Computers and Electrical Engineering 72 (2018) 274–282

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Deep convolutional neural networks for diabetic retinopathy detection by image classification☆

Shaohua Wan[a,*], Yan Liang[a], Yin Zhang[a,b,*]

[a] School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China
[b] Hubei Key Laboratory of Medical Information Analysis & Tumor Diagnosis and Treatment, Wuhan 430074, China

ARTICLE INFO

Article history:
Received 31 January 2018
Revised 30 June 2018
Accepted 31 July 2018

Keywords:
Diabetic retinopathy
Fundus images classification
Convolutional neural networks
Transfer learning

ABSTRACT

Diabetic retinopathy (DR) is a common complication of diabetes and one of the major causes of blindness in the active population. Many of the complications of DR can be prevented by blood glucose control and timely treatment. Since the varieties and the complexities of DR, it is really difficult for DR detection in the time-consuming manual diagnosis. This paper is to attempt towards finding an automatic way to classify a given set of fundus images. We bring convolutional neural networks (CNNs) power to DR detection, which includes 3 major difficult challenges: classification, segmentation and detection. Coupled with transfer learning and hyper-parameter tuning, we adopt AlexNet, VggNet, GoogleNet, ResNet, and analyze how well these models do with the DR image classification. We employ publicly available Kaggle platform for training these models. The best classification accuracy is 95.68% and the results have demonstrated the better accuracy of CNNs and transfer learning on DR image classification.

1. Introduction

Diabetic retinopathy(DR), one of the most common retinal diseases, is a common complication of diabetes and one of the major causes of blindness in humans. Since the disease is a progressive process, medical experts suggest that diabetic patients need to be detected not less than twice a year in order to timely diagnose signs of illness. In the current clinical diagnosis, the detection mainly relies on the ophthalmologist examining the color fundus image and then evaluates the patient's condition. This detection is arduous and time-consuming, which results in more error. Furthermore, due to the large number of diabetic patients and the lack of medical resources in some areas, many patients with DR can not timely diagnosed and treated, thus lose the best treatment opportunities and eventually lead to irreversible visual loss, as well as even the consequences of blindness. Especially for those patients in early phase, if DR can be found and treated immediately, the deteriorated process can be well controlled and delayed. At the same time, the effect of manual interpretation is extremely dependent on the clinician's experience. Misdiagnosis often occurs due to the lack of experience of medical doctors. In the last ten years, deep CNNs have made the remarkable achievements in a large amount of computer vision and image classification, substantially surpassing all previous image analysis methodologies. Computer-aided diagnosis is desired because it allows for mass screening of the disease. Therefore, in order to achieve fast, reliable computer-aided diagnosis

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. M. M. Hassan.

* Corresponding author.
E-mail addresses: shaohua.wan@ieee.org (S. Wan), lyan95@zuel.edu.cn (Y. Liang), yin.zhang.cn@ieee.org (Y. Zhang).

https://doi.org/10.1016/j.compeleceng.2018.07.042
0045-7906/

Table 1
Classification dataset.

Class name	The degree of DR	Numbers
Class 0	Normal	25810
Class 1	Mild	2443
Class 2	Moderate	5292
Class 3	Server	873
Class 4	Proliferative	708

and treatment, the application of CNNs to automatically and accurately process and analyze DR fundus images is still a very necessary and urgent task.

The motivation of this paper is to implement an automatic diagnosis of DR using fundus images classification. We work on classifying the fundus images by the severity of DR, so that an end-to-end real-time classification from fundus image to the condition of patients can be achieved. Instead of the doctors' manual operation with experience, it relieves their pressure on the diagnosis and treatment for DR in an automatic and high-accuracy way. For this task, we are employing various image preprocessing methods to extract many important features and then classify to their respective classes. We adopt CNNs architecture to detect the DR in 5 data sets. We evaluate the sensitivity, specificity, accuracy, Receiver Operating Characteristic (ROC) and Area Under Curve (AUC) of the models. The contributions of this paper are summarized as follows:

- In order to get the finest mass image dataset to train with for our models, we take preprocessing steps, like data augmentation will increase the number of training examples, and data normalization will denoise to precisely predict classification.
- We could train the latest CNNs model (AlexNet, VggNet, GoogleNet and ResNet) to recognize the slight differences between the image classes for DR Detection.
- Transfer learning and hyper-parameter tuning are adopted and the experimental results have demonstrated the better accuracy than non-transferring learning methodology on DR image classification.

The remainder of this article is organized as follows. Section 2 presents related work of Convolutional Network Architectures on DR image classification. Section 3 introduces Kaggle data sets and describes the image preprocessing methods. Section 4 overviews the CNNs models, AlexNet,VggNet,GoogleNet and ResNet. Section 5 presents experimental analysis and evaluation metrics. Finally, we conclude this article in Section 6

2. Related work

Automated detection of DR images has such benefits that DR can be diagnosed at early levels efficiently. Early detection and treatment are really important for delaying or preventing visual degradation. For an overview of such methods referred [1,2]. More recently, deep learning techniques have extraordinary revolutionized the computer vision field. Especially leveraging CNNs to perform image classification has attracted many researchers. Research in this field include segmentation of these features, as well as blood vessels [3,4]. Deep CNNs structures were originally presented for the solution of natural image classification, and recent research has made rapid progress in working on DR fundus images classification. Wang et al. [5] adopt a CNN (LeNet-5) model to extract image features for addressing blood vessel segmentation. These methods has some limitations. Firstly, because of the features of dataset are extracted manually and empirically, their accuracy can't be guaranteed. Secondly, the data sets are small in size and low in quality, usually only a few hundred or even dozens of fundus images with relatively single collection environment, bringing difficulties to compare the performance of algorithms in the experiment. Since Alex et al. [6,7] presented AlexNet architecture for remarkable performance improvements at the 2012 ILSVRC competition, the widespread applications of deep CNNS in computer vision have formally mushroomed. After a number of excellent CNNs architectures have been proposed, such as VggNet [8,9], GoogleNet [10]. As one of the most important network models, ResNet [11] was proposed in 2015, which further enhances the performance of CNNs in image classification. Since it is arduous and time-consuming to build a model from zero, transfer learning and hyperparameter tuning are used in this paper. These architectures can be proposed in [12–24]. We employ transferring learning to expedite the learning time and compare the performance with AlexNet,VggNet, GoogleNet and ResNet, which finally provides an automatic and accurate detection so that visual damage could be minimized to the minimum degree. Compared to the previous methods, our work has the following improvements over the convergence time for the large scale experimental datasets and a better performance on classification.

3. Dataset and preprocessing

Our dataset is from an publicly Kaggle [25] website, which tries to develop a model for DR detection. Data set consists of high resolution eye images and graded by trained professionals in 5 classes(0–4) which is according to below Table 1and Fig. 1. The data set contains 35,126 high resolution RGB images with a resolution of about 3500x3000 in multiple varieties of imaging situations. The labels are provided by professionals who rank the presence of DR in each image by a scale of 0, 1, 2,