Automatic Optic Disc Abnormality Detection in Fundus Images: A Deep Learning Approach

Hanan S. Alghamdi[12], Hongying Lilian Tang[2] , Saad A. Waheeb[3], and Tunde Peto[4]

[1] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[2] Department of Computer Science, University of Surrey, Guildford, United Kingdom
[3] King Faisal Specialist Hospital, Jeddah, Saudi Arabia
[4] National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom

Abstract. Optic disc (OD) is a key structure in retinal images. It serves as an indicator to detect various diseases such as glaucoma and changes related to new vessel formation on the OD in diabetic retinopathy (DR) or retinal vein occlusion. OD is also essential to locate structures such as the macula and the main vascular arcade. Most existing methods for OD localization are rule-based, either exploiting the OD appearance properties or the spatial relationship between the OD and the main vascular arcade. The detection of OD abnormalities has been performed through the detection of lesions such as hemorrhaeges or through measuring cup to disc ratio. Thus these methods result in complex and inflexible image analysis algorithms limiting their applicability to large image sets obtained either in epidemiological studies or in screening for retinal or optic nerve diseases. In this paper, we propose an end-to-end supervised model for OD abnormality detection. The most informative features of the OD are learned directly from retinal images and are adapted to the dataset at hand. Our experimental results validated the effectiveness of this current approach and showed its potential application.

Keywords: Fundus Images, Optic Disc, Cascade Classifiers, Deep Learning, Convolutional Neural Networks

1   Introduction

Automated fundus image analysis received a lot of attention lately. The main goal of this approach is to enable accurate and timely analysis of large-scale image sets, such as those obtained either in DR screening programmes or in eye epidemiology studies. The identification of landmark features such as the OD is essential before analysing an image for pathological signs. During the past few years, several image analysis techniques have been proposed to detect OD [4, 5]. Early methods exploited the appearance characteristics of the OD such as shape and brightness. OD is relatively bright and circular in shape in a normal

<div align="center">

(a)   (b)   (c)   (d)   (e)   (f)   (g)

</div>
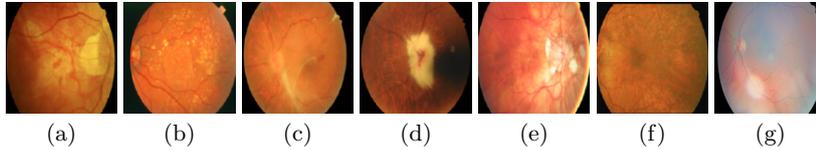
Fig. 1: Variable appearances of optic disc in fundus images

retina, however, might take different forms in pathological cases. These early methods were simple, fast and performed well in normal and high quality images [17, 2]. However, in large-scale image sets these methods might fail due to the variety of OD appearance, the presence of artefacts, pathological changes mimicking OD-like features and the bright lesions such as exudates or drusen. As blood vessels in the retina arise from the OD, more robust methods suggested the use of vessel structure in order to locate the OD. These methods also take advantage of the unique curved shape vascular arcade arising from the OD and encircling the macula in parabolic like curves [5]. The reported results show that vasculature-based methods are more robust in pathological images [2]. Complexity, computational cost and dependency on vessel segmentation are major disadvantages associated with these methods limiting their scalability. Moreover, several of these methods have only been tested on small datasets not containing sufficient variety of images to represent the wide range of ODs found in large-scale studies and screening programmes. Figure 1 shows examples of fundus images acquired during DR screening. In some images, OD might not appear as a bright or a complete circular object, and it could be partially invisible. It might be also impractical to take advantage of the main vascular arcade as it is not evident, covered or has been pathologically changed. Therefore, it is hard to define effective and reliable rules for OD characteristics, which emphasizes the need for an abstract and high-level understanding of OD features. The detection of OD abnormality indicates the presence of a different disease, not the one DR screening is set out to find. Glaucoma affects the optic disc and the area surrounding it. Cup to disc ratio measurement (CDR) is commonly used to indicate suspicion for glaucoma and to monitor the disease once diagnosed. In addition, the detection of veins and arteries being displaced towards the edge of the disc, peripapillary atrophy and optic disc hemorrhaeges are also helpful signs for determining abnormality. When these are found, the patient is referred for assessment of other signs and symptoms of glaucoma before the final diagnosis is made. In this work, we propose an end-to-end learning based model for OD localization and abnormality detection without adopting any prior assumptions. The ability to perform accurately on previously unseen data, known as generalization, is one the main advantages of deep learning models over rule-based or other machine learning models trained with handcrafted features. Convolutional neural networks (CNNs) constitute one such class of models that learn features in the form of convolutional filters. CNNs are now the dominant approach for effective and robust image classification in large data sets [22]. Our proposed

model constitutes of two successive deep learning architectures. The OD global features were learned through efficient multiple classifier and deep CNNs to extract the OD regions. These extracted regions were then passed to a second layer optimized deep CNNs to learn abnormality feature representations. The rest of the paper is organized as follows. In Section 2 we presents relevant previous literature in OD abnormality detection. Section 3 explains our proposed approach. Experimental evaluation is described in section 4. The conclusions are presented in section 5.

2   Related Work

Currently, the most common approach for optic disc abnormality detection aims to estimate the cup to disc ratio (CDR). The cup intensity is the primary feature used by many authors. In [13] the CDR was measured by means of vertical profiles obtained around the center of the OD and smoothed to calculate the cup edge by thresholding. In [14] CDR and ISNT ratio were estimated and passed to SVM classifier. In [15] CDR, ISNT ratio and the distance between optic disc center and optic nerve head features were estimated and passed to a feed-forward neural network classifier. The main disadvantage of such methods is their dependency on cup intensity, which may be vague and indefinite in many images due to the pathological changes that might surround the disc and the fuzzy cup-disc and disc-retina boundaries. Some suggested to exploit more characteristics such as vessel kink integration [17], superpixels features [16, 12], and pixel-level segmentation [18, 21]. While these approaches are more robust than the intensity-based ones, they all focused on estimating CDR to diagnose OD-s showing one sign of potential glaucoma. However, CDR is not the only consideration by many ophthalmologists to determine the abnormalities and there is a significant disagreement in estimating the CDR even between expert human graders [19]. The proposed approach aims to address all above-mentioned limitations by dynamically learning discriminative OD features from the data.

3   Proposed Methodology

3.1   Preprocessing

Before training every image in the training set is normalized by subtracting the mean image and dividing by the standard deviation. Dataset whitening is then conducted to improve the performance. Whitening normalizes the features scale and makes them easier to model [9]. Figure 2 shows the result of the pre-processing stage of OD and non-OD candidate regions. The mean image extracted from the training set is shown in Figure2-g.

3.2   Model Overview

Our approach consists of two successive learning based models for OD localization and abnormality assessment. The overall architecture of our proposed

Fig. 2: Pre-processing example for OD localization CNN model. (a) Original non-OD region. (b) Normalized image of a. (c) Whitened image of a. (d) Original OD region. (e) Normalized image of d. (f) Whitened image of d. (g) Mean image.

model is shown in Figure 3. The OD region is obtained through exploiting features learned automatically from fundus images. CNNs are powerful models for feature learning. However, scanning full fundus image with a deep CNN by multiple scales to detect OD, result in a high computational cost and a slow detector. Hence and in order to achieve fast OD detection, we propose to integrate the cascade classifiers with CNN for efficient OD localization. The OD abnormality detector was trained to classify the detected OD region into three classes; normal, suspicious and abnormal.
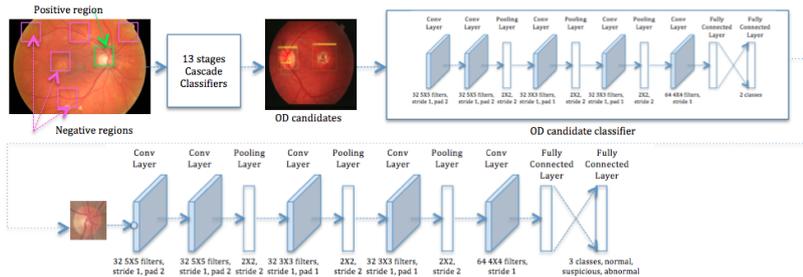
Fig. 3: Proposed Model Architecture

3.3   Cascade Classifiers

Cascade classifiers have been used successfully for object detection in real time [20]. It is based on the idea of weak classifiers ensemble being trained on simple features for rapid object detection. AdaBoost ensemble algorithm is used for training classifier and feature selection, as each classifier is restricted to select single rectangular feature that best classifies the positive and negative objects. The Haar-like features are extracted from effective image representation called integral image [20]. This image is scanned with rectangular sub-windows to extract the features. The features are used to train a cascade of classifiers into stages. At each stage, a classifier is trained to detect the positive examples and to reject the negative examples. The rejected sub-windows will not be processed

any further. Thus, the subsequent classifiers will consider only the candidate regions passed by previous classifier and hence become more and more complex. In this work, we used 3170 positive examples of OD and 9700 negative examples to train a cascade Adaboost classifier. The fully anonomysed images were collected from various resources to imitate the vast variety of fundus images.

3.4   Learning OD Feature Representation by CNN

CNNs use the same building units as classical artificial neural networks (ANN), i.e. neurons with learnable weights and biases. However, in CNNs features are learned into feature maps of replicated detectors. This results in fewer parameters to be learned in CNNs compared to ANN. This advantage results in easier training for deep model with hierarchical, complex and effective features. The performance of CNNs highly depends on their architecture, i.e. the number of layers and the depth and breadth of each layer. CNNs are composed of three layer types: convolutional, pooling and fully connected layers. The convolutional layer consists of replicated feature detectors organized into feature map. Every feature detector is connected to a small spatial region of its input layer and extends through the full depth of the input. Weights are shared between detectors in the same feature map. Pooling layer reduces the size a convolutional layer by dividing it into nonoverlapping windows and allocating a single value for each window. Maximum value and averaging are the most common approaches for pooling. The reduction in the convolutional layers sizes by pooling layers, allows more features to be learned with less computational cost. The fully connected layer is very similar to the ANN layer in which every neuron is connected to all neurons in the previous layer and each connection has a non-shared weight. The last fully connected layer in the CNN architecture combines all convolutional and pooling results into a class probability vector. In this work, two CNNs were trained to classify RGB images into OD and non-OD classes and then for the OD-class into normal/suspicious and abnormal classes. The architectures are shown in Figure 3. The network consists of five convolutional layers. Max pooling is performed after each convolutional layer to downsample the data. The fully connected layer combines all layer results to predict the output. The training starts by random initialization of the weights vectors. The input images are then forwarded into batches of size of a 100 via the CNN layers. The last layer outputs the class probabilities used to determine the image class by thresholding. The error between the ground truth and the prediction is used to update the weights by gradient descent back-propagation. The objective of this process is to minimize the mean output error. To visualize the discriminative ability of the proposed model, the outputs of the first convolutional layers of the OD localization model and the abnormality detector are shown in figure 4.

Fig. 4: Visualization of applying first layer convolutional filters to a test image. (a) OD image. (c) non-OD image. (b,d) Output of 1st convolutional filters in the OD localization model. (e) normal OD image. (g) abnormal OD image. (f,h) Output of 1st convolutional filters in the abnormality detector.

4   Experimental Evaluation

4.1   Databases

The proposed OD localization approach was evaluated on the following datasets. DRIVE [8], STARE [9], DIARETDB1 [10] and MESSIDOR [11] are publically available datasets. While, HAPIEE, KENYA and PAMDI are datasets collected from various population-based studies in collaboration with Moorfields Eye Hospital Reading Centre with permission from the individual studies. Particularly, HAPIEE was collected from Lithuania (Eastern Europe), KENYA from Kenya (Africa); and PAMDI from Italy (Europe). KFSH is a dataset collected from Saudi Arabia (Middle East), in collaboration with King Faisal Specialist Hospital. A total of 5781 images were used to evaluate the OD localization model. The ODs in PAMDI and HAPIEE are annotated by three class labels; normal, suspicious and abnormal and are used in this work to train and evaluate the abnormality detector.

4.2   Results and Discussions

The cascade classifier was trained in 13 stages on a 2.9 GHz Intel Core i7 using Matlab. The total training time for all stages was 17.3 minutes. However, The average predicting time for a previous unseen image was only 0.034 seconds. From the conducted experiments, it is observed that the cascade classifier is sufficient and performs well on datasets consisting of good quality normal images. However, the performance degrades clearly in other datasets that exhibit variable image conditions, which confirms the need for learning more discriminative features through the CNNs as shown in Figure 4(a-d). Figure 5 shows examples of OD detections by the proposed approach. Table 1 present the evaluation result for the OD detection approach. As shown in figure 4(e-h), the model is able to learn
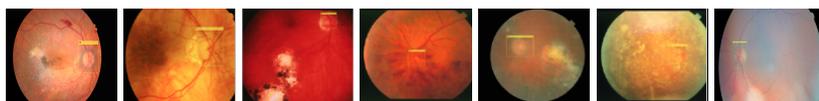
Fig. 5: Obtained Optic Disc by the proposed OD localization

high level complex features for OD abnormality feature representation. Results shown in table 2, demonstrate the discriminative ability of the deep learned features. The achieved sensitivity validates the effectiveness of our method and shows its potential application.

5   Conclusions

A new approach for OD abnormality detection is proposed. Our detector is fully supervised and utilizes fast to compute features and deep CNN filters. Current results are promising and show fast and powerful OD localization and high sensitivity in the evaluated datasets. In the future we aim to improve the specificity of the model by exploiting other CNN and auto-encoder architectures. We also aim to investigate the influence of pre-training and data augmentation and considering multiple classifiers.

Dataset	# images	Resolution	Abnormality	Accuracy
DRIVE	40	565X584	17.50 %	100.0 %
DIARETDB1	89	1500X1152	94.38 %	98.88 %
MESSIDOR	1200	2240X1488	54.50 %	99.20 %
STARE	402	700X605	91.04 %	86.71 %
KENYA	1125	850X565	24.98 %	99.53 %
HAPIEE	1951	3072X2048	11.48 %	98.36 %
PAMDI	907	1024X1024	18.19 %	98.13 %
KSSH	67	858X570	100 %	92.00 %
Total	5781			

Table 1: OD localization Accuracy obtained on different image sets

Dataset	Sensitivity%	Specificity%	Accuracy%
HAPIEE	96.42	86.00	86.52
PAMDI	94.54	98.59	97.76

Table 2: Performance of the proposed OD abnormality detector

References

1. Almazroa, A. et al. *Optic Disc and Optic Cup Segmentation Methodologies for Glaucoma Image Detection: A Survey.* , Journal of Ophthalmology, ID: 180972, 2015.
2. Raja, J. Benadict R., C.G. *Optic Disc Localization Methdodologies in Retinal Images: A review* ,Journal of Chemical Information and Modeling, vol. 53 pp. 1689-1699, 2015.
3. Yaniv T., et al.*DeepFace: Closing the Gap to Human-Level Performance in Face Verification.* Conference on Computer Vision and Pattern Recognition, p. 8, 2014.

4. Chuang W., et al. *Level Set Segmentation of Optic Discs from Retinal Images* Journal of Medical and Bioengineering, vol. 4, no. 3, pp. 213220, 2015.

5. Dongbo Z. and Yuanyuan Z., *Novel Accurate and Fast Optic Disc Detection in Retinal Images with Vessel Distribution and Directional Characteristics* IEEE Journal of Biomedical and Health Informatics, vol.20, no.1, pp. 333342, 2014.

6. Alex K., Ilya S., and G. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks,* Advances In Neural Information Processing Systems, pp. 19, 2012.

7. Adam C., Ann A., and Andrew Y Ng, *An Analysis of Single- Layer Networks in Unsupervised Feature Learning,* Aistats 2011, pp. 215223, 2011.

8. Staal, J J Abramoff, M D Niemeijer, M Viergever, M a Van Ginneken, B *Ridge based vessel segmentation in color images of the retina,* IEEE Transactions on Medical Imaging, vol. 23, no. 4, pp. 501509, 2005.

9. Adam Hoover, *Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response,* IEEE Transactions on Medical Imaging, vol. 19, no. 3, pp. 203210, 2000.

10. T. Kauppi, et al*DIARETDB1 diabetic retinopathy database and evaluation protocol,* Procedings of the British Machine Vision Conference 2007, vol. 1, pp. 15.115.10, 2007.

11. Decencire et al *Feedback on a publicly distributed database: the Messidor database,* Image Analysis and Stereology, v. 33, n. 3, p. 231-234, 2014.

12. H. Alghmdi, H. Tang, M. Hansen, A. OShea, L. Al Turk, and T. Peto *Measurement of optical cup-to-disc ratio in fundus images for glaucoma screening,* Computational Intelligence for Multimedia Understanding (IWCIM), 2015

13. H. Yuji, et al. *Vertical cup-to-disc ratio measurement for diagnosis of glaucoma on fundus images,* Medical Imaging, vol. 7624, 2010.

14. K. Narasimhan and K. Vijayarekha, *An efficient automated system for glaucoma detection using fundus image,* Journal of Theoretical and Applied Information Technology, vol. 33 no.1, pp. 104110, 2011.

15. J. Nayak, et al. *Automated diagnosis of glaucoma using digital fundus images,* J. Med. Syst., vol. 33, no. 5, pp. 337346, 2009

16. J. Cheng, et al. *Superpixel classification based optic cup segmentation,* Lect. Notes Comput. Sci, vol. 8151, pp. 421428, 2013

17. Almazroa A.,et al *Optic Cup Segmentation Based on Extracting Blood Vessel Kinks and Cup Thresholding Using type-II Fuzzy Approach* IEEE,978-1-4673-7519-1/15, 2015

18. Lim, Gilbert, et al *Integrated Optic Disc and Cup Segmentation with Deep Learning* EEE 27th International Conference on Tools with Artificial Intelligence Integrated, 2015

19. Muramatsu, C.et al *Automated determination of cup-to-disc ratio for classification of glaucomatous and normal eyes on stereo retinal fundus images* Journal of Biomedical Optics, vol.16, no.9, 2011

20. P Viola and M Jones *Rapid object detection using a boosted cascade of simple features* Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. I511-I518, 2001

21. Zilly, et al. *Boosting convolutional filters with entropy sampling for optic cup and disc image segmentation from fundus images* Ophthalmic Medical Image Analysis (OMIA), pp. 153160, 2015

22. LeCun, Y. Bengio, Y. Hinton, G. *Deep learning* Nature, vol. 521, pp. 436-444, 2015