COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 124 (2016) 45–57

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Exploiting ensemble learning for automatic cataract detection and grading

Ji-Jiang Yang[a], Jianqiang Li[b,*], Ruifang Shen[a], Yang Zeng[c], Jian He[b], Jing Bi[b], Yong Li[b], Qinyan Zhang[c], Lihui Peng[a], Qing Wang[d]

[a] Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing, China
[b] School of Software Engineering, Beijing University of Technology, Beijing, China
[c] Automation School, Beijing University of Posts and Telecommunications, Beijing 100876, China
[d] Research Institute of Application Technology in Wuxi, Tsinghua University, Jiangsu, China

ARTICLE INFO

Article history:
Received 24 March 2015
Received in revised form
5 October 2015
Accepted 14 October 2015

Keywords:
Cataract detection
Fundus image classification
Ensemble learning
Support vector machines
Neural network

ABSTRACT

Cataract is defined as a lenticular opacity presenting usually with poor visual acuity. It is one of the most common causes of visual impairment worldwide. Early diagnosis demands the expertise of trained healthcare professionals, which may present a barrier to early intervention due to underlying costs. To date, studies reported in the literature utilize a single learning model for retinal image classification in grading cataract severity. We present an ensemble learning based approach as a means to improving diagnostic accuracy. Three independent feature sets, i.e., wavelet-, sketch-, and texture-based features, are extracted from each fundus image. For each feature set, two base learning models, i.e., Support Vector Machine and Back Propagation Neural Network, are built. Then, the ensemble methods, majority voting and stacking, are investigated to combine the multiple base learning models for final fundus image classification. Empirical experiments are conducted for cataract detection (two-class task, i.e., cataract or non-cataractous) and cataract grading (four-class task, i.e., non-cataractous, mild, moderate or severe) tasks. The best performance of the ensemble classifier is 93.2% and 84.5% in terms of the correct classification rates for cataract detection and grading tasks, respectively. The results demonstrate that the ensemble classifier outperforms the single learning model significantly, which also illustrates the effectiveness of the proposed approach.

1. Introduction

Cataract is defined as a clouding of the lens inside the eye leading to impaired vision; it is considered the most common cause of blindness [1–4]. The longer a patient has an untreated cataract, the more severe is the vision impairment. According to a WHO report [2], the estimated number of people in the world who are visually impaired is 285 million: 39 million are blind, and 246 million have impaired vision; 33% of cases of visual impairment and 51% of cases of blindness are caused by cataracts. In low and middle income counties and regions,

* *Corresponding author.*
E-mail addresses: yangjijiang@tsinghua.edu.cn (J.-J. Yang), lijianqiang@bjut.edu.cn (J. Li), ruifangshen@163.com (R. Shen), zak47y@163.com (Y. Zeng), jianhee@bjut.edu.cn (J. He), bijing@bjut.edu.cn (J. Bi), li.yong@bjut.edu.cn (Y. Li), lqflucky@163.com (Q. Zhang), lihuipeng@tsinghua.edu.cn (L. Peng), qing.wang@tsinghua.edu.cn (Q. Wang).

http://dx.doi.org/10.1016/j.cmpb.2015.10.007
0169-2607/

the prevalence of cataracts is even higher because of lower investment in health [2].

Although it is widely accepted that early detection and treatment can reduce the suffering of cataract patients and prevent visual impairment from turning into blindness, people in less developed areas still do not receive timely treatment because of poor eye care services and lack of professional ophthalmologists. On the other hand, common methods for cataract diagnosis require a slit lamp (e.g., the lens opacity classification system (LOCS III) [4] for clinical assessment or the Wisconsin cataract grading system (Wisconsin System) [5] for photographic grading), which is complicated and expensive for many patients. Therefore, reducing costs and simplifying the process for early cataract diagnosis is a crucial means of improving eye care service in less developed areas and bringing light to cataract patients.

We know that ophthalmologists can diagnose cataracts by checking the clear degree of fundus images [3,7,8]. Fig. 1 shows the gradings of cataract fundus images. Fig. 1(a) is an image with no cataract, where the optic disc and large and small blood vessels can be seen very clearly. There are fewer blood-vessel details in the Mild cataract image in Fig. 1(b), while only the large blood vessels and optic disc are visible in the moderate cataract image in Fig. 1(c). Furthermore, virtually nothing can be seen in Fig. 1(d), the severe cataract image. Thus, the task of fundus image classification is to build a classifier to simulate fundus image checking activities for automatic cataract detection and grading [1]. The fundus image classification task is based on the assumption that image blurring is not secondary to technical errors in image acquisition.

The research on fundus image analysis has been widespread [1,7], and several studies on automatic cataract detection and classification have been reported [9–18,42–46]. However, the proposed techniques use a single learning model for fundus image classification.

It is widely accepted that by combining multiple learning models, ensemble learning has greater potential to achieve more accurate classification than any of the constituent learning models [18]. This paper presents our experimental work on the ensemble learning framework for improved cataract detection and grading. To achieve classifier diversity (the key to high quality classification results in ensemble learning [19]), three independent feature sets, i.e., wavelet- [20], sketch- [21], and texture-based [22] features, are extracted from each fundus image. For each feature set, two base learning models, i.e., Support Vector Machine (SVM) [9,23] and Back Propagation Neural Network (BPNN) [24], are used to construct the base classifiers. Then, the ensemble methods [19], majority voting and stacking, are investigated to combine the multiple base classifiers for the final fundus image classification.

The contribution of this paper can be summarized as follows: (1) An ensemble learning-based approach is proposed for fundus image classification, where three independent feature extraction methods (i.e., wavelet-, sketch-, and texture-based methods) and two base learning models are investigated for base classifier building; then, two widely used ensemble approaches, i.e., stacking and majority voting, are employed to combine the multiple base classifier for improved cataract detection and grading. (2) The empirical experiments on the real-world datasets are presented; they illustrate that the best performance of the ensemble classifier can achieve 93.2% and 84.5% accuracy in terms of classification rates for cataract detection (two-class task) and cataract grading (four-class task), respectively. The experiments also demonstrate that the ensemble learning approach outperforms the single learning model significantly. We believe that our experimental study can serve as an important reference for eye disease diagnosis based on fundus image analysis.

This paper is organized as follows. Section 2 discusses the related work. Section 3 describes the details of the proposed ensemble learning approach of fundus image classification. Experiments and evaluation results are reported in Section 4. Section 5 concludes the paper.

2. Related work

Studies on fundus image analysis have been conducted for years. Segmentation and location of retinal structures, such as retinal lesions [26–28], vessels [20,28], optic discs [30–33], and aneurysms [32], have been widely studied. Based on these techniques, researchers are also trying to develop diagnostic systems for specific retina-related diseases, including microaneurysms [25], diabetic retinopathy [6, 29, 34–36], age-related macular degeneration [7], glaucoma [37–41], and cardiovascular diseases [41]. These studies are relevant but not closely related to the work in this paper.

From the point of view of automatic cataract detection and grading, the work in this paper is closely related to [8] and [10]. For example, Li et al. made an effort to classify and diagnose specific cataracts automatically by split images and retro-illumination images, including nuclear cataracts [42–45], cortical cataracts [45], and posterior sub-capsular cataracts [46]. Yang et al. [8] used an improved Top-bottom hat transformation to pre-process the retinal image and extracted the luminance and texture message as features; they then constructed the classifier with a two-layered back propagation neural network [24]. The work in [10] proposed a fundus image classification approach for cataract detection using the wavelet transform- and sketch-based methods for feature extraction and multi-class Fisher discriminant analysis [47] for classifier building. All of these studies built a single learning model for fundus image classification. In this paper, however, we adopt an ensemble learning framework for automatic cataract detection, where multiple heterogeneous learning models are combined for more accurate prediction of cataract classification.

3. Ensemble learning-based fundus image classification

The basic process of ensemble learning is to construct multiple base learning models and combine them to solve the same problem [18,38]. If each base learning model is viewed as an expert, multiple experts may be better than any single expert if their individual judgments are appropriately combined. Because the ensemble idea has great potential for reducing the learning bias of base learning models, it may demonstrate better performance in many classification tasks than any single constituent base model [18].