Glaucoma Detection Based On Shape Features and SMOTE Algorithm

Arwa A. Gasm Elseid and Alnazier O. Hamza

Abstract---Glaucoma is a chronic eye condition in which the optic nerve is progressively damaged. As the disease progresses, more optic head damage due to loss of peripheral vision and a resultant in gradually vision loss. Glaucoma progression precedes some structural damage in the retina are signs of glaucoma appear as changes in size, structure, shape, and color of the optic disc and optic cup, which suffer from the subjectivity of human due to experience, fatigue factor etc. There are increasing demands for medical image-based Computer-Aided Diagnosis (CAD) systems based on fundus image for glaucoma detection, thus the human mistakes, other retinal diseases like Age-related Macular Degeneration (AMD) affecting in early glaucoma detection, and the existing medical devices like Optical Coherence Tomography (OCT) and Heidelberg Retinal Tomography (HRT) are expensive. This paper proposes a novel algorithm by extract 13 shape features from disc and cup. Next, the best features selected using the student t-test method and balanced using SMOTE algorithm. The evaluation of the proposed algorithm is performed using a RIM_ONE database, the average accuracy 91.3%, maximize the area under the curve (AUC) 0.92, using Ensembles RUSBoosted tree. Future works suggested designing a complete, automated CAD system using a different type of features.

Keywords---Digital Fundus Image, Glaucoma, Ensembles RUSBoosted Classifier, SMOTE Algorithm, Shape Features.

I. INTRODUCTION

SHAPE features are important in the CAD system because they provide an alternative to describing an object, using its most important characteristics. Shapes features example are: center of Gravity/centroid, axis of least inertia, digital bending energy, eccentricity, circularity ratios, elliptic variance, rectangularity, convexity, solidity, Euler number, profiles, and whole area ratio, thus the glaucoma is dangerous an ocular disease and the second cause of blindness in the world with about 60 million glaucomatous cases [1], and it is responsible for 5.2 million cases of blindness with more than 90% of the patients unaware of the condition mentioned [2] early diagnosis and managed are needed. In glaucoma, the optic nerve is progressively damaged with the disease progresses, and resultant in gradually blindness. As the optic nerve damage is Irreversible, early detection of glaucoma is essential to prevent from blindness.

Generally, there are three ways for detecting glaucoma: assessment of the abnormal visual field, assessment of intraocular pressure (IOP) and assessment of optic nerve damage. Visual field test requires special equipment present only in hospitals and is a subjective examination needs patient cooperation and test complete. Sometime, and the information obtained may not be reliable for kids or ICU patient.

In the second method, a large proportion of glaucoma patients have a normal level of IOP like normal tension glaucoma. Thus, IOP measurement some time-specific nor sensitive for early glaucoma. The assessment of optic nerve damage is the best compared to the other two methods [3]. The optic nerve can be assessed by trained specialists or through 3D imaging techniques such as Heidelberg Retinal Tomography (HRT) and Ocular Computing Tomography (OCT). However, optic nerve evaluation by specialists is subjective and the availability of HRT and OCT equipment is limited due to the high cost involved. That lead to there is still no systematic and economic way of detecting glaucoma. There is a need for automatic and economic system for detection of glaucoma in an accurate way, using the digital color fundus image as Figure [1]

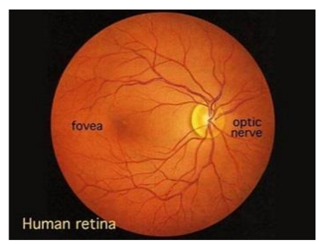


Fig.1 shows an example of the Fundus Image obtained from Fundus Camera [4].

Features extraction technique in fundus images is classified based on the type of features. The types of features are divided into two groups namely morphological and non-morphological [5], the morphological features based on optic disc and optic cup segmentation then calculate features like RNFL, CDR, PPA and ISNT rule Sharanagouda at [6] proposed a method unlike past works which relies on a single color channel for extracting the Optic Disk (OD) and Optic Cup (OC) used in CDR calculation, they propose a novel combined color channel and ISNT rule-based automated glaucoma detection, and find that the proposed method betters single channel based giving an overall efficiency of 97%.

Babu at [7] proposed for the measurement of CDR as a sign for the diagnosis of glaucoma and 90% accuracy is obtained.

Manuscript received on November 09, 2018, review completed on December 11, 2018 and revised on December 13, 2018.

Arwa Ahmed Gasm Elseid is with the Department of Biomedical Engineering, Sudan University of Science and Technology, Khartoum, Sudan, Phone Number: 00966536472733, E-Mail: roro19842016@gmail.com.

Alnazier Osman Hamza is with the Department of Radiology, Medical Sciences and Technology University, Khartoum, Sudan. E-Mail: alnazierhamza@gmail.com.

Digital Object Identifier: DIP102018003.

Yuji at [8] proposes a method to measure the cup-to-disc ratio using a vertical profile on the optic disc. Via these steps: the blood vessels were removed, the edge of optic disc detected by the canny edge detection filter, extract twenty profiles around the disc center in the blue channel, the profile was smoothed by averaging these profiles, and the cup area calculated by thresholding technique. Lastly, the CDR ratio was calculated using seventy-nine images, contain twenty-five glaucoma images, and obtained a sensitivity of 80% and a specificity of 85%.

Non-Morphological features are whole image features without segmentation like color, shape, and texture are the type of features that captured from the existing image. As Bock at [9] developed the previous research by modifying the feature extraction method used. The feature extraction methods were used such as the value of intensity pixel rows, FFT and B-spline methods, then reduced the size of the features dimensions using PCA. This research made the Glaucoma Risk Index (GRI) system using two-stage in the first stage classification combine different features (fraw, ffft, fspline), then in the second stage, the features normalized. The algorithm tested with 575 images (239 glaucoma and 336 normal images) and achieved an accuracy of Area under Convergence (AUC) 80% and GRI 88%.

The use of texture feature extraction has done by Karthikeyan [10] using two methods: histogram and Gray Level Co-occurrence Matrix (GLCM). The results obtained using the Sequential Forward Floating Selection (SFFS). Experimental fundus images obtained using the fundus camera Topcon TRC50 EX with image size 1900x1600 pixels derived from the Aravind eye hospital, Madurai. The results of the sensitivity, specificity, and accuracy reached respectively 96%, 94% and 95% with 32 GLCM quantize level.

Therefore, this study provides an expert system for real-time fundus image analysis, which gives an opportunity to improve the image interpretation. The proposed system will improve the diagnosis of glaucoma and therefore, it will minimize the miss detection rate and help in early diagnosis and treatment, which can significantly improve the chance of managing the glaucoma disease.

II. Material and Method

Data Acquisition from An Open Retinal Image Database for Optic Nerve Evaluation (RIM-ONE) , which contain 169 fundus ONH images, where the images are classified as follows: normal 118 images, early glaucoma 12 images, moderate glaucoma 14 images, severe glaucoma 14 images and ocular hypertension (OHT) 11 images. In this study, the RIM-ONE database is used to test the algorithm validation [11]. Figure [2] an example of normal and abnormal glaucomatous images taken from the RIM-ONE database.

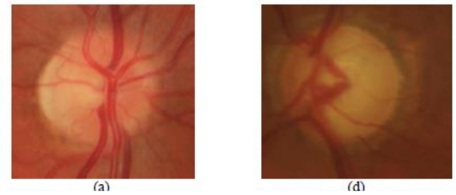


Fig.2 an example of RIM-ONE Database Images: (a) Normal Image, (d) Glaucoma Image (Abnormal)

Proposed algorithm provides an automated glaucoma detection computer-aided system used for early diagnosis of glaucoma patients with high accuracy. The algorithm takes a preprocessed fundus image and segment optic cup and optic disc followed by extracting shape features and classify it. Result used to classify the image as glaucoma and healthy. Figure [3] illustrates the complete methodology.

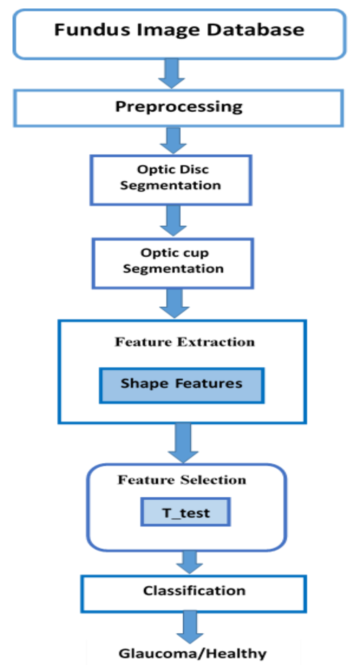


Fig.3 Proposed Methodology

A. *Preprocessing Step*

Image pre-processing steps are, resizing the image to 256 X 256 pixels so that it has the specified number of rows and column to reduce computational time, denoising stage by median filters (5×5) Before any procedures made, red channel has been extracted because it appears as good boundary and less blood vessel, which is affecting segmentation accuracy as to facilitate intensity analysis, as in figure [4].

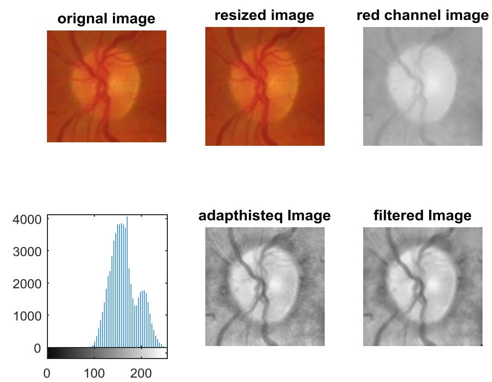


Fig.4 shows the Steps of An Image Preprocessing.

B. Optic Disc Segmentation

The Optic Disc Segmentation algorithm developed and tested in the rim-one public database. The proposed method can be divided into 4 steps, the **first step** is vessel remove to get accurate segmentation, and it was done by the morphological operation. The **Second step** is applying thresholding level 180 to segment disc region from the red channel which appears the best contrast to the disc region, after that in the **third step** is the boundary was smoothed and cleared from unconnected object and binary image obtained at the **final step** circle image was construct based on the radius and center of the detected region to minimize segmentation errors resulting from the main blood vessels and PPA surrounded the optic disc, using the formula and the result shown in figure[5]

$$(x - h)2 + (y - k)2 = r2 \quad (1)$$

Where r= the radius from segmented object, h, k = the center from segmented.

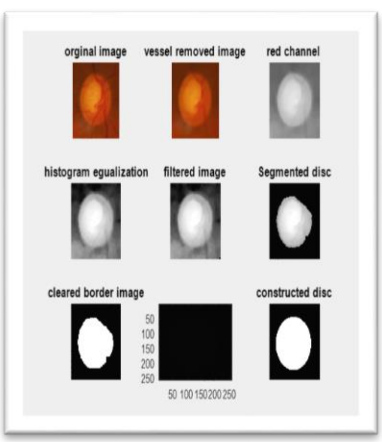


Fig.5 shows Fundus Image OD Segmentation Steps

C. Cup Segmentation

The optic cup segmentation method proposed, mainly depend on optic disc thresholding at level 240, which is the best level differences between the disc and cup parts then clear border, smoothing boundary and binaries the final image,

Fixed thresholding is of the form:

$$\mathrm{OC}(x, y) = \sum_{1 \;\; f(x,y) >= t}^{0 \;\; f(x,y) < t} \quad (2)$$

Where t is thresholding level= 240.

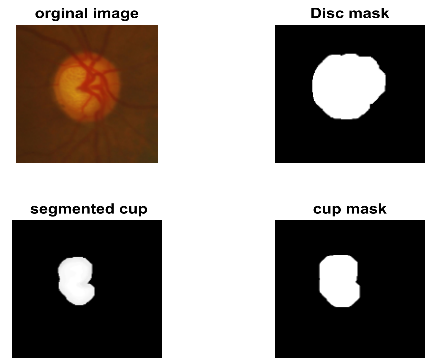


Fig.6 shows a Cup Segmentation and Binarization from the Disc Part.

D. Shape Feature Extraction Step

Glaucoma can be diagnosed by examination of size, structure, shape, and color of the optic nerve head. Based on this theory a shape features were applied to disc and cup to classify glaucoma from non-glaucoma cases, the features are, in this paper measures, several properties describe shape for each disc and cup within an image via Matlab built-in function (region prop) [12].

Region props compute Area, Euler Number, Centroid, Filled Area, Filled Image, Extent, Major Axis Length, Minor Axis Length, Solidity and Perimeter, these measurements applied for disc and cup masks. For measuring these features a binary image was obtained

$$f(x, y) = \begin{cases} 1 \; if (x, y) \in D \\ 0 \; otherwise \end{cases} \quad (3)$$

a) The Centroid

The centroid determines the object center points, which illustrate by this formula:

Cntroid$(\vartheta x, \vartheta y)$ is:

$$\vartheta x = \frac{1}{N} \sum_{i}^{N} = 1xi \quad (4)$$

$$\vartheta y = \frac{1}{N} \sum_{i}^{N} = 1yi$$

N is the number of point in the shape

b) Eccentricity

Eccentricity is the measure of aspect ratio, it’s the ratio of the length of the major axis to the minor axis which calculated by principal axes method or minimum bounding rectangular box.

c) Solidity

The proportion of the pixels in the convex hull that is also in the region returned as a scalar. Computed as Area/Convex Area. Solidity describes the extent to which the shape is convex or concave and it is defined

$$Solidity = A_s / H \quad (5)$$

Where, As is the area of the shape region and H is the convex hull area of the shape. The solidity of a convex shape is always 1.

d) Area

The actual number of pixels in the region. The equivalent diameter of a circle with the same area as the region returned as a scalar. Computed as

$$\text{Area} = (4*Area/pi) \quad (6)$$

e) Major Axes

Length of the major axis of the ellipse that has the same normalized second central moments as the region in pixels.

f) Minor Axes

Length of the minor axis of the ellipse that has the same normalized second central moments as the region in pixels.

g) Extent

The ratio of pixels in the region to pixels in the total bounding box returned as a scalar. Computed as the Area divided by the area of the bounding box Perimeter. The perimeter is the length of the entire outside boundary of a region, Distance around the boundary of the region by calculating the distance between each adjoining pair of pixels around the border of the region.

E. Feature Selection Step

Feature selection is the process of selecting the more relevant features for use in model construction and CAD building.

The objective of this step is three-fold: improving the prediction performance of the predictors, providing faster and more cost-effective predictors, and providing a better understanding of the underlying process that generated the data.

In this paper T-test methods were applied with $p<0.05$ that result in 5 features, where a t-test's statistical significance used to compare the means of glaucoma and healthy samples.

F. Classification Step

Many classification techniques are used like SVM, KNN, and ensembles which use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone. To get more model accuracy Ensemble learning used to improve machine learning results by combining several models that is ensembles RUS Boost an excellent technique for learning from imbalanced data (one class outnumbers other class by a large proportion), where Synthetic Minority Oversampling Technique (SMOTE) algorithm creates artificial data based on feature space (rather than data space) similarities from minority samples. Therefore, it generates a random set of minority class observations to shift the classifier learning bias towards minority class, to generate artificial data, using bootstrapping and k-nearest same technique used before at [13], [14], and [15].

III. Results and Discussion

The performance of the proposed glaucoma detection method is evaluated in terms of sensitivity, specificity, accuracy, and the ROC curve. Those values are defined as follows:

$$Sensitivity = TP/(TP+FN) \quad (7)$$

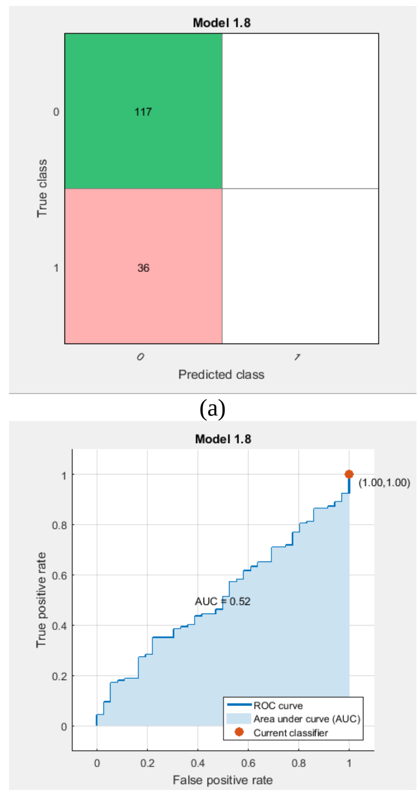
$$specificity = TN/(FP+TN) \quad (8)$$

$$Accuracy = TP + TN/(TP+TN+FP+FN) \quad (9)$$

The ROC curve: shows true positive rate versus false positive rate for the currently selected trained classifier.

The marker on the plot shows the performance of the currently selected classifier. The marker shows the values of the False Positive Rate (FPR) and the True Positive Rate (TPR) for the currently selected classifier.

The selected features are cup minor axes and disc solidity, these features are evaluated by many classifiers like SVM, KNN, ensembles bagging classifier and ensembles boosting classifier the classification result in Figure [7] show the best result using SVM, which are accuracy 76.5% and AUC 0.052.

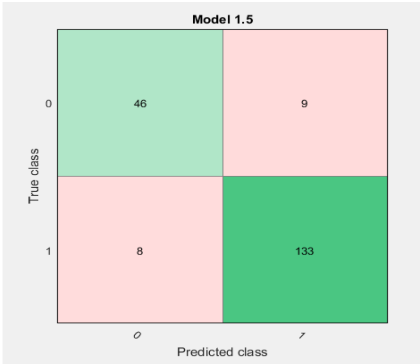


(a)

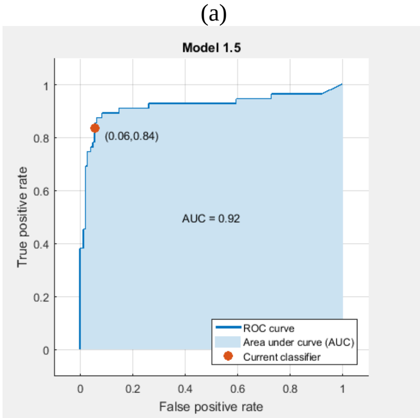
(b)

Fig.7 shows (a) Confusion Matrix (b) ROC Curve Results from the Shape Selected Features.

From the above results TP, TN, FP and FN=117, 0, 0 and 36 respectively, noticed that the healthy images are correctly classified (117) and all glaucoma image are misclassified (36 of 36) due to imbalanced features (When examples of one class in a training data set vastly outnumber examples of the other class, traditional data mining algorithms tend to create suboptimal classification models), to solve this problem the smote a logarithm was applied to balance the features, here the balanced features are classified again at cross-validation value 10 to get more accuracy, the classification results are shown in Figure:[8] are accuracy 91.3% and AUC 0.92 by ensemble RUSBOOSTED tree classifier and final results shown in table (I).



(a)



(b)

Fig.8 shows (a) Confusion Matrix (b) ROC Curve Results from the Shape Balanced Features at 10 Folds Cross-Validation.

TABLE I: SHOWS THE FINAL PROPOSED SYSTEM EVALUATION PARAMETERS VALUES.

Sensitivity	Specificity	Accuracy	AUC
85.1%	**93.6%**	**91.3%**	**92%**

The shape features classification error rate was an 8.7%, because the Optic Disc (OD) and Optic Cup (OC) segmentation accuracy due to main blood vessel and big size of Parapapillary Atrophy (PPA) which affect in disc boundary and cup segmentation, but the obtained result is good compared with the previous study.

IV. CONCLUSION AND FUTURE WORK

In conclusion, this paper proposed a glaucoma detection algorithm based on the analysis of digital fundus images using shape features (cup minor axes, and disc solidity) classified by ensemble RUSBOOSTED tree classifier. The proposed method achieved an accuracy of 91.3%. The key contribution in this work is proposed features are suitable for glaucoma detection with high accuracy.

Future work suggests designing a complete, integrated, automated system to classify all different types of glaucoma and used for follow up and apply different types of features to improve the accuracy and test different classifiers.

REFERENCE

[1] Lim. R, I. Goldberg, mph, franzco, franzco, fracs , paul n. schacknow, john r. samples: A Practical Evidence-Based Approach to Patient Care, the Glaucoma Book, Chapter 1, 2010, p. 3.

[2] Zhang Zhuo, Ruchir Srivastava, Huiying Liu, Xiangyu Chen, Lixin Duan1, Damon Wing Kee Wong, Chee Keong Kwoh, Tien Yin Wong3 and Jiang Liu: A survey on computer aided diagnosis for ocular diseases, BMC Medical Informatics and Decision Making, Copyright @ www.biomedcentral.com/1472-6947/14/80, 2014, p. 2-29.

[3] Zhang Z, Khow CK, Liu J, Cheung YLC, Aung T, et al.: Automatic Glaucoma Diagnosis with mRMR-based Feature Selection. J Biomet Biostat, S7:008, 2012, doi:10.4172/2155-6180.S7-008.

[4] Fundus image, (http: //medical4help. blogspot.com/p/ fundus - camera.html), visited at Jan 2018.

[5] Anindita Septiarini,Agus Harjoko: Automatic glaucoma detection based on the type of features used: a review. Journal of theoretical and applied information technology 28th february 2015. vol.72 no.3 2005 -jatit & lls. issn: 1992-8645,2015.

[6] Sharanagouda Nawaldgi and Y, S. Lalitha., a Novel Combined Color Channel and ISNT Rule Based Automatic Glaucoma Detection from Color Fundus Images, Indian Journal of Science and Technology,2017,Vol. 10(13), DOI: 10.17485/ijst/2017/v10i13/111722.

[7] BabuT.R &S.Shenbagadevi: Automatic detection glaucoma using fundus image, European journal of scientist research, 2011, Vol.59, No.1, pp.22-32.

[8] Yuji Hatanaka, Atsushi Noudob: Computer-Aided Diagnosis. Edited by Nico Karssemeijer, Ronald M. Summers. Proc. of SPIE Vol. 7624, 76243C, 2010, © 2010 SPIE · CCC code: 1605-7422/10/$18, doi: 10. 1117/12.843775.

[9] Bock, J. Meier, L. G. Nyúl, J. Hornegge: Glaucoma risk index: Automated Glaucoma Detection from Color Fundus Iimages, Medical Image Analysis, 2010, Vol. 14, No. 3. pp. 471–481.

[10] Karthikeyan and N. Rengarajan: Performance Analysis of Gray Level Co-Occurrence Matrix Texture Features for Glaucoma Diagnosis, American Journal of Applied Sciences, 2014, Vol. 11, No. 2, pp. 248–257.

[11] Fumero1.F, S. Alayon1, J.L. Sanchez1, RIM-ONE: An Open Retinal Image Database for Optic Nerve Evaluation, Proceedings of the IEEE Symposium on Computer-Based Medical Systems, (July 2011), DOI:10.1109/CBMS.2011.5999143·

[12] Matlab website, https://octave .sourceforge.io/image /function/regionprops.html, visited at Jun 2017.

[13] Kumar Arun M.N, H.S. Sheshadri , Building Accurate Classifier for the Classification of Micro calcification, (IJCSIT) International Journal of Computer Science and Information Technologies, 2012, Vol. 3 6, p.5346-5350.

[14] Alghamdi M, Al-Mallah M, Keteyian S, Brawner C, Ehrman J, Sakr S, Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford ExercIse Testing (FIT) project, 2017, Vol. 12, No.7, e0179805, https://doi. org/10.1371/journal. pone.01 79805.

[15] Yuan Sui, Ying Wei, and Dazhe Zhao, Computer-Aided Lung Nodule Recognition by SVM Classifier Based on Combination of Random Undersampling and SMOTE, Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine, 2015, Article ID 368674, p. 13, http://dx.doi.org/10.1155/2015/368674.

Arwa A. Gasm Elseid was born in Sudan at 2/07/1984 studied her B.Sc., Master, and PHD in a Biomedical Engineering at the Sudan University of Science and Technology.

She was worked as a service engineer at William Tail Company at 2007, as a teaching assistant in Albayan Collage at 2013, and a Lecturer in The University of Science and Technology at 2014 in Khartoum, Sudan.

Dr. Gasm Elseid was specialize in Digital Image Processing and had several publishing in this field like:

[1] Arwa Ahmed Gasm Elseid, Mohamed Eltahir Elmanna, Alnazier Osman Hamza, Evaluation of Spatial Filtering Techniques in Retinal Fundus Images, American Journal of Artificial Intelligence. Vol. 2, No. 2, 2018, pp. 16-21. doi: 10.11648/j.ajai.20180202.11.

[2] Arwa Ahmed Gasm Elseid and Mohamed Eltahir Elmanna. "Optic cup segmentation using manual thresholding level technique", International Journal of Current Research in Life Sciences, 7, (10), 2769-2773.

[3] Arwa Ahmed Gasm Elseid1, Alnazier Osman Hamza2 and Ahmed Fragoon3, Developing A Real Time Algorithm For Diagnosing Glaucoma, ICTACT journal on image and video processing, august 2018, volume: 09, issue: 02, ISSN: 0976-9102 (online)