Received: July 20, 2023. Revised: September 5, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Tri-AlgoVision: A Multifaceted Approach for Automated Glaucoma Diagnosis

Hussein Mahdi[1] **Nidhal El Abbadi**[2]*

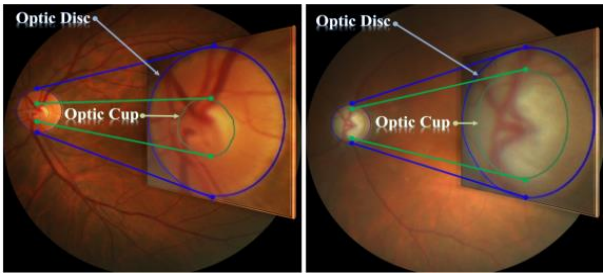
[1]*Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq*
[2]*Computer Techniques Engineering, Al-Mustaqbal College, Babylon, Iraq*
* Corresponding author's Email: nidhal.abass@fulbrightmail

Abstract: Glaucoma is a dangerous retina disease that can cause irreversible vision loss. It is called the "Silent Thief of Sight" because it has no symptoms in its early stages. A retinal fluid drainage imbalance causes elevated eye pressure. Over time, this pressure could damage the cells of the optic nerve that transmit visual data to the brain. Consequently, individuals with glaucoma can have progressive vision loss, eventually leading to total blindness if left untreated in the earliest stages. In recent years, there has been a rising interest in utilizing computer vision and machine learning techniques for developing an efficient automated diagnosis system for analyzing the retina fundus images to assist optometrists in diagnosing glaucoma at the early stages and reducing the time, cost, and human errors. All these issues motivate us to suggest an automatic glaucoma screening system as decision support for ophthalmologists at the early stages. The system based on combination decisions comes from three different approaches. Each approach uses different kinds of features for analyzing images. The first approach extracts medical features using the YOLOv8 algorithm to detect the optic disc and cup, then classifies features using the logistic regression model. The second approach extracts texture features using deep wavelet scattering and then employs the long short-term memory layer and sigmoid function to classify features. The third approach extracts deep features using convolutional neural network and then employs the attention layer and sigmoid function to classify features. The proposed system attained average results in testing images for the REFUGE, RIM-ONE-DL, DRISHTI-GS1, and ORIGA datasets, 96.21% accuracy, 98.65% specificity, 83.13% recall, 92% precision, 87.34% F1- score, 97.5% area under curve, and the average time is 0.4 seconds for classifying each image.

Keywords: Glaucoma, Deep wavelet, CNN classification, Optic disc detection, YOLO, Medical image processing.

1. Introduction

Glaucoma, often referred to as the "Silent Thief of Sight," is recognized as the second leading cause of irreversible blindness worldwide [1]. This insidious disease can progress gradually in its early stages, often without presenting noticeable symptoms [2]. It primarily affects the optic nerve, resulting in potential blindness if not promptly addressed. Glaucoma is currently the primary cause of irreversible vision loss and is typically attributed to elevated intraocular pressure exerting pressure on the optic nerve within the eye. As a consequence of this damage, the optic cup region expands, and the inferior rim surrounding the optic nerve undergoes thinning (see Fig. 1), indicative of glaucoma-related changes [3].



(a) (b)

Figure. 1 Retina images: (a) Normal and (b) Glaucomatous

According to the world health organization (WHO), the number of people suffering from Glaucoma is estimated at 64 million in 2016 [3], 80 million in 2020 [4], and this number will increase to 95 million by 2030 [3] and 110 million by 2040. It is worth mentioning that glaucoma disease is

responsible for about 12% of all people who lose their sight around the world [5].

Early glaucoma diagnosis is a problem in ophthalmology. Several measures are taken to diagnose for preventing progressive vision loss. However, traditional glaucoma diagnosis has various limitations: (1) Prone to human error: as the "cup to disc ratio" (CDR) test is typically measured manually. (2) Time-consuming: most physical retinal tests take time. (3) Costly: optical coherence tomography (OCT) is a high-cost test only available in specialized institutions. In order to address these problems, computer-aided techniques have emerged to assist experts in diagnosing at the early stages and reduce the time, cost, and errors [6].

Precise segmenting of the optic disc and cup is a significant challenge for accurate glaucoma diagnosis. Unfortunately, many existing segmentation methods encounter several limitations. Firstly, variations in image devices and illumination often result in inconsistent intensity across fundus images, leading to intensity inhomogeneity. Secondly, the spatial relationships between the optic disc and optic cup are frequently overlooked, such as the fact that the optic cup is typically located within the optic disc region and the optic cup without boundaries. Consequently, these limitations significantly diminish the effectiveness of segmentation approaches [7].

To overcome these problems and challenges, we suggest an automatic glaucoma screening system to assist in diagnosing glaucoma as a decision support for ophthalmologists at the early stages and reduce the time, cost, and errors.

The new aspects and strong points of the proposed method

Ensemble learning is based on three different models with exhaustive diverse features**:** our suggested method's theoretical foundation is inspired by the principle of ensemble learning, where the decision comes from combining the three different models' predictions. These combining methods are analogous to how ophthalmologists discuss complex cases and come to an acceptable decision. However, each model analyses retinal fundus images independently using distinct feature sets. Approach-1 extracts medical features relating to structural changes in the optic disc and cup, Approach-2 extracts deep features to identify complex patterns, and Approach-3 extracts texture features associated with finer details. Thus, the ensemble technique improves the diagnosis of glaucoma by capturing more helpful information.

The main advantages of the proposed method:

The suggested method is less error-prone. It outperformed conventional glaucoma diagnosis methods by 96.21% accuracy on many datasets. Our automated system improves screening. It reduces glaucoma diagnosis costs by reducing resources. YOLOv8 replaced the segmentation problem by accurately detecting OD and OC at 100%, reducing time. The proposed method automatically computes CDR quickly and accurately. The suggested approach diagnoses each image in 0.4 seconds.

The remainder of the paper is structured as follows: section 2 outlines related articles, while Section 3 briefly reviews essential concepts utilized in work and datasets used. The suggested method will be described in section 4, while the algorithm's results are introduced in section 5. The sixth and final section concludes the procedures and results.

2. Related works

Glaucoma diagnosis is essential to prevent blindness. Thus, several approaches have been created recently. This section describes some:

Sreng et al. (2020) [3] proposed a method for diagnosing glaucoma in retinal fundus images based on deep features. The approach segmented the OD by combining MobileNet and DeepLabv3, then feed-forward it to classify by two paths: ensemble transfer learning CNNs and CNNs with support vector machine. The method achieved an accuracy of 95.59% with REFUGE, 97.37% with RIM-ONE-v1, 99.53% with ACRIMA, 83.59% with ORIGA, and 92.06% with DRISHTI-GS1. The method's drawbacks are that it takes between 43.7 and 56.7 seconds to classify each image, and the best results for classifying datasets come from different steps.

Singh et al. (2021) [8] presented a method that relied on cropped OD and then classified it by CNN. The crop step enhanced the red channel, divided it into 50 super-pixel parts, averaged intensity values of each area, and thresholded by converting regions above 254 to 255 and below to 0. Then the method cropped the OD in the center of the largest part. The method used DenseNet-201 to classify OD. The approach achieved an accuracy of 90.1% with the DRISTI-GS dataset and 93.4% with the RIM-ONE-v2 dataset. The method's drawbacks are that it localized OD manually for some images, used a constant threshold, and built the system with different steps for different datasets, where the localization step was done on DRISHTI‑GS only.

Pascal et al. (2022) [9] proposed a method based on deep and medical features. The approach used

multi-task learning (MTL) network. MTL consisted of U-Net and VGG-16 Net. Parallel task learning is used for both classification and segmentation in MTL. OD and OC are segmented, while the image is classified simultaneously based on the fully connected layer. The OD and OC segmentation results are used for another classification by computing the VCDR. Finally, the method combines the two classification outcomes in one output result. The method's drawbacks are that it used only REFUGE dataset in training and testing the approach, which may limit its generalizability, and it did not work on ONH image datasets.

Adnan et al. (2023) [10] suggested a method for glaucoma detection based on medical features. The technique contains two significant steps. It segmented the OD and OC using a CNN net called "Feature-blending-based shallow segmentation network," then computed the CDR and classified images. The approach attained an accuracy of about 95% for REFUGE and RIM-ONE-v3 datasets. The method's drawbacks include constructing separate segmentation models for each dataset, using varying input sizes to segment OD and OC.

3. Background

This section provides an overview of some essential concepts utilized in this study.

3.1 YOLO "You only look once"

YOLO (You only look once) is a game-changing object detection algorithm in computer vision. It performs object detection in a single pass through a neural network, predicting bounding boxes and class probabilities directly. YOLO's holistic approach considers the entire image, enabling it to capture context and make accurate predictions. This real-time algorithm has found applications in autonomous driving, surveillance, and video analysis due to its speed and accuracy. YOLO has revolutionized object detection by achieving impressive results with its single-shot methodology [11].

3.2 Wavelet scattering

Wavelet scattering is a mathematical technique used to analyze and extract features from images or signals. Wavelet scattering helps characterize image texture since it captures spatial and frequency characteristics by transforming the image to the wavelet domain [12].

Wavelet scattering is similar to CNN in general steps such as convolution, nonlinearity, and pooling

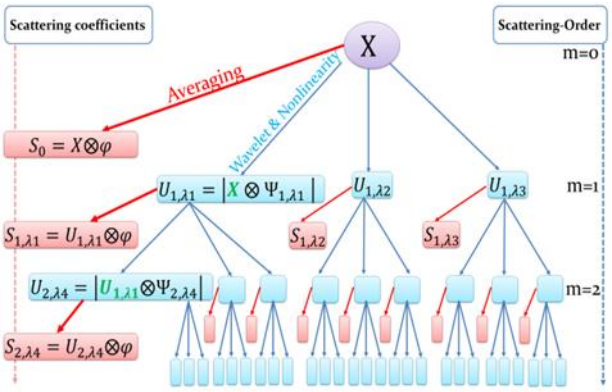


Figure. 2 Wavelet scattering decomposition

or averaging. However, in a wavelet scattering network, the filters are not learned as in CNN but rather set in advance and the features are not just the output of the last convolution layer but rather the combination of all those layers in different order depth [12]. The wavelet scattering decomposition is shown in Fig. 2.

Deep wavelet scattering is the result of a series of processes involving convolution, nonlinearity, and averaging [12], the discretion of them as follows:

Convolution: This step transforms the image to a wavelet domain by convoluting (⊗) the bank of Morlet bandpass filters (Ψ) with the input image (X) in different depths (m) at different scales (J) and rotations (L).

Nonlinearity: This step removes the complex values from the frequency bands that come from the wavelet transform to obtain modulus coefficients ($U_{m,\lambda}$) by the absolute process for each sequence ordered (λ) generated in the path of the wavelet network (λ depends on scales (J), rotations (L), and order depth (m)).

Averaging: This step averages modulus coefficients ($U_{m,\lambda}$) by low pass filter (φ) to obtain the scattering coefficients ($S_{m,\lambda}$).

3.3 Evaluation metrics

Multiple metrics are utilized to evaluate the performance of algorithms. High scores for various metrics are used to determine the best algorithm. Accuracy (Acc), Specificity (Spe), Recall (Rec), Precision (Pre), F1-score (F1), and "Area Under Curve" (AUC) are some of the metrics used as follows [13]:

$$Acc = [TP + TN]/[TP + FN + TN + FP] \quad (1)$$

$$Rec = Tp / [TP + FN] \quad (2)$$

$$Spe = TN / [TN + FP] \quad (3)$$

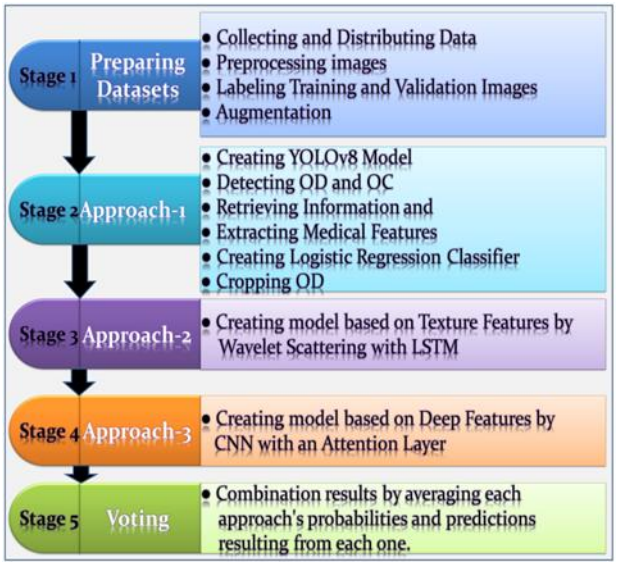


Figure. 3 Block diagram of the proposed method

$$Pre = TP / [TP + FP] \quad (4)$$

$$F1 = [2 \times TP] / [2 \times TP + FP + FN] \quad (5)$$

Where TN is the true negative, FN is the false negative, TP is the true positive, and FP is the false positive.

The receiver operating characteristic (ROC) curve exhibits true positives against false positive rates with different classification thresholds, and this metric quantifies the entire AUC [13].

3.4 Retinal fundus image datasets

Retinal fundus images have become popular for analysis and glaucoma diagnosis. The most used datasets in this field are:

REFUGE Dataset: "Retinal fundus glaucoma challenge" was accepted as a challenge during the international Conference "Medical imaging and computer assisted intervention" (MICCAI) 2018 (Granada, Spain). It consisted of 1200 retina fundus images. It is divided into three parts; training, testing, and validation. Each component has 400 retina fundus images (40 healthy and 360 glaucomatous). The dataset provided joint segmentation for OC and OD [14].

RIM-ONE-DL dataset: "Retinal image database for optic nerve evaluation deep learning" was gathered in three Spanish hospitals: "Hospital universitario de canarias," "Hospital universitario miguel servet," and "Hospital clinico san carlos." It is version four for the RIM-ONE dataset. It contains 485 retinal fundus images, (172) glaucoma, and (313) healthy. The images in the dataset are ONH images (only the region containing the OD). It provides segmentation images for optic cup and disc regions [15].

DRISHTI - GS1 dataset: It was gathered at aravind eye hospital, India. It contained 101 retina fundus images classified into 70 glaucomatous and 31 healthy. It provided separated segmentation for the optic cup and optic disc regions for each image determined manually by four experts, CDR values by four experts, and disc center [16].

ORIGA Dataset: "An online retinal fundus image database for glaucoma analysis and research" was gathered from the "Singapore malay eye study" (SiMES) by Singapore eye research institute. It contained 650 retina fundus images, 482 healthy, and 168 glaucomatous [17].

G1020 Dataset: It was collected at a private medical clinic in Kaiserslautern, Germany, between 2005 and 2017. It consists of (1020) high-resolution retinal fundus images. It has 296 glaucomatous images and 724 healthy images. The dataset provides JSON files containing ground truth annotations for OC and OD [18]. Unfortunately, the ground truth is incomplete for a group of images.

4. Proposed method

The main objective of this research is to develop an automated glaucoma screening system that can assist in detecting glaucoma in its early stages and significantly reduce screening time, costs, and errors by utilizing retina fundus images and three types of features; medical, deep, and texture features.

The key idea behind the proposed model is to employ three approaches, each performing its unique analysis of the retinal fundus image utilizing a distinct set of features. The approaches independently extract features and classify images, then the proposed method aggregates these results to make a final decision. The major stages of the proposed method are depicted in Fig. 3.

4.1 Preparing datasets

There are several steps that the model needs to prepare the datasets, especially the images in the training and validation parts.

4.1.1. Collecting and distributing data

The first stage collects the datasets that use in the proposed model. The dataset used in this proposal is REFUGE, RIM-ONE-DL, DRISHTI-GS1, ORIGA, and G1020 datasets. These datasets are selected because they meet three criteria; first, they are publicly available, well-known, and published by scientific researchers from specialized hospitals. The

Table 1. The datasets used in the model

Datasets	Ref.	No. of images	No. of images used for					
			Train		Validation		Test	
			H	G	H	G	H	G
REFUGE	[14]	1200	360	40	360	40	360	40
RIM-ONE	[15]	485	219	120	62	34	32	18
DRISHTI	[16]	101	21	49	6	14	4	7
ORIGA	[17]	650	337	117	96	33	49	18
G1020	[18]	718	529	45	133	11		
Total		3154	1466	371	657	132	445	83

* H symbolizes healthy images, and G symbolizes glaucoma images

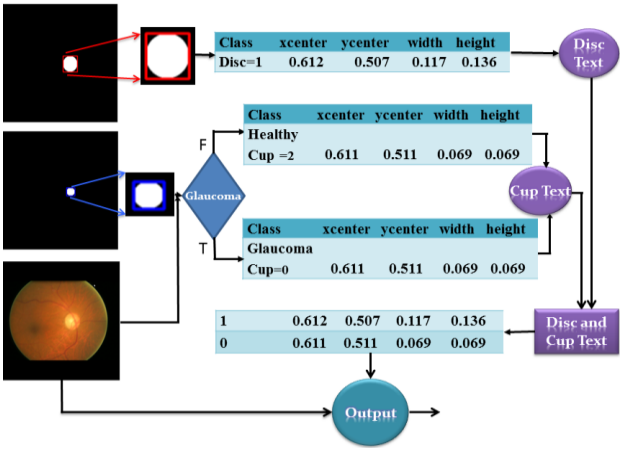


Figure. 4 Example for labeling glaucoma image

Table 2. Augmented images in the training set

Datasets	Original	Crop (x*3)	Rate	Aug.	Original+ Aug.
REFUGE/H	360	1080	3	3240	3600
REFUGE/G	40	120	10	1200	1240
RIM-ONE/H	219		3	657	876
RIM-ONE /G	120		10	1200	1320
DRISHTI/H	21	63	3	189	210
DRISHTI/G	49	147	10	1470	1519
ORIGA/H	337	1011	3	3033	3370
ORIGA/G	117	351	10	3510	3627
G1020/H	529	1587	3	4761	5290
G1020/G	45	135	3	405	450
Total					**21502**

* H symbolizes healthy images, G symbolizes glaucoma images and Aug. is augmentation.

second criterion, these datasets contain some challenges such as some of the images suffering from degradation. The third criterion of these datasets is providing ground truth for OD and OC segmentation images, and this is important because it determines the OD and OC boxes necessary for the detection algorithms such as the YOLOv8 algorithm which is used in this proposal.

After collecting the datasets, the datasets are distributed; RIM-ONE, DRISHTI-GS, and ORIGA datasets are divided into train 70%, validation 20%, and test 10%. The REFUGE dataset is not distributed because the author has already split the dataset into train, validation, and test. In the G1020 dataset, some images have no ground truth, so a group of images has been selected and allocated to train and validate without a test part. Table 1 shows the datasets used in training, validating, and testing the model.

4.1.2. Preprocessing images

Preprocessing step performs two processes, squaring and resizing images. The squaring process adds zero values in the surrounding area of each image as necessary to make the images square. The resizing process resizes the images to 640 * 640 pixels.

4.1.3. Labeling training and validation images

The labeling or annotating stage provides the text in the xywh YOLOv8 format for training and validation images with three classes: Disc=1, Glaucoma Cup=0, and Healthy Cup=2. Each image must contain text with two rows: class, x-center, y-center, width, and height. The first row is related to the disc, and the second row is associated with the cup (because each image contains a cup and a disc). This information is normalized depending on the image's dimensions. After squaring and resizing, OD and OC segmentation masks images (datasets that give ground truth mask segmentation images) can be converted into text. This process is essential because it saves time and effort and is more accurate than the prior methods that used the standard procedure called the Roboflow application for labeling. Fig. 4 shows an example of labeling a glaucoma image.

4.1.4. Augmentation

The images used to train the model must be a large number to some extent. Augmentation is the key to fixing this issue. The augmentation process includes cropping, horizontal and vertical flips, and adjusting the image's brightness, hue, contrast, saturation, and blur. Table 2 shows the augmentation, its rate, the number of augmented images for each dataset, and the total number of images.

4.2 Approach-1 (Medical-based features)

Approach-1 is the backbone of the proposed method because it includes two main jobs: detect and crop the OD, which is used for the other approaches, and classify the image. This approach has been categorized as a medical-based features approach

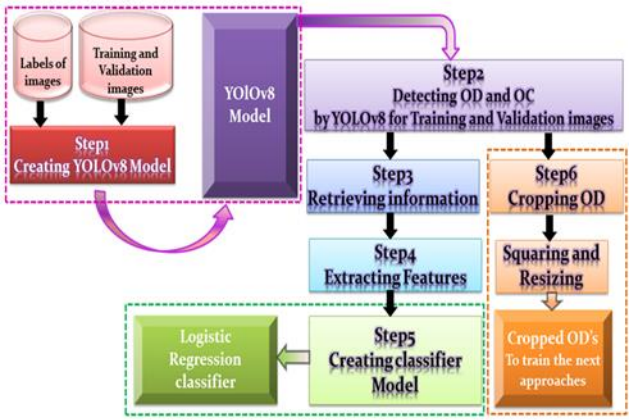


Figure. 5 Diagram for the training phase of Approach-1

Table 3. Information on the maximum detection (three classes for each image) of YOLOv8x

name	class	x-center	y-center	width	height	confidence
Disc	1					
GCup	0					
HCup	2					

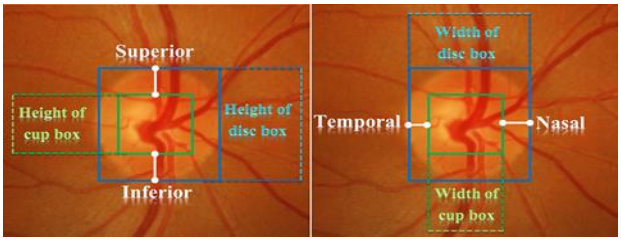


Figure. 6 Regions of the retina fundus image

because it classifies the images depending on changes in the optic nerve's structure. These changes generate modifications in some of its features. One of these features is the Cup to Disk ratio (CDR), which is considered the primary medical indicator of glaucoma infection. Fig. 5 shows the main steps of Approach-1.

Step 1: Creating YOLOv8 model

Approach-1 uses the images (training and validation) with their text information about OD and OC and the YAML file (that contains the location of the data, the number of classes, and the names of classes) for training to create the YOLOv8x model. YOLOv8x model trains with images size 640*640 pixels, batch 48, and 200 epochs.

Step 2: Detecting OD and OC

The constructed YOLOv8x model detects OD and OC by identifying them within bounding boxes. This step detects the OD and OC of training and validation images (as a new input) by using the YOLOv8 model (built in the previous step) to extract information from bounding boxes of them. YOLOv8 algorithm has been trained on three classes, it can detect a maximum of three different objects or classes for each image: Disc=1, Glaucoma cup (GCup)=0, and Healthy cup (HCup)=2.

Step 3: Retrieving information

The next step after detection of the OD and OC is to retrieve the information of boxes surrounding the OD and OC. Each object in a single image will be detected by one box and carry the following information: x-center, y-center, width, height, confidence, class, and name. The retrieved information is the information results from the YOLOv8x which are visualized in Table 3.

Step 4: Extracting features

In this step, we suggested extracting six features from the information which is retrieved in the previous step for bounding boxes of OD and OC. Theoretically; we know that the relation between disc diameter and cup diameter regards as a crucial parameter that changes according to the development of Glaucoma. So, all these six features are focused on the relation between the disc and cup diameters. The six proposed features are:

First: Vertical cup to disc ratio (VCDR): determined by using Eq. (6).

$$VCDR = \frac{Height\ Cup\ box\ has\ the\ largest\ confidence}{Height\ Disc\ box} \quad (6)$$

Second: Horizontal Cup to Disc ratio (HCDR): determined by using Eq. (7).

$$HCDR = \frac{Width\ Cup\ box\ has\ the\ largest\ confidence}{Width\ Disc\ box} \quad (7)$$

Third: Superior plus Inferior to Disc ratio (SIDR): determined by using Eq. (8).

$$SIDR = \frac{Superior+Inferior}{Height\ Disc} = \frac{Height\ Disc - Height\ Cup}{Height\ Disc} \quad (8)$$

Fourth: Temporal plus nasal to disc ratio (TNDR): determined by using Eq. (9).

$$TNDR = \frac{Temporal + Nasal}{Width\ Disc} = \frac{Width\ Disc - Width\ Cup}{Width\ Disc} \quad (9)$$

Note: the terms superior, inferior, temporal, and nasal represent the distance between the disc box boundary and cup box boundary as shown in Fig. 6.

Fifth: The confidence of the glaucoma cup (GCupC): the probability of predicting the cup is a Glaucoma cup.

Sixth: The confidence of the healthy cup confidence (HCupC): the probability of predicting the cup is a healthy cup.

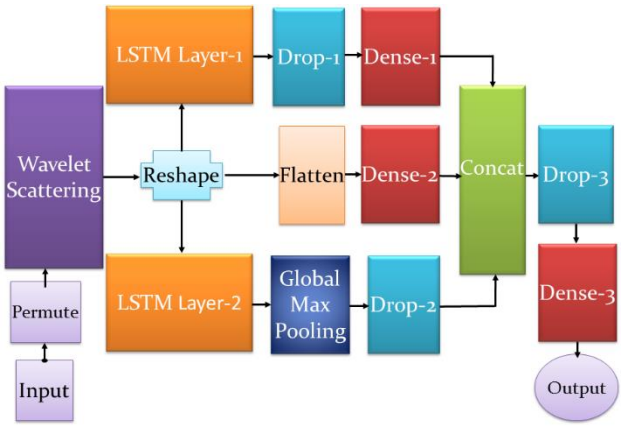


Figure. 7 The architecture of Approach-2

Step 5: Creating a logistic regression classifier

This step is the classifier step. A Logistic Regression classifier is suggested to use as a classifier. This algorithm is trained on the data results from the previous step using 5-fold cross-validation. In this step, and according to the features input the algorithm classified the image as healthy or glaucoma.

Step 6: Cropping OD

The important thing that should be noted is that after the information is retrieved, the discs (OD's) are cropped. The OD is resized to 50×50 pixels for all the training and validation images in the next approaches. The cropped disc is used as input for the other approaches.

4.3 Approach-2 (texture-based features)

In addition to OC which is affected by Glacuma is the optic nerve. The optic nerve is part of OD, so in this approach, we focus on the recognition of glaucoma from the OD which reduces the amount of data for diagnosis of glaucoma.

Approach-2 uses deep wavelets with two long short-term memory (LSTM) layers for classifying the retina (healthy or glaucoma). Approach-2 takes images of size 50x50 with three color channels (cropped OD) as input and outputs a binary classification. The approach utilizes the sigmoid function to generate probability classification as the output of each image and then determines the prediction. If the probability is greater than or equal to 0.5, the prediction equals one; otherwise, it equals zero.

The OD images are divided into training and testing images, as in Table 1. The training process uses 5-fold cross-validation. The architecture of Approach-2 is shown in Fig. 7.

Three essential parameters in wavelet scattering are used: J determines the number of wavelet scales. With J = 2, the image will be analyzed into two scales, and the wavelet transform will contain information from both scales, providing more image detail. L is the wavelet transformation angle count. Deep wavelets gather image information by rotating each wavelet. Since L = 3, each wavelet transformation turned three orientations. One at 0, 120, and 240 degrees. Thus, three copies of each scale will reflect particular features in one direction. Max-order expresses the maximum computable order or depth of scattering coefficients. Max-order = 2 refers to applying the wavelet scattering in zero, first, and second orders.

4.4 Approach-3 (deep-based features)

This approach uses a CNN with an Attention Layer. It extracts deep features from the image and then classifies it. The name of Approach-3 is deep-based features because the image features are extracted through the deep CNN.

A 50x50 Discs image results from Approach-1 is used as input in this approach. Output discs are used for training in this approach. 5-fold cross-validation is used for training. During training, the best model for each fold is stored so that it can later be used for final selection.

The inputs of the approach are OD images of size 50x50 with three color channels, and the outputs are binary classification. The approach gets the probability classification output of each image from the binary classification, and then it determines the prediction. If the probability is greater than 0.5, the prediction is one; otherwise, it is zero.

The model architecture consists of several convolution and pooling layers, two Attention layers that compute attention weights for different input parts, and finally, a dense layer with a sigmoid activation function that produces the binary classification output. Fig. 8 shows the architecture of Approach-3.

4.5 Voting or combination

As we mentioned, each approach will produce two values representing the image's classification: the probability (real value from 0 to 1) and the prediction (integer value either 0 or 1).

In the last stage, the proposed method employs a voting or combination procedure to extract a final probability and prediction by averaging each approach's probabilities and predictions. The final probability (PF) and final prediction (DF) is computed as:

$$PF = [P1 + D1 + P2 + D2 + P3 + D3]/6 \quad (10)$$

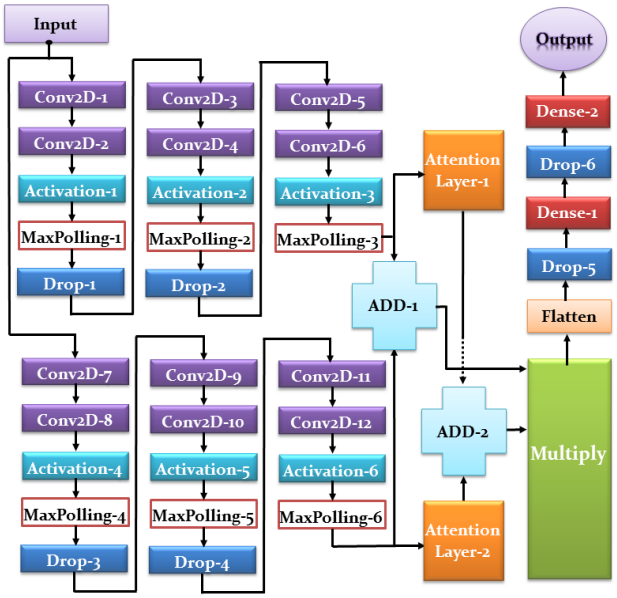


Figure. 8 The architecture of Approach-3

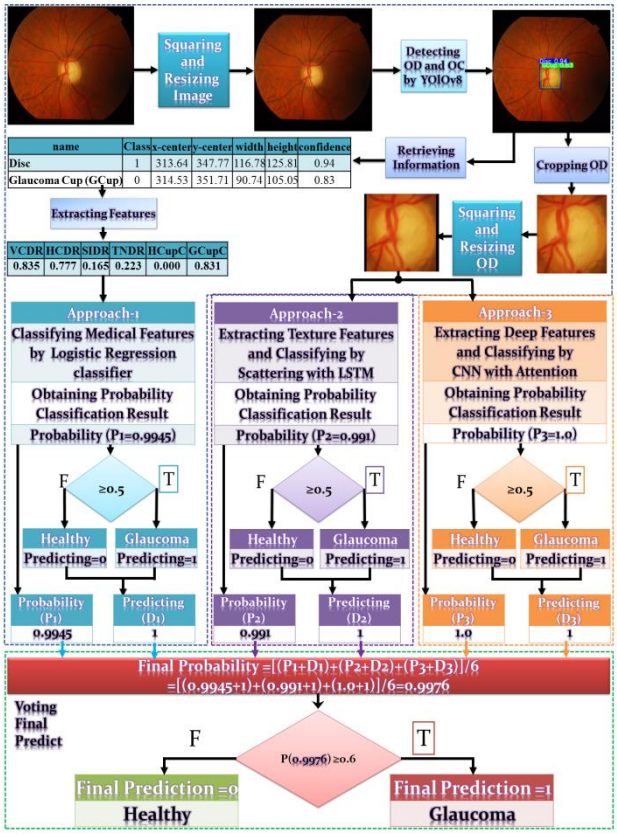


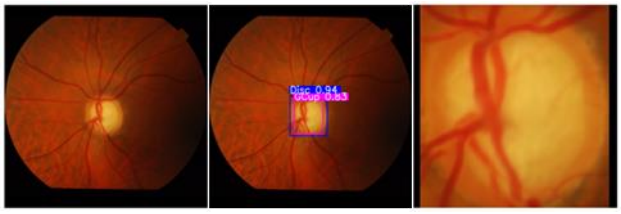
Figure. 9 An example of testing an image by the proposed system

$$DF = \begin{cases} 1 & if\ PF \geq 0.6 \\ 0 & otherwise \end{cases} \quad (11)$$

Where P1, P2, and P3 are the classification probability results, D1, D2, and D3 are the classification prediction results for Approach-1, Approach-2, and Approach-3 respectively.

Table 4. Sample of the retrieved information for boxes of OD and OC from one image

name	class	x-center	y-center	width	height	confidence
Disc	1	313.64	347.77	116.8	125.81	0.94
GCup	0	314.53	351.71	90.74	105.05	0.83



(a) (b) (c)

Figure. 10 Sample for localizing and cropping: (a) input image (b) localizing the OD and OC and (c) cropping OD

Table 5. Sample of the extracted features

VCDR	HCDR	SIDR	TNDR	HCupC	GCupC
0.835	0.777	0.165	0.223	0.000	0.831

Fig. 9 explains the voting stage and all testing steps in the proposed method by utilizing an example for testing a single glaucoma image.

5. Experimental results

The suggested system utilizes the colab to create and test the system. It is evaluated on testing images of four datasets (DRISHTI-GS1, ORIGA, REFUGE, and RIM-ONE) as shown in Table 1. Various evaluation metrics are employed when evaluating a system, including accuracy, specificity, recall, precision, F1-Score, AUC, and the average time required to classify each image.

In this section, we try to show the details of the proposed model performance. Many tests implemented to prove the efficiency of this model.

First: The first test was focused on measuring the performance of the YOLOv8 algorithm which is used for localizing the OD and OC. The YOLOv8 algorithm succeeded in localizing OD and OC for all the images used in the test. The accuracy of localizing was 100%.

After the detection of the OD and OC, there are two processes we have to do. The first one is to find the features of Approach-1 by retrieving information from the bounding boxes results from YOLOv8. The other process is to crop the ODs to use in Approaches 2 and 3. Fig. 10 shows a sample of the result of localizing and cropping processes. Table 4 shows an example of retrieved information for boxes of OD and OC.

Second: From Table 4, some of the features are extracted as shown in Table 5. The value of the

Table 6. Result of classification test images by Approach-1 (unit: %)

Datasets	Acc	Spe	Rec	Pre	F1	AUC
REFUGE	95.5	97.78	75	78.95	76.92	96.94
RIM-ONE	98	96.88	100	94.74	97.3	100
DRISHTI	100	100	100	100	100	100
ORIGA	94.03	97.96	83.33	93.75	88.24	96.03
Average	**95.64**	**97.75**	**84.34**	**87.5**	**85.89**	**97.89**

Table 7. Result of classification test images by Approach-2 (unit: %)

Datasets	Acc	Spe	Rec	Pre	F1	AUC
REFUGE	95.75	98.89	67.70	87.1	76.06	94.53
RIM-ONE	96	93.75	100	90	94.74	98.96
DRISHTI	90.91	75	100	87.5	93.33	100
ORIGA	85.07	89.8	72.22	72.22	72.22	87.41
Average	**94.32**	**97.3**	**78.31**	**84.42**	**81.25**	**95.17**

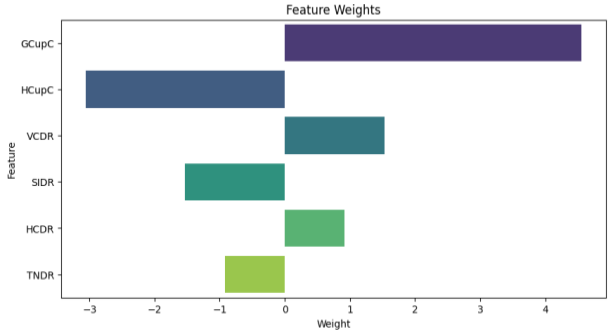


Figure. 11 Features weights relevance on classification

features listed in Table 5 has a high effect on the classification.

Fig. 11 shows the feature's importance and relevance in classification for features extracted in Table 5, all feature values are positive, and the negative side indicates a negative relationship. Increasing the weight of features increases the impact of features in classification. The relationship between the feature on the positive and negative sides is negative. The relation between the probability of classifying an input image as having glaucoma is positive, with the feature on the positive side

Third: The performance of Approach-1 was measured, and we measured the confusion matrix of each dataset when testing with. The performance of testing is listed in Table 6.

Fourth: Also, Approach-2 performance is tested, as mentioned previously this approach used the OD for detection of glaucoma based on the deep wavelet and LSTM. The result of testing Approach-2 is shown in Table 7.

Fifth: Approach-3 performance is tested with many datasets. The measuring test performance is listed in Table 8.

Table 8. Result of classification test images by Approach-3 (unit: %)

Datasets	Acc	Spe	Rec	Pre	F1	AUC
REFUGE	96.75	99.17	75	90.91	82.19	96.12
RIM-ONE	98	96.88	100	94.74	97.3	99.48
DRISHTI	90.91	75	100	87.5	93.33	100
ORIGA	89.55	93.88	77.78	82.35	80	91.04
Average	95.83	98.2	83.13	89.61	86.25	96.33

Table 10. Result of classification test images by proposed method (unit: %)

Datasets	Acc	Spe	Rec	Pre	F1	AUC
REFUGE	96.75	99.17	75	90.91	82.19	96.84
RIM-ONE	98	96.88	100	94.74	97.3	100
DRISHTI	100	100	100	100	100	100
ORIGA	91	95.92	77.78	87.5	82.35	94.47
Average	**96.21**	**98.65**	**83.13**	**92**	**87.34**	**97.5**

Table 11. Comparison of glaucoma detection using the DRISHTI-GS1 dataset (unit: %, S: time in seconds)

Authors	S	Acc	Spe	Rec	Pre	F1	AUC
Sreng et al. (2020) [3]	43.7	92.06					90
Singh et al. (2021) [8]	7.5	90.1	71	98.6			91
Natarajan et al. (2022) [19]	28	97.05	90.32	100			99.90
Thanki (2023) [20]		99		100	98.6	99.3	100
Proposed system	**0.4**	**100**	**100**	**100**	**100**	**100**	**100**

Table 12. Comparison of glaucoma detection using the ORIGA dataset (unit: %, S: time in seconds)

Authors	S	Acc	Spe	Rec	Pre	F1	AUC
Sreng et al. (2020) [3]	47.3	88.86					83.59
Chaudhary and Pachori (2021) [21]	4.98	85.38	90	73.81			
Hussain et al. (2022) [22]		83	93	43		26	83
Thanki (2023) [20]		76.2		76.2	73.2	73	66.1
Proposed system	**0.4**	**91**	**95.92**	**77.78**	**87.5**	**82.35**	**94.47**

Sixth: Finally, the results of the proposed method for image classification using three types of features are shown in Table 10.

The proposed method is compared for glaucoma detection and previous studies based on retinal fundus images. The studies used medical, textures, deep and hybrid features. The comparisons for DRISTHI-GS1, ORIGA, REFUGE, and RIM-ONE-DL are shown in Table 11, Table 12, Table 13, and Table 14, respectively.

According above results, it can notice that:

Table 13. Comparison of glaucoma detection using the Refuge dataset (unit: %, S: time in seconds)

Authors	S	Acc	Spe	Rec	Pre	F1	AUC
Sreng et al. (2020) [3]	47.3	95.59					95.10
Bao et al. (2021) [23]		94.5		50	90	64.5	
Pascal et al. (2022) [9]	0.557						96.76
Adnan et al. (2023) [10]		95.25	98.33	67.5		80.05	
Proposed system	**0.4**	**96.75**	**99.17**	75	**90.91**	**82.19**	**96.84**

Table 14. Comparison of glaucoma detection using the RIM-ONE-DL dataset (unit: %, S: time in seconds)

Authors	S	Acc	Spe	Rec	Pre	F1	AUC
Fumero et al. (2020) [15]		93.15		100			98.67
Mahrooqi et al. (2022) [24]						91.7	93.08
Carlos et al. (2023) [25]		90.17	93.11	83.61			9119
Agboola and Zaccheus (2023) [26]	0.895	98					
Proposed system	**0.4**	**98**	**96.9**	**100**	**94.74**	**97.3**	**100**

- At first glance, it can notice that all the compared methods used only some of the evaluation metrics that give a complete and clear understanding of the method's efficiency. In contrast, the proposed method provides six metrics for results.
- The proposed method is as fast as it exceeds all comparing methods, as the average execution speed for each image was 0.4 seconds.
- The proposed method attained results better than previous methods.
- The average results of the proposed method outperformed the average results of Approach-1, Approach-2, and Approach-3 individually.

6. Conclusion

The paper presented a new automatic glaucoma diagnosis method based on three distinct sets of features (medical, textures, and deep) that provide more information to detect the disease.

By combining the findings from the three methods (probability predictions of Approach-1, Approach-2, and Approach-3), the proposed method was able to attain average results an AUC of 97.5%, an accuracy of 96.21%, a specificity of 98.65%, a recall of 83.13%, a precision of 92%, and an F1-score of 87.34% with an average time 0.4 seconds for each image on well-known REFUGE, DRISHTI-GS1, RIM-ONE, and ORIGA datasets. Notably, OD and OC detection rates were both 100%.

Promising results were obtained using the suggested ensemble method, surpassing conventional glaucoma diagnosis techniques and the individual approaches (Approach-1, Approach-2, and Approach-3). These findings strengthened the theoretical justification for combining multiple models with varying strengths to improve model performance is analogous to benefiting from the various experiences of numerous specialists in enhancing disease diagnosis.

Based on the findings, the goals were achieved by increasing the efficiency of early glaucoma diagnosis while decreasing screening time, cost, error, and screening resources.

Achieving the goals of the proposed method demonstrate that it is more dependable, advantageous, and portable. Consequently, it can be used to diagnose glaucoma in impoverished countries, remote areas, and communities displaced by wars and natural disasters such as earthquakes, which are prevalent in our time. On the other hand, a proposed method can benefit glaucoma examinations in all age groups, especially children, newborns, and elderly individuals. This method will ease the glaucoma tests for these groups.

Specifically, the significant contributions of the suggested method are employing YOLOv8 in retrieved medical features after detecting OD and OC, utilizing three sets of distinct features to classify glaucoma images. To our knowledge, no previous approaches have employed these techniques in glaucoma diagnosis with different datasets.

In future works, we will develop a model for identifying several diseases that infect the retina fundus image rather than just one disease.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization, Nidhal, and Hussien; methodology, Nidhal; software, Hussien; validation, Nidhal, Hussien; formal analysis, Hussien; investigation, Hussien; resources, Hussien; data curation, Hussien; writing—original draft preparation, Hussien; writing—review and editing, Nidhal; visualization, Nidhal; supervision, Nidhal; project administration, Nidhal.

References

[1] R. Sharma and A. Grover, “Myocilin-associated glaucoma: a historical perspective and recent

research progress", *Molecular Vision*, Vol. 27, pp. 480-493, 2021.

[2] A. Sharma, M. Agrawal, S. D. Roy, and V. Gupta, "Automatic glaucoma diagnosis in digital fundus images using deep CNNs", *S. Jain, M. Sood, S. Paul, (eds) Advances in Computational Intelligence Techniques. Algorithms for Intelligent Systems. Springer*, pp. 37-52, 2020.

[3] S. Sreng, N. Maneerat, K. Hamamoto, and K. Y. Win, "Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images", *Applied Sciences*, Vol. 10, No. 14, p. 4916, 2020.

[4] S. Pathan, P. Kumar, R. M. Pai, and S. V. Bhandary, "Automated segmentation and classification of retinal features for glaucoma diagnosis", *Biomedical Signal Processing and Control*, Vol. 63, p. 102244, 2021.

[5] R. Hemelings, B. Elen, J. Breda, S. Lemmens, M. Meire, S. Pourjavan, E. Vandewalle, S. Van, M. Blaschko, P. Boever, and I. Stalmans, "Accurate prediction of glaucoma from color fundus images with a convolutional neural network that relies on active and transfer learning", *Acta Ophthalmol.*, Vol. 98, No. 1, pp. e94–e100, 2020.

[6] R. Mahum, S. U. Rehman, O. D. Okon, A. Alabrah, T. Meraj, and H. T. Rauf, "A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging", *Electronics*, Vol. 11, No. 1, p. 26, 2021.

[7] W. Zhou, Y. Yi, Y. Gao, and J. Dai, "Optic disc and cup segmentation in retinal images for glaucoma diagnosis by locally statistical active contour model with structure prior", *Computational and Mathematical Methods in Medicine.*, Vol. 2019, 2019.

[8] H. Singh, S. Saini, and V. Lakshminarayanan, "Rapid classification of glaucomatous fundus images", *JOSA A*, Vol. 38, No. 6, pp. 765–774, 2021.

[9] L. Pascal, O. J. Perdomo, X. Bost, B. Huet, S. Otalora, and M. A. Zuluaga, "Multi-task deep learning for glaucoma detection from color fundus images", *Scientific Reports*, Vol. 12, No. 1, p. 12361, 2022.

[10] A. Haider, M. Arsalan, C. Park, H. Sultan, and K. R. Park, "Exploring deep feature-blending capabilities to assist glaucoma screening", *Applied Soft Computing*, Vol. 133, p. 109918, 2023.

[11] J. Terven and D. C. Esparza, "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond", *arXiv Preprint, arXiv2304.00501*, 2023.

[12] B. Soro and C. Lee, "A wavelet scattering feature extraction approach for deep neural network based indoor fingerprinting localization", *Sensors*, Vol. 19, No. 8, p. 1790, 2019.

[13] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation", *BMC Genomics*, Vol. 21, No. 1, pp. 1–13, 2020.

[14] J. I. Orlando, H. Fu, J. B. Breda, K. V. Keer, D. R. Bathula, A. D. Pinto, R. Fang, P. A. Heng, J. Kim, and J. Lee, "Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs", *Medical Image Analysis*, Vol. 59, p. 101570, 2020.

[15] F. J. F. Batista, T. D. Aleman, J. Sigut, S. Alayon, R. Arnay, and D. A. Pereira, "RIM-ONE DL: A Unified Retinal Image Database for Assessing Glaucoma Using Deep Learning", *Image Analysis and Stereology*, Vol. 39, No. 3, pp. 161–167, 2020.

[16] J. Sivaswamy, S. Krishnadas, G. Joshi, M. Jain, and A. Tabish, "Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation", In: *Proc. of 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 53–56, 2014.

[17] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, "Origa-light: An online retinal fundus image database for glaucoma analysis and research", In: *Proc. of 2010 Annual International Conf. of the IEEE Engineering in Medicine and Biology*, pp. 3065–3068, 2010.

[18] M. N. Bajwa, G. A. P. Singh, W. Neumeier, M. I. Malik, A. Dengel, and S. Ahmed, "G1020: A benchmark retinal fundus image dataset for computer-aided glaucoma detection", In: *Proc. of 2020 International Joint Conf. on Neural Networks (IJCNN)*, pp. 1–7, 2020.

[19] D. Natarajan, E. Sankaralingam, K. Balraj, and S. Karuppusamy, "A deep learning framework for glaucoma detection based on robust optic disc segmentation and transfer learning", *International Journal of Imaging Systems and Technology*, Vol. 32, No. 1, pp. 230–250, 2022.

[20] R. Thanki, "A deep neural network and machine learning approach for retinal fundus image classification", *Healthcare Analytics*, Vol. 3, p. 100140, 2023.

[21] P. K. Chaudhary and R. B. Pachori, "Automatic diagnosis of glaucoma using two-dimensional Fourier-Bessel series expansion based empirical

wavelet transform", *Biomed. Signal Proce. and Control*, Vol. 64, p. 102237, 2021.

[22] S. Hussain, F. Guo, X. Shi, W. Li, and Z. Shen, "Compact scattering features for glaucoma detection", *Journal of Physics: Conf. Series*, Vol. 2253, No. 1, p. 12031, 2022.

[23] Y. Bao, J. Wang, T. Li, L. Wang, J. Xu, J. Ye, and D. Qian, "Self-Adaptive Transfer Learning for Multicenter Glaucoma Classification in Fundus Retina Images", In: *Proc. of Ophthalmic Medical Image Analysis: 8th International Workshop, OMIA 2021, Held in Conjunction with MICCAI 2021*, pp. 129–138, 2021.

[24] A. A. Mahrooqi, D. Medvedev, R. Muhtaseb, and M. Yaqub, "GARDNet: Robust multi-view network for glaucoma classification in color fundus images," *Inter. Work. on Ophthalmic Medical Image Analysis*, pp. 152–161, 2022.

[25] C. Vasquez, M. Martinez, H. Avila, G. Romo, P. Luque, and M. Miranda, "Multi-Stage Ensemble-Based System for Glaucomatous Optic Neuropathy Diagnosis in Fundus Images", *Electronics*, Vol. 12, No. 4, p. 1046, 2023.

[26] H. A. Agboola and J. E. Zaccheus, "Wavelet image scattering based glaucoma detection", *BMC Biomed. Eng*, Vol. 5, No. 1, p. 1, 2023.