Neurocomputing 312 (2018) 285–295

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Fast and effective optic disk localization based on convolutional neural network

Xianjing Meng[a], Xiaoming Xi[a], Lu Yang[a], Guang Zhang[c], Yilong Yin[b,*], Xinjian Chen[d]

[a] School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China
[b] School of Computer Science and Technology, Shandong University, Jinan 250101, China
[c] Hospital of Qianfoshan, Jinan 250014, China
[d] School of Electronic and Information Engineering, Soochow Univeristy, Suzhou 215006, China

ARTICLE INFO

Article history:
Received 12 June 2017
Revised 19 March 2018
Accepted 29 May 2018
Available online 6 June 2018

Communicated by Dr Tianzhu Zhang

Keywords:
Optic disc localization
Convolutional neural network
Rectangular ROI
Multi-stage sampling
Probability guided search

ABSTRACT

As a benchmark structure in the retina, the precise localization of the Optic Disc (OD) has received considerable attention from researchers owning to its importance for ophthalmic image analysis and retinopathy screening. In most literature works, although OD localization has achieved fairly good performance, rather less attention has been paid to supervised techniques. As the ODs are common in characteristics and form a fixed pattern, we propose in this paper an OD localization method based on Convolutional Neural Networks (CNNs). In data preparation, the blue channel in the fundus image, which is always empty or noisy and tends to be uninformative, is replaced by coarsely segmented vasculature maps (Red channel, Green channel and vasculatures are denoted as RGV). In addition, the input of the CNN structures is rectangular instead of square in order to involve more context information of the vasculatures. For training, a two-stage process is designed to alleviate the class-imbalance problem by choosing large number of positive samples at the first stage and choosing samples at the center of OD to fine-tune the fist-stage model at the second stage. Then a probability guided search algorithm is developed to improve the efficiency of OD localization by selectively searching patches with large chances to be an OD and their neighbors. Finally, the precise OD position is located as center of the highest probability regions after weighted voting. Extensive experiments on four publicly available databases demonstrate the effectiveness and efficiency of the proposed method, with an average detection rate of 99.11% and detection time less than 15 s per image.

1. Introduction

Optic Disc (OD), one of the principal structures in the retina, is the region of posterior pole where the blood vasculatures and nerve axons enter and leave the eye [1]. Due to the absence of pigmented epitheliums, a typical OD is bright yellowish circular or oval object and is in high contrast to the backgrounds in healthy retinal fundus images, as is shown in Fig. 1 (a). The position of the OD is of great importance in automated ophthalmic pathology diagnosis [2], for example, the OD regions must be avoided when analyzing diabetic retinopathy pathologies due to their resemblance. In addition, the OD is a reference structure in the localization of other important structures, such as the fovea. What is more, precise localization of the OD is a prerequisite operation of OD segmentation, which is essential to the diagnosis of eye diseases associated with OD, such as glaucoma, etc. The OD must be first located before segmentation to ensure accuracy. Recently, OD localization has received increasing interests from researchers and numerous works have been published.

Although the characteristics of OD in healthy retina are prominent, automatic localization of OD still remains a challenging task in abnormal images. The main difficulties come from the variety of pathologies and different capture settings, as shown in Fig. 1 (b) and (c). As a consequence, the ODs in abnormal retinal images exhibit quite different appearances comparing to normal retinal images. The resemblance of color in pathologies and ambiguity of vasculatures in OD make automated localization a serious class overlapping problem.

Extensive OD localization methods have been reported to locate ODs in retinal images and to deal with the above-mentioned problems, which can be traditionally categorized as appearance-based methods and anatomy-based methods. The first category methods focus on a local region around OD, and extracts different types of OD-specific image characteristics, such as intensity, texture and

* Corresponding author.
E-mail address: ylyin@sdu.edu.cn (Y. Yin).

https://doi.org/10.1016/j.neucom.2018.05.114
0925-2312/

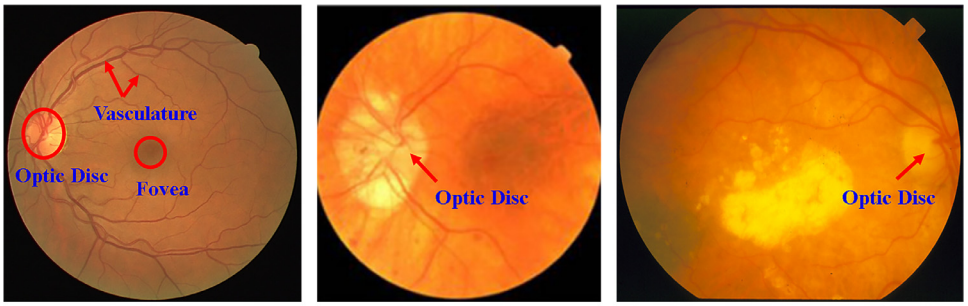

Fig. 1. Examples of different appearances of the optic disc. (a) A typical OD is bright yellowish and circular in shape with surface vessels spread; (b) and (c) are cases when eye diseases exist, the shape of OD in (b) is non-circular and pathologies similar in color with OD exist in (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

shape. For instance, in literature [3] the authors calculate the variance for each candidate sub-window, then position of maximum variance is taken as the center of OD. Osareh et al. [4] average the OD region of 25 color-normalized images as templates to detect the OD. In [5,6], Principal Component Analysis (PCA) is employed to locate optic disk. In [7], feature match is employed, and an Approximate Nearest Neighbour Field (ANNF) algorithm is proposed to find the correspondence between a chosen optic disk reference image and any given query image. In [8–10], features of OD or vasculature edges are extracted to represent the OD. Authors in [11,12] detect OD by modeling it as a circular shape. These techniques can perform well on normal retinal fundus images, but often fail when in the presence of pathological regions, the existence of confusing elements and changes of Field Of View (FOV).

The second category focuses on a more global view by making use of the anatomical structures around the OD, such as the macula and the retinal vessels. This approach tends to be the more effective and reliable, even in the presence of retinal diseases. The pioneering works are Fuzzy convergence [13] which is a voting-type algorithm developed by Hoover et al. and the one [14] model main vasculatures as parabolic course with the OD as the center in a retinal image. In [15], oriented filter is used to model the vessels. In [16], adaptive mathematical morphology is designed to detect the OD. In [17], entropy of vascular directions is calculated to assess the position of the OD. In [18], the authors determine the OD location candidates using template matching at different resolutions. Subsequently, the vessel characteristics of the OD are used to determine the final OD position. In [19], the authors utilize symmetry of main vessels to realize optic disc localization. Muangnak et al. [20] propose a new method suitable for low-quality images based on exploiting the convergence of the blood vessels to the OD.

To simultaneously combine local appearance and global anatomy, hybrid methods have also been proposed. In [21], the robustness of the method for OD localization is improved by constraining the search for maximal values of entropy to image areas with high intensities. In [22], A feature combining three vessel distribution characteristics, i.e., local vessel density, compactness, and uniformity, is designed to find possible horizontal coordinate of OD. Then, according to the global vessel direction characteristic, a general Hough Transformation is introduced to identify the vertical coordinate of OD. In [23], a cumulative sum fields based method is proposed, the final optic disc localization is determined by a vessel convergence algorithm using its two most relevant features: high vasculature convergence and high intensity values. In literature [24], the authors designed both global and local vessel models to detect the OD, the detection rates reach 100% on all involved databases.

Although good performances have been achieved, most of literature works suffer from a number of typical common limitations. First, most OD localization methods define a correct localization if the detected OD center lies within the OD boundary, they failed to report the exact center with pixel errors. However, the initial position of OD is essential in OD segmentation algorithms. Second, as a prerequisite operation, most OD localization methods are too time-consuming to be utilized in a real-time automatic diagnosis system. On the other hand, as a kind of fixed pattern, supervised pattern recognition methods have been neglected to some extent.

OD has some intrinsic characteristics, which make it a favorable pattern to be recognized using supervised pattern recognition methods. The common characteristics can be summarized as follows. First, the OD is a bright yellowish circular or oval object with surface vessel spread on it. Second, the OD center is where the vasculatures originate, at the near locality of OD there are vasculatures at multiple directions with relatively large widths. Third, globally the OD is on a direction to which the vascular bifurcation merges. These characteristics make OD different with the backgrounds.

To the best of our knowledge, supervised pattern recognition methods are not a mainstream method for OD localization, nor does it perform well on this problem. Some of the most existing methods are as follows. In [5,6], the authors utilize PCA for feature analysis and the methods are tested by private databases. Tobin et al. [25] predict the location of the optic nerve in the retina using a two-class Bayesian classifier, detection performance of 81% is reported on a personal Diabetic Retinopathy (DR) database. In [26], the authors define the problem of localizing structures in a retinal image as a regression problem, a k-Nearest Neighbor (kNN) regressor is utilized and the detection rate is 99.4% and 93% for normal and abnormal retinal images, respectively. In [27], a novel scale-embedded dictionary-based method is presented to pose the problem of OD localization as that of classification, with a detection rate of 95% on the DRIVE database. Zang et al. [28] detect the OD using AdaBoost face detectors, an average success rate of 97.8% is achieved.

Based on above-mentioned consideration, to attain the full capacity of pattern recognition methods, in this paper, we present an accurate and efficient OD localization method based on Convolutional Neural Networks (CNNs) [29]. First, the rectangular RGV (Red channel, Green channel and Vasculatues) training samples are collected in data preparation. Second, a two-stage training process is designed to alleviate the class-imbalance problem and to accurately localize the OD: at the first stage, the model is trained on a dataset with large number of positive samples around the OD; at the second stage, the model is finetuned with positive samples at the center of OD. Third, a probability guided search algorithm is developed to improve the efficiency of OD detection by selectively searching patches with large chance to be an OD and their neighbors. Finally, the precise OD position is located as center of highest probability regions after weighted voting.

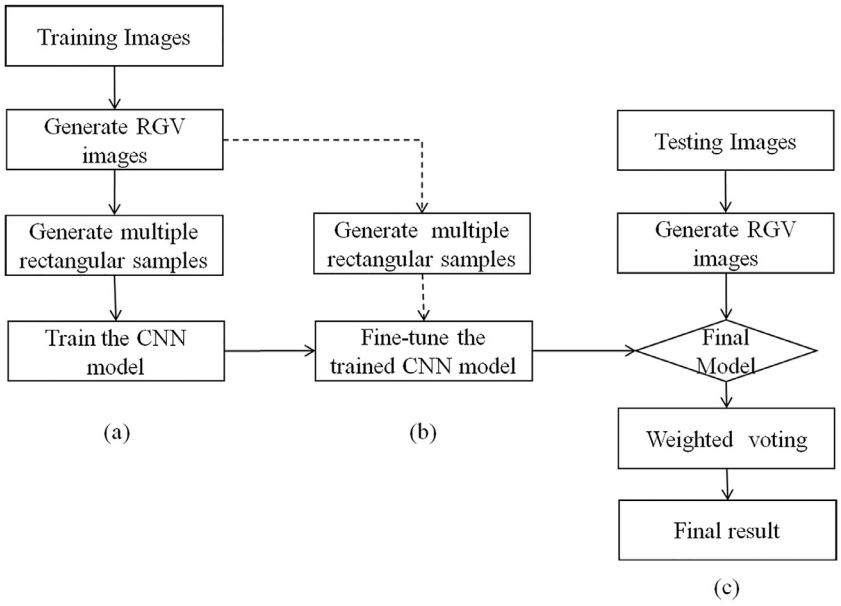

Fig. 2. Flowchart of the proposed OD localization method based on Convolutional Neural Network.The training process include 2 stages (a) and (b); (c) is the testing process.

The contributions of the proposed method are as follows: (i) we present a novel OD localization framework, of which the performance has been evaluated by extensive experiments. (ii) we design inputs of rectangular RGV regions to contain more context information, which can improve the discriminability of the following trained model. (iii) we provide a two-stage training process to alleviate the class-imbalance problem, the proposed model is first coarsely trained and then fine-tuned by samples from the very center of the ODs.

The rest of the paper is organized as follows. In Section 2, the proposed OD localization based on Convolutional Neural Network is described. In Section 3, the experimental comparisons are presented and discussed. Finally, in Section 4 conclusions are drawn and several issues for future works are indicated.

2. Methodology

In this part, the proposed OD localization method is first outlined, which is shown in Fig. 2. First, the data preparation and augmentation techniques are introduced. Second, the CNN model involved is presented, in which the two-stage training technique is adopted in order to alleviate the class imbalance problem. Then, candidate locations in a test image are determined using the probability guided search. Finally, the OD is localized as the most probable position after weighted neighborhood voting.

2.1. Data preparation and augmentation

As mentioned in the Introduction, the OD localization suffers from serious class overlapping problem. In the data preparation, we employ two techniques to alleviate this problem. First, we encode the red, green and coarsely segmented vasculature into RGB color images (denoted as RGV), as shown in Fig. 3. Second, contrast to most of the CNN models, we change the square inputs into rectangular ones to encode more discriminative information into the trained model.

In the fundus image analysis, the green and red channels are most frequently used. In the literature, the green channel usually be applied in vasculature segmentation and optic disc localization [10,13,22], the red channel tends to be saturated and usually be selected in OD or optic cup (OC) segmentation [18,27]. However, the blue channel is seldomly used independently because it tends to be empty and noisy [12]. In order to involve more information and encode the vasculature anatomy as the anatomy-based methods suggested, we replace the blue channel with coarsely segmented vessels using the algorithm proposed in our previously work [30] which is threshold based and efficient. As introduced in [31], information in different dimensions or feature maps can be encoded into an RGB-like image as inputs of the CNN models. The examples of original RGB and proposed RGV images are shown in Fig. 3.

In order to involve more morphological information of the OD, instead of selecting a square ROI region, a rectangular region is designed to better delineate the geometrical structure of the OD. The main branches of retinal vessels spread around the optic disc in a shape of star, and is large in density in the vertical direction. Rectangular inputs have also been designed in pedestrian detection, therefore different human body parts can be integrated [32]. The ratio of height and width of the rectangular region is set to 2:1. Rectangular is also adopted in [8,18], to capture more context information about the vasculatures.

In addition, detection or localization tasks suffer from typical class imbalance problems, in which the size of positive samples always too small. In order to boost the number of positive training samples, samples are selected in three scales, which is 28, 32 and 36 pixels in width, respectively. The rectangular and square ROI regions in comparison are illustrated in Fig. 3 on the RGB and RGV images, respectively. Note that, the inputs of the proposed method are all normalized to 56×28, and in order to save data preparation and localization time, the height of all the retinal images is resized to 200 pixels with the width changed correspondingly. In the evaluation process, the detected optic disc center is translated to the original image space.

2.2. CNN model and two-stage training

2.2.1. CNN architecture

We train our CNN model on a binary classification task, namely to classify an image patch to OD or non-OD class. Since the size of the input images is relatively small, the CNN model employed here need not to be too deep. Fig. 4 depicts the overall architecture of our CNN model, which contains six layers to be learned

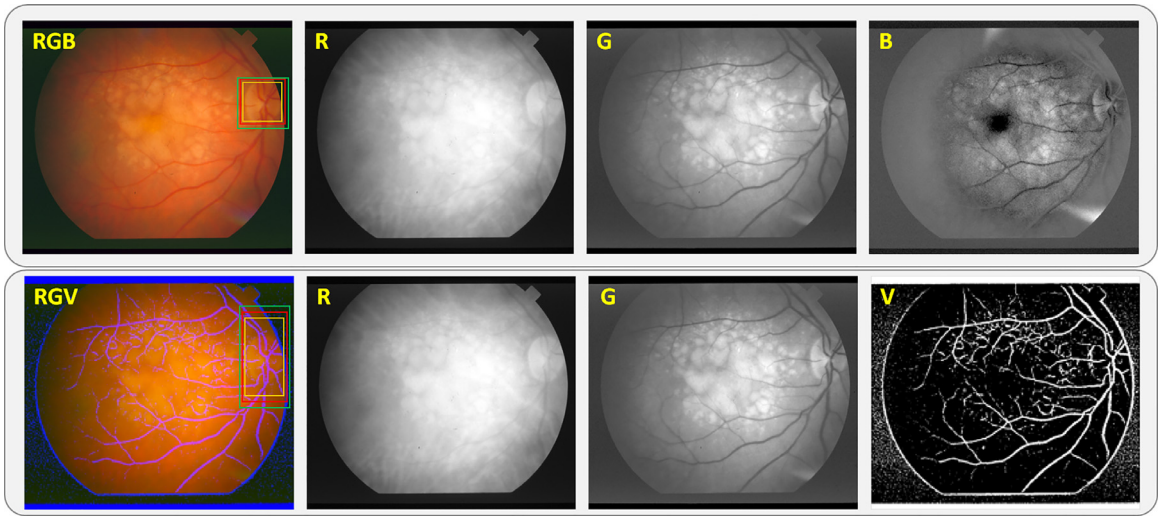

Fig. 3. The original RGB image and proposed RGV image. The first row shows the original RGB and the corresponding Red, Green and Blue channels. The second row demonstrates the proposed RGV image and the Red, Green and coarsely segmented Vasculature channels. The three sizes of square and rectangular patches are also illustrated in the RGB and RGV images, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

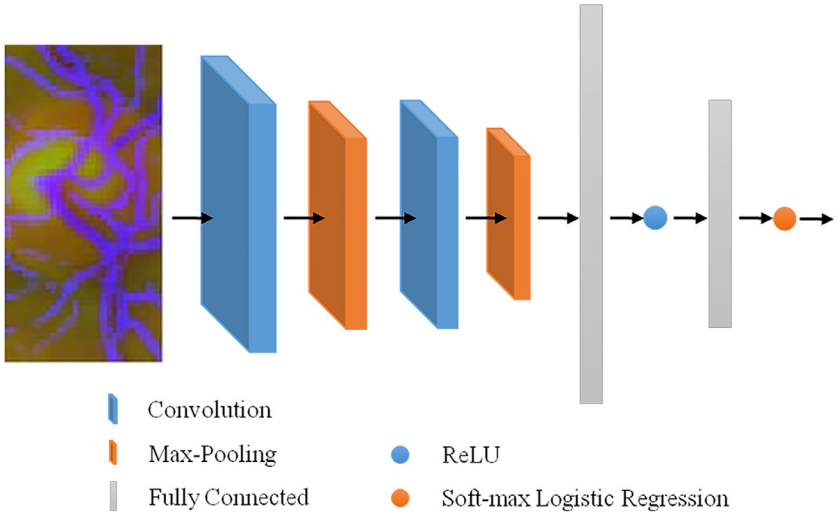

Fig. 4. Illustration of the CNN structures used in this paper. The detail configurations can be seen in Table 1.

Table 1
Architecture of the CNN model for OD localization.

Layer	Operation	Input Size	Details
Layer 1	Convolution	56×28	$5 \times 5 \times 3$, k=20
Layer 2	Max-pooling	52×24	2×2, a=2
Layer 3	Convolution	26×12	$5 \times 5 \times 20$, k=50
Layer 4	Max-pooling	22×8	2×2, a=2
Layer 5	Fully Connected	11×4	2200 nodes
Layer 6	Fully Connected	2200	500 nodes
Layer 7	Soft-max	500	2 classes

with the first fully connected layer followed by a rectified linear unit (ReLU). The architecture is a simple modification of the LeNet-5 network [33]. The modifications are as follows: First, the square gray-scale inputs are changed to rectangular RGV patches. Second, the multi-class classification output is changed to a binary classification. Hence, the feature maps generated after convolution or max-pooling are also in rectangular size. The final fully connected layer consists of 500 nodes followed by a soft-max logistic regression which outputs a score ranges from 0 to 1, we denoted as the probability to be the center of OD. The details of the proposed CNN model are listed in Table 1.

2.2.2. Two-stage training

Object detection is a typical class imbalance problem in which the number of samples of non-OD is tremendous. To simultaneously deal with the class imbalance problem and accurately locate the OD, a two-stage training strategy is adopted. The CNN model is first trained to assess the approximate candidate region of OD. In this process, to balance the OD and non-OD samples, OD samples within half of the radius is selected, with the same number of negative samples randomly selected from the background. Though the located position from the first-stage model is relatively not accurate, the class imbalance problem is solved. In the second stage, samples around the exact center of OD is selected as training examples, also with the same number of negative samples randomly selected from the background, and then the first-stage model is finetuned to locate the extract position of an OD. The detailed configurations of the two-stage training techniques are illustrated in Algorithm 1.

Algorithm 1: Pseudo-code of the two-stage training.

Input:
D represents the 1924 training images;
P is the labeled OD positions of the training images;
DB denotes the training samples.
Stage 1: $DB \leftarrow \Phi$
for *every image i in D* **do**
  for $w = -7 : 2 : 7$ **do**
    for $h = -7 : 2 : 7$ **do**
      if $(w^2 + h^2) \leq 49$ **then**
        Add samples centered at $(P_{i,1} + w, P_{i,2} + h)$ in three scales to DB with label 1;
        Randomly selected same number of samples in three scales from the background and add to DB with label 0;
Learning rate $\leftarrow$ 0.001; epoch $\leftarrow$ 30; batch size $\leftarrow$ 64;
Random initialize the CNN model;
Train the CNN model;
Output: the first-stage CNN model.
Stage 2: $DB \leftarrow \Phi$
for *every image i in D* **do**
  for $w = -2 : 1 : 2$ **do**
    for $h = -2 : 1 : 2$ **do**
      if $(w^2 + h^2) \leq 4$ **then**
        Add samples centered at $(P_{i,1} + w, P_{i,2} + h)$ in three scales to DB with label 1;
        Randomly selected same number of samples in three scales from the background and add to DB with label 0;
Learning rate $\leftarrow$ 0.0001; epoch $\leftarrow$ 10; batch size $\leftarrow$ 64;
Fine-tune the first-stage CNN model;
Output: the second-stage CNN model.

2.3. Probability guided search

In most cases, the OD candidate regions searched are backgrounds, but there is resemblance in neighborhood pixels because there will be large overlaps among the corresponding search windows. According to the above-mentioned considerations, a probability guided search is proposed. In this technique, the density of candidate OD regions is assumed to be consistent with their corresponding probabilities to be an OD. Hence when testing, the test image is first searched with a large step, and then regions around large probability candidates are thoroughly tested.

As the OD only accounted for a small fraction of the whole fundus image, the time efficiency of the proposed method can be boosted to almost $1/d^2$ of the original time consumption, d is the step size for searching. The effectiveness of probability guided search (PGS) is analyzed in Section 3.2.4. The pseudo code of the proposed techniques is presented in Algorithm 2 , where the probability threshold is set to 0.01.

2.4. Weighted voting

To robustly determine the precise position of the OD, the final position of the OD is located using the proposed neighborhood weighted voting, it is based on the observation that the center of each candidate region is overlapped by many around regions. And if the neighborhood regions are closer, the contributions of them are greater. The weighted voting is performed by convolution filter with Gaussian template, of which the size is fixed as 5×5 with standard deviation $\sigma = 1$. The voting operation is defined as follows:

Algorithm 2: Pseudo-code describing the PGS approach.

Input:
testimage: the test retinal image;
net: the trained CNN model;
datamean: the mean value of all the training images;
d: the step size for searching;
threshold: the minimum probability to be an OD.
$W \leftarrow$ width of the test image; $H \leftarrow$ height of the test image;
$P \leftarrow \Phi$; $p \leftarrow 0$;
while $j \leq (H - 14)$ **do**
  $j = j + d$
  while $i \leq (W - 14)$ **do**
    $i = i + 1$
    $p \leftarrow$ net.predict(patch centered at (i, j) in testimage, net, datamean)
    if $p > threshold$ **then**
      $P_{i,j} \leftarrow p$
      for $k \in \{j - d/2, \ldots, j + d/2\}$ **do**
        $P_{i,k} \leftarrow$ net.predict (patch centered at (i, k) in testimage, net, datamean)
    else
      $i = i + d$
Output: P

$$pmap' = pmap * G(x, y, \sigma) \tag{1}$$

$$P = arg\ max(pmap') \tag{2}$$

here, $pmap$ stands for the probability map of an image, $pmap'$ indicates the probability map after weighted voting. Eq. (2) reveals that the final OD center P is located as the position with the highest probability to be an OD.

3. Experiments

3.1. Experimental materials

The training set is collected from the public available databases, other than DIARETDB0 (DB0) [34], DIAREIDB1 (DB1) [35], STARE [36] and DRIVE [37] which are used as the testing set. Consequently, the number of images in training set and testing set are 1924 and 340, respectively. The information of the four databases used in testing is listed in Table 2.

The benchmark position of all the OD center in training set is manually labeled by an expert with experiences of more than three years. The position of the testing databases is manually labelled by an independent expert.

3.2. Evaluation

In this part, the performance of the proposed method is first reported. Then, the effects of data preparation techniques, two-stage training and time efficiency are analyzed, respectively. Finally, the proposed method is empirically compared with the existing methods.

The proposed method is evaluated by localization rate and pixel error. The localization is regarded as correct when the detected center is within half of the OD radius, which is traditionally adopted in OD localization. The localization rate is the ratio of correctly detected cases against all the tested cases. The pixel error is calculated by the following equation:

$$pixel_error = \sqrt{(p_x - p_x')^2 + (p_y - p_y')^2} \tag{3}$$

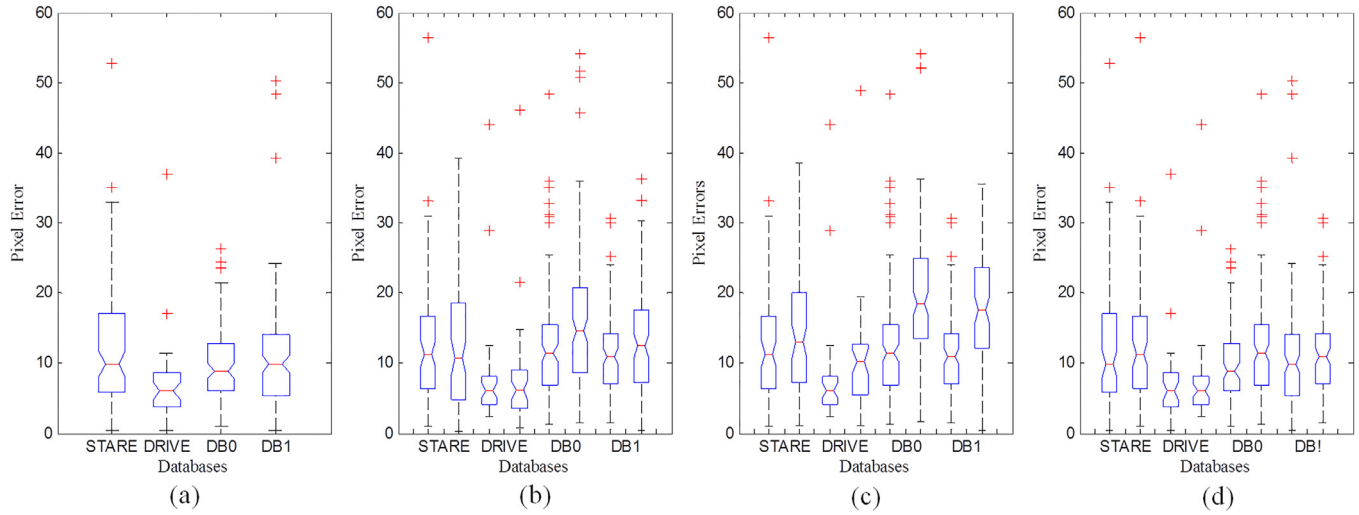

Fig. 5. The pixel errors analysis. (a) The performance of the proposed method. (b) Performance on four databases when RGV compare with RGB inputs. (c) Performance when Rectangular in comparison with square inputs. (d) The performance comparison of the first-stage and second-stage model.

Table 2
Details of the testing databases of the proposed method.

Database	DRIVE	DB0	DB1	STARE
Input size	565×584	1500×1152	1500×1152	700×605
Fundus camera	Canon CR5	unknown	Nikon F5	TopCon TRV-50
Field of view	45°	50°	50°	35°
Normal images	33	20	5	31
Diseased images	7	110	84	50
Total images	40	130	89	81

Table 3
Performance of the proposed method.

Database	Location rate (%)	Pixel error
STARE	98.77	$12.83(\pm 10.97)$
DRIVE	100	$6.97(\pm 5.91)$
DB0	98.46	$10.90(\pm 7.65)$
DB1	100	$10.01(\pm 5.79)$

here, (p_x, p_y) represents the detected optic disk center, $(p_x^{'}, p_y^{'})$ denotes the optic disk location labeled by experts.

3.2.1. Performance of the proposed method

The results of the proposed method are tabulated in Table 3 with Fig. 5 (a) exhibits the results in pixel errors by box-plots for better visualization and presentation. The localization rates are 98.77%, 100%, 98.46% and 100% respectively on the four databases, which is 99.12% (337/340) on all the tested images with 1 erroneous detection in the STARE database and 2 erroneous detection in the DB0 database, and is a relatively high performance. In Fig. 6, we give some example images in which the ODs are correctly localized. It can be seen that the proposed method can perform well on most of the notably hard cases with quite impressive OD centers located. The incorrect localization and hard samples can be seen in Fig. 7, with the probability maps attached below. Note that, the maximum probability to be an OD center of case (b) and (c) in Fig. 7 are just 0.06 and 0.13, respectively, which indicates that though erroneous correction may occur in our method, the extreme low probability of an image containing OD may feed it back to the researchers or medical workers for correction measures. Moreover, there are similarities among the erroneous images, i.e., the pathology regions may produce vascular-like responses or the image quality is not satisfactory. Since the purpose of the experiments is to explore the general performance of the proposed method, so here we have not spent much time in trying to use front-end processings to deal with the quality problems. Although doing so might still improve the performance. The average pixel errors are evaluated on images with OD correctly detected and on the STARE, DRIVE, DB0 and DB1 databases are 12.83, 6.97, 10.90 and 10.01, respectively, with the average standard deviations ± 10.97, ± 5.91, ± 7.65 and ± 5.79 in respective order. Fig. 5 (a) also reveals that most of the pixel errors are within 30 pixels, which indicates that most detected OD centers are within half of the OD diameters and accurate.

3.2.2. Analysis of data preparation techniques

In order to study whether the rectangular RGV channels employed in the proposed method is beneficial or not, two groups of comparisons are conducted in this part, which are RGV compared with RGB inputs and rectangular compared with square inputs, respectively, each comparison with all the other experimental configuration fixed. In these two comparisons, the results are all from the first-stage training models since we want to see whether the OD localization performance will be improved by the two data preparation techniques.

The detection rate and pixel errors of RGV inputs compared with RGB ones are tabulated in Table 4, with the box-plot depicted in Fig. 5 (b). It is obvious that the performance when using RGV channels is better than using original RGB images, with much lower average pixel errors and standard deviations on the four test databases. Typically, the localization rate of RGV images on the STARE is 7.41% higher than the RGB ones, which verifies the effectiveness of vasculature information. The undesirable results of RGB inputs may be caused by the pathology regions in the images, which is similar in RGB values with the OD region. There are also two localization errors of the RGV based approaches on the DB0 database. It is because the unbalanced illumination of these

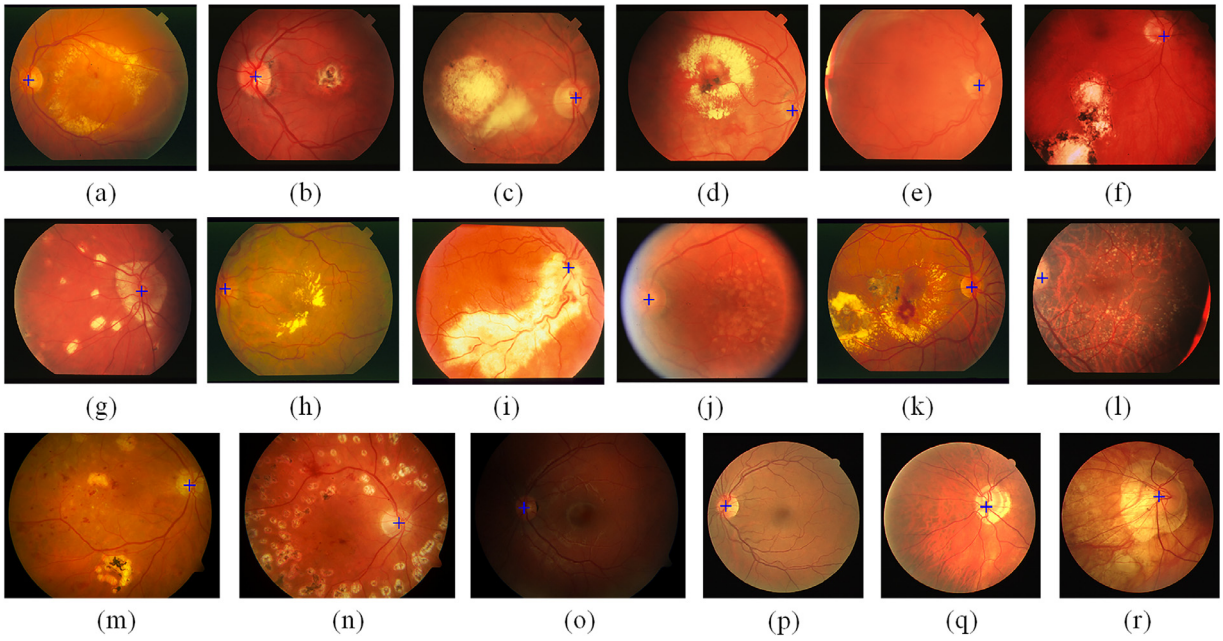

Fig. 6. Examples of correct OD localization. Images (a) to (l) are from the STARE database, (m) and (n) are from the DB0 database, (o) is from DB1, and the last three images are from the DRIVE database.

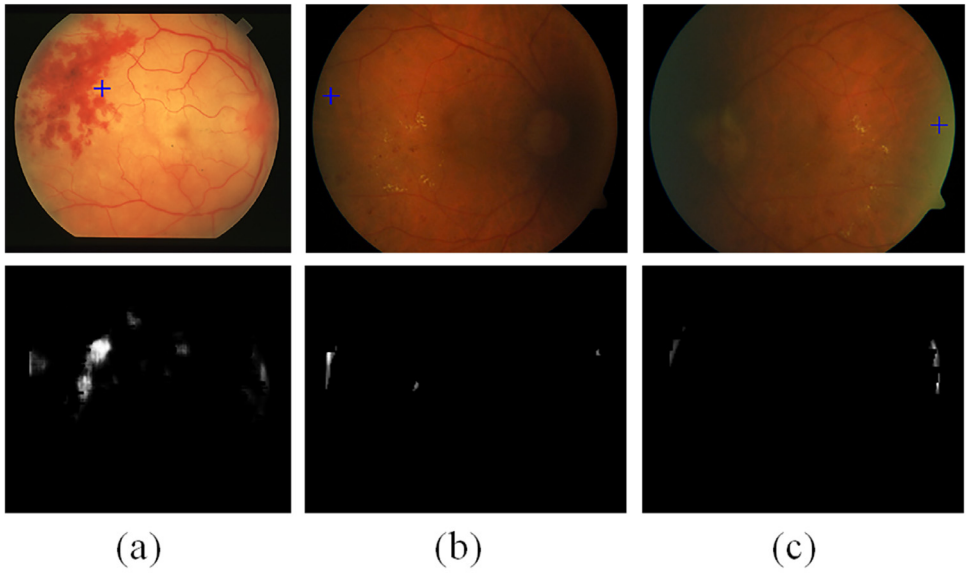

Fig. 7. The three erroneous cases of the proposed method, (a) is from the STARE database, (b) and (c) are from the DB0 database. The probability map is normalized to 0 to 255 for better visualization.

Table 4
Performance analysis: RGV compare with RGB inputs.

Database	Location rate (%)	Pixel error
STARE	97.53	12.68(±9.11)
	90.12	14.26(±13.53)
DRIVE	100	7.68(±7.39)
	100	7.76(±7.46)
DB0	98.46	12.99(±9.79)
	100	15.93(±9.86)
DB1	100	11.48(±6.11)
	100	13.16(±8.08)

Table 5
Performance analysis: rectangular compare with square inputs.

Database	Location rate (%)	Pixel error
STARE	97.53	12.68(±9.11)
	93.83	15.14(±11.77)
DRIVE	100	7.68(±7.39)
	100	10.45(±8.00)
DB0	98.46	12.99(±9.79)
	98.46	19.84(±11.09)
DB1	100	11.47(±6.11)
	100	17.57(±8.35)

images, the OD boundary and vasculatures are largely corrupted, which is already discussed in Section 3.2.1.

Inspired by the anatomy-based methods, the rectangular RGV inputs are used for the CNN model. The effects of rectangular region compared to square ones are shown in Table 5 and Fig. 5 (c),

Fig. 8. An example of two-stage training. (a) original image. (b) Probability map generated by first-stage model. (c) Probability map generated by second-stage model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Performance analysis: Fine-tune vs. direct training.

Database	Location rate (%)	Pixel error
STARE	98.77	12.10/12.83(±8.88/10.97)
	97.53	12.68(±9.11)
DRIVE	100	6.97(±7.62)
	100	7.68(±7.39)
DB0	98.46	11.20(±7.38)
	98.46	12.99(±9.78)
DB1	100	10.27(±5.97)
	100	11.47(±6.11)

respectively. From the results, it can be figured out that results on rectangular regions are generally lower in pixel errors and higher in localization rate. The localization rate with rectangular regions on the STARE database is 3.7% higher than with square ones, which demonstrates the helpfulness of anatomy information.

In summary, the observations reported in this subsection suggested that vasculatures and more context information can improve the performance of OD localization, especially when retinal images suffer from severe pathologies.

3.2.3. Analysis of two-stage training

The performance of the two-stage training, which aims to alleviate the class imbalance problem and achieve high localization accuracies is analyzed in this part. The results of using samples from the extract center is not detailed here, the location rate on the STARE database is around 92.59%, which is unsatisfactory. The results of first-stage and second-stage model are presented in Table 6 and Fig. 5 (d), respectively. It can be found from the results that the localization rate and pixel errors from the first-stage model are acceptable, and are further improved after fine-tuning in the second stage. Two average pixel errors are listed for the proposed method on the STARE database which is before and after localization rate improvement to better describe the improvement of average pixel errors. An example with probability maps for the two-stage training is given in Fig. 8, which is consistent with and confirms the above-mentioned results. It can be found that the probability map for OD and pathology region in the first stage is relatively large, but in the second stage, probability of pathologies is apparently reduced with the candidate region of OD concentrating much more to the center. Fig. 5 (d) also show that although the average pixel error is improved in the DB1 database, the abnormal cases in DB1 is worse after fine-tuning. This may be caused by the image quality problems mentioned in Section 3.2.1.

3.2.4. Efficiency analysis

Reducing time consumption of automatic diagnostic system is important for improving the work efficiency for medical workers. In this part, the time efficiency of the proposed method is analyzed with the average time consumption per image illustrated in Fig. 9. Since the probability guided search finds the candidate regions of OD by introducing a step size d, the time cost will apparently be improved and assumed to be $1/d^2$ of the original time cost. It can be found in Fig. 9 and positions searched illustrated in Fig. 10 that when step size is small the average time cost decreases exponentially, while when step size is large and the OD region counts for a large fraction of searched positions, the average time cost decrease is not that obvious. Note that probability map in the OD sub region depicted in Fig. 10 also reveals that the proposed search technique will not degrade the result of detection.

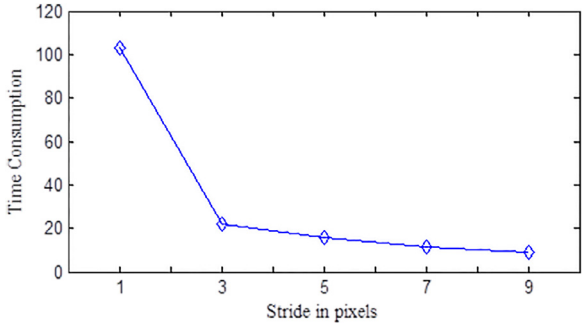

Fig. 9. Average time consumption varies according to step size d.

The above-mentioned experiments are implemented on a personal PC with the NVIDIA GeForce GTX 745 GPU. The model is trained by Caffe, and the localization process are implemented using Python. The time consumption is around 15 s with $d = 5$. The time consumption may be further reduced by fine-tuning the candidate OD to a much smaller region or by adopting parallel computing or a much more powerful GPU.

3.3. Comparison with existing methods

The proposed method is finally compared with state-of-the-art methods on the four public databases. Two evaluation protocols are used, including localization/detection rates (DR) and average pixel errors (PE). The quantitative results of three kind of different methods are tabulated in Table 7, in which the proposed method is categorized as appearance-based method. It can be seen from Table 7 that the proposed method achieves comparable detection rates and lowest pixel errors. For the DRIVE database, since the pathologies interfering with the OD localization are relatively small, the detection rates are generally high with only one appearance-based method achieved a detection rate of 97.50%. For databases DB0 and DB1, the issues related to OD are still uncomplicated, 7 out of 10 methods have achieved acceptable results with not more than three error images. Our method has only 2 errors on

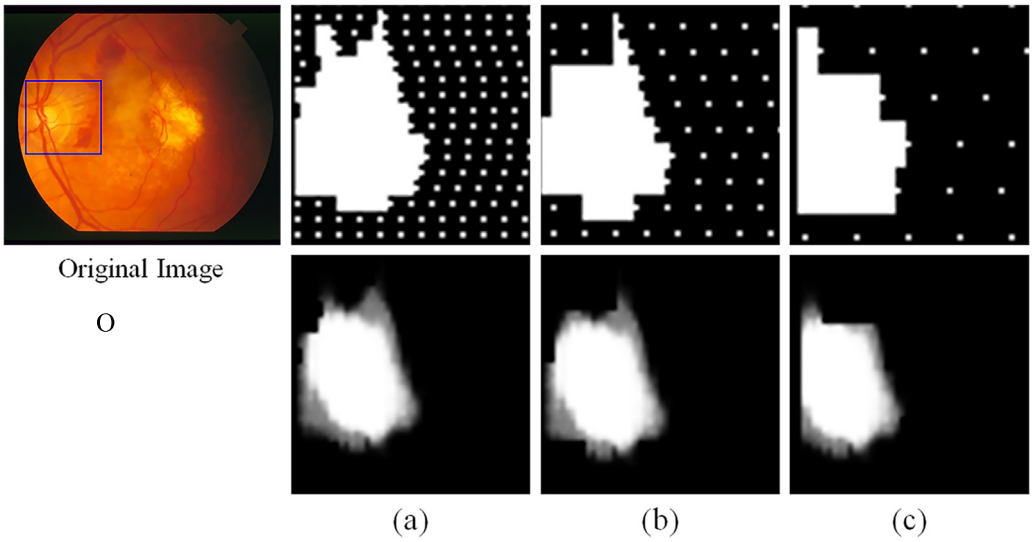

Fig. 10. Examples of probability guided search, the first row indicates the detected region centers, the second row is the corresponding probability maps, (a) (b) (c) correspond to $d = 3, 5, 7,$ respectively.

Table 7
Performance compared with existing methods, DR denotes detection rate, PE stands for pixel error.

Type	Method	STARE		DRIVE		DB0		DB1	
		DR (%)	PE	DR (%)	PE	DR (%)	PE	DR (%)	PE
Hybrid	Wu et al. [24]	100	–	100	–	100	–	100	–
-based	Soares et al. [23]	98.77	–	100	–	100	–	100	–
Methods	Zhang and Zhao [22]	98.77	18.28	100	14.33	98.46	23.51	98.88	24.08
Anatomy	Welfer et al. [16]	97.53	–	100	–	–	–	97.75	–
-based	Youssif et al. [15]	98.77	26.00	100	17.00	–	–	–	–
	Our method	98.77	12.83	100	6.97	98.46	11.20	100	10.27
Appearance	Ramakanth and Babu [7]	93.83	–	100	–	98.46	–	98.88	–
-based	Lu and Lim [9]	96.30	25.00	97.50	–	99.23	–	98.88	–
Methods	Mahfouz and Fahmy [8]	92.59	14.00	100	11.00	98.46	–	97.75	–
	Hoover and Goldbaum [13]	88.89	–	100	–	–	–	97.75	–

- means the data item is not provided.

these two databases, the most probably reason is the low quality of these images, we believe the problem can be solved by adding front-end image enhancement operations. The STARE database is the most complicated database, except the hybrid method in literature [24], the detection rates of most of the existing methods have decreased. The proposed method has only one error detection. The average pixel errors of the proposed method on the four databases are 12.83, 6.97, 11.20 and 10.27, respectively. Note that, all the pixel errors are absolute values. Since the image size in databases DB0 and DB1 is large, it is obvious that the localization of the proposed method is relatively accurate.

4. Conclusion and future work

Optic disk is where the blood vasculatures and axons converge and leave the eye. The precise localization of its center is of great importance in the diagnosis of OD-related diseases, such as nerve head atrophy and glaucoma. This paper presents an efficient and accurate OD localization method based on one of the most popular image processing models CNN. In the data preparation, RGV channels and rectangular regions are employed for inputs of the CNN model. Experiments results in Tables 4, 5 and Fig. 5 (b) and (c) demonstrate the effectiveness of these techniques. Both the localization rates and accuracies of OD center are significantly improved on the four databases. During model training, a two-stage training process is designed to deal with the class imbalance problem and to accurately detect the OD. Experiment results in Table 6 indicate the effectiveness of the proposed training method, the detection rate of the proposed method is comparable to the existing methods. On the other hand, the second stage model achieves a much more lower localization pixel error. The best average time consumption per image is less than 15 s under the probability guided search. The time efficiency and localization accuracy imply the applicability of the proposed method.

The proposed method is designed for OD localization, region-based CNN model is deployed limited by the existing datasets. We expect even lager improvements on automatic retinopathy diagnosis on much larger-scale training sets and more complicated CNN networks in the future.

Acknowledgment

This work is supported by the NSFC Joint Fund with Guangdong of China under Key Project (Grant No. U1201258), the Natural Science Foundation of China (Grant No. 61573219, 61701280, 61703235), the Natural Science Foundation of Shandong Province (Grant No. ZR2018BF012, ZR2016FQ18), the Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions, Shandong Provincial Key Research and Development Plan (2017CXGC1504) and the Key Research and Development Project of Shandong Province under Grant 2018GGX101032.

The authors would like to thank the researchers who made the public databases available.

References

[1] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, L. Kennedy, Optic nerve head segmentation, IEEE Trans. Med. Imaging 23 (2) (2004) 256–264.
[2] M.R.K. Mookiah, U.R. Acharya, C.K. Chua, C.M. Lim, E.Y.K. Ng, A. Laude, Computer-aided diagnosis of diabetic retinopathy: a review, Comput. Biol. Med. 43 (12) (2013) 2136–2155.
[3] C. Sinthanayothin, J.F. Boyce, H.L. Cook, T.H. Williamson, Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images, Br. J. Ophthalmol. 83 (8) (1999) 902–910.
[4] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, Comparison of colour spaces for optic disc localisation in retinal images, in: Proceedings of the Sixteenth International Conference on Pattern Recognition, 2002, p. 10743.
[5] H. Li, O. Chutatape, Automatic location of optic disk in retinal images, in: Proceedings of the International Conference on Image Processing, 2001, pp. 837–840. Vol.2.
[6] H. Li, O. Chutatape, Automated feature extraction in color retinal images by a model based approach, IEEE Trans. Biomed. Eng. 51 (2) (2004) 246–254.
[7] S.A. Ramakanth, R.V. Babu, Approximate nearest neighbour field based optic disk detection, Comput. Med. Imaging Gr. Off. J. Comput. Med. Imaging Soc. 38 (1) (2013) 49–56.
[8] A.E. Mahfouz, A.S. Fahmy, Fast localization of the optic disc using projection of image features, IEEE Trans. Image Process. 19 (12) (2010) 3285–3289.
[9] S. Lu, J.H. Lim, Automatic optic disc detection from retinal images by a line operator, IEEE Trans. Biomed. Eng. 58 (1) (2011) 88–94.
[10] M. Lalonde, M. Beaulieu, L. Gagnon, Fast and robust optic disc detection using pyramidal decomposition and Hausdorff-based template matching, IEEE Trans. Med. Imaging 20 (11) (2001) 1193.
[11] X. Zhu, R.M. Rangayyan, Detection of the optic disc in images of the retina using the hough transform, in: Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, pp. 3546–3549.
[12] S. Lu, Accurate and efficient optic disc detection and segmentation by a circular transformation, IEEE Trans. Med. Imaging 30 (12) (2011) 33–2126.
[13] A. Hoover, M. Goldbaum, Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels, IEEE Trans. Med. Imaging 22 (8) (2003) 951–958.
[14] M. Foracchia, E. Grisan, A. Ruggeri, Detection of optic disc in retinal images by means of a geometrical model of vessel structure, IEEE Trans. Med. Imaging 23 (10) (2004) 1189–1195.
[15] A.R. Youssif, A.Z. Ghalwash, A.R. Ghoneim, Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter, IEEE Trans. Med. Imaging 27 (1) (2008) 8–11.
[16] D. Welfer, J. Scharcanski, C.M. Kitamura, M.M.D. Pizzol, L.W. Ludwig, D.R. Marinho, Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach, Comput. Biol. Med. 40 (2) (2010) 124–137.
[17] A.M. Mendonca, F. Cardoso, A.V. Sousa, A. Campilho, Automatic Localization of the Optic Disc in Retinal Images Based on the Entropy of Vascular Directions, International Conference Image Analysis and Recognition, Springer, 2012, pp. 424–431.
[18] H. Yu, E.S. Barriga, C. Agurto, S. Echegaray, M.S. Pattichis, W. Bauman, P. Soliz, Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets, IEEE Trans. Inf. Technol. Biomed. A Publ. IEEE Eng. Med. Biol. Soc. 16 (4) (2012) 644–657.
[19] A. Giachetti, L. Ballerini, E. Trucco, Accurate and reliable segmentation of the optic disc in digital fundus images, J. Med. Imaging 1 (2) (2014) 024001.
[20] N. Muangnak, P. Aimmanee, S. Makhanov, B. Uyyanonvara, Vessel transform for automatic optic disk detection in retinal images, IET Image Process. 9 (9) (2015) 743–750.
[21] A.M. Mendonça, A. Sousa, L. Mendonça, A. Campilho, Automatic localization of the optic disc by combining vascular and intensity information, Comput. Med. Imaging Gr. Off. J. Comput. Med. Imaging Soc. 37 (5-6) (2013) 409–417.
[22] D. Zhang, Y. Zhao, Novel accurate and fast optic disc detection in retinal images with vessel distribution and directional characteristics, IEEE. J .Biomed. Health. Informat. 20 (1) (2014) 333–342.
[23] I. Soares, M. Castelo-Branco, A.M.G. Pinheiro, Optic disc localization in retinal images based on cumulative sum fields, IEEE J. Biomed. Health Inf. 20 (2) (2016) 574–585.
[24] X. Wu, B. Dai, W. Bu, Optic disc localization using directional models, IEEE Trans. Image Process. 25 (9) (2016) 1.
[25] K.W. Tobin Jr, E. Chaum, V.P. Govindasamy, T.P. Karnowski, O. Sezer, Characterization of the optic disc in retinal imagery using a probabilistic approach, Med. Imaging Int. Soc. Opt. Photonics (2006). 61 443F–61 443F.
[26] M. Niemeijer, M.D. Abràmoff, B.V. Ginneken, Fast detection of the optic disc and fovea in color fundus photographs, Med. Image Anal. 13 (6) (2009) 859–870.
[27] N. Sinha, R.V. Babu, Optic disk localization using l1 minimization, in: Proceedings of the Nineteenth IEEE International Conference on Image Processing (ICIP), IEEE, 2012, pp. 2829–2832.
[28] P. Zang, B. Wei, Y. Xu, Y. Yin, Q. Zhang, X. Meng, Optic disc detection based on classification for retinal image, in: Proceedings of the Seventh International Conference on Biomedical Engineering and Informatics (BMEI), IEEE, 2014, pp. 890–895.
[29] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Netw. Off. J. Int. Neural Netw. Soc. 61 (2015) 85–117.
[30] X. Meng, Y. Yin, G. Yang, Z. Han, X. Yan, A framework for retinal vasculature segmentation based on matched filters, Biomed. Eng. Online 14 (1) (2015) 94.
[31] H.C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Trans. Med. Imaging 35 (5) (2016) 1285–1298.
[32] W. Ouyang, X. Wang, Joint deep learning for pedestrian detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 2056–2063.
[33] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[34] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, J. Pietilä, Diaretdb0: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms, Machine Vision and Pattern Recognition Research Group, 73, Lappeenranta University of Technology, Finland, 2006.
[35] R.V.J.P.H. Kälviäinen, H. Uusitalo, Diaretdb1 diabetic retinopathy database and evaluation protocol, Med. Image Underst. Anal. 2007 (2007) 61.
[36] A. Hoover, M. Goldbaum, Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels, IEEE Trans. Med. Imaging 22 (8) (2003) 951–958.
[37] J. Staal, M.D. Abràmoff, M. Niemeijer, M.A. Viergever, B.V. Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Trans. Med. Imaging 23 (4) (2004) 501–509.

Xianjing Meng received her Ph.D. degree from Shandong University, Jinan, China in 2016. She is now the lecturer of Shandong University of Finance and Economics. Her research interests include machine learning, medical image analysis and biometrics.

Xiaoming Xi is the lecturer of Shandong University of Finance and Economics. He received his Ph.D. degree in 2015 from Shandong University. His research interests include machine learning, data mining, and medical image analysis.

Lu Yang is the lecturer of Shandong University of Finance and Economics. He received his Ph.D. degree in 2016 from Shandong University. His research interests include machine learning, data mining, and pattern recognition.

Guang Zhang received his M.S. degree from the School of Computer Science and Technology, Shandong University, Jinan, China, in 2009. He is now the director of the medial examination center in Qianfoshan hospital. His main research interests are machine learning and medial image analysis.

Yilong Yin received the Ph.D. degree from Jilin University, Changchun, China, in 2000. He is the Director of the Machine Learning and Applications Group and a Professor with Shandong University, Jinan, China. From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Electronic Science and Engineering, Nanjing University, Nanjing, China. His research interests include machine learning, data mining, and biometrics.

Xinjian Chen received the Ph.D. degree from the Center for Biometrics and Security Research, Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2006. After completing the graduation, he joined Microsoft Research Asia, where he was involved in research on handwriting recognition. From 2008 to 2012, he conducted the Post-Doctoral Research at several prestigious groups, such as the Medical Image Processing Group, University of Pennsylvania, the Department of Radiology and Image Sciences, National Institutes of Health, and the Department of Electrical and Computer Engineering, University of Iowa. In 2012, he joined the School of Electrical and Information Engineering, Soochow University, as a Full Professor. He has authored over 70 high quality international journal/conference papers. His research interests include medical image processing and analysis, pattern recognition, machine learning, and their applications. He is a Distinguished Professor with Soochow University, and serves as the Director of the University Level Laboratory-Medical Image Processing, Analysis and Visualization Laboratory.