Knowledge-Based Systems 60 (2014) 20–27

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

An ensemble-based system for automatic screening of diabetic retinopathy

Bálint Antal *, András Hajdu

University of Debrecen, Faculty of Informatics, POB 12, 4010 Debrecen, Hungary

A R T I C L E I N F O

Article history:
Received 9 September 2013
Received in revised form 22 December 2013
Accepted 26 December 2013
Available online 20 January 2014

Keywords:
Diabetic retinopathy
Ensemble learning
Decision making
Machine learning
Automatic screening

A B S T R A C T

In this paper, an ensemble-based method for the screening of diabetic retinopathy (DR) is proposed. This approach is based on features extracted from the output of several retinal image processing algorithms, such as image-level (quality assessment, pre-screening, AM/FM), lesion-specific (microaneurysms, exudates) and anatomical (macula, optic disk) components. The actual decision about the presence of the disease is then made by an ensemble of machine learning classifiers. We have tested our approach on the publicly available Messidor database, where 90% sensitivity, 91% specificity and 90% accuracy and 0.989 AUC are achieved in a disease/no-disease setting. These results are highly competitive in this field and suggest that retinal image processing is a valid approach for automatic DR screening.

1. Introduction

Diabetic retinopathy (DR) is a consequence of diabetes mellitus which manifests itself in the retina. This disease is one of the most frequent causes of visual impairment in developed countries and is the leading cause of new cases of blindness in the working age population. In 2011, 366 million people were diagnosed with diabetes and a further 280 million people were having risk to develop it. At any point in time, approximately 40% of diabetic patients suffer from DR, out of which an estimated 5% face the sight-threatening form of this disease. Altogether, nearly 75 people go blind every day as a consequence of DR even though treatment is available.

Automatic computer-aided screening of DR is a highly investigated field [3]. The motivation for creating reliable automatic DR screening systems is to reduce the manual effort of mass screening [18], which also raises a financial issue [28]. While several studies focus on the recognition of patients having DR [18,4] and considering the specificity of the screening as a matter of efficiency, we show how both sensitivity and specificity can be kept at high level by combining novel screening features and a decision-making process. Especially, our results are very close to meet the recommendations of the British Diabetic Association (BDA) (80% sensitivity and 95% specificity [1]).

The basis for an automatic screening system is the analysis of color fundus images [2]. The key to the early recognition of DR is the reliable detection of microaneurysms (MAs) on the retina, which serves as an essential part for most automatic DR screening systems [4,19,9,24]. The role of bright lesions for DR grading has also been investigated with positive [17] and negative outcomes [4] reported. Besides lesions, image quality assesment [25,16] is also considered to exclude ungradeable images. As a new direction, in Agurto et al. [5] an image-level DR recognition algorithm is also presented.

The proposed framework extends the state-of-the-art components of an automatic DR screening system by adding pre-screening [11] and the distance of the macula center (MC) and the optic disk center (ODC) as novel components. We also use image quality assessment as a feature for classification rather than a tool for excluding images. The comparison of the components used in some recently published automatic DR screening systems can be found in Table 1.

Regarding decision making, automatic DR screening systems either partially follow clinical protocols (e.g. MAs indicate presence of DR) [19,9,25,16] or use a machine learning classifier [3,17,5]. A common way to improve reliability in machine learning based applications is to use ensemble-based approaches [21]. For medical decision support, ensemble methods have been successfully applied to several fields. In West et al. [29] the authors have investigatedthe applicability of ensembles for breast cancer data classification. The prediction of response to certain therapy is improved by the use of a classifier ensemble [22]. In Eom et al. [15] the authors used an ensemble of four classifiers for cardiovascular disease prediction. Ensemble methods are also provided

* Corresponding author. Tel.: +36 52512900.
E-mail addresses: antal.balint@inf.unideb.hu (B. Antal), hajdu.andras@inf.unideb.hu (A. Hajdu).

0950-7051/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.12.023

Table 1
Comparison of components of the automatic screening system.

Screening system	Image quality	Red lesion	Bright lesion	AM/FM	Pre-screening	MC-ODC
Abramoff et al. [4]		X				
Jelinek et al. [19]		X				
Antal and Hajdu [9]		X				
Philip et al. [25]	X	X				
Fleming et al. [16]	X	X	X			
Agurto et al. [5]				X		
Proposed	X	X	X	X	X	X

improvement over single classifiers in a natural language processing environment [13].

Ensemble systems combine the output of multiple learners with a specific fusion strategy. In Abramoff et al. [4] and Antal and Hajdu [9], the fusion of multiple MA detectors has proven to be more efficient than a single algorithm for DR classification. The proposed system is ensemble-based at more levels: we consider ensemble systems both in image processing tasks and decision making.

In this paper, a framework for the automatic grading of color fundus images regarding DR is proposed. The approach classifies images based on characteristic features extracted by lesion detection and anatomical part recognition algorithms. These features are then classified using an ensemble of classifiers. As the results show, the proposed approach is highly efficient for this task. The flow chart of our decision making protocol can be seen in Fig. 1, as well.

We have tested our approach on the publicly available dataset Messidor (see http://messidor.crihan.fr), where it has provided a 0.989 area under the ROC curve (AUC) value in a disease/no disease setting, which is a relatively high figure compared with other state-of-the-art techniques.

The rest of the paper is organized as follows: in Section 2, we present the image processing components of our system. Section 3 presents the details of the presented ensemble learning framework. Our experimental methodology and results can be found in Sections 4 and 5, respectively. Finally, we draw conclusions in Section 6.

2. Components of an automatic system for diabetic retinopathy screening

In this section, the components we used for feature extraction are described. They can be classified as image-level, lesion-specific, and anatomical ones.

2.1. Image-level components

2.1.1. Quality assessment

We classify the images whether they have sufficient quality for a reliable decision with a supervised classifier, where the box count values of the detected vessel system serve as features [7]. For vessel segmentation we use an approach proposed in Kovács and Hajdu [20] based on Hidden Markov Random Fields (HMRF). Here, the authors extend the optimization problem of HMRF models considering the tangent vector field of the image to enhance the connectivity of the vascular system consisting of elongated structures.

2.1.2. Pre-screening

During pre-screening [11], we classify the images as severely diseased (abnormal) ones or to be forwarded for further processing. Each image is split into disjoint regions and a simple texture descriptor (inhomogeneity measure) is extracted for each region. Then, a machine learning classifier is trained to classify the images based on these features.

2.1.3. Multi-scale AM/FM based feature extraction

The Amplitude-Modulation Frequency-Modulation (AM/FM) [6] method extracts information from an image, decomposing the green channels of the images into different representations which reflect the intensity, geometry, and texture of the structures with signal processing techniques. The extracted information are then filtered to establish 39 different representations of the image. The images are classified using these features with a supervised learning method. More on this approach can be found in Agurto et al. [6].

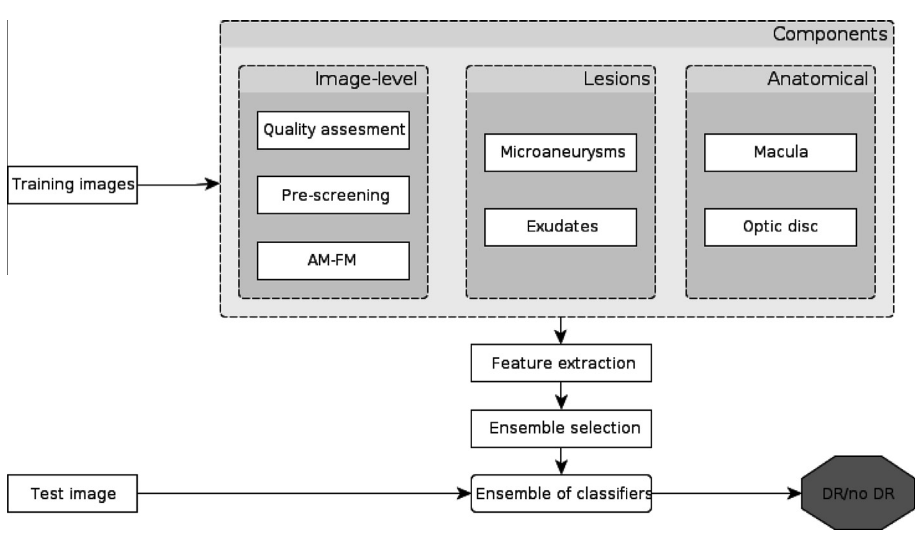

Fig. 1. Flow chart of the proposed decision support framework.