An Explainable AI-Based Computer Aided Detection System for Diabetic Retinopathy Using Retinal Fundus Images

Kind, Adrian; Azzopardi, George

Published in:
International Conference on Computer Analysis of Images and Patterns

DOI:
10.1007/978-3-030-29888-3_37

An Explainable AI-Based Computer Aided Detection System for Diabetic Retinopathy Using Retinal Fundus Images

Adrian Kind[1(✉)] and George Azzopardi[2(✉)]

[1] University of Malta, Msida, Malta
kindadrian@gmail.com
[2] University of Groningen, Groningen, The Netherlands

Abstract. Diabetic patients have a high risk of developing diabetic retinopathy (DR), which is one of the major causes of blindness. With early detection and the right treatment patients may be spared from losing their vision. We propose a computer-aided detection system, which uses retinal fundus images as input and it detects all types of lesions that define diabetic retinopathy. The aim of our system is to assist eye specialists by automatically detecting the healthy retinas and referring the images of the unhealthy ones. For the latter cases, the system offers an interactive tool where the doctor can examine the local lesions that our system marks as suspicious. The final decision remains in the hands of the ophthalmologists. Our approach consists of a multi-class detector, that is able to locate and recognize all candidate DR-defining lesions. If the system detects at least one lesion, then the image is marked as unhealthy. The lesion detector is built on the faster R-CNN ResNet 101 architecture, which we train by transfer learning. We evaluate our approach on three benchmark data sets, namely Messidor-2, IDRiD, and E-Ophtha by measuring the sensitivity (SE) and specificity (SP) based on the binary classification of healthy and unhealthy images. The results that we obtain for Messidor-2 and IDRiD are (SE: 0.965, SP: 0.843), and (SE: 0.83, SP: 0.94), respectively. For the E-Ophtha data set we follow the literature and perform two experiments, one where we detect only lesions of the type micro aneurysms (SE: 0.939, SP: 0.82) and the other when we detect only exudates (SE: 0.851, SP: 0.971). Besides the high effectiveness that we achieve, the other important contribution of our work is the interactive tool, which we offer to the medical experts, highlighting all suspicious lesions detected by the proposed system.

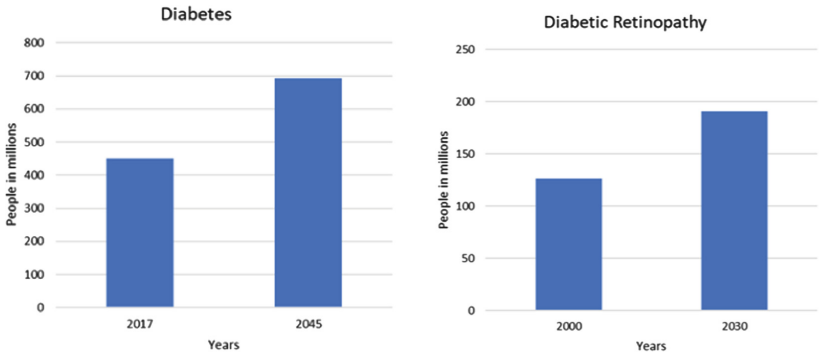
Keywords: Computer-aided detection · Diabetic retinopathy · Object detection · Convolutional neural networks · ResNet

M. Vento and G. Percannella (Eds.): CAIP 2019, LNCS 11678, pp. 457–468, 2019.
https://doi.org/10.1007/978-3-030-29888-3_37

1 Introduction

Diabetic retinopathy (DR) is responsible for the visual loss and blindness of millions of people world-wide. Statistics show that such diabetic retinal diseases are on the increase [16]. The early diagnosis and treatment of ocular diseases reduces progression of such illnesses and provides better quality of life. Currently, mass-screening programs are very expensive to conduct and are laborious and prone to error. Ophthalmologists and trained specialists are hard pressed to keep up with the large demands imposed by such a labour-intensive procedure.

Globally, it was estimated that in 2017 there were 451 million people with diabetes worldwide. These figures were expected to increase to 693 million by 2045 [4], Fig. 1a. As a consequence, the number of people with DR will grow from 126.6 million in 2010 to 191.0 million by 2030 [18], Fig. 1b. Such statistics indicate that diabetic retinopathy cases are on the increase and that it is difficult to have enough medical specialists running mass screening programs.



(a) (b)

Fig. 1. Worldwide projections of (a) diabetes and (b) diabetic retinopathy [18].

Most vision problems are often asymptomatic at the early stages, and the affected persons only realize there may be a problem when their conditions worsen. Such pathology may lead to visual impairment and even to blindness in just a few years. More effort is required to address these issues in order to reduce the unacceptable amount of unnecessary blindness.

There is, therefore, a growing need to semi-automate mass-screening programs for the identification of diabetic retinopathy at an early stage followed by regular monitoring. Automating such a diagnostic system would reduce waiting lists and would also result in mass-screening that is more efficient, more effective and financially feasible. The current lengthy process delays any intervention that needs to be performed to prevent or slow down the disease, the effects of which may lead to vision impairment or even blindness.

We propose a computer-aided detection system that automatically classifies retinal images into healthy and unhealthy. The system draws bounding boxes around lesions that define diabetic retinopathy and provides an interactive tool to the medical experts that allows them to conduct a further in-depth examination. Ultimately, it is the ophthalmologist who makes the decision of whether the concerned retina is healthy or not. For each processed image, the system that we propose generates a report including the list of lesions detected and an image showing the location of the detected lesions. This allows medical experts understanding the reason for the healthy and unhealthy classification and taking better informed decisions. This software is presented as an online web solution and can be easily used from a browser.

1.1 Related Work

Deep learning is a term coined for data driven learning mechanisms composed of multiple layers to learn representations of data with numerous levels of abstraction [10]. Convolutional Neural Networks (CNNs) are special types of networks that learn vision-based and data-driven classification models by means of deep learning. AlexNet [9] was the winning CNN in the image classification and object recognition competition called ImageNet Large Scale Visual Recognition Challenge [15] for the first time in 2012. The error rate was nearly cut in half from the previous year. From then onwards new and more sophisticated architectures have been introduced. Such architectures include GoogLeNet [17], and ResNet (Deep Residual Learning for Image Recognition) by He et al. [8], among others.

There are various state-of-the-art works that propose automatic methods for the detection of DR from retinal fundus images and the best results have been achieved from those that are based on deep learning. Image classification using convolutional neural networks is the main architecture used to diagnose fundus images and detect DR. Some of the best works published using these techniques include that of Gulshan et al. [7]. They applied deep learning to develop an algorithm for automated detection of diabetic retinopathy and diabetic macular edema in retinal fundus images. A deep convolutional neural network was trained using a data set of 128,175 retinal images, which were graded three to seven times by a panel of 54 US licensed ophthalmologists and ophthalmology senior residents. The final algorithm was validated using two separate data sets (EyePACS-1 and Messidor-2), both graded by at least seven US board-certified ophthalmologists.

Abramoff et al. [12] presented the IDx-DR commercial product providing a fully autonomous system capable of making the medical decisions safely and effectively. They provided an automated diabetic retinopathy diagnostic solution which has been authorised by the US Food Drug Administration (FDA) for use in clinics. IDx-DR is a commercializing partnership with IBM Watson Health Europe. It consists of a hybrid screening algorithm using both supervised and unsupervised deep learning techniques. That work is developed to detect referable DR (RDR), defined as moderate and worse diabetic retinopathy.

Gargeya et al. [6] developed a data-driven deep learning algorithm for the problem at hand. Their algorithm processes colour retinal fundus images and classifies them as healthy (no retinopathy) or having DR. A total of 75,137 publicly available fundus images were used to train and test their model. Their method was evaluated on the Messidor-2 and E-Ophtha data sets.

The highest sensitivity from automatic DR detection algorithms on public data sets has been measured by Rakhlin et al. [13]. In 2017 they obtained a sensitivity of 0.99 at a specificity of 0.71. They made use of CNNs to diagnose retinal fundus images. For the evaluation, they used the Kaggle and Messidor-2 data sets.

Desbiens et al. [5] also proposed a CNN-based solution for the diagnosis of DR from retinal images. They achieved a sensitivity of 0.93 at a specificity of 0.99. In particular, they built a multi-phase automatic grading system for diabetic retinopathy.

Most of the papers on automated DR detection use deep learning image classifiers, which are trained to give a label to a retinal fundus image as a whole; healthy or unhealthy. In this respect, such systems are considered as black boxes, in that it is difficult to interpret why a given image is labelled in a certain way. The system that we propose is more interpretable, in that it outputs bounding boxes around each suspicious lesion that are characteristic of diabetic retinopathy. As a result, our system allows the medical experts to take better-informed decisions on the images that the system marks as unhealthy.

2 Methods

2.1 Overview

In the following we describe in detail the new approach that we propose. The novelty component of our work is based on learning a multi-lesion detector that can localize the DR-defining lesions. Figure 2 illustrates an overview of the proposed pipeline. The software runs on a browser, and the whole process is initiated on the upload of a retinal fundus image. The system crops the field-of-view (FOV) and performs other pre-processing steps to enhance features and to reduce noise. The resulting image is then processed by an object detection model, which provides the classification, bounding boxes and confidence scores for the detected lesions. A decision function considers the scores of all detected lesions and classifies the given fundus image as Healthy or Unhealthy. Finally, the result is presented to the user on the browser where an interactive tool allows the user to zoom in and to examine both the pre-processed and the original images.

2.2 Pre-processing

We apply a couple of pre-processing steps before analysing the images for diabetic retinopathy. Firstly, we crop the FOV of the retina from the background. This is achieved by first applying contrast-limited adaptive histogram equalization

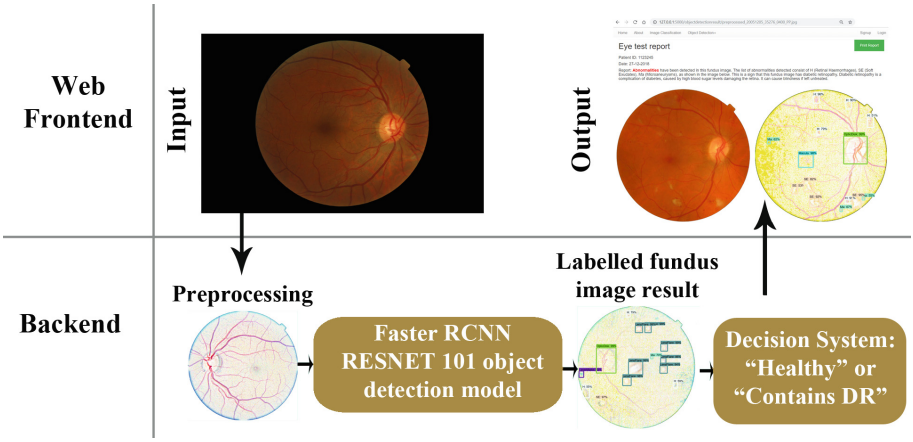


Fig. 2. A high-level overview of the proposed pipeline.

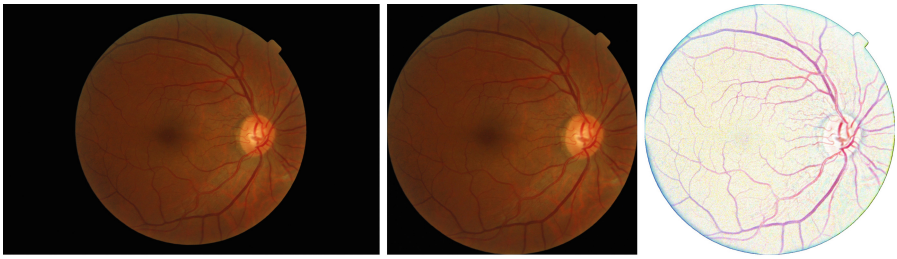


Fig. 3. Pre-processing steps. (a) Input image, (b) segmentation of the field-of-view (FOV), and (c) the Retinex result.

(CLAHE) followed by Gaussian blurring in order to improve the contrast and to smoothen out noise in the given images, respectively. Finally, we consider the outer edges detected by the Canny edge operator [3] in order to delineate the FOV from the background, Fig. 3b.

Secondly, we use the Retinex algorithm with four iterations for further enhancement of the segmented coloured retinal fundus images [11], Fig. 3c. This step provides a means of sharpening the fundus images, improves their colour consistencies regardless of variations in illumination and achieves dynamic range compression.

2.3 Classification Model

The classification model that we propose is one that can distinguish between various retinal healthy and unhealthy features. In particular, we use the Faster R-CNN ResNet 101 COCO as it is very robust in the localization and recognition of patterns of interest [8]. The Faster R-CNN [14] is an object detection

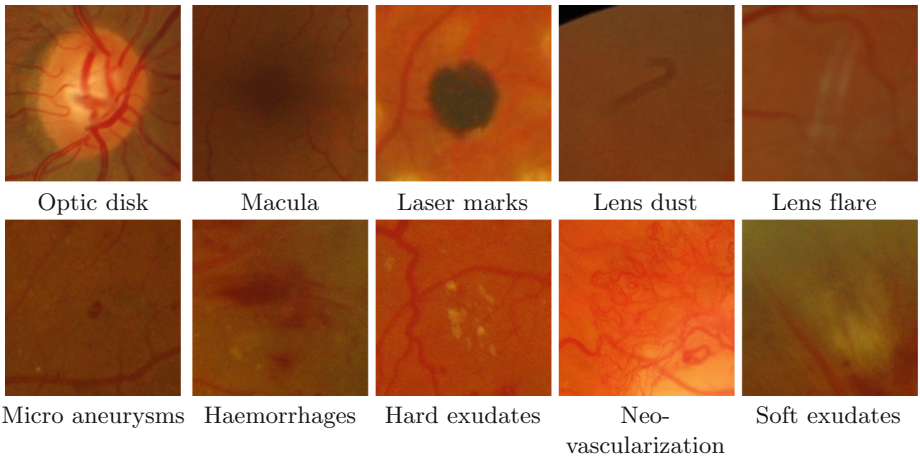


Fig. 4. Examples of all (top row) retinal features and (bottom row) diabetic retinopathy features found in fundus images.

algorithm with Region Proposal Networks (RPN). The RPN process uses the last layer of the CNN and slides a 3×3 window across the feature map and maps it to a lower dimension. For each window multiple possible regions are generated. If such a region has a score above a certain threshold the proposed region is selected. For each of these rectangular regions, a CNN is applied. The output of each region is passed to a soft-max layer to classify the region and a regression algorithm tightens the bounding box around the detected object. The architecture of the pre-trained model consists of 101 layers and it was originally trained on the COCO data set. In our work we use transfer learning by using the publicly available pre-trained model[1] as a feature extractor and only re-learn the last layer of the neural network. We also change the output classes of the network with the following 10 labels: optic disk, Macula, laser marks, lens dust and lens flare and the other five are DR-defining lesions, namely microaneurysms (Ma), haemorrhages (H), hard exudates (HE), soft exudates (SE), and neovascularization (NVE). The first five labels are retinal features that are commonly found in healthy retinas and the remaining labels are DR-defining features. Figure 4 shows examples of each of these 10 classes. We identify the DR lesion classes from the study in [2].

For the transfer learning of our model we use the retinal and DR features from 590 images, which we first preprocess using the above mentioned steps and then we manually extract (by cropping) and label 5,243 retinal and DR-defining features. Figure 5 illustrates the output of our system to an image of unhealthy retina.

[1] https://tinyurl.com/yxlmyq38.

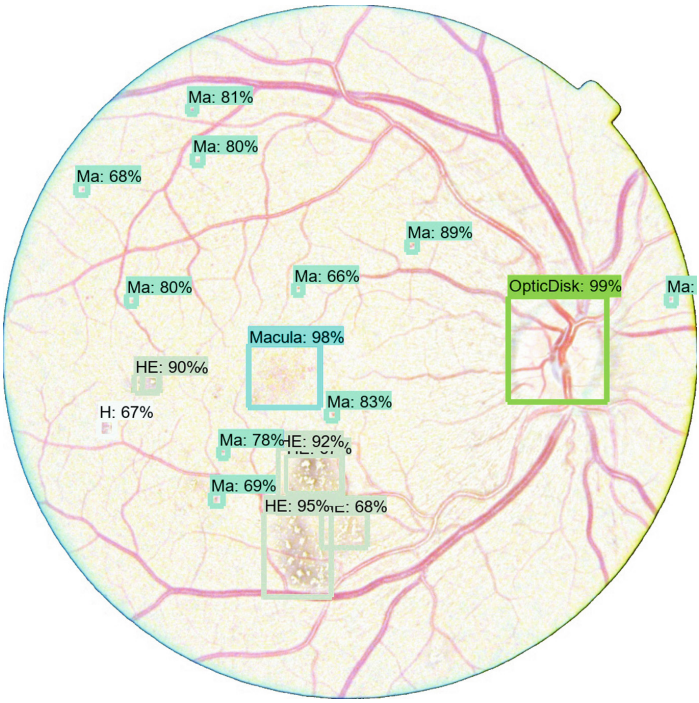


Fig. 5. Example of the output of our system to a retinal image with signs of DR. Each bounding box shows the name of the detected feature along with the confidence score.

2.4 Decision Criteria

A retinal image is marked as unhealthy if the classification model detects at least one DR feature out of the five mentioned above. Such a strict decision criteria is implemented in order to refer all images to the medical experts that contain even very small signs of DR.

2.5 System-Generated Report

For each processed image, our system generates a report containing whether the image is healthy or abnormal, the retinal image and the detected lesions, if any, along with some descriptive text for the medical expert. An example of such a report is shown in Fig. 6. It contains some descriptive text along with both the input image and the pre-processed image with the detected features indicated by bounding boxes and confidence scores.

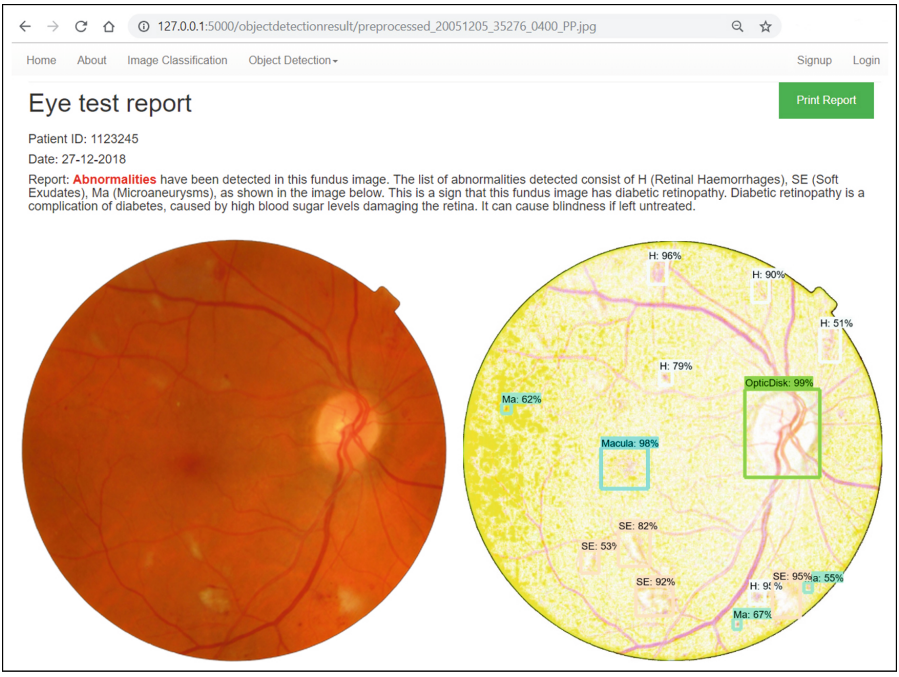


Fig. 6. The final report for a single image generated by our system.

3 Data Sets

We evaluate the proposed system on the retinal fundus images of three public benchmark data sets found online, namely Messidor-2[2] (1748 images), IDRiD[3] test set (103 images) and E-Ophtha[4] (463 images). In total, they contain 2,314 images. We do not use the EyePACS-1 data set mentioned above as it is not publicly available.

The training data, which we used to learn our classification model described above, consist of the retinal and DR features, which we manually labelled from 590 images. We obtained these training images from the public data sets IDRID (training set), KAGGLE[5], DiaRetDB1 V2.1[6] and STARE[7] data sets. Each training fundus image contains between 1 to 30 labelled retinal and DR features, which in total add up to 5,243 samples.

The Messidor-2 data set contains 1,748 images, which are divided into five grades. Grade 0 contains images with No DR, and the remaining grades represent

[2] http://www.adcis.net/en/Download-Third-Party/Messidor.html.
[3] https://idrid.grand-challenge.org/.
[4] http://www.adcis.net/en/Download-Third-Party/E-Ophtha.html.
[5] https://www.kaggle.com/c/diabetic-retinopathy-detection/data.
[6] http://www.it.lut.fi/project/imageret/diaretdb1_v2_1/.
[7] http://cecas.clemson.edu/~ahoover/stare/.

images with mild, moderate, severe and proliferative DR, respectively. Of the 1,748 images, 1,017 are grade 0 (i.e. healthy), 270 are grade 1, and the remaining ones have what are known as referable diabetic retinopathy (RDR) with a total of 457 images.

The IDRiD data set is also separated into the same five grades. It contains 103 fundus images where 34 are healthy images and 69 images contain DR.

E-Ophtha contains two sub data sets namely E-Ophtha-MA (Micro aneurysms), and E-Ophtha-EX (Exudates). The two data sets are separated in other two subsets: the first containing the fundus images without DR lesions and the other containing the ones with lesions. There are 381 fundus images labelled with micro aneurysms. Therefore the complete E-Ophtha data set consists of 463 images. The Micro aneurysms data set consists of 148 fundus images and 233 fundus images without Micro aneurysms but may contain other lesions, such as Exudates. On the other hand, the Exudates subset contains 47 images with Exudate lesions and 35 images without Exudates lesions but possibly containing other DR lesions, such as Micro aneurysms and Hemorrhages.

4 Experiments and Results

In the following we describe the experiments that we carried out and report the obtained results. Similar to the related works, for each experiment we count the number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN), and subsequently we quantify our results in terms of sensitivity (SN) and specificity (SP).

$$SN = \frac{TP}{FN + TP}, \quad SP = \frac{TN}{TN + FP} \tag{1}$$

An image is counted as TP if the system correctly marks it as unhealthy, while it is counted as FP if the system incorrectly marks it as unhealthy. A TN is counted when the system correctly marks an image as healthy. If the system incorrectly marks an image as healthy then it is counted as FN. Sensitivity and specificity measure the abilities of our method in detecting the truly unhealthy images (aka true positive rate), and the truly healthy ones, respectively.

In Table 1 we report the results that we achieve for the concerned data sets. Similar to the related works we eliminate the mild DR (grade 1) and treat the remaining grades of unhealthy retinas as one unhealthy class.

5 Discussion

We compare our results on the Messidor-2 data set to those of related works in Table 2. The large size and its popularity motivate our decision to focus our comparison on Messidor-2. We compare our results with the best performing methods in the literature. The fundamental difference between our approach and the others lies in the interpretability of the output. Our approach uses an object

Table 1. Experimental results in terms of sensitivity and specificity.

		Predicted healthy	Predicted unhealthy	Sensitivity	Specificity
Messidor-2 1474 images	Healthy	TN: 857	FP: 160	0.965	0.843
	Unhealthy	FN: 16	TP: 441		
IDRiD 103 images	Healthy	TN: 32	FP: 2	0.830	0.940
	Unhealthy	FN: 11	TP: 53		
E-Ophtha-MA 381 images	Healthy	TN: 191	FP: 42	0.939	0.820
	Unhealthy	FN: 9	TP: 139		
E-Ophtha-EX 82 images	Healthy	TN: 34	FP: 1	0.851	0.971
	Unhealthy	FN: 7	TP: 40		

Table 2. Comparison of our results to those of the related works on the Messidor-2 data set. The ground truths labelled as **Grades** and **Ophthalmologists** refer to the standard grades as provided by the Messidor-2 data set, and to proprietary ground truth labelled by ophthalmologists, respectively.

Author	Year	Ground truth	Sensitivity	Specificity
Ours	**2019**	**Grades**	**0.965**	**0.843**
Desbiens et al. [5]	2018	Ophthalmologists	0.929	0.989
Gargeya et al. [6]	2017	Grades	0.930	0.870
Rakhlin et al. [13]	2017	Grades	0.990	0.710
Abramoff et al. [1]	2016	Grades	0.968	0.870
Gulshan et al. [7]	2016	Ophthalmologists	0.961	0.939

detection model, which locates the DR-defining features in a given image and determines their confidence scores. This is in contrast to the other approaches, which address the problem from an image classification point of view and label a given image as healthy or unhealthy without indicating the involved features. Our method, therefore, provides more information - in terms of bounding boxes, labels and confidence scores - to the medical experts as to why the referred images are flagged as unhealthy. The application of the retinex preprocessing algorithm is also a novel introduction to automatic DR detection, which was not encountered in the related work. It provides colour enhancement and makes retinal features much more visible from the background.

For future works we aim to involve medical experts in order to have a bigger and more accurate labelled training set, which will contribute in achieving a more robust data driven classification model. Moreover, we aim to investigate the fusion of our object detection approach with an image classification method. By combining local and global decisions we expect to improve the results even more. Another direction for future work is to investigate a computational model that can also identify the extent (i.e. the grade) of diabetic retinopathy detected

in the given images. Such an approach would further improve the efficiency of a mass screening program as the flagged images could then be referred to specific ophthalmologists based on the grade that would be predicted by the system.

6 Conclusion

The proposed computer-aided detection system for the detection of diabetic retinopathy from retinal fundus image is very effective. The novel component of our work is the application of an object detection faster R-CNN ResNet model that is able to localize and recognize all features related to diabetic retinopathy. In this way, our approach assists the medical experts to take better-informed decisions on the referred images.

References

1. Abràmoff, M.D., et al.: Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest. Ophthalmol. Vis. Sci. **57**(13), 5200–5206 (2016)
2. Agurto, C., et al.: Multiscale AM-FM methods for diabetic retinopathy lesion detection. IEEE Trans. Med. Imaging **29**(2), 502–512 (2010)
3. Bao, P., Zhang, L., Wu, X.: Canny edge detection enhancement by scale multiplication. IEEE Trans. Pattern Anal. Mach. Intell. **27**(9), 1485–1490 (2005)
4. Cho, N., et al.: IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diab. Res. Clin. Pract. **138**, 271–281 (2018)
5. Desbiens, J., Gupta, S., Stevenson, J., Alderman, A., Trivedi, A., Buehler, P.: Deep annotated learning, harmonic descriptors and automated diabetic retinopathy detection (2018)
6. Gargeya, R., Leng, T.: Automated identification of diabetic retinopathy using deep learning. Ophthalmology **124**(7), 962–969 (2017)
7. Gulshan, V., et al.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA **316**(22), 2402–2410 (2016)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
10. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436 (2015)
11. McCann, J.: Lessons learned from mondrians applied to real images and color gamuts. In: Color and Imaging Conference, vol. 1999, pp. 1–8. Society for Imaging Science and Technology (1999)
12. Michael, D.A., Renato, A., Michael, F.C., Peter, A.K., Ursula, S.E.: IDx transforming healthcare through automation. https://www.eyediagnosis.net. Accessed 30 May 2018
13. Rakhlin, A.: Diabetic retinopathy detection through integration of deep learning classification framework. bioRxiv, p. 225508 (2018)

14. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems, pp. 91–99 (2015)
15. Russakovsky, O., et al.: ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. (IJCV) **115**(3), 211–252 (2015). https://doi.org/10.1007/s11263-015-0816-y
16. Shah, S.: Blindness and visual impairment due to retinal diseases. Commun. Eye Health **22**(69), 8 (2009)
17. Szegedy, C., et al.: Going deeper with convolutions. In: CVPR (2015)
18. Zheng, Y., He, M., Congdon, N.: The worldwide epidemic of diabetic retinopathy. Indian J. Ophthalmol. **60**(5), 428 (2012)