ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250116094958

Original scientific paper
Received: 2025-01-16, Accepted: 2025-08-26

Glaucoma Stages Detection Using Feature Selection Techniques

Dimah Alkasem*, Hazem Issa

Abstract: Glaucoma, also known as the "silent thief of sight", is a chronic eye disease caused by increased intraocular pressure, leading to irreversible vision loss without the affected individual ever being aware of it. Over time, it results in damage to the optic nerve, which carries visual images from the eye to the brain. It slowly destroys the eyes of the patients unsuspectingly. There are no early symptoms of glaucoma. Therefore, the challenge lies in detecting the early stages of the disease. This paper puts forth a methodology of feature extraction and relevant feature selection. Feature selection is a much important Stage because it cuts the time required for training time and the features extracted by selecting only worthy and important features while neglecting all unnecessary or non-important features that helps in classifier boosting accuracy. ANN were the methods used in classifying. The result obtained from the proposed method was evaluated with accuracy, specificity and sensitivity. The experimental results of the proposed system achieved an accuracy of 96.12 % with the use of the Harvard dataset and 91.67 % using the Rim-One1 dataset. The experimental proved that the proposed model outperform advanced methods in diagnosing early stages of glaucoma.

Keywords: ANN (Artificial Neural Network); CatBoost; feature selection; glaucoma; HOG (Histogram of Oriented Gradients)

1 INTRODUCTION

Glaucoma is a progressive neurodegenerative disease that causes permanent vision impairment and ranks as the second leading cause of blindness with an estimated prevalence of 80 million individuals in 2020 and is expected to increase to 111 million in the year 2040 [1]. Glaucoma can also go unnoticed until its later stages because of its silent nature; thus, its timely diagnosis becomes very difficult and essential to hinder irreversible vision loss [2]. Conventional modes of diagnosis like the measurement of intraocular pressure (IOP), cup- to disc ratio (CDR), retinal nerve fibre layer (RNFL) and the detection of visual field (VF) abnormalities by optical coherence tomography (OCT) are commonly used [3].Nevertheless, these tests have such drawbacks as the inter-observer variability, extreme high cost, as well as limited accessibility in low-resource environments, which substantiates the necessity to develop automated, reliable, and scalable solutions of diagnostic methods .

Lately, AI and, in particular, ML demonstrated its enormous possibilities in analysing medical images and enables defining the presence of ocular diseases more precisely and within a small amount of time. Research articles have tried out the method utilising support vector machines (SVMs), random forests, and convolutional neural networks (CNNs). SVMs and random forests have shown good performance in the handcrafted features cases but also tend to perform poorly in high-dimensional data and generic transfer across datasets. Conversely, the CNN based deep learning methods have demonstrated high performance with automatic discrimination of features, but they usually demand huge annotated datasets and enormous computing resources which may not always be practical in practice in medical context. These shortcomings demonstrate the necessity of finding an alternative, which would strike a balance between the interpretability and efficiency of hand engineered features, and the robustness of contemporary machine learning classifiers--a solution that this research will attempt to find.

This paper seeks to build and test an automatic glaucoma detection system that combines Histogram of Oriented Gradients (HOG) to extract features and CatBoost to select features and an Artificial Neural Network (ANN) to classify. The scientific value of the present work resides in: (i) shows how to efficiently and effectively integrate handcrafted features with modern ML-based feature selection to diminish dimensionalities with preservation of diagnosis power, (ii) verifying the suggested model on two publicly available datasets (Harvard and RIM-ONE) to guarantee generalizability, (iii) showing superior results in terms of accuracy at various stages of glaucoma as compared to numerous approaches, previously described in the literature.

This research will help in guiding clinicians and researchers in developing credible, accessible and scalable methods of early-stage diagnosis or screening of potential vision loss subjects, which lessen the global burden of vision loss [4].

The article was organized as follows: Section 2 will describe the papers related to the work; Section 3 will explain the methodology used for glaucoma detection; Section 4 will describe the databases used and discuss the evaluation criteria used; Section 5 is for discussion; and Section 6 is the conclusion.

2 RELATED WORKS

In recent years, a large number of studies have been devoted to the application of machine learning methods to automatic glaucoma diagnosis. The researchers develop a group of machine learning algorithms that are intended for the development of an automated glaucoma detection system. Today, machine learning has supplanted other tools when it comes to the prediction and treatment of this dangerous illness. Hatanaka et al. [5] proposed a method for the calculation of the CDR. Through the assessment of the second derivative of the vertical profile curve with respect to

the area under the curve, it was evaluated if the patient was affected by glaucoma. The researchers Parashar and Agrawal have presented a novel technique for the automatic classification of glaucoma stages on a 2D wavelet transform (2D-T-EWT) [6]. More than half of the approaches related to the automated recognition of glaucoma problems seized on these existing elements (CDR and RDR) [7, 8] when they worked individually. Classification of glaucoma, which depends on the optical boundaries, can sometimes be misleading based on the fact that it might generate falsely incorrect glaucoma classifications. To defeat these shortcomings, scientists started using pipelines in multiple stages, including the preprocessing, feature extraction, and classification. To give an example, Sumaya et al. [9] presented a multi-stage and automatic model of diagnosis of glaucoma, and Alice et al. [10] accepted image feature descriptors as a contribution to the mixture of machine learning classifiers, resulting in an accuracy of 80.43%. The same case applies to Latif et al. [11], who applied a filter to remove noise, followed by feature extraction. The important features were then selected using the Grey Wolf Optimizer, and classification was performed using several classifiers. The results showed that the feature selection process combined with the SVM classifier achieved the highest accuracy. Singh et al. [12] segmented the optic disc using preprocessing, after which the HOG features were extracted, and found ANN to have the highest performance as compared to several other classifiers. Similarly, Rahouma et al. [14] used a similar ANN based on statistical texture features extracted with both GLCM and GLRLM, obtaining a very impressive accuracy rate of 99 percent on early glaucoma detection. This goes to show the potential of ANN on glaucoma detection.

Recent research still finds that an important value lies in the combination of hand-crafted feature extraction and ensemble classifiers. The modified Pelican Optimization Algorithm was coupled with Extreme Learning Machine (ELM) by Muduli et al. [15] and combined with curvelet-based feature extraction and dimensionality reduction (FDCT-WRP + PCA/LDA). The method resulted in high accuracy (93.25% and 96.75% on G1020 and ORIGA, respectively) and was able to use explainable AI methods as a means of model explanation. Parallel to it, Zhang et al. [16] have developed an efficient feature selection-based clinical decision support system that, besides increasing the classification accuracy, offered lower training costs, which reveals the relevance of the dimensionality reduction in glaucoma detection. On the whole, there is a definite evolution evident in these studies: the initial dependence on the structural parameters (CDR, RDR), then usage of handcrafted features (HOG, texture descriptors), and, finally, ANN-based specialized methods with feature selection and optimization algorithms to produce a robust and efficient glaucoma detection.

3 PROPOSED METHODOLOGY

A number of important steps will be included in the suggested strategy. The principal five steps of the proposed method are depicted in Fig. 1.

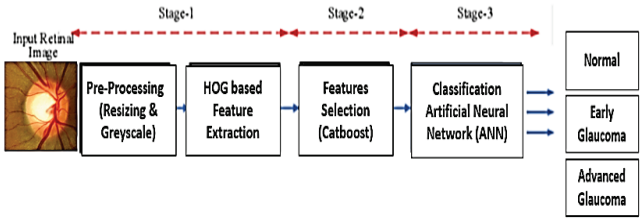


Figure 1 The Main Steps of my Proposed Process

3.1 Input Image

Two types of retinal images have been collected from different datasets. Each of the input variables is in RGB colour. Herein, data collection and pre-processing steps may be singled out in Section 3.

3.2 Image Pre-Processing

The preprocessing step acts as a key stage in image processing and involves two components. Color images are resized to 128 × 128 to match the image's pixel resolution and the resized images are then converted from RGB to grayscale.

3.3 Features Extraction

It is a technique that aims to find important data from necessary data. These features play an important part in distinguishing different classes. However, this method selects the most important features or things that make one sample distinct from others.

3.3.1 The Process of Computing the Histogram of Oriented Gradient (HOG) Feature

In the suggested system, the HOG has been used for the feature extraction from the image. The histogram of oriented gradients is technique is counting the number of times that the sign of the gradient is changing in such localized regions of the picture. The picture will consist of the small groups of cells which are connected and, for the pixels of each cell the histogram of the gradient directions is made. After that the feature vector is eventually the concatenation of these histograms. Fig. 2 demonstrates the process of computing the Histogram of Oriented Gradient (HOG) Feature [17].

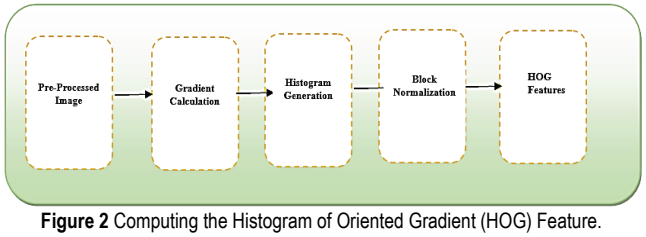


Figure 2 Computing the Histogram of Oriented Gradient (HOG) Feature.

1) Gradient calculation: The gradient is formed out of the picture magnitude and angle. $G(x)$ and $G(y)$ are estimated at each of the pixels based on a block of dimension 3 × 3. The calculation is made using the Eq. (1) and is given below per pixel value [12].

$$G_x(y,x) = H(y,x+1) - H(y,x-1)$$
$$G_y(y,x) = H(y-1,x) - H(y+1,x) \quad (1)$$

Where $G_x(y, x)$ represents the horizontal gradient of pixel (y, x) and $G_y(y, x)$ represents the vertical gradient of pixel (y, x), H represents the image, y represents rows and x represents coloums. After calculating G_x and G_y, magnitude and angle of each pixel is calculated using the Eq. (2) mentioned below [12].

$$Magnitude(\mu) = \sqrt{G_x^2 + G_y^2}$$
$$Angle(\theta) = \left| \tan^{-1}\left(\frac{G_y}{G_x}\right) \right| \quad (2)$$

2) Histogram generation: For every cell, a histogram is formed using that magnitude and direction information previously acquired.The vector gradients are then divided into groups and all the necessary groups are obtained, as well as the cells selected from the cluster are precisely collected so that the vectors are evenly produced.

3) Block normalization: In this step, the data resulting from the histograms for each block are normalized, which is predicated on the variations in the lighting and background of the image. But by normalizing the gradients, we can lessen this disparity in illumination. Therefore, after feature extraction is finished. The normalization of the feature vector, $\boldsymbol{v}$, is done through the L2-norm method which is standardized to normalize the blocks and calculated using the Eq. (3) mentioned below.

$$\boldsymbol{v} \rightarrow \frac{v}{\sqrt{\|\boldsymbol{v}\|_2^2 + \varepsilon^2}} \quad (3)$$

where ε is the small constant is mostly 1and $\boldsymbol{v}$ is a normalized vector.

3.4 Features Selection

In this section, feature importance was calculated for the data set used by a tree-based classification model such as Cat Boost. All the classifiers use various features that are pre-calculated and different thresholds to maximize accuracy.

3.4.1 Finding the Important Features Using CatBoost

CatBoost is a machine learning algorithm used to categorize data by means of a certain category [18]. It also gives details on the relevance of any given feature in decision-making [19]. The data gave the best accuracy after only 1000 iterations. Here, Cat Boost may be described as follows:

Assuming we have a training dataset containing N samples and M features as (x_i, y_i), since x_i is a vector of M features and y_i is the target variable, then the goal of CatBoost is to learn the function which is computed using the Eq. (4) mentioned below.

$$G(x) = G_0(x) + \sum_{m=1}^{M} \sum_{i=1}^{N} g_m(x_i) \quad (4)$$

There is function referred to as $G(x)$ that Cat Boost Try's to Learn Fully. This function accepts an input vector x and results in the target variable y, $G_0(x)$ refers to the first guess or initial assumption. In this case, the guess is typically selected as the median value of the target variable over the training dataset. $\sum_{m=1}^{M}$ stands for the summation throughout the ensemble of trees. M denotes the number of trees. $\sum_{i=1}^{N}$ denotes the sum of the training samples, N = the number of training samples, $g_m(x_i)$ represents is the predicted value of m tree for i training sample and Each tree inside the forest predicts the label for each training set and combines its results.

The equation is a summary of how the final prediction $G(x)$ can be calculated as a summation of the initial guess $G_0(x)$ and the predictions of each tree, $g_m(x_i)$ for each training sample. This average value is computed for all trees and on all training samples.

3.5 Classification

Machine learning can be subdivided into two folds: the first being supervised learning and the second one is called unsupervised learning. The supervised-learning method is the notion of the output formation from the inputs. It can facilitate the process of classification and help in the complicated diagnosis of cases. Being more accurate, supervised learning is more widely applied in research studies Classification from our standpoint has been done by ANN method, supervised learning tool.

3.5.1 Artificial Neural Network (ANN)

The suggested framework includes the input layer, three fully connected layers and the output layer [12, 20]. Features of the dataset would be represented as the input layer, whereas the three hidden layers would have 512, 256, and 128 neurons. The first fully connected layer employs the ReLU (Rectified Linear Unit) activation function to hasten the training procedure. ReLU takes values of inputs that are negative into zero whereas values that are positive are not altered and combined via Eq. (5) below [21].

$$f(x) = \max(0, x) \quad (5)$$

where x is the input value.

The output layer contains a number of neurons equal to the number of classes in the dataset, using the Softmax activation function for multi-class classification [22] was used in this research, as given in Eq. (6) below:

$$P = \frac{e^{zi}}{\sum_{j=1}^{n} e^{zj}} \tag{6}$$

In which P is a probability, e^{zi} is exponent of input vector, and e^{zj} is the exponent of the output vector.

The network uses Categorical-Cross Entropy loss to reduce the difference between predicted and actual distribution and improves accuracy [22]. Loss function provided in Eq. (7) below.

$$L = -\frac{1}{N}\sum_{i=1}^{n}\sum_{j=i}^{C} y_{i,j} \cdot \log(P_{i,j}) \tag{7}$$

Where N represents the number of samples, C denotes the number of classes, $y_{i,j}$ represents the true distribution of sample i belonging to class j, $P_{i,j}$ represents the predicted distribution by the model that sample i belongs to class j, and L denotes the loss function. The classification layer determines the number of classes from the size of the preceding layer, with the fully connected layer output predicting the final classification as $[1\times1\times N]$, where N is the number of classes.

In order to optimize the model, the Adaptive Moment Estimation (Adam) optimizer was employed on account of its success and memory efficiency in complicated tasks. Adam uses momentum and scaling in combination with dynamic learning rates [22]. The weight update rule for Adam is given by Eq. (8) below.

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \tag{8}$$

where $\hat{m}_t$ is the first moment mean of gradients, $\hat{v}_t$ is the second moment mean of squared gradients, ε prevents division by zero, and η is the learning rate for minimizing the cross-entropy loss.

4 RESULTS

4.1 Experimental Settings

Here, we shall discuss the experimental findings and compare the performance of our basically recommended model with that of some other methods.

4.1.1 Datasets

The initial images of the Harvard dataset were obtained from Kim's Eye Hospital [23]. The dataset includes 788 normal images, 289 early glaucoma images, and 467 advanced glaucoma images. The data is already preprocessed. Some images of this database are shown in Fig. 3.

The RIM-ONE DB, on the other hand, has approximately 169 images. The pictures are assigned to different classes: all 118 normal images, 12 early glaucoma images, 14 moderate glaucoma images, 14 advanced glaucoma images, and 11 ocular hypertension (OHT) images. Fig. 4 presents photographs of healthy individuals and glaucoma patients with early, moderate, and severe cases of the glaucoma as examples of the RIM-ONE dataset [24].

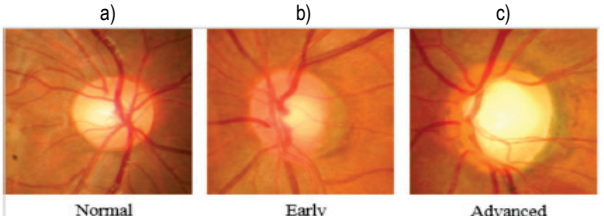


Figure 3 The normal image (a), early glaucoma (b), and advanced glaucoma (c) from the Harvard dataset

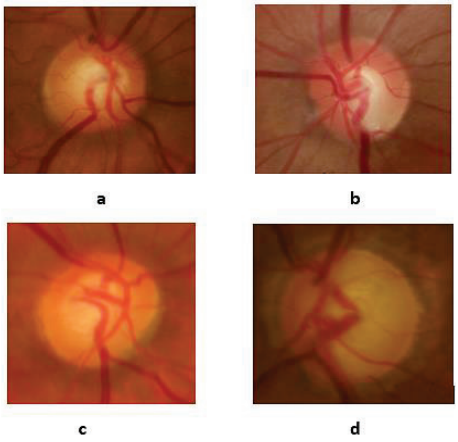


Figure 4 The normal image (a), mild glaucoma (b), moderate glaucoma (c), and severe glaucoma (d) from the RIM-ONE dataset

4.1.2 Implementation Details

It trained the model with the batch size of 32, and all images included in the dataset were resized to 128 × 128 pixels. The test set detection accuracy was registered at 100 epochs of model training.

4.1.3 Evaluating the performance of Machin Learning

The described models' performance is evaluated using accuracy, sensitivity, specificity. Diagnosis model accuracy is called the ability of the model to precisely distinguish between patients with and without disease status Eq. (9) shows how to obtain accuracy [12].

$$Accuracy\,(ACC) = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

The sensitivity of a diagnosis model is interpreted as the probability that the model will correctly recognize the sick among all sick patients. This ratio is calculated as shown by Eq. (10) [12].

$$Sensitivity\,(SN) = \frac{TP}{TP + FN} \tag{10}$$

Specificity is the ability to differentiate between tested persons without disease. This ratio is calculated as shown by Eq. (11) [12].

$$Specificity\,(SP) = \frac{TN}{TN + FP} \tag{11}$$

4.2 Multi-stage Glaucoma

During our experiments, the data is divided into two sets—a training set and a testing set. The training set is comprised of a round of 80 percent, with the rest being the testing set of 20 percent.

In the case of the model, we used the Adam optimizer to tune the parameters based on the suggested default values, i.e., the parameters were fixed as $\beta_1 = 0.9$ and $\beta_2 = 0.999$ in the training of the artificial neural network (ANN). Also, the learning rate was endowed with 0.001 in order to arrive at the ideal value of categorical cross-entropy loss [21] to the ANN.

The training and validation accuracy curves for the Harvard and RIM-ONE datasets were plotted as shown in Figs. 5 and 6, respectively

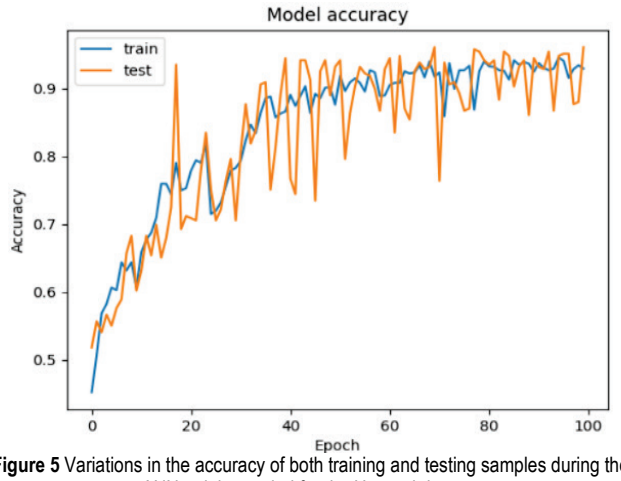


Figure 5 Variations in the accuracy of both training and testing samples during the ANN training period for the Harvard dataset

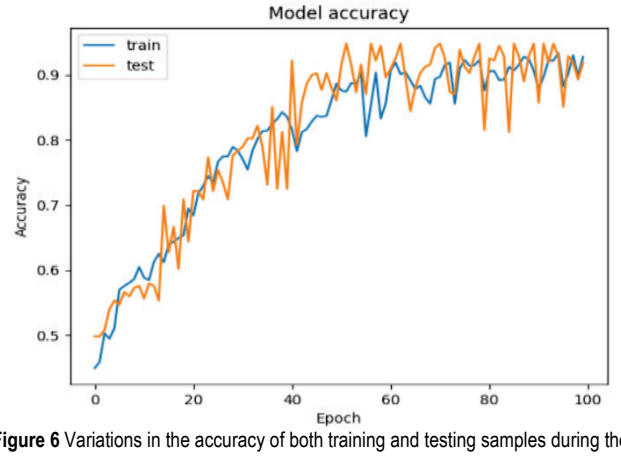


Figure 6 Variations in the accuracy of both training and testing samples during the ANN training period for the RIM-ONE dataset

For the Harvard dataset, the training accuracy achieved was 96.12%, while for the RIM-ONE dataset, the training accuracy achieved was 91.67%.

At the beginning of training, accuracy values are low but gradually improve as training progresses in classifying the stages of glaucoma (multi-class classification) for both datasets. In the early stages, the test accuracy is higher than the training accuracy, then the values converge as training continues.

In the case of the Harvard data (Fig. 5), training and test accuracy were low initially and improved as the epochs progressed, ending with an accuracy of 96.12% during training. The accuracy of the testing was close to the training curve, showing a high level of generalization.

Training performance and test performance trained on the RIM-ONE data set Fig. 6 were initially low but continued to increase consistently across the durations of the epochs. Though the test curve had more convulsions as the sample was small, the overall pattern fitted close to the training curve, with a convergence and high degree of generalization as pertains to the case of the Harvard dataset .The loss and validation curves for the Harvard and RIM-ONE datasets were plotted as shown in Figs. 7 and 8, respectively.

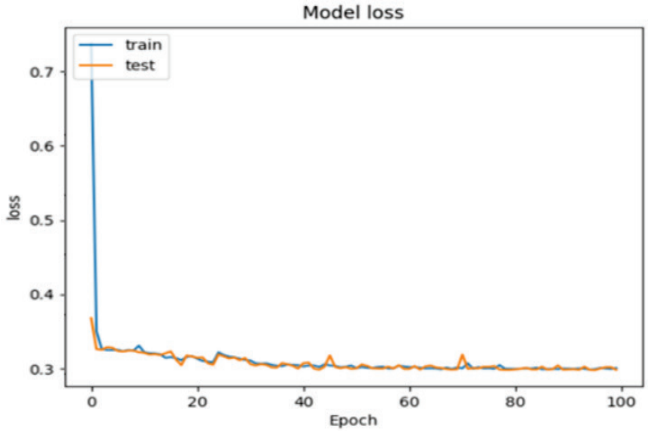


Figure 7 Variations in the loss of both training and testing samples during the ANN training period for the Harvard dataset

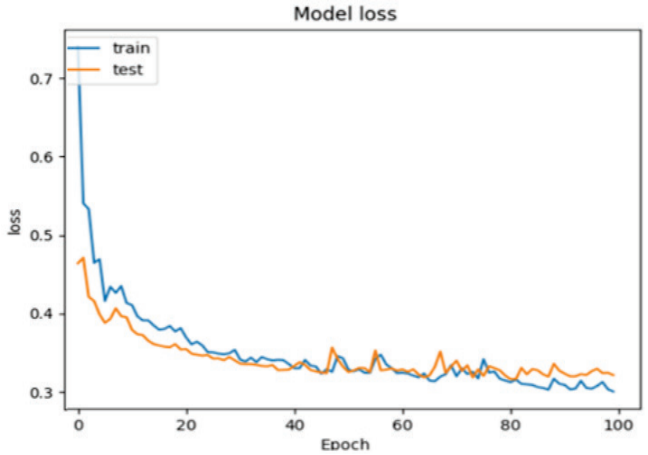


Figure 8 Variations in the loss of both training and testing samples during the ANN training period for the RIM-ONE dataset

Concerning the loss function, it begins by taking a high value before reducing at a sharp rate, depicting that the model

initially had a tricky time with the classification, and then it began to act better. The loss values of the test samples are less than that of the training samples and then tend to come closer and closer with training.

For the Harvard dataset Fig. 7, the training loss started at a high value and gradually decreased as the epochs progressed. The loss values of the test samples were initially lower than those of the training samples, and with further training, the two curves converged, indicating that the model was successfully minimizing the categorical cross-entropy loss while maintaining generalization.

In the case of RIM-ONE data, Fig. 8 the training loss exhibited an initial steep decline as the model takes time to pick up on its accuracy. The losses of the test samples were converted more strongly than the training samples, but as the training proceeded, the two curves slowly approached, which is another evidence that ANN was capable of the generalization encountered even though the size of the dataset was small.

To evaluate the model's performance and its ability to separate and distinguish between categories, the following metrics were used: Sensitivity, Specificity and Accuracy. In Tabs. 1 and 2 Sensitivity, Specificity and Accuracy rates for the proposed model.

Table 1 Multiple classification on the Harvard dataset using the proposed model

The datasets (Classes)		Sensitivity	Specificity	Accuracy
Harvard	Advanced	94.95%	99.52%	96.12%
	Early	96.72%	95.97%	
	Normal	96.64%	99.38%	

Harvard dataset results show a high performance of the model in all glaucoma stages. Sensitivity and specificity are always high and reach their maximum in the Advanced as well as Normal category (96.12 percent), showing that they are highly classified. Markedly, also the model got a high sensitivity at the early stage (96.72%), an aspect that holds its potential in effective identification of glaucoma at an early stage.

Table 2 Multiple classification on the RIM-ONE dataset using the proposed model

The datasets (Classes)		Sensitivity	Specificity	Accuracy
RIMONE	Advanced	100%	100%	91.67%
	Early	100%	97.14%	
	Normal	100%	81.82%	
	Moderate	33.33%	100%	
	OHT	66.67%	100%	

The performance on the RIMONE dataset indicates the above model makes perfect sensitivity (100%) and high specificity in detecting advanced, early, and normal stages of glaucoma. Still, sensitivity on the Moderate and OHT stages is significantly lower (33.33% and 66.67%, respectively), which may be explained by fewer samples and the less obvious nature of these stages; that is why they can be detected less precisely.

4.3 Confusion Matrix

The confusion matrix delivers a comprehensive assessment of the classification accuracy in terms of the comparison of actual labels and predicted classes. It provides a glimpse into the effectiveness with which the various glaucoma levels and normal scenarios are discriminated by the proposed system. Confusion matrices for the Harvard and RIM-ONE datasets are attached as Figs. 9 and 10, respectively.

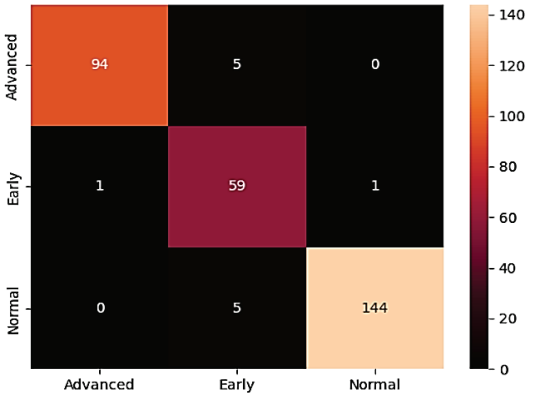


Figure 9 Confusion matrix for identifying types of glaucoma using ANN for the Harvard dataset

94 advanced cases are precisely classified as being advanced, whereas a few cases are misclassified between early and normal, which shows that the system is robust with little confusion in borderline cases.

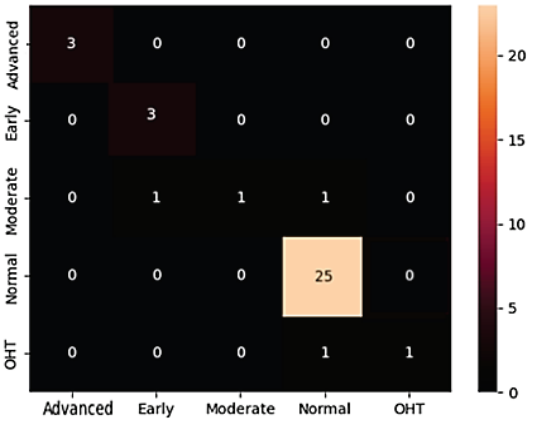


Figure 10 Shows the confusion matrix for identifying types of glaucoma using ANN for the RIM-ONE dataset

The model has shown acceptable classification even with a reduced data size, but only some normal cases (25) can be classified accurately, and there is minor confusion with mild or normal classes. This implies that although the model has high generalization ability, sensitivity as a differentiator of faint glaucoma stages may be hampered by the small quantity of samples in some of the categories.

4.4 Performance Comparison

In our search, we classified the RIM ONE database into five stages of glaucoma (advanced, early, moderate, healthy, and ocular hypertension). While we did not find any research with such a classification to compare the results of our research with.

We compared glaucoma identification with existing methods, taking into account the Harvard data set, and the results are presented in Tab. 3.

Different CNN models [25] : such as pre-trained CNN networks (ResNet50, AlexNet, VGG19, DenseNet-201, and Inception-ResNet-v2), were used for feature extraction and classification.These models achieved moderate performance, with accuracy up to 85.43%.They did not do feature selection but directly employed deep features of the fundus images.

U-Net and Xcep-Dense [26]: networks were combined, with U-Net being used for segmenting the optic disc region and Xcep-Dense utilized for both feature extraction and classification, reaching 87.58% accuracy.

the GS-Net [27]: GS-Net is a CNN that applied a Global Self-Attention Module (GSAM). The GSAM assists the model to concentrate on those most significant aspects of the retina only to consider the features of the picture about them and also to locate them so that the model could see the whole picture of the retina. Conversely, our model is based on HOG features, which capture the distribution of edge orientations and gradients in the retinal images, emphasizing the most informative aspects of the images and achieving high accuracy in glaucoma stage classification.

HOG_ANN: Initially, all extracted HOG features were used for classification without selecting the most important ones, resulting in an accuracy of 84.14%.

The proposed model: The model uses HOG features with CatBoost for selecting only the most important features, followed by classification with ANN. the model achieved improved performance with 96.12% accuracy, 97.10% sensitivity, and 98.29% specificity, showing the importance of feature selection in improving the outcome of the classification process.

Table 3 Comparative for glaucoma classification in Harvard dataset

Method	SP (%)	SEN (%)	ACC (%)
Different CNN models [25]	92,14%	80,80%	85,43%
U-Net and Xcep-Dense [26]	-	-	87.58%
GS-Net [27]	-	-	84.91%
HOG_ANN	92.01%	86%	84.14%
Proposed model	**98.29%,**	**97.10%**	**96.12%**

The results of the comparison make clear that the proposed solution is superior to those that can be implemented in practice, especially thanks to the effective feature selection step, as it added to the overall enhancement of the performance. As evidenced in the comparison table, the combination of the CatBoost algorithm during the selection of the most relevant features enabled the model to concentrate on important information, thus making it less redundant and offering it a better chance of distinguishing between various glaucoma stages.

The proposed method reveals a more focused and effective use of the given information as compared to its counterparts that tended to rely on greater numbers of raw features. Such augmentation further emphasizes the role of a combination of advanced feature selection methods with powerful models of classification, the provision of a more obvious and less subjective framework of glaucoma-stage provisions, and the entering of the possible prospects of clinical practice.

5 DISCUSSION

The findings of the study illustrated that feature extraction resulted in accuracy, sensitivity, and specificity values higher than 96%. The experiments in this study also show that the feature extraction technique and the selection of significant features for disease diagnostics give good results. There was a comparison made with the results of this study and other literature on glaucoma screening. The outcomes are contrasted with the help of all the information presented in the tables above. The same datasets are used in the study we discuss. It can be concluded, that our approach is more accurate compared to other studies. The benefit of this research is that it uses traditional machine learning methods like ANN, but at the same time, it provides an accuracy, sensitivity, and specificity of more than 96%.

Here, we apply the CatBoost algorithm, which is critical in extracting features that distinguish the stages of glaucoma.

6 CONCLUSION AND FUTURE WORK

In this paper, we proposed an AI-based framework for the automatic detection of glaucoma using Histogram of Oriented Gradients (HOG) for feature extraction, CatBoost-based feature selection, and Artificial Neural Networks (ANN) for classification. The method was tested against two test data sets, Harvard (1,544 images) and RIM-ONE (169 images).

The results of our method were also quite good, with the accuracy, sensitivity, and specificity being 96.12%, 97.10%, and 98.29% on the Harvard dataset and 91.67%, 80.00%, and 95.79%, respectively, on the RIM-ONE dataset. Such research evidence shows that the problem of detecting glaucoma automatically, using feature extraction, can prove to be extremely useful for researchers and doctors in the future.

The principal asset of this research is in the combination of the efficient feature selection and ANN classification technique, which minimizes irrelevant data and maintains the accuracy of the diagnostic data. In contrast to most of the previous efforts, the framework demonstrates a high generalization potential in several collections of information that have various properties.

In the future, there are a few directions that can be taken that would augment this study. First, the proposed model can be put into practice in real-life clinical trials, which will confirm its appropriateness in daily medical practice. Second, by growing the datasets to an increased number and variety of populations, across imaging modalities and acquisition settings, one would strengthen robustness and generalization. Thirdly, the combination of the multi-modal data (i.e., OCT, fundus photography, and patient clinical records) might present a broader diagnosis tool. Lastly, the framework could be generalized to identify and classify other diseases of the eye, like diabetic retinopathy or macular degeneration, to increase the clinical significance. Overall, such work is also important to the science of computer-assisted diagnosis that identifies glaucoma and introduces the robust fundamentals to develop AI-assisted ophthalmology devices that could be deployed in the practical world of health care.

7 REFERENCES

[1] Tham, Y. C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., & Cheng, C. Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology, 121*(11), 2081-2090. https://doi.org/10.1016/j.ophtha.2014.05.013

[2] Sommer, A., Tielsch, J. M., Katz, J., Quigley, H. A., Gottsch, J. D., Javitt, J., & Singh, K. (1991). Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans: the Baltimore Eye Survey. *Archives of ophthalmology*, *109*(8), 1090-1095 https://doi.org/10.1001/archopht.1991.01080080050026

[3] Shabbir, A., Rasheed, A., Shehraz, H., Saleem, A., Zafar, B., Sajid, M., ... & Shehryar, T. (2021). Detection of glaucoma using retinal fundus images: A comprehensive review. *Mathematical Biosciences and Engineering, 18*(3), 2033-2076. https://doi.org/10.3934/mbe.2021106

[4] Chan, Y. M., Ng, E. Y. K., Jahmunah, V., Koh, J. E. W., Lih, O. S., Leon, L. Y. W., & Acharya, U. R. (2019). Automated detection of glaucoma using optical coherence tomography angiogram images. *Computers in biology and medicine*, *115*, 103483. https://doi.org/10.1016/j.compbiomed.2019.103483

[5] Hatanaka, Y., Fukuta, K., Muramatsu, C., Sawada, A., Hara, T., Yamamoto, T., & Fujita, H. (2009). Automated measurement of cup-to-disc ratio for diagnosing glaucoma in retinal fundus images. In *World Congress on Medical Physics and Biomedical Engineering*, September 7-12, 2009, Munich, Germany: Vol. 25/11 *Biomedical Engineering for Audiology, Ophthalmology, Emergency & Dental Medicine* (pp. 198-200). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03891-4_53

[6] Parashar, D., & Agrawal, D. K. (2020). Automated classification of glaucoma stages using flexible analytic wavelet transform from retinal fundus images. *IEEE Sensors Journal*, *20*(21), 12885-12894. https://doi.org/10.1109/JSEN.2020.3001972

[7] Cheng, J., Liu, J., Xu, Y., Yin, F., Wong, D. W. K., Tan, N. M., ... & Wong, T. Y. (2013). Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE transactions on medical imaging*, *32*(6), 1019-1032. https://doi.org/10.1109/TMI.2013.2247770

[8] Mohamed, N. A., Zulkifley, M. A., Zaki, W. M. D. W., & Hussain, A. (2019). An automated glaucoma screening system using cup-to-disc ratio via simple linear iterative clustering superpixel approach. *Biomedical Signal Processing and Control, 53*, 101454. https://doi.org/10.1016/j.bspc.2019.01.003

[9] Pathan, S., Kumar, P., Pai, R. M., & Bhandary, S. V. (2023). An automated classification framework for glaucoma detection in fundus images using ensemble of dynamic selection methods. *Progress in Artificial Intelligence*, *12*(3), 287-301. https://doi.org/10.1007/s13748-023-00304-x

[10] Alice, K., Deepa, N., Devi, T., BeenaRani, B. B., & Nagaraju, V. (2023). Effect of multi filters in glucoma detection using random forest classifier. *Measurement: Sensors, 25*, 100566. https://doi.org/10.1016/j.measen.2022.100566

[11] Latif, J., Tu, S., Xiao, C., Bilal, A., Ur Rehman, S., & Ahmad, Z. (2023). Enhanced Nature Inspired-Support Vector Machine for Glaucoma Detection. *Computers, Materials & Continua, 76*(1). https://doi.org/10.32604/cmc.2023.040152

[12] Singh, L. K., Garg, H., & Khanna, M. (2022). Histogram of Oriented Gradients (HOG)-Based Artificial Neural Network (ANN) Classifier for Glaucoma Detection. *International Journal of Swarm Intelligence Research (IJSIR), 13*(1), 1-32. https://doi.org/10.4018/IJSIR.309940

[13] Al-Sarem, M., Saeed, F., Boulila, W., Emara, A. H., Al-Mohaimeed, M., & Errais, M. (2021). Feature selection and classification using CatBoost method for improving the performance of predicting Parkinson's disease. In *Advances on Smart and Soft Computing: Proceedings of ICACIn 2020* (pp. 189-199). Springer Singapore. https://doi.org/10.1007/978-981-15-6048-4_17

[14] Rahouma, K. H., Mohamed, M. M., & Hameed, N. S. A. (2019). Glaucoma Detection and Classification Based on Image Processing and Artificial Neural Networks. *Egyptian Computer Science Journal, 43*(3).

[15] Muduli, D., Kumari, R., Akhunzada, A., Cengiz, K., Sharma, S. K., Kumar, R. R., & Sah, D. K. (2024). Retinal imaging based glaucoma detection using modified pelican optimization based extreme learning machine. *Scientific Reports, 14*(1), 29660. https://doi.org/10.1038/s41598-024-79710-7

[16] Singh, L. K., Khanna, M., Garg, H., & Singh, R. (2024). Efficient feature selection based novel clinical decision support system for glaucoma prediction from retinal fundus images. *Medical Engineering & Physics, 123*, 104077. https://doi.org/10.1016/j.medengphy.2023.104077

[17] Shamna, N. V., & Aziz Musthafa, B. (2023). Feature extraction method using hog with ltp for content-based medical image retrieval. *International journal of electrical and computer engineering systems*, *14*(3), 267-275. https://doi.org/10.32985/ijeces.14.3.4

[18] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems, 31*.

[19] Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:1810.11363.

[20] Gershenson, C., (2003), Artificial neural networks for beginners. *arXiv preprint cs/0308031*.

[21] Nair, V. & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807-814.

[22] Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A. & Chavez-Urbiola, E. A. (2023). Loss functions and metrics in deep learning. *arXiv preprint arXiv:2307.02694*.

[23] Kim, U. (2018). Machine learn for glaucoma. *Harvard Dataverse*.

[24] Fumero, F., Alayón, S., Sanchez, J. L., Sigut, J., & Gonzalez-Hernandez, M. (2011, June). RIM-ONE: An open retinal image database for optic nerve evaluation. In *24th IEEE international symposium on computer-based medical systems (CBMS2011)* (pp. 1-6). https://doi.org/10.1109/CBMS.2011.5999143

[25] Velpula, V. K., & Sharma, L. D. (2023). Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion. *Frontiers in Physiology*, *14*, 1175881. https://doi.org/10.3389/fphys.2023.1175881

[26] Desiani, A., Priyanta, S., Ramayanti, I., Suprihatin, B., Halim, M. R., Geovani, D., & Rayani, I. (2023). Multi-Stage CNN: U-Net and Xcep-Dense of Glaucoma Detection in Retinal Images. *Journal of Electronics, Electromedical Engineering, and Medical Informatics, 5*(4), 211-222. https://doi.org/10.35882/jeeemi.v5i4.314

[27] Das, D., & Nayak, D. R. (2023, October). Gs-net: Global self-attention guided cnn for multi-stage glaucoma classification. In *IEEE International Conference on Image Processing (ICIP2023)* (pp. 3454-3458). https://doi.org/10.1109/ICIP49359.2023.10222689

Authors' contacts:

Dimah Alkasem, Master's degree student
(Corresponding author)
Department of Computer Engineering,
College of Electrical and Electronic Engineering, University of Aleppo,
12212 Aleppo, Syria
dimaalkasem581@gmail.com

Hazem Issa, PhD
Department of Computer Engineering,
College of Electrical and Electronic Engineering, University of Aleppo,
12212 Aleppo, Syria
hazemisaa17@gmail.com