Optic Disk Detection in Fundus Image Based on Structured Learning

Zhun Fan, *Senior Member, IEEE*, Yibiao Rong ⓘ, Xinye Cai, Jiewei Lu, Wenji Li, Huibiao Lin, and Xinjian Chen ⓘ, *Senior Member, IEEE*

Abstract—**Automated optic disk (OD) detection plays an important role in developing a computer aided system for eye diseases. In this paper, we propose an algorithm for the OD detection based on structured learning. A classifier model is trained based on structured learning. Then, we use the model to achieve the edge map of OD. Thresholding is performed on the edge map, thus a binary image of the OD is obtained. Finally, circle Hough transform is carried out to approximate the boundary of OD by a circle. The proposed algorithm has been evaluated on three public datasets and obtained promising results. The results (an area overlap and Dices coefficients of 0.8605 and 0.9181, respectively, an accuracy of 0.9777, and a true positive and false positive fraction of 0.9183 and 0.0102) show that the proposed method is very competitive with the state-of-the-art methods and is a reliable tool for the segmentation of OD.**

Index Terms—**Edge detection, fundus image, optic disk, structured learning.**

I. Introduction

NOWADAYS, some of the most common causes of visual impairment and blindness are diabetic retinopathy, glaucoma, hypertension and macular degeneration [1], [2]. These eye diseases manifest themselves in the retina [3] and all of these diseases can be detected through a direct and regular ophthalmologic examination. However, many factors, such as population growth, aging, are contributing to the increase of the patients with these diseases, which makes the number of

Z. Fan, J. Lu, W. Li, and H. Lin are with the Key Laboratory of Digital Signal and Image Processing of Guangdong Provincial, College of Engineering, Shantou University, Shantou 515063, China (e-mail: zfan@stu.edu.cn; 12jwlu1@stu.edu.cn; liwj@stu.edu.cn; 13hblin@stu.edu.cn).

Y. Rong and X. Chen are with the School of Electrical and Information Engineering, Soochow University, Suzhou 215006, China (e-mail: ybrong@stu.edu.cn; xjchen@suda.edu.cn).

X. Cai is with the School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: xinye@nuaa.edu.cn).

Digital Object Identifier 10.1109/JBHI.2017.2723678

ophthalmologists needed for evaluation by direct examination becomes a limiting factor [4]. As a result, a computer aided diagnosis system which can significantly reduce the burden on the ophthalmologists and may alleviate the inter and intra observer variability [5] is desired.

In the process, OD detection plays an important role, which has attracted extensive attention from clinicians and researchers. OD detection is often a key step for the detection of other anatomical structures [6], [7]. For example, the OD location helps to prevent false positive detection of exudates incurred by diabetic retinopathy, since both OD and exudates are formed by bright regions in the fundus image [4]. Besides, the vessels, which are of direct importance in assessing vascular condition, radiate from the OD, which is the starting point of some algorithms for tracking vessels [7]. Moreover, the ratio of the size of the OD over the size of the optic cup has been widely utilized for glaucoma diagnosis. A high cup-to-disc ratio will indicate that a fundus is suspicious of glaucoma [4].

However, detecting OD automatically is challenging due to the variations of the OD's shape, size, colour and so on. Many OD detection algorithms have been introduced in the literature. The review of the methods for OD segmentation can be seen in [3] and [8]. Morales *et al.* [4] divided these methods into three categories, namely template based methods [9]–[11], deformable model based methods [7], [12], [13] and morphology based methods [4], [14], [15].

In the category of the template based methods, the method presented in [9] extracts the candidate regions of the OD by thresholding. Then Hough transformation is performed to outline the candidates as circles and the one which has the highest average intensity is selected as the OD. In [10], Prewitt edge detector is applied to obtain OD boundary candidates and circle Hough transform is employed to finish the final OD boundary segmentation. In [11], Lalonde *et al.* employed the Canny edge detector to extract the edge of OD and then Hausdorff-based template matching is applied to fit the OD boundary by a circle.

For the deformable model based methods, in [12], colour mathematical morphology is applied to remove the blood vessels and provide a homogenous OD region for the snake to lock onto. In [13], the principal component analysis (PCA) is employed to locate OD, and the boundary of OD is detected by a modified active shape model. In [7], Lowell *et al.* proposed an OD segmentation method using a deformable contour model. They exploited the specific characteristics of the OD to address
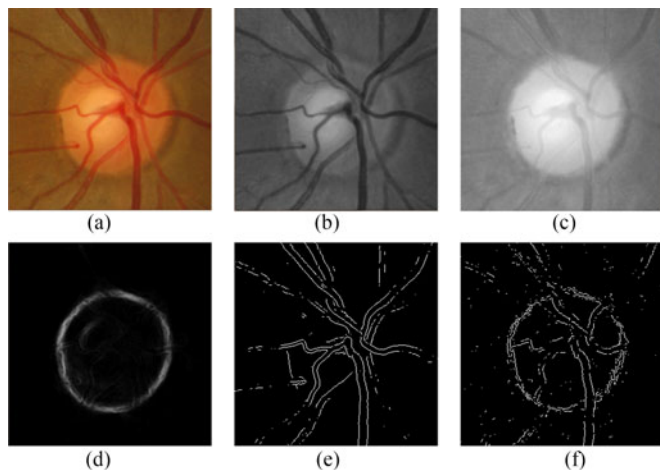
Fig. 1. A typical example to illustrate the different edge detection results. (a) Fundus image. (b) The green channel of the fundus image. (c) The red channel of the fundus image. (d) Edges obtained by structured learning. (e) Edges obtained by applying the Prewitt operator on the green channel. (f) Edges obtained by applying the Prewitt operator on the red channel.

the problem including strong distractors along the pallor and vessel edges, weakness of the rim and peripapillary atrophy.

For the morphology based methods, Walter *et al.* [14] extracted the OD using watershed transformation based on the assumption that the OD represents a bright region. In [4], the gray image obtained by PCA is chosen as the input. The stochastic watershed is applied to extract the watershed regions then region discrimination is performed to select the pixels which belong to the OD based on the average intensity of the region. In [15], Welfer *et al.* proposed an adaptive method for the segmentation of the OD using adaptive morphological approach.

All these methods introduced above belong to unsupervised methods. They were designed based on some assumptions, such as, the OD appears as a bright region in the fundus image [4], [14]. In this paper, we propose a method for OD detection based on structured learning which belongs to a supervised method to avoid making assumptions. The proposed method utilizes the edge information of the fundus image to detect the OD. It is different from the traditional method which applied the traditional edge detector, such as Prewitt edge detector [10], to capture the edge information. Since the vascular edges on the fundus image are very strong, when the traditional edge operator is applied to detect the OD edge, many vascular edges are detected besides OD edge.

In addition, for the traditional edge operator, using different channels as input may lead to different results, which makes choosing which channels as input is a critical decision. In [16], the green channel of the original fundus image is used. In [15], the red channel is used. In [10], a combination of the red channel and green channel is used. However, due to the variability of the fundus image, any individual method does not guarantee an optimal result [4].

A typical example to illustrate the disadvantages of the traditional edge operator is given in Fig. 1. It can be observed that when the Prewitt edge detector is performed on the green channel, only vascular edges are detected, which is shown in
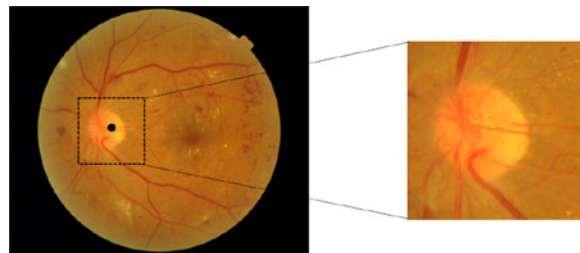

Fig. 2. The process of cropping the sub image.

Fig. 1(e). Besides, Fig. 1(e) and (f) illustrate that using different channels as input leads to different results.

To address these disadvantages, we employ the structured learning to detect the OD in this paper. Because structured learning belongs to a supervised method, we can train the edge detector to capture the special edge information, such as the OD edge information in this work. Moreover, we can simply take the original fundus image as the input of the edge detector to be trained, thus avoiding the need to consider which channels of the original image should be chosen. The reason for this is because Random Forest is employed as the edge detector for structured learning in this work, which has the capability to automatically select an optimal set of features from the original fundus image including all three (green, red and blue) channels, so that the resulting detector can capture the edge most properly. Fig. 1(d) gives an example where OD edge is obtained by the proposed method. Compared with the Prewitt operator [10], when structured learning is applied to detect the OD edge, there are hardly any blood vessel edges but the obvious OD edge map in the detected result.

The rest of the paper is organized as follows: in the Section II, the main steps of the proposed algorithm are described and explained. Section III presents the empirical study of parameter setting and experimental results obtained using three public databases MESSIDOR, DRIONS and ONHSD. Comparisons with other state-of-the-art methods, including [4], [7] and [10] are also made. We have discussion and conclusion in Section IV. A preliminary version of this work has been reported in [17].

II. METHOD

To realize the segmentation of OD, we crop the sub images which include the OD from the fundus images first, which can be done by locating the OD. The review of the methods for OD location can be seen in [8]. We employ the correlation filter [7] which uses a Laplacian of Gaussian template to match the key elements of OD structure to detect a point located in the OD. A 300×300 sub image which include the OD is then cropped based on the detected point. The process of cropping a sub image from the fundus image can be illustrated by Fig. 2. It is noteworthy to point out that subsequent operations are performed on the sub images.

The overall flowchart for OD segmentation is shown in Fig. 3. In the training stage, structured learning is used to train the model. In the test stage, given an unseen image, the trained model is employed to achieve the edge map of the OD. Thresholding is then performed on the edge map to obtain a binary
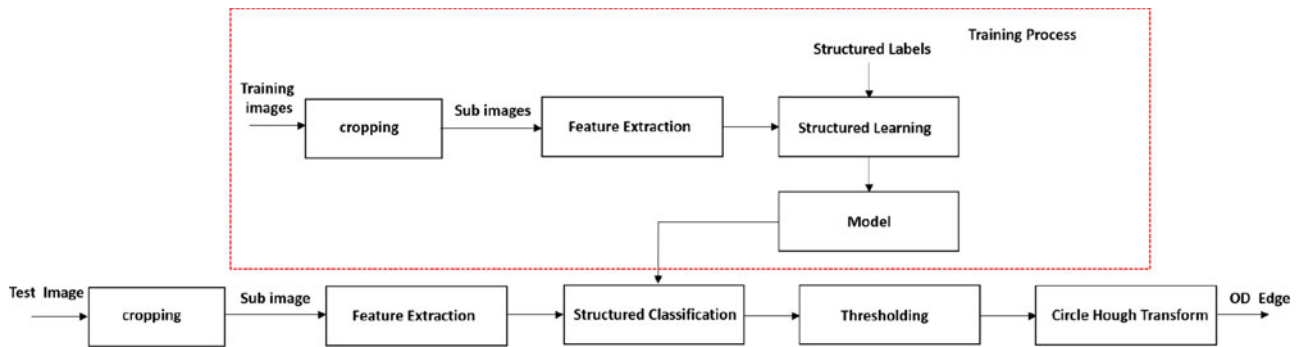
Fig. 3. The flow chart of The proposed algorithm. In training stage, structured learning is used to obtain the model. Then we use the trained model to achieve the edge map of OD. Thresholding is performed on the edge map thus a binary image of OD is obtained. Finally, circle Hough transform is carried out to approximate the boundary of OD.
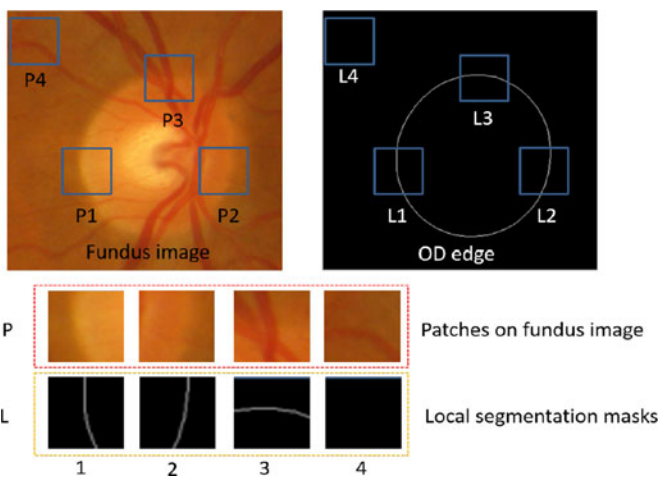

Fig. 4. The OD boundary is consisted of local structures, these structures are usually characterized by arcs. Since the output space of the structured learning can be arbitrarily complex representing, the problem of OD edge detection can be formulated as predicting local segmentation masks given fundus image patches.

image. Finally, circle Hough transform is applied to approximate the boundary of OD.

A. Obtaining Edge Map by Structured Learning

1) Problem Formulation: Structured learning or structured prediction is a term for supervised machine learning techniques that involve predicting structured objects, rather than scalar discrete or real values. It addresses the problem of learning a mapping where the input or output space may be arbitrarily complex representations, such as strings, sequences, graphs, object pose, bounding boxes etc. [18]–[20]. Similar to commonly used supervised learning techniques, structured prediction function are typically trained by means of observed data in which the ground truth is used to adjust model parameters. To learn more about structured learning, we refer readers to [21] for a comprehensive reading.

Similar with general images, the OD edges in local patches are highly interdependent [22]. They often contain well-known patterns, such as arc. Fig. 4 gives some examples where P1, P2, P3 are the patches with OD edges, P4 is the patch without

OD edge. L1, L2, L3 and L4 are corresponding local segmentation masks. Since the output space of the structured learning can be arbitrarily complex representations, the problem of OD edge detection using structured learning can be formulated as predicting local segmentation masks given input image patches. Concretely, given a patch on the fundus image (such as P1 in Fig. 4.) as input, the desired output of the trained model is a local segmentation mask (such as L1 in Fig. 4.).

2) OD Edge Detection: In this work, we employ the Structured Edge Detection Toolbox[1] developed by Dollár [18] to obtain the edge map of OD. Here, we give an explanation of how to apply the structured forests for OD edge detection. For more details of the structured forests, please refer to paper [18].

The pipeline of the method using structured forests to detect the OD edge can be illustrated by Fig. 5. A patch is cropped from the fundus image. A feature vector is then extracted to represent the patch so that the trained random forest can recognize the pattern in the patch. The random forest consists of different trees and the output of each tree is a local segmentation of the OD edge. The final output of the random forest is averaged over the local segmentation on each tree. A sliding window with default size 32×32 slides across the fundus image and the trained structured forest makes a prediction for each window to achieve the OD edge map.

For the details of feature extraction for each patch, please refer to the paper [18]. Here, we only give a brief explanation. Fig. 6 demonstrates the process of feature extraction for each patch. 13 channels, including 3 color, 2 magnitude and 8 orientation channels, are generated first. A patch (32×32) is cropped from each channel and downsampled by a factor of 2, resulting in $32 \times 32 \times 13/4 = 3328$ candidate features. In addition, the pair-wise features are extracted for the patch. Concretely, the patch is downsampled to a resolution of 5×5. Sampling all candidate pairs and computing their differences yield an additional $C_{5 \times 5}^2 = 300$ candidate features for each patch in each channel. Thus the total number of features for each patch is $13 \times 300 + 3328 = 7228$.

It should be noted that the edge map (as shown in Fig. 1(d) as an example) obtained by the structured random forest has a

---
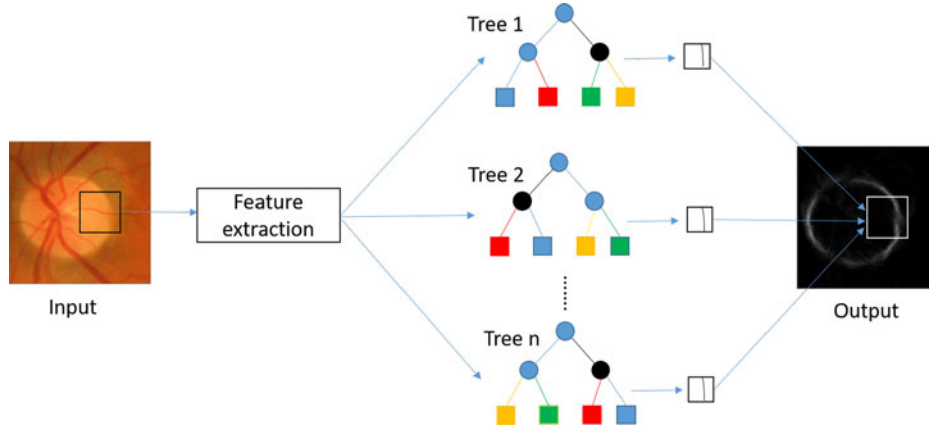
[1]https://github.com/pdollar/edges

Fig. 5.   The process of detecting the edge map of OD by structured forest.
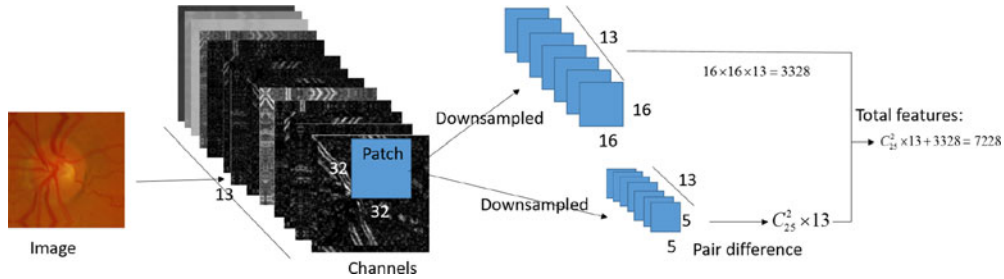

Fig. 6.   The schematic of feature extraction for each patch.

phenomenon of diffusion. To address this phenomenon, we first employ the thresholding to create a binary image, and then apply circle Hough transform on the binary image to approximate the OD boundary.

B. Thresholding

Given a grayscale image I , thresholding can be used to create a binary image BW. The threshloding method replaces each pixel in an image with a black pixel if the intensity $I(x,y)$ is less than some fixed value T, or a white pixel if the image intensity is greater than that value. This can be defined as:

$$BW(x,y) = \begin{cases} 1 & \text{if } I(x,y) > T \\ 0 & \text{others} \end{cases} \quad (1)$$

The key step of thresholding is to select the threshold T. In this work, we employ the Otsu's method [23] to decide the threshold. To protect more information of the OD boundary, we reduce T by multiply a factor f less than 1. Fig. 7 shows the binary image when f is equal to 1, 0.618, 0.4, respectively. We can see that when f is smaller, more OD edge information is protected. But at the same time, more false edge may be included. So a tradeoff has to be made. In our work, we set f is to 0.618 by empirical experiments.

C. Circle Hough Transform

Circle Hough transform is a popular method which used to find the circle patterns in the image [24]. The procedure of circle

Fig. 7.   Different results when f is set to different values. (a) f is equal to 1, (b) f is equal to 0.618, (c) f is equal to 0.4.

Hough transform can be defined as [10] :

$$(c_x, c_y, r) = CHT(I_{\text{BW}}, r_{\min}, r_{\max}) \quad (2)$$

where I_{BW} is a binary image and $(r_{\min}, r_{\max})$ is the search range of the radius. (c_x, c_y) and r are respectively the center position and radius obtained by circle Hough transform.

The $r_{\min}$ and $r_{\max}$ are key parameters which used to limit the search range of the circle Hough transform. Assume that the size of a fundus image is $s_x \times s_y$, the following equations are used for setting $r_{\min}$ and $r_{\max}$ automatically based on the empirical study in this work:

$$r_{\min} = 0.013 \times s_{\max} - 0.01 \times s_{\min} + 37 \quad (3)$$

$$r_{\max} = 0.013 \times s_{\max} - 0.01 \times s_{\min} + 77 \quad (4)$$

where $s_{\max} = \max[s_x, s_y]$, $s_{\min} = \min[s_x, s_y]$. For example, if the size of the fundus image is 640×480 , then $s_{\max} = 640$,

Fig. 8. An example of fitting the OD boundary by a circle. (a) A circular approximation of the OD boundary. (b) The effect of superimposing the circle on the fundus image.
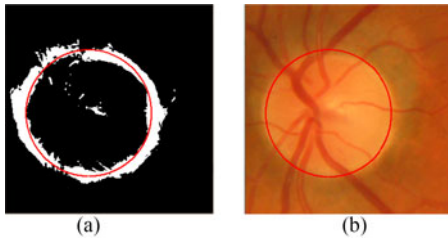
TABLE I
VISUAL CHARACTERISTIC OF DRIONS DATABASE

Characteristics	Number of Images
Light Artifact	3
Rim Blurred or Missing	5
Moderate Peripapillary Atrophy	16
Concentric Peripapillary Atrophy	20
Strong Pallor Distractor	6

$s_{\min} = 480$. Fig. 8 gives an example where the red circle is the approximation of the OD boundary detected by circle Hough transform.

III. RESULTS

The proposed algorithm has been evaluated on three public databases: MESSIDOR [25], DRIONS [26], [27] and ONHSD [7].

MESSIDOR database contains 1200 eye fundus color images. The images were acquired by 3 ophthalmologic departments using a color video 3CCD (Charge Coupled Device) camera on a Topcon TRC NW6 non-mydriatic retinpgraph with 45 degree field of view. The images were captured using 8 bits per color plane at 1140×1960, 2240×1488 or 2304×1536 pixels. 800 images were acquired with pupil dilation and 400 without dilation [25]. The OD boundaries of these 1200 images have been segmented and are currently available online [28].

DRIONS database consists of 110 colour digital fundus images with their OD manually segmented by two different specialists. The images were acquired with a colour analogical fundus camera and digitized using a HP-PhotoSmart-S20 high-resolution scanner. The resolution is 600×400 and 8 bits per pixel. In this database, the mean age of the patients was 53.0 years (s.d. 13.05), with 46.2% male and 53.8% female and all of them were Caucasian ethnicity. 23.1% patients had chronic simple glaucoma and 76.9% eye hypertension. Some of the 110 images contain visual characteristics related to potential problems that may distort the detection process of the OD contour. These visual characteristics are listed in Table I [26]. The OD boundaries of these 110 images have been segmented by two specialists. To facilitate comparing with other algorithms on this database, in this work, the images segmented by the first specialist have been chosen as the golden standard, the same choice has been made by [4].

TABLE II
VISUAL CHARACTERISTIC OF ONHSD DATABASE

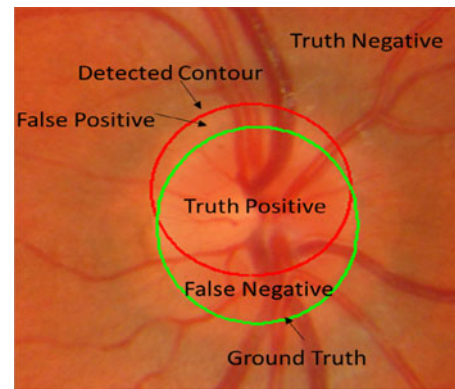| Characteristic	Number of images
No detectable optic nerve head	4
Severe Cataract	8
Moderate Cataract	2
Exudates or laser scars	7
Light artifacts	7
Easily visible choroidal vessels	20
Some of rim blurred or missing	27
Severe peripapillary atrophy	6
Moderate peripapillary atrophy	29
Concentric peripapillary atrophy	23
Strong pallor distractor	13

Fig. 9. The sketch map used to demonstrate the basic metrics TP, FP, TN and FN.

TABLE III
THE DEFAULT VALUE OF EACH PARAMETER

Parameters	Default Values
The max depth of the tree	64
The number of the trees	8
The number of patches	10^6
The number of features	1932
The dimensionality of the binary vector	256
The size of the patch	32×32
The size of the local segmentation	16×16
The number of principle components	5

ONHSD database contains 99 fundus images with a resolution of 640×480 taken from 50 patients randomly sampled from a diabetic retinopathy screening programme. 96 images have discernable optic nerve head (ONH) and 90 images are nominated for evaluating segmentation of the algorithm. These images were acquired using a Canon CR6 45MNf fundus camera, with a field angle lens of 45 degrees. The subjects are from various ethnic backgrounds (Asian 20%, Afro-Caribbean 16%, Caucasian 50%, Unknown 14%). 19 have type 2 diabetes mellitus, while the diabetes status was unavailable for the remaining 31. In this database, there is considerable variation in the images, with many characteristics that can affect the algorithm. They are summarized in Table II [7]. In this database, the ONH center
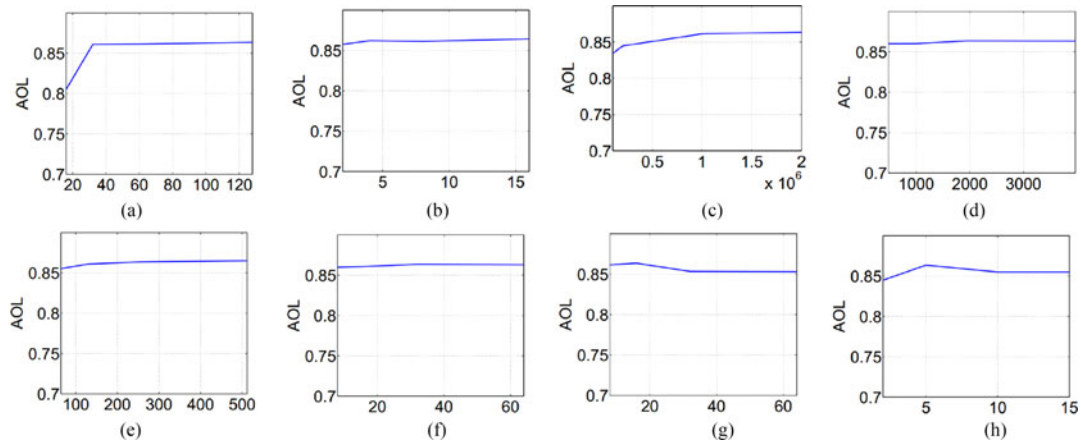
Fig. 10. Empirical study of parameter setting. (a) The max depth of the tree. (b) The number of the trees. (c) The number of patches used to train the forest. (d) The number of features used to train each tree. (e) The dimensionality of the binary vector. (f) The size of the patch. (g) The size of the local segmentation (LS). (h) The number of principle components (PC).

has been marked up by a clinician, and four clinicians marked the ONH edge which intersects with radial spokes (at 15 degree angles) radiating from the nominated center. The average of the edges marked by the four experts has been used to generate the golden standard.

The performance of the methods has been evaluated according to different metrics. We use Fig. 9 to illustrate the concepts. Let TP represent True Positive, FP False Positive, TN True Negative, and FN False Negative. Based on these basic metrics, Area Overlap (AOL), Dices (S) coefficients, Accuracy (Ac), True Positive Fraction (TPF) and False Positive Fraction (FPF) are defined as follows:

$$AOL = TP/(TP + FN + FP) \tag{5}$$

$$S = 2TP/(2TP + FN + FP) \tag{6}$$

$$Ac = (TP + TN)/(TP + TN + FP + FN) \tag{7}$$

$$TPF = TP/(TP + FN) \tag{8}$$

$$FPF = FP/(FP + FN) \tag{9}$$

Since the proposed algorithm belongs to supervised method, we employ K-fold cross validation [29] to evaluate the algorithm. (K is equal to 12 for the MESSIDOR database since the images have been divided into 12 groups (each group contains 100 images) by the authors [25], K is equal to 10 for DRIONS and 9 for ONHSD to ensure the images in the database can be averagely divided into K groups.

It is noteworthy to point out that the accuracy of locating the OD employed the correlation filter [7] are 99%, 98%, 97% on ONHSD, DRIONS and MESSIDOR datasets respectively. For those images that the OD is located by correlation filter successfully, we employ the method demonstrated in Fig. 2 to crop the sub images automatically. For those images that correlation filter fails to locate the OD, we crop the sub images by hand. As a result, all the sub images used for evaluating the performance of the proposed method include the OD for segmentation.

	MESSIDOR	DRIONS	ONHSD
AOL	0.8636(0.1268)	0.8473(0.0964)	0.8346(0.1003)
S	0.9196(0.1019)	0.9137(0.0634)	0.9032(0.0740)
Ac	0.9770(0.0284)	0.9760(0.0166)	0.9895(0.0077)
TPF	0.9212(0.1213)	0.8957(0.0927)	0.9077(0.0991)
FPF	0.0106(0.0129)	0.0098(0.0092)	0.0055(0.0052)

AOL–Area Overlap; S–Dices coefficients; Ac–Accuracy; TPF–True Positive Fraction; FPF–False Positive Fraction.

A. Empirical Study of Parameter Setting

The proposed method for OD detection is based on structured forests, whose performance may be influenced by parameters, including the max depth of the tree, the number of decision trees, the number of patches used to train the forest, the number of features used to train the tree, the dimensionality of the binary vector, the size of the patch, the size of the local segmentation and the number of principle components [18]. We employ the MESSIDOR database to study the influence of the parameters on the performance of the proposed method. We set a default value for each parameter. Then we allow one change when the others are equal to the default values. As a result, we can explore how varying the parameter may affect the performance of the proposed method. The default value of each parameter is obtained by empirical study and listed in Table III.

The experiments are performed on a PC equipped with an Intel (R) Core (TM) i-5 4210 M CPU at 2.60 GHZ and 4 GB of RAM capacity using MATLAB. Once we have obtained the trained structured forest (It takes about 30 minutes to train a tree when the number of the patches used to train the tree is 10^6), the average computational time obtained for OD segmentation is 1.7494 s with a standard deviation of 0.2987 s.

Fig. 10 shows how varying the parameters affects the AOL, a metric which can reflect the overall performance of the proposed
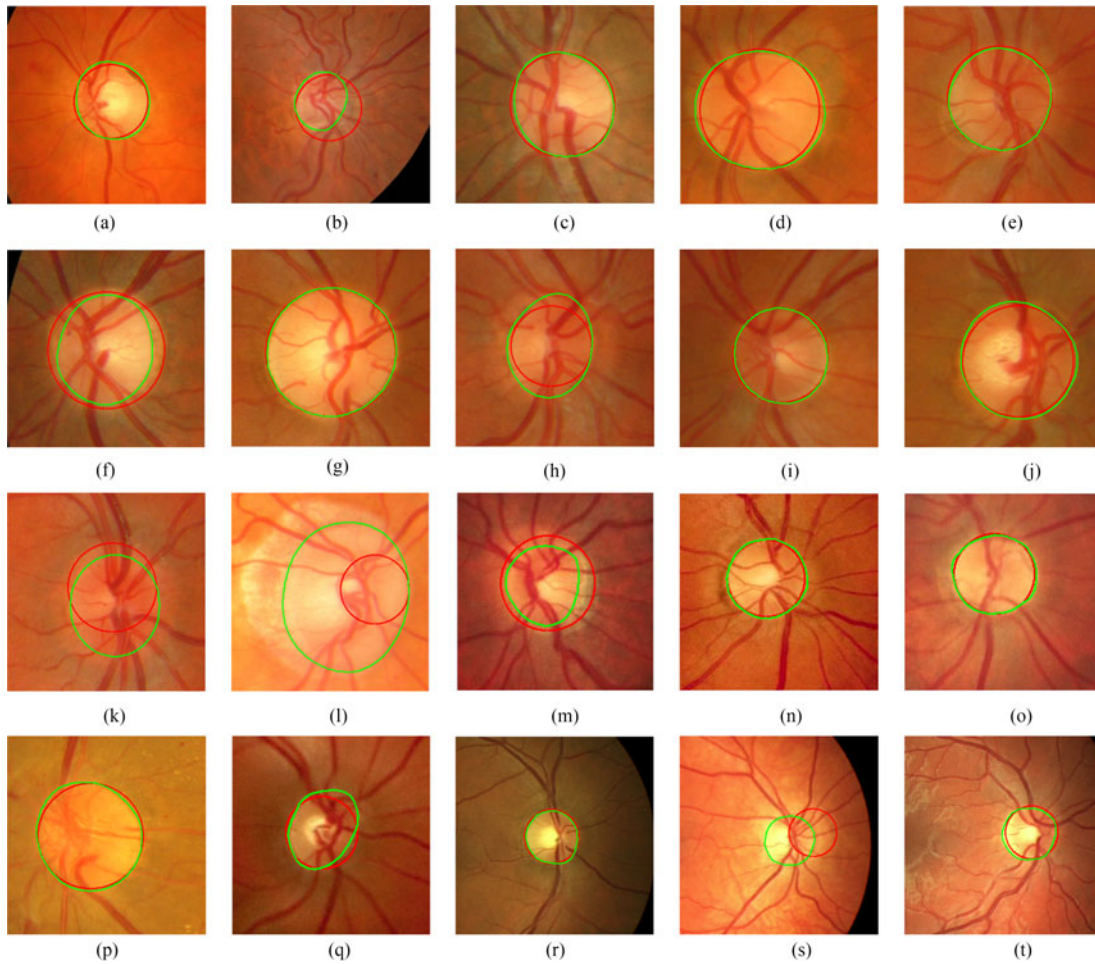
Fig. 11. Some examples where green circles are the ground truth and red circles are the circle detected by the proposed algorithm. (a) AOL = 0.8983, S = 0.9464, Ac = 0.9880, TPF = 0.9200, FPF = 0.0031. (b) AOL = 0.6102, S = 0.7579, Ac = 0.9646, TPF = 0.9340, FPF = 0.0334. (c) AOL = 0.9339, S = 0.9658, Ac = 0.9854, TPF = 0.9820, FPF = 0.0137. (d) AOL = 0.9173, S = 0.9569, Ac = 0.9743, TPF = 0.9260, FPF = 0.0042. (e) AOL = 0.9513, S = 0.9751, Ac = 0.9894, TPF = 0.9976, FPF = 0.0127. (f) AOL = 0.7673, S = 0.8683, Ac = 0.9361, TPF = 1, FPF = 0.0810. (g) AOL = 0.9815, S = 0.9907, Ac = 0.9938, TPF = 0.9946, FPF = 0.0066. (h) AOL = 0.7438, S = 0.8531, Ac = 0.9555, TPF = 0.7438, FPF = 0. (i) AOL = 0.9636, S = 0.9814, Ac = 0.9934, TPF = 0.9902, FPF = 0.0059. (j) AOL = 0.9185, S = 0.9575, Ac = 0.9783, TPF = 0.9185, FPF = 0. (k) AOL = 0.6059, S = 0.7545, Ac = 0.9175, TPF = 0.7039, FPF = 0.0355. (l) AOL = 0.2478, S = 0.3971, Ac = 0.7154, TPF = 0.2478, FPF = 0. (m) AOL = 0.6678, S = 0.8008, Ac = 0.9399, TPF = 1, FPF = 0.0683. (n) AOL = 0.9217, S = 0.9593, Ac = 0.9899, TPF = 0.9217, FPF = 0. (o) AOL = 0.9411, S = 0.9697, Ac = 0.9919, TPF = 0.9479, FPF = 0.0011. (p) AOL = 0.9326, S = 0.9651, Ac = 0.9847, TPF = 0.9523, FPF = 0.0006. (q) AOL = 0.7969, S = 0.8870, Ac = 0.9761, TPF = 0.8661, FPF = 0.0105. (r) AOL = 0.9370, S = 0.9675, Ac = 9963, TPF = 0.9836, FPF = 0.0030. (s) AOL = 0.2358, S = 0.3816, Ac = 0.9414, TPF = 0.0362, FPF = 0.0283. (t) AOL = 0.8540, S = 0.9213, Ac = 0.9911, TPF = 0.9723, FPF = 0.0037.

method. Fig. 10(a) gives the relationship between the max depth of the tree and the AOL. It is noted that the performance of the algorithm improves with the increase of the max depth of the tree until it reaches 32. When the max depth of the tree is greater than 32, the performance of the algorithm tends to be stagnant. A similar trend can be observed for other parameters as shown in Fig. 10(b)–(f). In addition, Fig. 10(g) and (h) show how varying the size of local segmentation and the number of principle components affect the AOL respectively. It is noted that the performance of the algorithm improves with the increase of the size of the local segmentation in the range of 0 to 16 . After that, the performance decreases slightly and then becomes stable again. A similar trend can be observed for the number of principle components as shown in Fig. 10(h).

We set the parameters equal to the default values as shown in Table III for the following experiments. Table IV shows the

TABLE V
COMPARISONS WITH THE MORPHOLOGY BASED METHODS ON DRIONS DATABASE

	Proposed	Morales [4]	Walter [14]
AOL	**0.8473(0.0964)**	0.8424(0.1174)	0.6227(0.3695)
S	**0.9137(0.0634)**	0.9084(0.0982)	0.6813(0.3854)
Ac	0.9760(0.0166)	**0.9934(0.0051)**	0.9689(0.0492)
TPF	0.8957(0.0927)	**0.9281(0.1177)**	0.6715(0.3980)
FPF	0.0098(0.0092)	**0.0040(0.0041)**	0.0210(0.0417)

AOL–Area Overlap; S–Dices coefficients; Ac–Accuracy; TPF–True Positive Fraction; FPF–False Positive Fraction.

average values and the standard deviations of the various performance metrices achieved on the MESSIDOR database, DRIONS database and ONHSD database.

TABLE VI
COMPARISONS WITH THE METHOD PROPOSED BY MORALES ON MESSIDOR AND ONHSD DATABASES

	MESSIOR		ONHSD	
	Proposed	Morales [4]	Proposed	Morales [4]
AOL	**0.8636(0.1268)**	0.8228(0.1384)	**0.8346(0.1003)**	0.8045(0.1175)
S	**0.9196(0.1019)**	0.8950(0.1056)	**0.9032(0.0740)**	0.8867(0.0776)
Ac	0.9770(0.0284)	**0.9949(0.0050)**	0.9895(0.0077)	**0.9941(0.0042)**
TPF	0.9212(0.1213)	**0.9300(0.1239)**	0.9077(0.0991)	**0.9310(0.1046)**
FPF	0.0106(0.0129)	**0.0035(0.0041)**	0.0055(0.0052)	**0.0043(0.0042)**

AOL–Area Overlap; S–Dices coefficients; Ac–Accuracy; TPF–True Positive Fraction; FPF–False Positive Fraction.

TABLE VII
COMPARISONS WITH THE TEMPLATE-BASED METHODS

	$AOL \geq 0.95$	$AOL \geq 0.90$	$AOL \geq 0.85$	$AOL \geq 0.80$	$AOL \geq 0.70$	$\overline{AOL}$
Elliptical Geometrical Min. [10]	2%	11%	20%	30%	51%	0.6700
Elliptical Algebraic Min1. [10]	2%	9%	17%	26%	48%	0.6600
Elliptical Algebraic Min2. [10]	2%	10%	19%	27%	47%	0.6500
Elliptical Algebraic Min3. [10]	2%	9%	18%	26%	48%	0.6600
Circular Hough. [10]	7%	46%	73%	84%	93%	0.8600
Proposed	7%	51%	78%	86%	93%	0.8636

Results in terms of percentage of images per area overlap interval and average area overlap of the whole messidor database.

Fig. 11 gives some examples where red circles are the OD boundaries detected by the proposed algorithm, and green circles are the ground truth. It is noted that the proposed method can detect the OD boundaries properly for most given cases. In addition, the proposed method can exclude the affections coming from bright lesions, which is demonstrated by Fig. 11(p) whose original figure is shown in Fig. 2. In this case, the proposed method can detect the OD boundary properly even though the original fundus image is affected by bright lesions. However, the proposed method may fail for some cases, such as the ones with severe peripapillary atrophy (Fig. 11(l)), with OD rim missing (Fig. 11(s)). The reason is that the proposed method relies on the edge information to detect the OD. However, severe peripapillary atrophy may cause false edge, and OD rim missing may cause edge information missing. Both of them may make the edge detector difficult to detect the true edge.

From these cases, we can see that when the detected boundary are included in the ground truth, such as Fig. 11(l), an ideal value of FPF is achieved even though the detected result is not perfect at all. When the ground truth are included in the detected boundary, such as Fig. 11(m), an ideal value of TPF is obtained even though the detected result is not perfect either. So, the indicators of TPF and FPF can only evaluate the performance of the algorithm one-sided.

In addition, we can draw the conclusion that AOL is more sensitive to the deteriorating results than Ac according the definition of AOL and Ac. This can also be illustrated by comparing Fig. 11(g) and Fig. 11(t), where AOL = 0.9815, Ac = 0.9938 in Fig. 11(g) and AOL = 0.8540, Ac = 0.9911 in Fig. 11(t), which shows that for a deteriorating result the AOL has more obvious change. It can also be illustrated in case of Fig. 11(s) when a

Fig. 12. The examples of each category where green circles are the ground truth and red circles are the results detected by the proposed algorithm. (a) Excellent. (b) Good. (c) Fair. (d) Poor.

bad result is obtained, AOL can reflect this more obviously than Ac (AOL = 0.2358 while Ac = 0.9414 in Fig. 11(s)).

B. Comparisons with Other Algorithms

As was introduced in Section I, the methods for segmentation of OD can be divided into three categories: template based

TABLE VIII
THE AVERAGE VALUES AND STANDARD DEVIATIONS OF THE METRICS IN EACH RESULT GROUP

	AOL	S	Ac	TPF	FPF
Excellent	0.9027(0.0235)	0.9487(0.0129)	0.9936(0.0023)	0.9542(0.0315)	0.0036(0.0020)
Good	0.8307(0.0292)	0.9027(0.0175)	0.9896(0.0023)	0.9646(0.0445)	0.0087(0.0037)
Fair	0.7016(0.0748)	0.8223(0.0548)	0.9808(0.0085)	0.9448(0.0768)	0.0170(0.0100)
Poor	0.3385(0.1892)	0.4916(0.2061)	0.9493(0.0299)	0.6748(0.4600)	0.0345(0.0059)

AOL–Area Overlap; S–Dices coefficients; Ac–Accuracy; TPF–True Positive Fraction; FPF–False Positive Fraction.

TABLE IX
SUBJECTIVE CLASSIFICATION OF PERFORMANCE ON ONHSD DATASET

	Excellent	Good	Fair	Poor	Excellent-Good	Excellent-Fair
Morales [4]	28%	36%	31%	6%	64%	94%
Circular Hough [10]	40%	39%	18%	3%	79%	97%
Temploral Lock [7]	42%	31%	10%	17%	73%	83%
Simple [7]	9%	8%	30%	53%	17%	47%
DV-Hough [7]	39%	22%	20%	19%	61%	81%
Proposed	33%	40%	25%	2%	73%	98%

methods, morphology based methods and deformable model based methods. In this work, we select one or more methods of each category for comparisons.

1) Comparisons with Morphology Based Methods: Two morphology based methods are selected for comparisons. One is proposed by Morales *et al.* [4] which is based on PCA and mathematical morphology, another one is proposed by Walter *et al.* [14] which is based on marker controlled watershed transformation.

Table V gives the results on the DRIONS dataset. It should be noted that the proposed method improves the performance of the method proposed by Walter significantly and is very competitive with the method proposed by Morales, which is one of the best performing methods for OD segmentation up to now.

To have a more comprehensive evaluation, Table VI shows the comparisons with the method proposed by Morales on the MESSIORD and ONHSD databases. It can be noted that the AOL and S obtained by the proposed method are better than those obtained by the method proposed by S. Morales in all three databases. For the Ac, TPF, FPF, the results obtained by the proposed method are not as good as those of the S. Morales. But the difference is not significant. In addition, as we illustrate before, the Ac is less sensitive to deteriorating results, and the TPF, FPF can only evaluate the performance of the algorithm one-sided. So from this perspective, the proposed algorithm may be better.

2) Comparisons with Template Based Methods: Table VII summarizes the comparisons with one circular template-based method and four elliptical template-based methods [10] which have been validated on MESSIDOR database. The average area overlap $\overline{AOL}$ obtained by the proposed method is 0.8636 on MESSIDOR database, which is a very competitive result to the circular template-based method proposed in [10]. In addition, the proposed method outperform all the elliptical template-based approaches proposed in [10].

3) Comparisons with Deformable Based Methods: In order to compare with the method based on deformable models [7], Our algorithm has also been tested on ONHSD database. With the aim of obtaining a fair comparison, We use the same indicator as the used in [7], which is discrepancy δ^j defined as:

$$\delta^j = \sum_i \frac{m_i^j - \mu_i^j}{\sigma_i^j + \varepsilon} \tag{10}$$

where μ_i^j and σ_i^j summarize the clinician's choice of rim location on spoke i of image j. m_i^j is segmentation location on spoke i of image j. Spokes are points belonging to the OD boundary. 24 spokes were considered each taking 15 degree with its adjacent ones. $\varepsilon = 0.5$ is a small factor to prevent division by zero when the clinicians are in exact agreement with the OD boundary. Each image is classified as Excellent, Good, Fair, or Poor depending on the discrepancy value (up to one, two, five or more respectively).

Fig. 12 shows the examples of each category, where we can see that the detected OD edge is almost coincident with the ground truth for the excellent case. Small deviation is observed for the good case and obvious deviation for the fair case. For the poor case, the detected OD edge is deviated from the ground truth completely. Table VIII summarizes the average values and standard deviations of the metrics (AOL, S, Ac, TPF and FPF) in each result group.

Table IX lists the performance on the subjective scale for three compared algorithms [7] (the third to fifth row) and the method proposed in this paper (the last row). As it can be noted, with the best deformable model approach (Temploral Lock [7]), 9% more of excellent segmentations are obtained than those obtained with the proposed method. However, the proposed method obtains a same result as the Temploral Lock [7] in the Excellent-Good range. In addition, the proposed method provides a significant enhancement in the percentage of Excellent-Fair range. The best

TABLE X
DATABASE CROSS VALIDATION

	MESSIDOR	DRIONS	ONHSD
AOL	0.7462(0.2479)	0.8627(0.0679)	0.8343(0.0892)
S	0.8205(0.2272)	0.9246(0.0421)	0.9031(0.0689)
Ac	0.9515(0.0568)	0.9781(0.0110)	0.9893(0.0078)
TPF	0.7975(0.2547)	0.9365(0.0675)	0.9172(0.0872)
FPF	0.0156(0.0281)	0.0148(0.0105)	0.0063(0.0063)

AOL–Area Overlap; S–Dices coefficients; Ac–Accuracy; TPF–True Positive Fraction; FPF–False Positive Fraction.

approach of [7] has Excellent-Fair performance in 83% of all cases, while our method obtains it in 98% of them.

Table IX also summarizes the performance on this subjective scale for morphology based method [4] (the first row) and template based method [10] (the second row). It can be observed that within the Excellent-Good range, the result (73% images belong to the Excellent-Good range) obtained by the proposed method is the second best among the listed methods. Within the Excellent-Fair range, the proposed method obtains the best result among these methods.

IV. DISCUSSION & CONCLUSION

The proposed method in this paper employs structured learning to capture the OD edge information. Because the proposed algorithm belongs to supervised methods, the trained edge detector plays an important role in the performance of the proposed method. To illustrate the importance of the trained edge detector, we do database cross validation. Concretely, we train the model using the images on the MESSIDOR database, and test on the DRIONS database and ONHSD database. In turn, the images on the DRIONS and ONHSD are used as training set, and the images on the MESSIDOR are used test set. Table X summarizes the results.

It should be noted that when we train the model using the images on MESSIDOR database, and test on the DRIONS and ONHSD, very promising results are obtained (The last two columns in the Table X). But not vice versa. The reason is that the number of the images on MESSIDOR are larger than the number of images on DRIONS and ONHSD, thus more different local structures are contained on MESSIDOR. If we use the images on DRIONS and ONHSD to train the model, many local structures on MESSIDOR can not be captured during training and the trained model is unable to synthesize novel labels. That is why we obtain worse results when we train the model using DRIONS and ONHSD databset, and test on the MESSIDOR dataset.

However, it dose not mean that the more number of patches used to train the structured forest, the better. It can be illustrated by Fig. 10(c) where the performance of the proposed method tends to become stagnant when the number of patches used to train the forest is larger than 10^6. Thus, the deep learning, which has the potential to improve the performance by increasing the number of the samples used to train the learner [30], will be considered for training the edge detector in our future work.

Recently, deep learning based methods have been successfully applied in different fields, such as edge detection [31],

[32], and retinal image analysis [33], [34]. Many researchers have also employed deep learning for OD segmentation (e.g., [35], [36]) and obtained very promising results (AOL of 0.888 obtained by [35] in Messidor dataset and AOL of 0.913 obtained by [36] in a local dataset). In both [35] and [36], the authors labeled the patches based on the pixel level. In this work, we employ structured labels to train the edge detector which can take advantage of the neighborhood information of a pixel because the adjacent pixels are highly interdependent. In the future work, we will combine the advantages of deep learning and the structured labels to train the edge detector.

In summary, in this paper we proposed an algorithm for the automatic detection of OD based on structured learning, which can train a model to detect the specified edge. The weighted average of the results (according to the number of images of each database: AOL = 0.8605, S = 0.9181, Ac = 0.9777, TPF = 0.9183, FPF = 0.0102) show that the proposed algorithm is reliable since it works properly on different databases.

REFERENCES

[1] *Action Plan for the Prevention of Avoidable Blindness and Visual Impairment, 2009–2013*, World Health Organization, Geneva, Switzerland, 2010.

[2] D. Pascolini and S. P. Mariotti, "Global estimates of visual impairment: 2010," *Brit J. Ophthalmol.*, vol. 96, pp. 614–618, 2012.

[3] M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 169–208, Sep. 2010.

[4] S. Morales, V. Naranjo, J. Angulo, and M. Alcañiz, "Automatic detection of optic disc based on PCA and mathematical morphology," *IEEE Trans. Med. Imag.*, vol. 32, no. 4, pp. 786–796, Apr. 2013.

[5] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "Computer-aided diagnosis of diabetic retinopathy: A review," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2136–2155, 2013.

[6] S. Lu, "Accurate and efficient optic disc detection and segmentation by a circular transformation," *IEEE Trans. Med. Imag.*, vol. 30, no. 12, pp. 2126–2133, Dec. 2011.

[7] J. Lowell *et al.*, "Optic nerve head segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 256–264, Feb. 2004.

[8] N. Patton *et al.*, "Retinal image analysis: Concepts, applications and potential," *Progress Retinal Eye Res.*, vol. 25, no. 1, pp. 99–127, 2006.

[9] M. Park, J. S. Jin, and S. Luo, "Locating the optic disc in retinal images," in *Proc. IEEE Int. Conf. Comput. Graph., Imag. Vis.*, 2006, pp. 141–145.

[10] A. Aquino, M. E. Gegúndez-Arias, and D. Marín, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1860–1869, Nov. 2010.

[11] M. Lalonde, M. Beaulieu, and L. Gagnon, "Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching," *IEEE Trans. Med. Imag.*, vol. 20, no. 11, pp. 1193–1200, Nov. 2001.

[12] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Comparison of colour spaces for optic disc localisation in retinal images," in *Proc. IEEE 16th Int. Conf. Pattern Recognit.*, 2002, vol. 1, pp. 743–746.

[13] H. Li and O. Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 2, pp. 246–254, Feb. 2004.

[14] T. Walter, J.-C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1236–1243, Oct. 2002.

[15] D. Welfer, J. Scharcanski, C. M. Kitamura, M. M. Dal Pizzol, L. W. Ludwig, and D. R. Marinho, "Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach," *Comput. Biol. Med.*, vol. 40, no. 2, pp. 124–137, 2010.

[16] M. Niemeijer, M. D. Abràmoff, and B. Van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs," *Med. Image Anal.*, vol. 13, no. 6, pp. 859–870, 2009.

[17] Z. Fan *et al.*, "Detecting optic disk based on structured learning," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2015, pp. 1127–1132.

[18] P. Dollár and C. L. Zitnick, "Fast edge detection using structured forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 1, pp. 1558–1570, Aug. 2015.

[19] B. Taskar, V. Chatalbashev, D. Koller, and C. Guestrin, "Learning structured prediction models: A large margin approach," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 896–903.

[20] M. B. Blaschko and C. H. Lampert, "Learning to localize objects with structured output regression," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 2–15.

[21] S. Nowozin and C. H. Lampert, "Structured learning and prediction in computer vision," *Found. Comput. Graph. Vis.*, vol. 6, no. 3/4, pp. 185–365, 2011.

[22] J. J. Lim, C. L. Zitnick, and P. Dollár, "Sketch tokens: A learned mid-level representation for contour and object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 3158–3165.

[23] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.

[24] J. Illingworth and J. Kittle, "The adaptive hough transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 5, pp. 690–698, Sep. 1987.

[25] M. T.-V. Project, "Messidor: Digital retinal images france," 2008. [Online]. Available: http://messidor.crihan.fr/download-en.php

[26] "Drions-db: Digital retinal images for optic nerve segmentation database," 2008. [Online]. Available: http://www.ia.uned.es/ ejcarmona/DRIONS-DB.html

[27] E. J. Carmona, M. Rincón, J. García-Feijoó, and J. M. Martínez-de-la Casa, "Identification of the optic nerve head with genetic algorithms," *Artif. Intell. Med.*, vol. 43, no. 3, pp. 243–259, 2008.

[28] U. Huelva, "Expert system for early automated detection of od by analysis of digital retinal images project website," Sep. 24, 2012. [Online]. Available: http://www.uhu.es/retinopathy/muestras.php

[29] S. Rogers and M. Girolami, *A First Course in Machine Learning*. Boca Raton, FL, USA: CRC Press, 2011, pp. 20–25.

[30] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[31] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1395–1403.

[32] X. Sui *et al.*, "Choroid segmentation from optical coherence tomography with graph-edge weights learned from deep convolutional neural networks," *Neurocomputing*, vol. 237, pp. 332–341, 2017.

[33] P. Arbeláez and L. Van Gool, "Deep retinal image understanding," in *Proc. Int. Conf. Med. Image Computing Comput. Assisted Intervention*, 2016, pp. 140–148.

[34] H. S. Alghamdi, H. L. Tang, S. A. Waheeb, and T. Peto, "Automatic optic disc abnormality detection in fundus images: A deep learning approach," in *Proc. Int. Conf. Med. Image Computing Comput. Assisted Intervention*, 2016, pp. 17–24.

[35] G. Lim, Y. Cheng, W. Hsu, and M. L. Lee, "Integrated optic disc and cup segmentation with deep learning," in *Proc. IEEE Int. Conf. Tools Artif. Intell.*, 2015, pp. 162–169.

[36] R. Srivastava, J. Cheng, D. W. K. Wong, and J. Liu, "Using deep learning for robustness to parapapillary atrophy in optic disc segmentation," in *Proc. Int. Symp. Biomed. Imag.*, 2015, pp. 768–771.