Measurement 125 (2018) 586–591

Contents lists available at ScienceDirect

Measurement

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

A retinal vessel detection approach using convolution neural network with reinforcement sample learning strategy

Yanhui Guo[a,*], Ümit Budak[b], Lucas J. Vespa[a], Elham Khorasani[a], Abdulkadir Şengür[c]

[a] *High Performance Computing Lab, Department of Computer Science, University of Illinois, Springfield, IL, USA*
[b] *Engineering Faculty, Department of Electrical-Electronics Engineering, Bitlis Eren University, 13000 Bitlis, Turkey*
[c] *Electrical and Electronics Engineering Department, Firat University, Elazig, Turkey*

A R T I C L E I N F O

Keywords:
Computer-aided detection
Retinal vessels
Convolution neural network
Image segmentation

A B S T R A C T

Computer-aided detection (CAD) provides an efficient way to assist doctors to interpret fundus images. In a CAD system, retinal vessel (RV) detection is an important step to identify the retinal disease regions automatically and accurately. However, RV detection is still a challenging problem due to variations in morphology of the vessels on a noisy background. In this paper, we formulate the detection task as a classification problem and solve it using a convolutional neural network (CNN) as a two-class classifier. The proposed model has 2 convolution layers, 2 pooling layers, 1 dropout layer and 1 loss layer. The contributions of the algorithm are two-fold. First, a new model of CNN is designed to automatically extract features and classify the retinal vessel region. Compared to traditional classification procedures, it is fully automatic and does not need preprocessing and manual extraction and description of features. Second, a novel reinforcement sample learning scheme is proposed to train the CNN with fewer iterations of epochs and less training time. The proposed model is trained and tested using the Digital Retinal Images for Vessel Extraction (DRIVE) and Structured Analysis of the Retina (STARE) data sets. The proposed CNN achieves better performance and significantly outperforms the state-of-the-art for automatic retinal vessel segmentation on the DRIVE data set with 91.99% accuracy and 0.9652 AUC score (area under ROC), and on the STARE data set with 92.20% accuracy and 0.9440 AUC value. We further compare our result with several state-of-the-art methods based on AUC values. The comparison shows that our proposal yields the second best AUC value. This demonstrates the efficiency of the proposed method without pre-processing and with high accuracy and training speed.

1. Introduction

Monitoring retinal vessels in fundus images is crucial in diagnosing ophthalmologic diseases and treatment of cardiovascular. This crucial task is also challenging and time-consuming if detection of retinal vessels is done manually. Automatic segmentation of retinal vessels in fundus images is in demand because manual segmentation is quite challenging due to variations in morphology of the retinal vessels against lighting conditions and background noise. In addition, manual segmentation of the retinal vessels necessitates highly skilled physicals. This challenging task is also quite important to examine the condition of the vessels which plays a significant role in early detection of several diseases such as hypertension, arteriosclerosis, and diabetes [1]. Diabetic retinopathy (DR) is a well-known disease which is one of the causes of blindness [2]. Statistics show that there are up to 342 million diabetic patients in the worldwide, which means high medical expenses and highly work-loaded physicians [3]. Thus, examining the width, tortuosity, and branching attributes of retinal vessels by automatic segmentation can either help an early detection of DR and can also reduce the medical expenses and workload of physicians.

Many methods have been proposed for segmentation of the retinal vessels in fundus images. These methods can be broadly categorized into two groups. In the first group, the segmentation of the retinal vessels was accomplished by pixel classification [4]. Generally, machine learning algorithms were adopted to classify the fundus images into the vessel and non-vessel regions [5]. To this end, neural networks (NN), Adaboost, support vector machines (SVM) and ensemble based learning methods were used in vessel segmentation [6–9]. In the second group, region growing, Gaussian and matched filters and morphological operations based methods were used for retinal vessel segmentation [4]. Some of the related efficient methods are reviewed in the following couple of paragraphs.

Soares et al. used pixel intensities and 2-D Gabor wavelet features for supervised retinal vessel segmentation [10]. The authors adapted

* Corresponding author.
E-mail address: yanhui.guo@aggiemail.usu.edu (Y. Guo).

https://doi.org/10.1016/j.measurement.2018.05.003
Received 21 November 2017; Received in revised form 2 March 2018; Accepted 2 May 2018
Available online 09 May 2018
0263-2241/

Gaussian mixture models for determining the class conditional probability densities. A Bayesian classifier was then used for supervised classification. Another supervised vessel segmentation method was proposed by Franklin et al. [11]. Authors combined Gabor and moment-based features for efficient segmentation of the blood vessels in fundus images. A multi-layer perceptron NN classifier was considered by the authors. Marin et al. proposed a retinal vessel segmentation scheme where a seven-dimensional feature vector was constructed based on gray scale intensities and invariant moments [6]. The authors used NN classifier in the supervised classification stage of the scheme. Zolfagharnasab et al. proposed a novel matched filter for segmentation of the retinal vessels [12]. The authors introduced a new kernel function which was based on the Cauchy distribution. As the authors mentioned, the presented new matched filter highly improved the segmentation accuracy. Rodrigues et al. proposed a methodology for segmentation of the optic disc and retinal vessels in fundus images [13]. The proposed methodology employed wavelets and mathematical morphology for optic disk segmentation. Moreover, the vessel segmentation was carried out by tabular characteristics. Chen et al. proposed a new scheme for retinal vessel segmentation [14]. The proposed method segmented the retinal vessel with 1-pixel width by employing the global graph-based decision. In the proposed method, the disconnected vessels were then connected and spurious ridges were discharged based on the shortest path algorithm.

Recently, as deep learning has gained much attention in object detection and image classification applications, there have been several attempts where convolutional neural networks (CNN) based schemes were used for retinal vessel segmentation. Liskowski et al. proposed a deep learning based method for segmentation of the retinal vessels in fundus images [15]. Authors considered two types of CNN models. One was a standard CNN architecture with 9 layers and the other only consisted of convolution layers. Wang et al. presented a retinal vessel segmentation method where CNN and random forests were combined for improved performance [16]. The authors employed the CNN as the feature extractor and the random forest was used for classification. Maji et al. presented an ensemble based methodology for retinal vessel segmentation [17]. Authors considered 12 deep CNN models for constructing the classifier structure. The mean operation was used for the output of all networks for the final decision. Lahiri et al. proposed another ensemble based deep model for retinal vessels segmentation [18]. The considered architecture was based on stacked autoencoders. The final decision was the combination of all stacked autoencoder outputs passed through a softmax layer.

In this paper, we formulate the retinal vessel detection task as a classification problem and solve it using the CNN as a two-class classifier. To increase the training speed and validation accuracy, a novel learning strategy, reinforcement sample learning is proposed in this paper. The proposed method achieves better performance and significantly outperforms the state-of-the-art for automatic retinal vessel segmentation on the DRIVE data set.

The remainder of the paper is organized as follows. In the next sections, we present the proposed CNN model using a novel learning scheme. Then the experimental results are discussed and the conclusions are presented in the final section.

2. Proposed method

2.1. CNN construction

In this model, the input to the CNN is a ROI image, and the output is the classification result of the pixel located at the center of the ROI. In contrast to traditional classification approaches, CNN is able to extract different features automatically through adaption of its multi-layer and feed-forward structure. Five types of layers are employed to construct the CNN network in the proposed method. These are convolution layer, pooling layer, dropout layer, fully connected layer and loss function layer. The dropout layer is employed to increase the generalization ability of the network.

In the convolutional layer, let $x^{l-1}(m)$ be the *mth* input feature at the layer *l-1,* and $W^l(m,n)$ the weight of filter which connects nth feature of the output layer to *mth* feature of the input layer, and $b^l(n)$ a bias. The values $x^l(\mathrm{n})$ in the *lth* convolutional layer is computed as:

$$x^l(\mathrm{n}) = \mathrm{f}\left(\sum_m \left(x^{l-1}(m)*W^l(m,n) + b^l(n)\right)\right) \tag{1}$$

$$\mathrm{f}(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

where $*$ is the convolutional operation and f is a nonlinear sigmoid function. The weights of the filter $W^l(m,m)$ are initialized randomly and then updated by a backpropagation algorithm.

The pooling layer is used to reduce the spatial size of the feature maps. It can alleviate the parameters numbers in the network and avoid overfitting problem. The values $x^l(\mathrm{n})$ in the pooling layer l are computed as:

$$x^l(\mathrm{n}) = \mathrm{pool}(x^{l-1}(n)) \tag{3}$$

where $\mathrm{pool}(\cdot)$ is a sampling function.

The fully connected layer at the last layer of the network works as a classifier. In the proposed method, a softmax function is used in the fully connected layer as:

$$\mathrm{P}(\ \mathrm{y}= 1|\mathrm{x};\mathrm{w}) = \frac{1}{1 + e^{-w^T x}} \tag{4}$$

where y is the class label and w is the weight.

In the proposed model, a cross entropy loss function is used in the loss layer to measure the error at the final softmax layer as:

$$\mathrm{L(W)} = -\left[\sum_{i=1}^{M}\sum_{j=1}^{N} \#(y(i) = r)log\frac{e^{W^T(r)x(i)}}{\sum_{j=1}^{N} e^{W^T(j)x(i)}}\right] \tag{5}$$

where $\#(\cdot)$ is a indicator function. M is the total number of the samples, and N is the total number of the classes. W is the weight vector for each layer.

The proposed model has two convolution layers, two pooling layers, one fully connected layer, one dropout layer and a loss layer. In the first and second convolutional layers, we use 32 and 64 filters of size 5 × 5, respectively. In the pooling layers, a 2 × 2 max pooling procedure is employed. In the loss layer, the softmax loss is adopted. The detailed structure of the CNN layers is described in Table 1 and its architecture is illustrated in Fig. 1.

2.2. Reinforcement sample learning strategy

The idea of reinforcement sample learning is to reinforce training the network on the samples with poor performance in the training procedure. The detailed steps are described as follows:

Table 1
Structure and configuration of CNN model (C: convolution layer; S: pooling layer; F: fully connected layer; D: dropout layer; L: loss function layer).

Layer number	Layer type	Filter size	Stride size
1	C	5 × 5 × 32	1
2	S	2 × 2	2
3	C	5 × 5 × 64	1
4	S	2 × 2	2
5	F		
6	D		
7	L		