Glaucoma-Deep: Detection of Glaucoma Eye Disease on Retinal Fundus Images using Deep Learning

Detection of Glaucoma by Abbas Q

Qaisar Abbas

College of Computer and Information Sciences,
Al Imam Mohammad Ibn Saud Islamic University (IMSIU),
Riyadh 11432, Saudi Arabia

Abstract—Detection of glaucoma eye disease is still a challenging task for computer-aided diagnostics (CADx) systems. During eye screening process, the ophthalmologists measures the glaucoma by structure changes in optic disc (OD), loss of nerve fibres (LNF) and atrophy of the peripapillary region (APR). In retinal images, the automated CADx systems are developed to assess this eye disease through segmentation-based hand-crafted features. Therefore in this paper, the convolutional neural network (CNN) unsupervised architecture was used to extract the features through multilayer from raw pixel intensities. Afterwards, the deep-belief network (DBN) model was used to select the most discriminative deep features based on the annotated training dataset. At last, the final decision is performed by softmax linear classifier to differentiate between glaucoma and non-glaucoma retinal fundus image. This proposed system is known as Glaucoma-Deep and tested on 1200 retinal images obtained from publically and privately available datasets. To evaluate the performance of Glaucoma-Deep system, the sensitivity (SE), specificity (SP), accuracy (ACC), and precision (PRC) statistical measures were utilized. On average, the SE of 84.50%, SP of 98.01%, ACC of 99% and PRC of 84% values were achieved. Comparing to state-of-the-art systems, the Nodular-Deep system accomplished significant higher results. Consequently, the Glaucoma-Deep system can easily recognize the glaucoma eye disease to solve the problem of clinical experts during eye-screening process on large-scale environments.

Keywords—Fundus imaging; glaucoma; diabetic retinopathy; deep learning; convolutional neural networks; deep belief network

I. Introduction

Glaucoma is the main reason of visual disability across the world [1] and it has no cure. At an early stage, if it is not detected then it can definitely be the cause of permanent blindness. There are cures to prevent the vision loss if it is recognized at an early stage. Since, it is a salient chronic eye disease that develops permanent blindness. In recent years, the glaucoma is rapidly increasing even in urban regions. By the year of 2020 [2], it was estimated that it might be affected 79 million people in the world. Thus, it is important to do eye screening for detecting of glaucoma. The eye screening process is expected to tedious and time consuming task due to checkup of every individual patients, which is generally large.

For ophthalmologists to do better eye screening process, the computer-aided diagnostics systems (CADx) [3] are developed to provide a cost effective solution to the patients. As a result, the automated CADx screening systems have capability to reduce time and effort wasted on the analysis of glaucoma eye disease. It is also important to develop CADX systems using image analysis [4] for clinical experts to differentiate between normal and glaucoma retinal images because it is difficult for ophthalmologists to make this discrimination as shown in Fig. 1. As shown in this figure, the CUP boundary is affected due to glaucoma eye disease. This effect can be determined through the cup to disc ratio (CDR). According to the literature survey, the classification of retinal fundus images into normal and glaucomatous stages is a significance problem due to large population of screening.

There is another indicator for glaucoma eye disease that is known as peripapillary atrophy (PPA). In case of PPA affect, a change in intensity adjoining the disc boundary can be observed clearly. The Fig. 2 depicts the effect of PPA on glaucomatous retinal fundus region-of-interest (ROI) image. It can observe from this figure that the PPA appears as a variation in the disc boundary and similar to rim thinning.
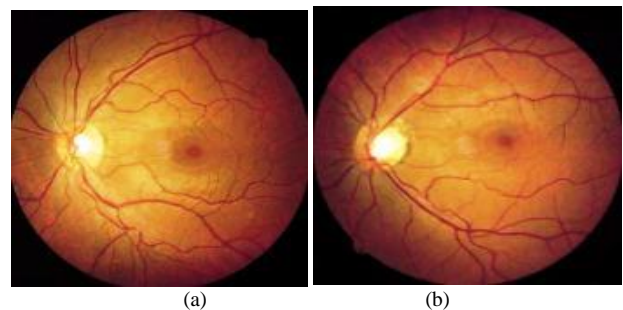

| (a) | (b) |

Fig. 1. An example of visually presented glaucoma: (a) normal retinal image, (b) glaucoma.
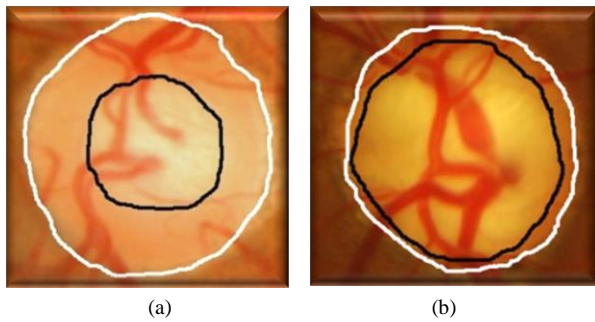
(a)   (b)

Fig. 2. Sample of normal (a) fundus image and glaucomatous (b) OD region-of-interest (ROI) image showing the effect of rim dilution due to the glaucoma eye disease.

In the past studies, the authors presented a solution to automatic detect and classify glaucoma through the cup to disc ratio (CDR) and extraction of texture features. Accordingly, it is very difficult for the CADx system to segment the CUP and OD regions in a robust manner. To extract features, there is a need to segment the OD and CUP regions so it needs lots of image processing techniques. After defining image features, there is also need a domain-expert knowledge to select most discriminative features. Therefore, the detection of glaucoma is a challenging task for ophthalmologists and CADx systems.

In contrast to segmentation-based approaches, the authors detect glaucoma eye disease through image features around OD region. In those approaches, the authors pre-assumed that the morphological changes in OD caused by the disease, can be encoded using the statistical image features, so removing the need to accurately recognize cup and disc borders. Hence in this paper, a deep-learning base features were automatically extracted, selected and classified using a modern deep-learning classifiers without the need of understanding about domain-expert knowledge in image processing tasks.

The rest of the paper is organized as: In Section II, the related work is presented. In the next Section III, the utilize dataset and the propose methods are discussed in detail whereas in Section IV, the experiments are described to analyze the impact on detection performance. Finally, the paper is concluded in Section V.

II. RELATED WORK

Glaucoma is detected through segmentation-based and features learning-based approaches. However in this paper, the features learning-based approaches are briefly described in the subsequent paragraphs. In particular, the modern deep-learning algorithms utilized in the past studies to automatically detect glaucoma eye disease are mentioned here.

In [5], the authors developed a convolutional neural network (CNN) architecture to automate the detection process of glaucoma. The CNN model is having multilayer architecture which belongs to the class of deep-learning algorithms. In that study, the authors make a clear differentiation between glaucoma and non-glaucoma patterns through hierarchical representation of features by CNN model. They used six multilayers of CNN model and divided into four convolutional layers along with two fully-connected layers. They performed experiments on ORIGA and SCES datasets and achieved

0.8321 and 0.887 of AUC values, respectively for detection of glaucoma eye disease. The authors reported that this developed system obtained accomplished better results compared to state-of-the-art systems. The smallest version of this paper [6] was also represented in conference.

Another version of deep-learning (DL) algorithm was developed in [7] to detect glaucoma eye diseased through the extraction of different features, such as 52 total deviation, mean deviation, and pattern standard deviation values. The authors utilized DL classifier such as a deep feed-forward neural network (FNN). However, the authors combined this DL classifier with other old machine learning classifiers such as random forests (RF), gradient boosting, support vector machine, and neural network (NN). Therefore, the authors presented deep ensemble solution for detecting of glaucoma disease. The authors reported that the 92.5% of AUC value was obtained through a deep FNN classifier.

There was another different approach presented in [8]. The authors utilized the structure of optic disc (OD) to analyze the glaucoma and other eye diseases such as retinal vein occlusion. The authors claimed that the OD is an essential to local the macula and main vascular arcade from retinal fundus images. In the past studies, the authors used OD properties and spatial relationship between OD and the main vascular arcade to diagnostic the eye disease. However in that paper, the authors used the structure of OD abnormalities and deep-learning algorithm to determine the glaucoma eye disease.

In [9], a novel algorithm was developed to detect glaucoma using support vector machine (SVM) instead of using advanced deep-learning algorithm along with hybrid feature set. The authors also detected color and texture features from retinal fundus images and some other properties of CDR of OD/CUP ratio to determine the severity-level of glaucoma eye disease. This developed approach was evaluated on 100 patients and they obtained the average sensitivity and specificity of are 100 and 87 %, respectively.

The authors developed an automatic solution for the detection of glaucoma by using features [10] learning through deep-learning algorithm on retinal fundus images. The authors utilized CNN model to learn the features with linear and nonlinear activation function. They used glaucoma and non-glaucoma patterns to differentiate for training of CNN model. They performed experiments on the ORIGA and SCES datasets and reported 0.838 and 0.898 of AUC values, respectively. In [11], the authors performed image processing techniques to automatic detection of glaucoma eye disease through ensemble machine learning classifiers. In that paper, the authors presented a system to segment optic disc, extract of texture feature in different color models and classified them by multilayer perceptron (MLP) model.

To improve OCT-based glaucoma diagnosis in [12], the authors used two ML learning algorithms, such as ANN and SVM with input on retinal nerve fiber layer thickness (RNFLT). In that study, the authors researched on 90 healthy persons and 62 glaucoma patients. It concluded that there was no statistically significant differences between ANNs and SVMs and reported the best ROCs for both ANN (0.982, 95% CI: 0.966–0.999) and VM (0.989, 95% CI: 0.979–1.0).
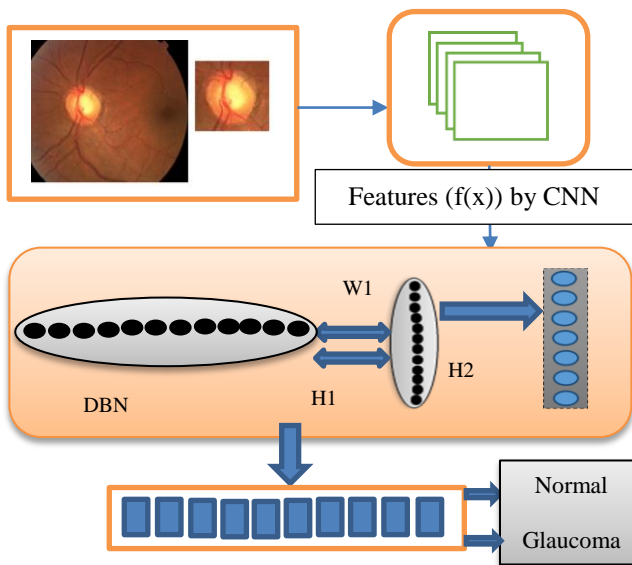
Fig. 3.   Systematic flow diagram of proposed Glaucoma-Deep system tested on 1200 retinal fundus images.

TABLE I.   SELECTED 1200 RETINAL FUNDUS IMAGES FOR DIAGNOSIS OF GLAUCOMA EYE DISEASES FROM DIFFERENT ONLINE AND PRIVATE SOURCES

No.	Glaucoma dataset		
	Dataset Name	*Normal*	*Glaucoma*
1	DRIONS-DB	50	60
2	sjchoi86-HRF	300	101
3	HRF-dataset	15	15
4	PRV-Glaucoma	235	424
	Total	**600**	**600**

III.   METHODOLOGY

The systematic flow diagram of proposed Glaucoma-Deep system is visually represented in Fig. 3. The Glaucoma-Deep system contains four main steps such as automatic detection of region=of-interest (ROI), extraction of deep features, optimization of features and classification phases. These steps are explained in detail in the upcoming subsections.

A.   Acquisition of Glaucoma-datasets

To test and evaluate the Glaucoma-Deep system, the dataset of 1200 retinal fundus images (normal of 600 and glaucoma of 600) were acquired from different public and private resources. The information about these data sources is mentioned in Table 1. The first 110 images were collected from online web source named as DRIONS-DB [13]. In DRIONS-DB dataset, there were 110 fundus images that were belonging to the Ophthalmology Service at Miguel Servet Hospital, Saragossa (Spain). Each image size was 600 x 400 and 8 bits/pixel and the ground truth was provided by two medical experts in the form of papillary contour points. Another publicly available dataset sjchoi86-HRF (High Resolution Fundus) [14] was used to detect and diagnose Glaucoma eye disease.

The sjchoi86-HRF dataset was normal of 300 and glaucoma of 101 retinal fundus images. The HRF-dataset was also used to automatic diagnosis the glaucoma disorder [15]. In this study, the 15 retinal fundus images for each category are considered in both normal and glaucoma. Instead of publically available data sources, the private dataset of PRV-Glaucoma was also collected from a central hospital. The PRV-Glaucoma dataset was contained 235 normal and 424 glaucoma images.

A clinical expert was requested to make the difference between these two categories and validate and provide ground truth about the image level annotation as normal and confirm case of glaucoma.

B.   Region-of-interest (ROI)

The color fundus images are acquired in the form of RGB image and the high intensity region-of-interest (ROI) image is extracted automatically from the green plane. This ROI region contains the OD and CUP regions. From each image, the ROI of size (300 x 300) pixels with radius 150 is utilized to extract deep features by using deep-learning algorithms.

C.   Extraction of deep features

The best variants of deep-learning algorithms have many applications in practice such as in bioinformatics, computer vision and medical image processing. For example, the convolutional neural network (CNN) [16] is having the framework of deep-learning algorithm used to extract the pixel-level features in a multi-layer form that are defined automatically from an image. In the past few years, the CNN model achieved best performance compare to manually tune classification algorithms. Yet, there are still many properties of CNNs that researchers are determined to recognize.

In recent years, the CNNs model is used to understand the feature maps represented by an image. In deep-learning models, the CNNs is describing the best feature-map without manually defined domain-expert methodologies of image processing algorithms. The features defined by CNN model is known as deep-invariant features that are acquired from different layers also known as visual features. The output layer is responsible for generating the final features-map that can be used in the training process for classification tasks. However, the features-map generated by CNN model is not optimized. In particular, a pre-trained CNN model is having capability to extract features that are invariant and achieves better results compared to handcrafted features.

Accordingly, the CNN model is used to extract the deep invariant features from the segmented region-of-interest (ROI) of size (300 x 300) pixels and then generate a features-map in different layers. In these experiments, the dataset of 1200 ROIs is used and therefore, the CNN model generated the huge amount of features-vector that is further needed to optimize.

D.   Optimization of deep features

These invariant features (IFs) that are extracted through convolutional neural network (CNN) deep-learning model are optimized through supervised deep-belief network (DBN) [17] deep-learning algorithm. The CNN algorithm is just used to describe the features in distributed fashion organize and then the DBN model is used to optimize the features for best

classification tasks. The DBN model is used to assist the final classification decision between normal and glaucoma eye disease on 1200 retinograph images.

For the optimization purposes, the three layers from deep belief nets (DBNs) are utilized along with step. In DBNs, the first layer is already trained from 40% ROI images contained normal and glaucoma category. On 1200 ROI images, the training of DBN is performed one layer at a time through learning the features that provided the bases for the upcoming layer and so on. In the DBNs model, there is also another fine-tuning step which is employed to improve the performance of features selection step. As shown in Fig. 3, there are many hidden (H1) layers of DBNs model that are stacked together to show a deep-neural network in a tree-like features representation of the training dataset.

E. Classification of deep features

For classification of deep and optimize features, the softmax linear classifier is used to best describe the differentiation between normal and glaucoma ROI images. The softmax classifier is generally used to recognize the classes in deep-learning algorithms. The softmax linear classifier is used in this paper to find out the differentiation between these two classes.

IV. EXPERIMENTAL RESULTS

The Glaucoma-Deep system was tested on a personal Laptop with an Intel core i7 CPU @ 3.35 GHz and 8 GB of random access memory (RAM). The Glaucoma-Deep program was implemented in MATLAB® 2016b and the results were also statistically calculated through MATLAB®.

The dataset of 1200 retinal fundus images including normal of 600 and glaucoma of 600 were acquired from different public and private resources to test and evaluate the performance of Glaucoma-Deep system. The information about these data sources is mentioned in Table 1. From 1200 retinal fundus images, the region-of-interest (ROI) was automatically segmented that serves as a feature extraction phase. The manual prediction was used to train and test the deep-learning classifiers that provided by an expert ophthalmologist as a ground truth. The proposed Glaucoma-Deep algorithm is compared to its corresponding ground truth images in two category problems, such as normal and glaucoma eye.

The statistical measures, such as sensitivity (SE), specificity (SP), accuracy (ACC) and precision (PRC) metrics were performed for achieving the numerical results. The performance of proposed Glaucoma-Deep algorithm is analyzed with the following parameters that are mathematically described in the following equations:

$$\text{Sensitivity (SE)} = TP/(TP+FN)) \qquad (1)$$

$$\text{Specificity (SP)} = TN/(TN+FP)) \qquad (2)$$

$$\text{Accuracy (ACC)} = (TP+TN)/(TP+FN+TN+FP)) \qquad (3)$$

$$\text{Precision (PRC)} = TP/(TP+FP) \qquad (4)$$

These performance parameters are calculated on the 1200 ROIs retinal fundus images based on ground truth and the results obtained by Glaucoma-Deep system. The achieved

results are listed in Table 2, where, TP denotes true positive, FP denotes false positive, FN is false negative and TN is true negative. True positive refers to the correctly identified normal and glaucoma ROI image. True negative refers to the wrongly identified normal and glaucoma classes. Whereas, false positive refers to the correctly identified classes and false negative refers to the wrongly identified classes. The entire Glaucoma-Deep algorithm was run on the 1200 retinal ROI images and results for normal and glaucoma categories were achieved. The results are reported in Table 2 based on 10-fold cross validation test and the dataset is divided into 40% for training and 60% for testing to check its viability in a large-scale environments.

Table 2 shows the performance comparisons of the proposed Glaucoma-Deep methodology in terms of sensitivity (SE), specificity (SP), accuracy (ACC) and precision (PRC) statistical measures. On average, the SE of 84.50%, SP of 98.01%, ACC of 99% and PRC of 84% values were achieved. As displayed in Table 2, the Glaucoma-Deep system obtained 85.17% of SE, 98.45% of SP, 99.00% of ACC and 82.01% of PRC in case of normal retinal eye. However, the Glaucoma-Deep system reported 83.50%, SP of 99.25%, ACC of 99.01% and 87.05% of PRC values in case of glaucoma eye disease. These results are comparable to state-of-the-art systems; the Nodular-Deep system is accomplished significant higher results. Consequently, the Glaucoma-Deep system can easily recognize the glaucoma eye disease to solve the problem of clinical experts during eye-screening process in a large-scale environment.

TABLE II. RESULTS OBTAINED BY GLAUCOMA-DEEP SYSTEM ON THE SELECTED 1200 RETINAL FUNDUS IMAGES FOR DIAGNOSIS OF GLAUCOMA EYE DISEASES

No.	Detection of Glaucoma eye disease				
	Category	*SE [a]*	*SP [b]*	*ACC [c]*	*PRC [d]*
1	Normal	85.17	98.45	99.00	82.01
2	Glaucoma	83.50	99.25	99.01	87.05
	Average	**84.50%**	**98.01%**	**99%**	**84%**

[a] Sensitivity, [b] Specificity, [c] Accuracy and [d] Precision

TABLE III. PERFORMANCE COMPARISION RESULTS WITH STATE-OF-THE-ART DEEP-LEARNING METHODS FOR DIAGNOSIS OF GLAUCOMA EYE DISEASES

Cited	Detection of Glaucoma eye disease		
	Methodologies	*Accuracy*	*Year*
[5]	CNN [a]	83.00%	2015
[7]	FFN [b]	92.00%	2016
[9]	SVM [c]	87.00%	2016
[11]	CNN [a]	90.00%	2016
Glaucoma-Deep (CNN, DBN [d], Softmax)		**99.0%**	**2017**

[a] Convolutional neural network, [b] Feed-forward neural network, [c] Support vector machine, [d] Deep-belief network

Table 3 reported the comparison results of state-of-the-art systems for classification of normal and glaucoma eye disease in the recent years. As displayed in this table, the proposed

Glaucoma-Deep system achieved higher statistical values compared to other four state-of-the-art computational methods for recognizing glaucoma eye disease. The reason behind is that the integration of CNN, DBN and softmax deep-learning classifiers can recognize the glaucoma without doing automatic or manual segmentation of OD and CUP area.

These experimental results indicate that the proposed Glaucoma-Deep system can be used to diagnosis glaucoma eye disease and assists the ophthalmologists to reduce the pressure of very large screening environment. The current studies focused on either local or global feature-based or segmentation of OD/CUP areas approaches. No current efforts have been made to integrate the deep-learning algorithms to combine the strengths of both the approaches. In this paper, the advanced form of automatic system is proposed to diagnose the glaucoma eye disease through deep-learning techniques.

The deep-learning based classification of glaucoma eye disease is provided best performance as reported in Table 2. In the previous approaches, the authors focused on segmenting and measuring the OD and CUP ratios. However in real, the segmentation of OD and CUP regions is a challenging task for automatic computerize system. Therefore, the deep-learning based methods may improve the performance of eye screening process and insist the ophthalmologists to correctly diagnosis the eye disease. The superior performance of this proposed the Glaucoma-Deep system which is achieved over CDR-based classification approaches, shows the importance of expanding the scope of analysis to multiple factors in glaucoma assessment.

V. CONCLUSIONS

In this paper, the advanced machine deep-learning algorithms are used to diagnose Glaucoma based on retinal fundus images without using segmentation-based hand-crafted features. To develop Glaucoma-Deep system, the convolutional neural network (CNN) unsupervised architecture is applied on 1200 images to extract the features through multilayer from raw pixel intensities. Afterwards, the deep-belief network (DBN) model is used to select most discriminative deep features based on the annotated training dataset. At last, the classification decision is performed by softmax linear classifier to differentiate between glaucoma and non-glaucoma images. To evaluate the performance of Glaucoma-Deep system, the sensitivity (SE), specificity (SP), accuracy (ACC), and precision (PRC) statistical measures were utilized. On average, the SE of 84.50%, SP of 98.01%, ACC of 99% and PRC of 84% values were achieved. Comparing to state-of-the-art systems, the Nodular-Deep system is accomplished significant higher results. Consequently, the Glaucoma-Deep system can easily recognize the glaucoma eye disease to solve the problem of clinical experts during eye-screening process in a large-scale environment.

Glaucoma-Deep system offers great promising results but still, there are certain areas open to do further research. In the future, the Glaucoma-Deep system will be tested on large-scale glaucoma-datasets to test its applicability in practice.

REFERENCES

[1] E. Dervisevic, S. Pavljasevic, A. Dervisevic, S.S. Kasumovic, "Challenges In Early Glaucoma Detection," Med Arch., vol. 70, pp. 203–207, June 2016.

[2] R. Bock, J. Meier, L. G. Nyl, G. Michelson, "Glaucoma risk index: automated glaucoma detection from color fundus images," Medical Image Analysis, vol. 14, pp. 471–481, 2010.

[3] J.E.W. Koh, U. Rajendra Acharya, Y. Hagiwara, U. Raghavendra, J. Hong Tan, S. Vinitha Sree et al, "Diagnosis of retinal health in digital fundus images using continuous wavelet transform (CWT) and entropies," Computers in Biology and Medicine, vol. 84, pp. 89–97, May 2017.

[4] A. Anton, M. Fallon, F. Cots, M. A Sebastian, A. Morilla-Grasa et al., "Cost and Detection Rate of Glaucoma Screening with Imaging Devices in a Primary Care Center," Clinical Ophthalmology, vol. 11, pp. 337–346, May 2017.

[5] X. Chen, Y. Xu, D. Wing Kee Wong, T. Yin Wong, J. Liu, "Glaucoma detection based on deep convolutional neural network," IEEE Eng Med Biol Soc, pp. 715–718, Annual Conf., August 2015.

[6] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong and J. Liu, "Glaucoma detection based on deep convolutional neural network," International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, pp. 715–718, 37th Annual Conference, 2015.

[7] R. Asaoka, H. Murata, A. Iwase, M. Araie, "Detecting Preperimetric Glaucoma with Standard Automated Perimetry Using a Deep Learning Classifier," Ophthalmology, vol. 123, pp. 1974–1980, September 2016.

[8] H.S. Alghamdi, H.L. Tang, S.A. Waheeb, T. Peto, "Automatic Optic Disc Abnormality Detection in Fundus Images: A Deep Learning Approach," OMIA3 (MICCAI 2016), pp. 10–17, Athens, October 2016.

[9] A. A. Salam, T. Khalil, M.U. Akram, A. Jameel and I. Basit, "Automated detection of glaucoma using structural and non structural features," Springerplus, vol. 5, pp. 1–22, 2016.

[10] X. Chen, Y. Xu, S. Yan, D. Wong, T. Wong, and J. Liu, "Automatic Feature Learning for Glaucoma Detection Based on Deep Learning," MICCAI (3), vol. 9351 of Lecture Notes in Computer Science, pp. 669–677, 2015.

[11] M. Claro, L. Santos, W. Silva Fl´avio Ara´ujo, N. Moura, "Automatic Glaucoma Detection Based on Optic Disc Segmentation and Texture Feature Extraction," Clei Eletronic Journal, vol. 19, pp. 1–10, August 2016.

[12] D. Bizios, A. Heijl, J. Leth Hougaard and B. Bengtsson, "Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT," Acta Ophthalmologica, vol. 88, pp. 44–52, January 2010.

[13] E.J. Carmona, M. Rincón, J. García-Feijoo and J. M. Martínez-de-la-Casa, "Identification of the optic nerve head with genetic algorithms," Artificial Intelligence in Medicine, vol. 43, pp. 243–259, 2008.

[14] sjchoi86-HRF dataset: https://github.com/sjchoi86/retina_dataset/tree/master/dataset. [Access date: 26/1/2017].

[15] High-Resolution Fundus (HRF) Image Database. https://www5.cs.fau.de/research/data/fundus-images/, [access date: 2/1/2016].

[16] G.W. Yang, and J. Hui-Fang, "Multiple Convolutional Neural Network for Feature Extraction," International Conference on Intelligent Computing, pp. 104–114. Springer International Publishing, 2015.

[17] A. Qaisar, "DeepCAD: A Computer-Aided Diagnosis System for Mammographic Masses Using Deep Invariant Features," Computers, vol.5,pp.1–15,2016.