GLAUCOMA DIAGNOSIS USING FEATURE LEARNING BASED ON CONVOLUTIONAL NEURAL NETWORK

BAIDAA AL-BANDER, WALEED AL-NUAIMY, MAJID AL-TAEE, YALIN ZHENG
Baidaa Al-Bander, Miss, **Waleed Al-Nuaimy**, Dr, **Majid Al-Taee,** Prof, and **Yalin Zheng,** Dr, University of Liverpool

ABSTRACT
Glaucoma is one of the common causes of blindness worldwide. It leads to deterioration in vision and quality of life if it is not cured early. This paper addresses the feasibility of developing an automatic feature learning technique for detecting glaucoma in coloured retinal fundus images using a deep learning method. A fully automated system based on convolutional neural network (CNN) is developed to distinguish between normal and glaucomatous patterns for diagnostic decisions. Unlike traditional methods where the optic disc features are handcrafted, the features are extracted automatically from the raw images by CNN and fed to the SVM classifier to classify the images into normal or abnormal. We demonstrate an accuracy, specificity and sensitivity of 88.2%, 90.8%, and 85%, respectively which compared favourably to the-state-of-the-art but at considerably lower computational cost. The obtained preliminary results clearly demonstrate that the proposed deep learning method is promising in automatic diagnosis of glaucoma.

Keywords: Deep learning; Convolutional neural network; Feature learning; Fundus images; Glaucoma

INTRODUCTION

Glaucoma is one of the chronic eye diseases, which results from the damaging of optic nerve head due to intra ocular pressure (IOP) of the eye. By 2020, it is predicted to affect around 80 million people in the worldwide (Quigley & Broman, 2006). In the early stages of the disease, the patients don't have symptoms of vision loss while with the disease progress the loss of peripheral vision occurs. In the advance stage of glaucoma, the patients suffer from the total blindness. Early detection and treatment of glaucoma helps in preventing the progression of the disease (Susanna Jr & Vessani, 2009). Considering the increasing number of people suffering from glaucoma, relying on the medical devices to diagnose and detect the disease may become unfeasible. Extensive research using various image processing techniques are applied to overcome the aforementioned problem for early detection of glaucoma.

The digital colour fundus image has been widely used in recent years to diagnose glaucoma where retinal fundus images are widely available in hospitals and can be obtained easily at a lower cost compared to optical coherence tomography (OCT). In the fundus images, the optic disc OD is divided into two regions: a peripheral zone called neuroretinal rim and a central bright region called the optic cup. One of the indicators of glaucoma is the enlargement of cup zone with respect to OD, which can be estimated by measuring the vertical cup to disc ratio (CDR) (Damms & Dannheim, 1993). Manual annotating of the optic cup and disc for each image is subjective and time consuming. Furthermore, extracting features that characterise glaucoma could be recognised only by expertise and training of the examiner.

To detect glaucoma in the retinal fundus images, the existing methods in the literature relied on handcrafted different features from images manually. Nayak et al. (2009) proposed a method based on artificial neural network (ANN) classifier and morphological features of optic disc nerve to detect the glaucoma. Furthermore, Bock et al. (2009) presented glaucoma Risk Index (GRI) approach to detect glaucoma by extracting features then feeding them to a principle component analysis (PCA) algorithm for dimensionality reduction and support vector machine (SVM) model to classify the images into either normal or pathological. A method based on texture and higher order spectral (HOS) features with random forest classifier has been proposed by Acharya et al. (2011) to diagnose the glaucoma in the images. Moreover, Dua et al. (2012) used wavelet features with a features selection technique to detect the glaucoma in the images using sequential minimal optimization (SMO) classifier. Recently, Acharya et al. (2015) proposed Gabor features along with SVM to detect the glaucoma in the images.

Nowadays, deep learning or deep neural networks (DNNs) have been an active research topic, which can learn to extract highly complex features from the input data automatically. Convolutional neural networks (CNNs) are one of the most common deep learning architectures that successfully applied in many classification applications

(Krizhevsky & Hinton, 2012; Lai et al., 2015). Recently, an automatic deep learning method for location detection of the fovea and optic disc (OD) in digital fundus retinal images was presented by Al-Bander et al. (2016).

In this work, we present an automated system to detect glaucoma in retinal fundus images. Unlike traditional methods, the features that discriminate the optic disc (OD) region from optic cup region are extracted by convolutional neural network CNN automatically without the need for human intervention.

The remainder of this paper is organized as follows. Section 2 presents a background of the proposed system stages. The dataset used in this study is described in Section 3 along with the proposed methodology. Next, the results of the developed method is presented and discussed in Section 4. Finally, the work is concluded in Section 5.

1. Background - Convolutional Neural Network and SVM

2.1 Convolutional Neural Network

The convolutional neural network CNN consists of convolutional layers, max-pooling layers, dropout layers, fully connected layers and activation functions. The convolutional layer is a sequence of filters to perform a 2D convolution on the input image where the output of this layer is called the feature map (LeCun et al., 1998). Max-pooling layer is a subsampling layer where the feature map is down-sampled (Ranzato et. al, 2007). Dropout layer is a regularization technique to reduce the overfitting (Srivastava et al., 2014). Fully connected layer is a final layer in CNN where each neuron completely connected to other neurons. Rectified linear unit (RELU) (Glorot et al., 2011) is usually used after convolution layers as an activation function. The CNN is usually trained under cross-entropy loss function that measures how compatible the network parameters with respect to the ground-truth values in the training dataset. Softmax layer is used to make the output of CNN lies between zero and one for a categorical target variable (i.e. the output is interpretable as posterior probabilities).

2.2 Support Vector Machine (SVM)

Support vector machines SVM are among the most widely used supervised classification methods in the field of machine learning. Hard margin classier considers the simplest kind of SVM, which solves an optimization problem to find the linear classification rule with maximal geometric margin. Thus, in the linearly separable case, the hard margin SVM builds the hyperplane that classifies all data correctly and maximizes the distance to the nearest training data points. In practice, datasets are usually not linearly separable, and therefore, the SVM optimization problem must be modified. This modification is necessary in order to achieve a trade-off between minimizing classification error on the training data points and maximizing geometric margin by applying soft margin idea. Soft margin builds a hyperplane that allows misclassification of difficult or noisy examples while still maximising the distance to the nearest cleanly separated data examples (Vapnik & Vapnik, 1998; Cristianini & Shawe-Taylor, 2000).

2. Materials and Methods

A publically available RIM-ONE (Fumero et al., 2011) a database for optic cup and disc segmentation evaluation, which comprises 455 high-resolution glaucoma and non-glaucoma images, is used to evaluate and test the proposed method. The images in this dataset are classified into 255 normal and 200 glaucomatous images. Figure 1 shows the stages of glaucoma grading proposed system.

3.1 Pre-processing

The images are resized to 227×227pixels. The images are applied to the network without any enhancement or further pre-processing step except image resizing to decrease the computational time. The data is randomly divided into 70% for training and 30% for implemented system validation and evaluation.

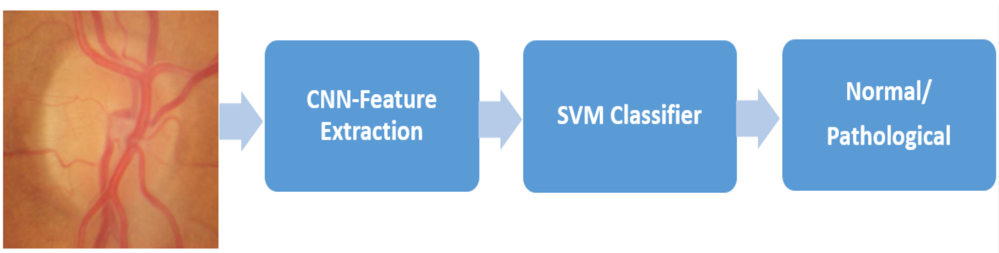

Figure 1. Proposed system stages of glaucoma grading

3.2 Feature Extraction using CNN

CNN is formed using linked layers of neurons like other neural networks with more complexity in the hidden layers. We use a pre-trained CNN model (Alexnet) which comprises 23 layers; convolution layers, max pooling layers, fully connected layers, softmax layer and output layer in the training stage as a feature extractor as shown in Table 1. The training data is passed through the CNN model to extract the features where the image higher-level features are available in the deeper layers of the CNN. These training features are extracted from the layer right before the classification layer 'fullyconnected7' and then fed into SVM. GPU is used in this stage rather than CPU to speed up the computation, as CNNs are highly computationally intensive.

Layer Name	**Layer Parameters Description**
'input'	227x227x3 images with normalization
'conv1'	96 11x11x3 convolutions with stride [4 4] and padding [0 0]
'relu1'	ReLU
'norm1'	Channel normalization
'pool1'	3x3 max pooling with stride [2 2] and padding [0 0]
'conv2'	256 5x5x48 convolutions with stride [1 1] and padding [2 2]
'relu2'	ReLU
'norm2'	channel normalization
'pool2'	3x3 max pooling with stride [2 2] and padding [0 0]
'conv3'	384 3x3x256 convolutions with stride [1 1] and padding [1 1]
'relu3'	ReLU
'conv4'	384 3x3x192 convolutions with stride [1 1] and padding [1 1]
'relu4'	ReLU
'conv5'	256 3x3x192 convolutions with stride [1 1] and padding [1 1]
'relu5'	ReLU
'pool5'	3x3 max pooling with stride [2 2] and padding [0 0]
'fc6'	4096 fully connected layer
'relu6'	ReLU
'fc7'	4096 fully connected layer
'relu7'	ReLU
'fc8'	1000 fully connected layer
'prob'	softmax
'classificationLayer'	cross-entropy

Table 1. Network Architecture

3.3 Train SVM using CNN Features and Predication

The features extracted from the previous step are fed to linear SVM classifier for training. Once the SVM model is learned, the test images are used to evaluate the performance of the proposed system by classifying the test data into either normal or pathological. SVM training algorithm with binary class predicts the labels of points in a test dataset by building a model for training dataset. Given a set of binary-labelled training vectors, SVMs learn a linear decision boundary to discriminate between the two classes. The resulted linear classification rule can be used to classify new test examples.

3. RESULTS AND DISCUSSION

We use NVIDIA GTX TITAN X 12GB GPU card with 3072 CUDA parallel-processing core to conduct the system implementation. MATLAB Toolboxes have been used to implement and train the proposed system. NVIDIA cuDNN library was used to accelerate the training and prediction processes. The confusion matrix that shows the prediction performance is shown in Figure 2. The proposed system was evaluated on 30% of 455 images where the network has achieved 88.2%, 85%, 90.8% of accuracy, sensitivity and specificity, respectively.

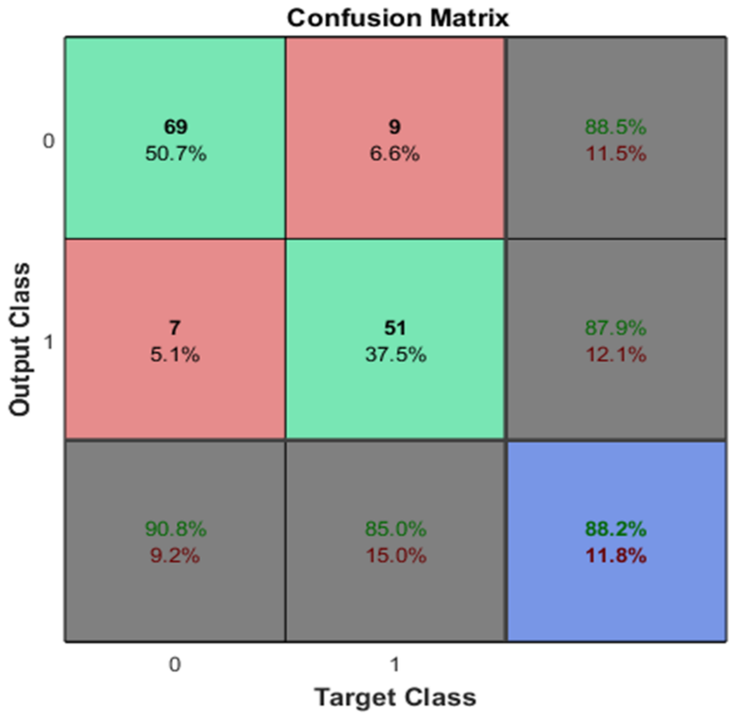

Figure 2. Prediction performance evaluation (0: normal; 1: pathological)

The obtained results seem very promising where the proposed system is implemented without applying any data augmentation or subsampling. Furthermore, the proposed approach does not require a prior knowledge of retinal image features like optic disc or blood vessels structure. Also, it is computationally simple where it doesn't require segmentation for the optic disc in order to get the features for the region of interest. Moreover, all the features that represent the region of interest are extracted automatically from the data itself by the CNN without the need for manual feature extraction.

4. CONCLUSIONS

In this paper, an automated approach to detect the glaucoma in retinal fundus images has been presented. The proposed method that is based on CNN demonstrated promising performance in diagnosing the glaucoma with considerably lower computational cost as compared to existing equivalent methods. The raw images are directly applied to the CNN without any enhancement or pre-processing except for mage resizing to minimise the computation cost. Key features of the disease are automatically extracted from stack layers of the filters convolved along the raw image.

The developed Glaucoma diagnosis prototype, however, is still open for further improvements and performance evaluation tests. For example, the dataset used in this study is limited to 455 high-resolution glaucoma and non-glaucoma images can be expanded to demonstrate the effectiveness of the proposed deep learning method in automatic feature learning and diagnosis of the glaucoma. In addition, a more efficient data augmentation and data sampling approaches can also be used to obtain higher performance measured in terms of accuracy, sensitivity, specificity and other metrics. These improvements and other are currently part of the on-going research of the authors.

REFERENCES

Acharya, U. R., Dua, S., Du, X., & Chua, C. K. (2011). Automated diagnosis of glaucoma using texture and higher order spectra features. IEEE Transactions on information technology in biomedicine, 15(3), 449-455.
Acharya, U. R., Ng, E. Y. K., Eugene, L. W. J., Noronha, K. P., Min, L. C., Nayak, K. P., & Bhandary, S. V. (2015). Decision support system for the glaucoma using Gabor transformation. Biomedical Signal Processing and Control, 15, 18-26.
Al-Bander B. , Al-Nuaimy W., A. Al-Taee M., Al-Ataby A. & Zheng Y. (2016). Automatic Features Learning Method for Detection of Retinal Landmarks. Proc. 9th International Conference on Developments in eSystems Engineering (DeSE ’2016), Liverpool & Leeds, England, 31st August – 2nd September 2016.
Bock, R., Meier, J., Nyúl, L. G., Hornegger, J., & Michelson, G. (2010). Glaucoma risk index: automated glaucoma detection from color fundus images. Medical image analysis, 14(3), 471-481.
Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge university press.
Damms, T., & Dannheim, F. (1993). Sensitivity and specificity of optic disc parameters in chronic glaucoma. Investigative ophthalmology & visual science, 34(7), 2246-2250.
Dua, S., Acharya, U. R., Chowriappa, P., & Sree, S. V. (2012). Wavelet-based energy features for glaucomatous image classification. IEEE Transactions on Information Technology in Biomedicine, 16(1), 80-87.
Fumero, F., Alayón, S., Sanchez, J. L., Sigut, J., & Gonzalez-Hernandez, M. (2011). RIM-ONE: An open retinal image database for optic nerve evaluation. In Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on (pp. 1-6). IEEE.
Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. In International Conference on Artificial Intelligence and Statistics (pp. 315-323).
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent Convolutional Neural Networks for Text Classification. In AAAI (pp. 2267-2273).
LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
Nayak, J., Acharya, R., Bhat, P. S., Shetty, N., & Lim, T. C. (2009). Automated diagnosis of glaucoma using digital fundus images. Journal of medical systems, 33(5), 337-346.
Quigley, H. A., & Broman, A. T. (2006). The number of people with glaucoma worldwide in 2010 and 2020. British journal of ophthalmology, 90(3), 262-267.
Ranzato, M. A., Huang, F. J., Boureau, Y. L., & LeCun, Y. (2007). Unsupervised learning of invariant feature hierarchies with applications to object recognition. In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on (pp. 1-8). IEEE.
Susanna Jr, R., & Vessani, R. M. (2009). Staging glaucoma patient: why and how. Open Ophthalmol J, 3(1), 59-64.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting.The Journal of Machine Learning Research, 15(1), 1929-1958.
Vapnik, V. N., & Vapnik, V. (1998). Statistical learning theory (Vol. 1). New York: Wiley.