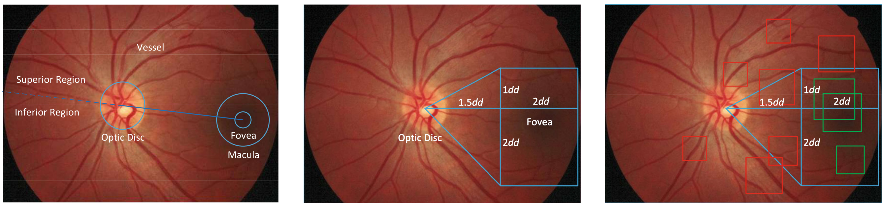
Keywords: Fovea localization · Faster R-CNN · Fundus photography · Optic disc · Physiological prior · Messidor

H. Fu et al. (Eds.): OMIA 2019, LNCS 11855, pp. 156–164, 2019.
https://doi.org/10.1007/978-3-030-32956-3_19

1 Introduction

In a human retina, the typical tissues are vessels, an optic disc, a macula, *etc.* An optic disc is a vertical oval in the nasal side. A macula is an oval-shaped pigmented area near the retinal center. A fovea is located in the macular center and it is a small pit that contains the largest concentration of cone cells, which is the most sensitive part to light. Once any macular lesion occurs, the vision will be significantly affected. Fast and accurate automatic fovea detection can greatly improve the diagnostic efficiency of retinal diseases, and it also can provide a basis for the large-scale screening of common retinal diseases.

A foveal center is located within an inferior macula without any vascular distribution. As shown in Fig. 1(a), in a color fundus photograph, the fovea is the lowest brightness area in the macular center, roughly located on a symmetrical axis which divides the upper and lower branches of the entire vascular network. From the optic-disc center to its temporal direction, a foveal center is located around 2–2.5 optic disc diameter (*dd*) away, with a small horizontal angle [15].



(a) A fundus example. (b) Macula Prior box. (c) Prior Verification.

Fig. 1. Prior physiological relationships. (a) Illustration of a macular, a fovea, and an optic disc in a fundus photograph. (b) A left eye example for macular restriction box. (c) Verification using Physiological Relationship Prior. (Color figure online)

By using different structure information, the fovea localization methods have three categories. One way is to locate the fovea using the appearance characteristics of a fovea itself [7,11,15]. Singh et al. [15] first proposed a pre-processing method to improve the image contrast between the macula and its surrounding areas in the red channel and then located the foveal center by searching the darkest area. Lu et al. [7] proposed a linear metric operator based on the brightness and shape features of the macula, which can search from different directions and reflect the gray gradient changes. The image enhancement by an adaptive manifold filter and mathematical morphological operations were applied to locate the fovea [11].

The second way is to use contexts such as the surrounding vessels or optic disc [2,5,10,18]. Niemeijer et al. [10] proposed a cost function based on global and local cues to find the correct foveal center. Zheng et al. [18] first segmented a vessel network and obtained the macular Region of Interest (ROI). Then, a circular region was fitted along vessel endings around the macula to locate the fovea. In [5], based on the optic-disc location, a one-dimensional scanned intensity profile analysis was applied for the fovea localization. Dashtbozorg *et al.* [2] presented an innovative super-elliptical filter to localize the optic disc and the fovea simultaneously.

The last way is to combine the fovea, the vessel network and the optic disc together [1,4,6,8,13,14,16,19]. In [16], an approximate macular position was obtained by using the optic-disc position, where the foveal center was estimated as the lowest response point from a directional matched filter. In [13], the main vessels were extracted first, and the macular ROI was obtained from the vessel tree. Then the foveal center was determined at a distance below the temporal direction of the optic disc. Gegundez-Arias et al. [4] utilized the information of an optic disc and a vascular tree to obtain an approximate location of the macula. In [1] visual and anatomical feature-based criteria were combined with respect to the optic disc and the vascular tree. Kao et al. [6] localized the optic-disc center first by the template matching, and determined the disc-fovea axis by searching the vessel-free region. Finally, the fovea center was detected by matching the fovea template around the center of the axis. In [19], another method was proposed to locate the macular ROI based on the optic-disc location first, and then the foveal center was determined by macular features and mathematical morphology. In [14], the vessel segmentation was applied first to locate the optic disc, and the fovea was localized by matching the expected directional pattern. Molina-Casado et al. [8] proposed a methodology combining intra- and inter-structure relational knowledge based on the candidate tuples validation to detect the optic disc, macula and vascular network in a unified framework.

With serious lesions, the fovea-based methods often fail. Accurate vessel tree segmentation also requires a heavy computation load. In a color fundus photograph, the macular characteristics are not obvious and accurate macular boundary is often difficult to distinguish. However, an optic disc is relatively brighter, which benefits its localization. Besides, retinal lesions have relatively less influence on locating an optic disc rather than that of a macula. A hemorrhage, macular edema, and other retinal lesions affect the macular appearance tremendously or even cover it completely leading to fovea locating failures. In this case, using the optic-disc position to locate the macula is the only option.

In this paper, we propose a simple but effective fovea localization algorithm by the Faster R-CNN deep learning framework with the prior structure relationship between a fovea and an optic disc. Our main contributions include: (1) The prior physiological relationship between a fovea and an optic disc in the human retina is sufficiently explored to eliminate unreasonable foveal candidates. (2) This prior knowledge can also help to build a hypothetic candidate in case of complete failure on the fovea localization, especially for retinal diseases with severe lesions, which is also difficult for a professional ophthalmologist. (3) The foveal annotation is supposed to be only a center point, not the regional bounding boxes. As a result, this case is more difficult than the case with accurate size information of the bounding boxes at the same time. The localization algorithm needs to determine the optimal size of the bounding boxes that will be fed into the Faster R-CNN network for training.

2 The Proposed Method

Our main objective is to automatically locate the foveal center in color fundus photographs. First, a bounding box of a fovea based on its center annotated by the ophthalmologists is fed to train a Faster R-CNN model. Then, the preferred candidate foveal areas of the top N rank predicted by the pre-trained Faster R-CNN model are screened as the foveal candidate areas. Finally, the candidates with the top N highest scores are further verified by using the prior physiological relationship between a fovea and an optic disc in the human retina, resulting in the final foveal center as the best output (Fig. 2).

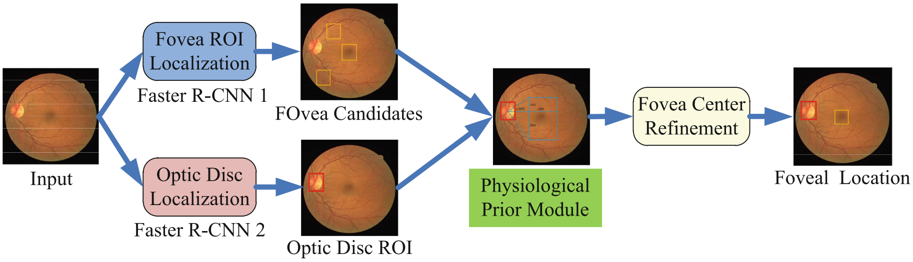


Fig. 2. Flowchart of our proposed fovea localization algorithm.

2.1 Optic Disc Localization

An optic disc provides useful information to locate the fovea according to their physiological relationships. A Faster R-CNN [12] deep learning network is applied to locate the optic disc first. Based on our evaluations, the predicted location of an optic disc in a high accuracy can be directly considered as its real center.

2.2 Fovea Localization

(1) Basic Network: The Faster R-CNN [12] is popular to locate the bounding box of an image object. It is a combination of the Fast R-CNN and the region proposal networks (RPN). We use it as our baseline network.
(2) Optimizing the Size of the Foveal Bounding Box for Training: To locate a fovea, the rough macular region should be detected first. A progressive macular area with vascular distribution exists around a macula, which also helps for the fovea localization. As a result, it is necessary to determine a suitable size of the bounding box for the target fovea to feed into the Faster R-CNN for training. Different sizes of foveal bounding boxes, which contain different context characteristics, affect the performance of the fovea localization. As a result, generally, its size can be optimized on the evaluation set. We try different box size for fovea localization model training from 120×120, 140×140, ..., to 240×240 in pixels, and finally 200×200 is verified as the best option.

2.3 Verification Using Physiological Relationship Prior

In the case that a macula can be roughly located successfully, the position of the optic disc can be applied to refine and optimize the foveal localization, where candidate bounding boxes are with top N highest scores from the Faster R-CNN prediction. The detailed steps are: (1) **Judging left or right eye.** In a typical fundus photograph of the posterior pole, the position of the optic disc is located in the nasal side. Taking the middle vertical line, if the predicted center of the optic disc is on the left side of the midline, it is a left eye and the macula is on the right side of the optic disc. Conversely, it is judged as the right eye, and the macular is on the left side of the optic disc. (2) **Determining macular candidates.** As shown in Fig. 1(b), an optic disc is on the left of the middle vertical line, and it is judged as a left eye. The macula is located in the right-down side of the optic disc. Based on our observations, candidate macular areas are supposed to be located within the blue restriction box as in Fig. 1(b). Its left edge is $1.5dd$[1] from the center of the optic disc. This restriction box is $2dd$ in width and $3dd$ in height. These parameters have been verified in the training sets of our private MF5K dataset (5,203 samples) and the public Messidor [3], DRIVE and STARE datasets (1,640 samples in total). (3) **Locating foveal center using prior knowledge.** First, filtering out unreasonable candidates from the top-N predicted bounding boxes of the macula when they are not completely contained within the restriction box of $3dd \times 2dd$, as the red bounding boxes in Fig. 1(c). The one from the remaining candidates (marked as green boxes) with the highest score from the Faster R-CNN is determined as a final predicted foveal bounding box. (4) **Failure Case Correction.** The prior relationship of $2.5dd$ from the center of the optic disc also helps to build a hypothetic candidate in the extreme case of complete failure on fovea localization with the severe lesions.

3 Evaluation

3.1 Experimental Setup

We validate our proposed method on two datasets. (1) **Messidor** [3]**:** a public dataset with 1,200 color fundus photographs of the posterior pole are acquired by 3 ophthalmologic departments. The foveal center annotations containing the pixel size of the disc diameter (dd) in each fundus photograph, which are provided by Niemeijer et al. [10], are applied in our evaluation. (2) **MF5K:** A private dataset with 5,203 fundus photographs is collected from a local partner hospital. The ophthalmologists manually annotate foveal centers, and the majority of photographs include diseases such as diabetic retinopathy (DR) in around 97%. Therefore, it is a very challenging dataset to locate the fovea due to the interference of the DR disease. In addition, the training set, validation set, and test set are partitioned randomly by 50%, 25%, and 25% respectively, resulting in 2,601, 1,301 and 1,301 fundus photographs for each subset.

[1] Empirically, dd is about 80 pixels in a 500×500 fundus photograph [17].

The experiments are conducted using a Pytorch deep learning framework under Ubuntu 14.04 with an NVIDIA 1080Ti GPU. The images in the MF5K are resized to 500×500 pixels. A carefully-verified 200×200 macular region is used for training the Faster R-CNN model. The shared convolution layer uses the VGG16 to extract features, scales are set as [8, 16, 32], ratios as [0.5, 1, 2], the momentum in the optimizer as 0.9, and the learning rate as 0.001. $N = 10$.

3.2 Evaluation Criteria

The evaluation is based on the Euclidean distance d between the predicted center of a fovea C and the center of ground truth C', that is $D = d(C, C')$. To obtain a binary output (correct or not), we allow an offset of the predicted foveal center from the center of the ground truth. This default offset is defined as the distance that is $1/2dd$ (1 dd is 80 pixels [17]) and the origin is the foveal center of ground truth. If the predicted foveal centers are within the maximum offset, they are considered as correct, and vice versa. An *accuracy* under the offset contexts is applied for our evaluations (the same as the *success rate* in other literature).

3.3 Results

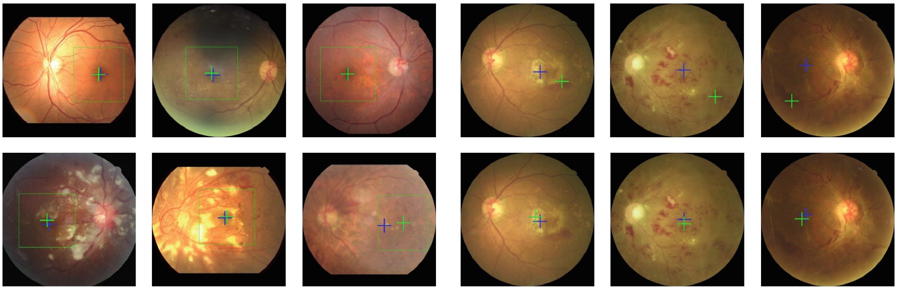
(1) Evaluation of the Prior Relationship Module

Evaluation on the MF5K data set: As in Table 1, in the MF5K dataset, the accuracy of our fovea localization method is 90.70% when applying prior physiological relationship or 89.93% if not (an accuracy improvement of 0.77%).

Table 1. Comparisons of the proposed fovea localization algorithm using prior physiological relationships (using-prior) or not on the private MF5K dataset. Joint Loc. means the jointly localize the fovea and the optic disc with one Faster R-CNN model.

Methods	Using-prior	Accuracy (%)($D \leq 1/3dd$)
Joint Loc. using faster R-CNN [9]	-	90.23
Faster R-CNN [12]	No	89.93
Our method	Yes	**90.70**

More precisely, for 1301 test images in the MF5K, the accuracy of the optic disc localization is 99.62% (5 images fails). Fovea localization before prior relationship refinement fails in 131 images (an accuracy of 89.93%), among which there are 20 images without any predicted candidate box at all due to severe lesions. Then, after applying the suggested hypothetic centers, 7 (out of 20) images success to infer correct foveal centers within $1/3dd$ offset. Another 3 fundus photographs with unreasonable candidate bounding boxes are removed, and the final outputs are corrected successfully by the relationship prior module.



(a) Predicting the foveal location. (b) Using prior relationship (or not).

Fig. 3. Examples of (a) predicting the fovea bounding boxes. (b) comparison of using prior physiological relationship module (the first row) or not (the second row). (Color figure online)

In addition, for the offset distance D of the predicted results, the mean is 0.16 dd, the variance is 0.098 dd. The maximum is 4.29 dd, and the minimum is 0.

Some examples of the predicted results are shown in Fig. 3(a). The green box is predicted ROI of the fovea and the green cross is the predicted fovea center. The blue cross is the foveal center of the ground truth. Further, three examples using prior relationships or not are compared in Fig. 3(b).
Evaluation on the Messidor data set: While, in the Messidor dataset, due to the relatively better quality of fundus photographs (no serious lesions), each image only has one predicted candidate within the restriction region. Actually, the prior relationship module has not yet been used at all, resulting in no difference between them in terms of accuracy (staying the same as 98.83%).

(2) Evaluation Different Methods on the Public Messidor Dataset
Further, we evaluate our proposed method and the existing methods on the Messidor dataset, as in Table 2. The accuracy of the optic disc localization here is 100%. The accuracy of our method is higher than that of the existing methods with the offset distance as 1/4 dd, 1/3 dd, and 1/2 dd respectively, which proves the effectiveness of our proposed method. In the case of the 1/2 dd offset, the accuracy of our proposed method is 0.71% higher than the existing method (98.87%) in [2]. In the case of the 1/4 dd offset, the accuracy of our proposed method is 0.67% higher than the existing method (96.83%) in [10].

Table 2. Comparisons of different fovea localization methods in the Messidor dataset.

Methods	Accuracy (%)(D: offset distance)				
	$D \leq \frac{1}{4}dd$	$D \leq \frac{1}{3}dd$	$D \leq \frac{1}{2}dd$	$D \leq dd$	Uncertain
Niemeijer et al. (2009) [10]	<u>96.83</u>	-	97.92	-	
Yu et al. (2011) [16]	95.00	-	-	-	
Gegundez-Arias et al. (2013) [4]	96.08	96.58	96.92	97.83	
Aquino et al. (2014) [1]	91.28	-	98.24	99.56	
Kao et al. (2014) [6]	-	-	97.80		
Dashtbozorg et al. (2016)[2]	93.75	-	<u>98.87</u>	<u>99.58</u>	
Molina-Casado et al. (2017)[8]	96.08	-	98.58	99.50	
Kamble et al. (2017)[5]	-	-	-	-	99.66
Pachade et al. (2019) [11]	-	98.66	-	-	
Our method	**97.50**	**98.83**	**99.58**	**100.0**	

4 Conclusion

In this paper, a simple and effective fovea localization method by the Faster R-CNN with physiological prior in fundus photographs is proposed. The different sizes of the input bounding box of a macular region for training a better Faster R-CNN model are investigated to obtain the optimal input size of the macular bounding box. At the same time, the predicted candidate bounding boxes of the fovea are re-validated with the help of the optic disc location, between which their prior physiological relationships are fully explored and utilized. Experiments on the public Messidor and a private MF5K dataset that is a relatively difficult task show that our proposed method is superior to the state-of-the-art methods, and it improves the prediction accuracy of the fovea localization by 0.71% in the case that the offset distance is less than 1/2 optic disc diameter.

Acknowledgements. This work is supported by the CSC State Scholarship Fund (201806295014), NSFC (No. 61672523, No. 61771468), CAMS Initiative for Innovative Medicine (2018-I2M-AI-001), Beijing Natural Science Foundation (No. 4192029), Beijing Hospitals Authority Youth Programme (QML20170206); The priming scientific research foundation for the junior researcher in Beijing Tongren Hospital, Capital Medical University (2018-YJJ-ZZL-052).

References

1. Aquino, A.: Establishing the macular grading grid by means of fovea centre detection using anatomical-based and visual-based features. Comput. Biol. Med. **55**, 61–73 (2014)
2. Dashtbozorg, B., Zhang, J., Huang, F., ter Haar Romeny, B.M.: Automatic optic disc and fovea detection in retinal images using super-elliptical convergence index filters. In: Campilho, A., Karray, F. (eds.) ICIAR 2016. LNCS, vol. 9730, pp. 697–706. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41501-7_78

3. Decenciére, E., et al.: Feedback on a publicly distributed image database: the messidor database. Image Anal. Stereology **33**(3), 231–234 (2014)
4. Gegundez-Arias, M.E., et al.: Locating the fovea center position in digital fundus images using thresholding and feature extraction techniques. Comput. Med. Imaging Graph. **37**(5), 386–393 (2013)
5. Kamble, R., et al.: Localization of optic disc and fovea in retinal images using intensity based line scanning analysis. Comput. Biol. Med. **87**, 382–396 (2017)
6. Kao, E.F., et al.: Automated detection of fovea in fundus images based on vessel-free zone and adaptive gaussian template. Comput. Methods Programs Biomed. **117**(2), 92–103 (2014)
7. Lu, S., et al.: Automatic macula detection from retinal images by a line operator. In: ICIP, pp. 4073–4076 (2010)
8. Molina-Casado, J.M., et al.: Fast detection of the main anatomical structures in digital retinal images based on intra- and inter-structure relational knowledge. Comput. Methods Programs Biomed. **149**, 55–68 (2017)
9. Nie, Y., et al.: Joint detection with faster R-CNN. In: ISITC (2016)
10. Niemeijer, M., et al.: Fast detection of the optic disc and fovea in color fundus photographs. Med. Image Anal. **13**(6), 859–870 (2009)
11. Pachade, S., Porwal, P., Kokare, M.: A novel method to detect fovea from color fundus images. In: Iyer, B., Nalbalwar, S.L., Pathak, N.P. (eds.) Computing, Communication and Signal Processing. AISC, vol. 810, pp. 957–965. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-1513-8_97
12. Ren, S., et al.: Faster R-CNN: towards real-time object detection with region proposal networks. NIPS **1**, 91–99 (2015)
13. Samanta, S., et al.: A simple and fast algorithm to detect the fovea region in fundus retinal image. In: EAIT, pp. 206–209 (2011)
14. Santhi, D., et al.: An efficient approach to locate optic disc center, blood vessels and macula in retinal images. Biomed. Eng. Appl. Basis Commun. **24**(05), 425–434 (2012)
15. Singh, J., et al.: Appearance-based object detection in colour retinal images. In: ICIP, pp. 1432–1435 (2008)
16. Yu, H., et al.: Fast localization of optic disc and fovea in retinal images for eye disease screening. SPIE Med. Imaging Comput.-Aided Diagn. **7963**, 317–328 (2011)
17. Zheng, S.H., et al.: A novel method of macula fovea and optic disk automatic detection for retinal images. J. Electron. Inf. Technol. **36**(11), 2586–2592 (2014)
18. Zheng, Y.: Research on macula area localization of fundus image based on the segmentation of retinal vessel ends. Ph. D. thesis, Huazhong University (2013)
19. Zhou, W., et al.: Detection of macula fovea in a retinal image. J. Image Graph. **23**(3), 442–449 (2018)