RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2529

Enhancing Glaucoma Detection in Fundus Images: A ResNet based Segmentation and Advanced ML Algorithms with Duck Pack Optimizer

Chinthakunta Manjunath [a], Archana Sasi [b], Smitha Chowdary Ch [c], C. Sharon Roji Priya [d], B Raveendra Naick [e], Harshini Macherla [f], Kranthi Kumar Lella [g,*]

[a] Department of Computer Science and Engineering (AI & ML), Vidyavardhaka College of Engineering, Mysore, India.
[b] Department of CSE, Faculty of Engineering and Technology, Jain University, Kannagapura Rd, Bengaluru, Karnataka 562112, India.
[c] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur 522302, India.
[d] Department of Computer Science and Engineering, CHRIST (Deemed to be University), Bangalore, India.
[e] Department of CSE (AI&ML), School of Computing, Mohan Babu University, Tirupati, India.
[f] Department of Information Technology, MLR Institute of Technology, Hyderabad 500043, India.
[g] School of Computer Science and Engineering, VIT-AP University, Vijayawada 522237, India
* Corresponding Author Email: kranthi1231@gmail.com

DOI: https://doi.org/10.54392/irjmt2529
Received: 05-07-2024; Revised: 22-02-2025; Accepted: 18-03-2025; Published: 24-03-2025

Abstract: Untreated glaucoma, a chronic eye illness, can cause irreversible vision loss if not caught early. The condition begins with abnormalities in the eye's drainage flow, leading to a rise in intraocular pressure. As the disease progresses, the optic nerve head deteriorates, resulting in vision loss. Ophthalmologists need extensive training and expertise to interpret findings accurately during medical follow-ups to examine the retina. To address this challenge, deep learning-based algorithms have been developed to screen for and diagnose glaucoma using images of the optic nerve, retinal structures, and retinal fundus. This research explores the use of classification and segmentation algorithms based on ResNet to identify glaucoma in fundus images. We fine-tuned the classifier using the DuckPack optimizer and employed XGBoost, LightGBM, and CatBoost algorithms for classification. The results were promising. The segmentation model based on ResNet effectively extracted features, aiding the classification models in accurately identifying glaucoma. All three algorithms performed admirably, though further fine-tuning is needed to determine the best one. Enhancing the model's performance was straightforward after using the DuckPack optimizer for fine-tuning. This study highlights the promising applications of deep learning and sophisticated machine learning algorithms in glaucoma detection. Its findings could inform the development of future diagnostic tools.

Keywords: Glaucoma Detection, Resnet Based Segmentation, Duck Pack Optimizer, Catboost, Lightgbm, Xgboost, Fundus Images

1. Introduction

Retinal ganglion cells deteriorate over time in glaucoma, an eye disease that, if left untreated, leads to irreversible vision loss. In 2020, it was estimated to affect 80 million people of all ages [1]. A significant risk factor for glaucoma is an imbalance between the drainage and flow of aqueous humor fluid, which can lead to elevated intraocular pressure and ultimately cause the disease [2]. Other risk factors include age, race, and family history, which may increase the likelihood of developing glaucoma. Critical diagnostic methods for glaucoma include thorough evaluation of the optic nerve head, visual field tests, and extensive tonometry testing of the eyes [3]. However, these examinations typically require specialized knowledge, costly equipment, and a significant investment of time. Despite these limitations, the use of algorithms for automated glaucoma identification using fundus images is on the rise [4].

One of the most important ways to detect glaucoma is through fundus imaging, a non-invasive technique that is both conveniently accessible and provides vital information about the eye and optic nerve head [5]. This imaging method offers detailed descriptions of important areas of the retina, including their size, shape, and color [6]. Several studies have noted high sensitivity and specificity rates for glaucoma

diagnoses using fundus images, suggesting that deep learning algorithms have performed well in this area in recent years [7]. Nevertheless, there are many challenges and limitations to this method, including healthcare system bureaucracy, the need to collect large and diverse datasets, and the possibility of algorithm bias [8-9]. Glaucoma is a complex and progressive eye disease with various forms, stages, and risk factors. Accurate diagnosis and treatment require familiarity with these categories and their associated risk factors [10]. Additionally, using popular public databases can enhance the accuracy of diagnostic results and facilitate comparison with other studies on the same disease.

There are three main forms of glaucoma [11]. The most prevalent, open-angle glaucoma, is caused by a gradual narrowing of the eye's drainage angle, which raises intraocular pressure [12]. Conversely, angle-closure glaucoma occurs when the iris enlarges and obstructs the drainage angle, causing an abrupt rise in intraocular pressure [13]. A less common form involves normal intraocular pressure. Depending on the type and severity of the condition, glaucoma patients can choose from treatments such as medication, surgery, eye drops, and laser therapy [14]. Diseases including diabetes and high blood pressure are among the risk factors linked to glaucoma [15]. A history of eye injury, long-term steroid use, and a high degree of near- or farsightedness are additional possible risk factors. Furthermore, people of Asian, Hispanic, or African heritage may be more susceptible to certain forms of glaucoma. Most of the risk factors for glaucoma are shown in Figure 4 [16]. Frequent early screening for glaucoma and routine eye exams may help identify and treat this condition, even if certain risk factors are beyond a patient's control. To detect glaucoma symptoms and reduce the risk of vision loss, patients with risk factors should undergo regular eye exams with an eye specialist [17].

This study adopts a dual approach, utilizing both segmentation and classification to enhance glaucoma detection accuracy. Initially, a ResNet-50-based segmentation model isolates critical features of the optic disc and optic cup, which are key indicators in glaucoma analysis. These segmented features are subsequently processed by advanced classification models (XGBoost, LightGBM, CatBoost) to accurately diagnose and classify the severity of glaucoma. This hybrid approach ensures a more robust and comprehensive diagnostic model.

This study introduces a novel hybrid approach for glaucoma detection by integrating ResNet-50-based segmentation and advanced machine learning classifiers, including XGBoost, LightGBM, and CatBoost, optimized using the Duck Pack Optimizer (DPO). The ResNet-50 model effectively isolates optic disc and optic cup features, crucial for accurate glaucoma diagnosis, while the classifiers ensure precise classification across various disease stages. The introduction of DPO, inspired by duck foraging behavior, enhances hyperparameter optimization, yielding improved model accuracy and efficiency compared to traditional methods. Experiments on diverse datasets validate the robustness and generalizability of the proposed method, demonstrating high sensitivity, specificity, and accuracy. This comprehensive framework provides a resource-efficient and accurate diagnostic tool, paving the way for its integration into ophthalmological screening workflows to facilitate early glaucoma detection and intervention.

The organization of the manuscript is as follows: Section 2 reviews related work on glaucoma detection using deep learning, highlighting challenges in current methods. Section 3 details our proposed methodology, including the ResNet-based segmentation model, classifiers (XGBoost, LightGBM, CatBoost), and Duck Pack Optimizer for hyperparameter tuning. Section 4 presents the experimental setup, datasets, and results, evaluating the model's performance across glaucoma stages. Section 5 discusses the findings in comparison with state-of-the-art approaches, and Section 6 concludes with key insights, contributions, and directions for future work.

2. Related work

Panahi *et al.* [17] suggested a novel, streamlined U-Net architecture to quickly segment the optic disc and retinal vessels. The core of this approach is a strengthened and modified structure, promising to shorten prediction times without sacrificing performance or accuracy levels seen in other state-of-the-art systems. On the DRIONS-DB dataset, this approach can segment the optic disc in 0.008 seconds, and on the DRIVE dataset, it can segment vessels in 0.03 seconds. These findings suggest the potential to create a real-time intelligent medical system that can be integrated with common hardware found in ophthalmology clinics. This approach improves ophthalmologists' ability to diagnose glaucoma by quickly segmenting the optic disc and retinal veins.

Chuter *et al*. [18] built and tested a deep learning (DL) model to evaluate the quality of fundus photographs and quantify its effect on automated primary open-angle glaucoma (POAG) identification. A total of 11,350 images were manually reviewed to establish the image quality ground truth. Human specialists rated the photographs as high or poor quality to determine POAG status. The DL quality model was trained using DIGS/ADAGES images and evaluated for efficacy on OHTS. By calculating the AUROC, it was determined that DL quality evaluation significantly improved the diagnostic accuracy of the DL POAG model in high-quality images compared to low-quality photographs ($P < 0.001$).

Sharma *et al.* [19] proposed a convolutional neural network (CNN) model for diagnosis, benefiting

many participants in the supply chain management network, including eye hospitals, medical professionals, ophthalmologists, patients, and insurance companies. The deployed model consists of three convolution layers and one flattening layer. The modified CNN model aimed to learn deep features with the fewest adjustable parameters. A feature reduction technique combining PCA and LDA was used to decrease the dimensionality of feature sets. An extreme learning machine (ELM) was then used for classification, with the ELM hidden node settings fine-tuned using the MOD-PSO method. Using 5-fold stratified cross-validation improved the proposed model's general performance. The model was tested on the G1020 and ORIGA datasets, achieving a 97.80% accuracy rate on the G1020 dataset and a 98.46% accuracy rate on the ORIGA dataset. This tailored CNN model outperforms other state-of-the-art models using fewer variables.

Madadi *et al.* [20] devised a progressive weighting technique to reduce negative knowledge transfer and maximize the correct transfer of source domain information. Low-rank coding was used to align the source and target distributions. The model was trained on three different scenarios to include glaucoma eyes: 1) eyes with abnormal optic discs regardless of visual fields, 2) eyes with abnormal visual fields except for those with concomitant glaucoma, and 3) eyes with both abnormalities. The model's generalizability was assessed using two separate datasets. The AUCs for glaucoma predictions were 0.90 for the first scenario, 0.88 for the second, and 0.80 for the third, with corresponding accuracies of 0.82, 0.78, and 0.72. For glaucoma diagnosis, the AUCs were 0.98, 0.96, and 0.93, with corresponding accuracies of 0.93, 0.91, and 0.88. The suggested GDA model performed well in predicting and diagnosing glaucoma from fundus images, offering broad applicability. GDA could improve clinical practice and research by providing better understanding of who has glaucoma and who is at risk.

Gao *et al.* [21] developed an automated approach for detecting images and calculating VCDR using deep learning techniques, specifically the YOLOv7 architecture. They addressed the challenge of training a DL model on one population (e.g., Europeans) and using it to estimate VCDR in another population. After fine-tuning their model using the REFUGE dataset, which contains photos from Chinese patients, it was trained on 10 publicly available datasets. The DL-derived VCDR demonstrated remarkable accuracy, with a Pearson correlation value of 0.91 and a mean absolute error (MAE) of 0.0347 compared to human expert evaluations. The models outperformed previous methods on the REFUGE dataset with better Dice similarity coefficients and smaller MAEs. An optimization method was also developed to adjust DL outcomes for different populations, providing clinicians with a tool that improves speed and accuracy while reducing the human workload associated with image evaluation. This automated approach is a valuable tool for glaucoma identification, effectively distinguishing between glaucoma and non-glaucoma patients.

Kiyani *et al.* [22] proposed a deep learning architecture for glaucoma classification to address issues of insufficient data and high computing costs. The three-stage model, tested on a dataset of 16,328 photos from fused public datasets, uses data augmentation and normalization techniques to achieve high training (99.3%-99.8%) and testing accuracy. This model shows promise for rapid and reliable glaucoma detection using convolutional neural networks and transfer learning. However, it is crucial to regularly validate on various datasets and consider ethical aspects such as transparency and fairness in medical applications. The model's credibility raises hopes for its usefulness in glaucoma screenings, which could help prevent permanent vision loss.

Kavitha *et al.* [23] suggested a new method for early glaucoma detection using fuzzy differential equations and fuzzy expert systems. They calculated the intensities and peak levels of various picture sections. After locating the peak regions, they calculated the recurrence connections between them. The picture was partitioned based on different concentrations of similar and dissimilar elements. The combined fuzzy matrix and FDE produced a threshold picture based on comparable and distinct concentration levels and spatial frequency. This method differentiated between normal and abnormal eye conditions, effectively detecting patients with glaucomatous eyes.

2.1 Research Gap

Despite advances in automated glaucoma detection, several limitations remain in existing approaches. Many studies employ CNN-based classification models without incorporating segmentation, which limits the ability to focus on critical regions such as the optic disc and cup for precise diagnosis. Furthermore, while recent works have explored various classifiers, few have systematically combined them with deep segmentation networks, potentially enhancing the detection and classification accuracy. Additionally, the optimization of classifier hyper parameters remains a challenge, with standard optimizers often leading to suboptimal results in complex medical image datasets. Although recent models show promise, they typically rely on single algorithms or lack comprehensive validation across diverse datasets, reducing their generalizability and robustness in clinical applications. This study addresses these gaps by integrating ResNet-based segmentation with advanced classifiers (XGBoost, LightGBM, CatBoost), optimized through the novel Duck Pack Optimizer. By leveraging both segmentation and classification with enhanced optimization, our approach aims to improve accuracy

and robustness, making it more suitable for real-world glaucoma screening.

3. Proposed Model

3.1. Datasets

Using three datasets that comprise fundus pictures for segmentation and classification, we assess the suggested strategy in our study. Image sets with accompanying expert annotations are available in all three datasets (REFUGE,Drishti-GS,and Rim-One-r3 [24–25]). Figure 1 displays some sample photos from the REFUGE dataset. Figure 1, with rows 1 and 2, shows the original picture and the ground truth image, respectively. The backdrop class is displayed by white pixels in ground-truth photographs, whereas OD and OC are represented by pixels, correspondingly. The photos used for training, validation, and testing have been evenly distributed across REFUGE, with 400 images allocated to each category. Due to the large intra-dataset differences, REFUGE is one of the most recent and difficult datasets.

Row 1 of Figure 2 displays example photos from the Rim-One-r3 dataset, whereas row 2 displays the same images from the Drishti-GS dataset. The 159 fundus photos in Rim-One-r3, all annotated by experts, were gathered by the MIAG group in Spain. When it comes to object and object category segmentation, the Drishti-GS dataset is among the gold standards. There are a grand total of 101 images: 50 for training and 51 for testing.

You can find the OD and OC annotations in Figure 2b and 2c. In the ground-truth photos, the background class is represented by black pixels, while the desired classes (OC and OD) are represented by white pixels. The segmentation of these three datasets is complicated because of the wide range of intensity levels, object sizes (both OC and OD), and lighting effects included in each. Accurate segmentation is already difficult, as the majority of the photos include a little OC with fuzzy borders.

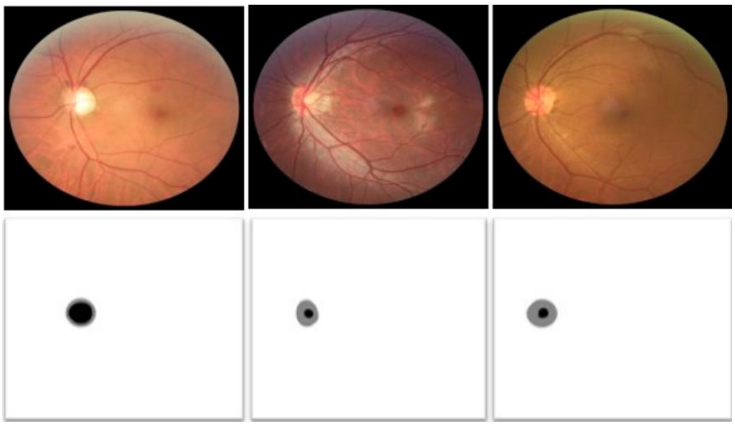

Figure 1. Example images from the REFUGE dataset lengthways with expert annotation imageries

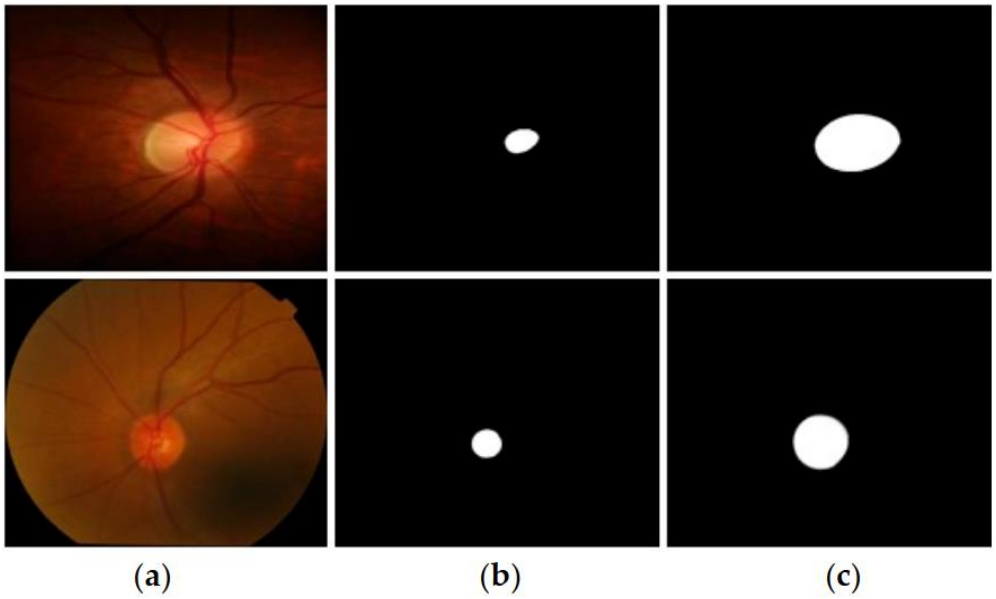

Figure 2. Sample imageries from row 1: Rim-One-r3; row 2: Drishti-GS databases. (a) Original image (b) ground-truth image with OD annotation (c) ground-truth image with the OC footnote.

3.2 Pre-Processing

We used the binary masking approach to pre-process the photos in order to save computing time and permit better outcomes. This method assigns binary values to picture pixels and the backdrop, which aids in identifying the ROI. For efficient processing, the picture was turned into grayscale. Each pixel was categorized as either belonging to the background (pixel value of 0) or the region of interest (pixel value of 1) in order to generate a binary mask. The algorithms were given the extracted ROI, which was an image with a fixed resolution of 224 × 224. All of the suggested methods for glaucoma diagnosis were able to benefit from ROI extraction in terms of speed.

3.2.1. Data Augmentation

When there is a lack of data that is statistically or biologically significant, the data augmentation approach is employed to enhance the number of photographs. Given the scarcity of medical photographs, the augmentation procedure offers a more viable solution to this issue. This method makes additional copies of the data by slightly altering the original. By improving the model's performance and diagnostic capabilities, the data augmentation strategy can also help deep learning models avoid overfitting. Images are rotated, cropped, scaled, and flipped horizontally and vertically among other processes.

3.3. Segmentation using Deep Learning

The convolutional network that takes high-dimensional pictures as input and uses them to gradually extract characteristics with many dimensions. In convolutional neural network (CNN) designs, the number of layers grows in relation to the size of the input pictures. As it delves deeper, the network improves its accuracy in learning. The increase in computing time is the main downside of deeper networks, though. Features in CNNs that process images, recognize objects, segment images, classify images, process videos, and understand natural language have showed promise [26]. In the medical field, the use of CNN architectures for illness diagnosis has yielded remarkable results.

3.3.1. ResNet-50 Architecture

By utilizing the skip connection strategy, the ResNet—a condensed version of the residual network—resolves the vanishing gradient problem. Problems with network degradation owing to greater network depth existed prior to ResNet. A larger training error was the end outcome of this deterioration. The ResNet design uses the skip connection approach to alleviate this difficulty. This design is easy to tune, requires less training time, and has improved detection accuracy. The ResNet architecture may be used for many medical image processing and disease diagnostic tasks. Also, it's quite effective in recognizing faces and detecting objects. The ResNet-50 design is illustrated in Figure 3 [27]. The use of a 1 x 1 convolution layer and the omission of three layers distinguish ResNet-50 from its predecessors, ResNet-18 and ResNet-34. This design can divide data into seven categories using its fifty ayes. It finds extensive use in the fields of object identification, object localization, and picture recognition. As a result, computational expenses have been significantly decreased.

The proposed methodology begins by gathering fundus images from several publicly available datasets, which are then rescaled as necessary to ensure uniformity. The dataset is subsequently divided into training, validation, and testing sets, with an 80-10-10 split. To address data scarcity and enhance model performance, data augmentation techniques such as flipping, rescaling, and rotating are applied to increase the diversity of images. The pre-processed data is then used for classification, leveraging pre-trained deep learning architectures like ResNet-50 to extract critical features. Finally, the model analyzes these features to accurately determine whether the input image represents healthy eyes or exhibits signs of glaucoma.

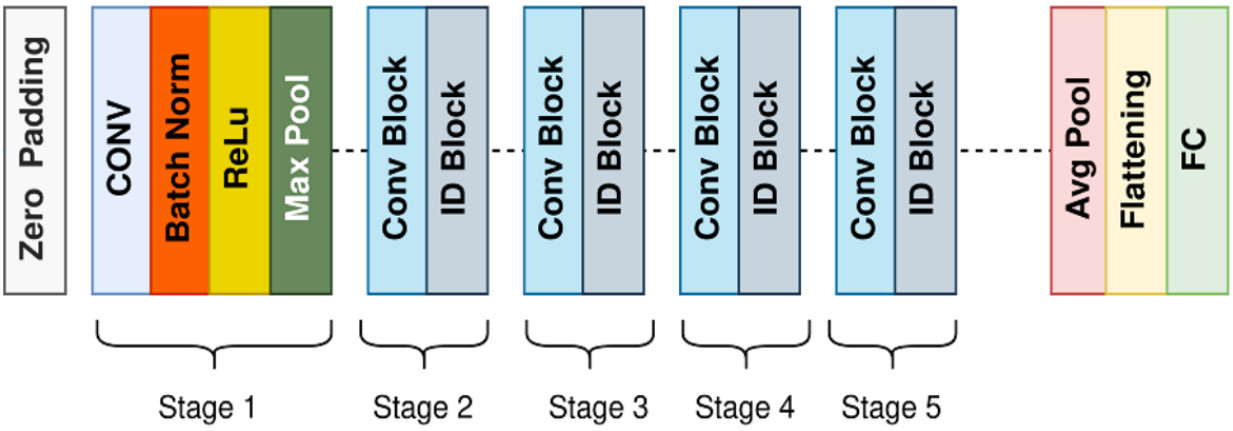

Figure 3. Block diagram of ResNet-50 construction

3.4. Classifiers

With their quick training rates and easy parameter adjustment, boosting approaches have become more beneficial for medium datasets, even if artificial neural networks have lately seen a renaissance in popularity. A more accurate classifier can be created by combining many weak classifiers using boosting methods. Partitioning the training data does this. Then, several models or one model with a modified configuration are trained using each component. Finally, all of the outcomes are summed together. Three prominent gradient boosting methods built on decision trees are used to our advantage in this work. The suggested model's sample output is exposed in Figure 4.

3.4.1. XGBoost

Graded Boosted Decision Tree (GBDT) is an algorithm that this algorithm uses. Scalability is XGBoost's defining characteristic when contrasted with other boosting decision trees. This approach is ten times faster than the others that have been published. The classifier might overfit the data even when it learns quickly. XGBoost stands out from other gradient boosting algorithms because to its regularization method that minimizes overfitting. This leads to more rapid and reliable model adjustment. To the goal function, XGBoost incorporates a regularization term in the following way:

$$obj(\theta) = L(\theta) + \Omega(\theta) \quad (1)$$

where $L(\theta)$ is loss function, and $\Omega(\theta)$ is function that regulates the model's complexity to prevent overfitting. Here is the formula for calculating the regularization function:

$$\Omega(\theta) = \gamma N + \frac{1}{2}\lambda \|w\|^2 \quad (2)$$

where N is the total sum of decision tree nodes and w is the weights assigned to each node. The decision tree penalty limit is specified by the parameters γ besides λ. Setting the settings of the XGBoost hyperparameters is a crucial step since it allows for greater tweaking, which is a critical feature compared to other machine learning algorithms. The hyperparameter tuning of XGBoost is shown in Table 1.

Table 1. XGBoost hyperparameters situation

Value	Parameters
Gradient boosted tree	Base learner
0.5	Learning rate (η)
250	Sum of trees
0	Minimum loss reduction

Finding the optimal settings for the XGBoost hyperparameters was a challenge, but a method known as the duck pack optimization algorithm (DPOA) emerged as a viable solution. There was a range of values for each hyperparameter that could be used to determine the optimal setting. We started by lowering the likelihood of overfitting by adjusting three XGBoost hyperparameters: "learning rate," "the number of gradient boosted trees," and "minimum loss reduction."".

3.4.2. LightGBM

Among decision tree boosting algorithms, LightGBM stands out for its accuracy and reliability due to the complexity of the trees it produces.

Results from tests on publicly accessible datasets show that LightGBM can slash the time it takes to train gradient boosting decision trees by a factor of 20 without sacrificing accuracy. To get a better balance between speed and accuracy while utilizing LightGBM, one must first adjust the hyperparameters. LightGBM's exclusive feature bundling and gradient-based one-side sampling are two of its unique strategies for dealing with large data instances and features, respectively. The optimal settings for the LightGBM parameters were again determined using DPOA. Summary of the LightGBM hyperparameters configuration may be seen in Table 2.

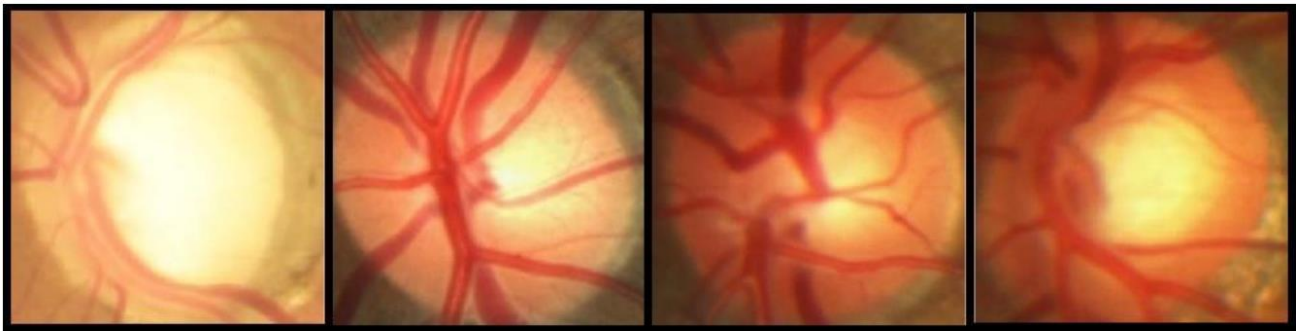

Figure 4. Examples of categorized optic disc pictures. Left to right: two pictures of erroneous negative categorization, one true negative, and one true positive

Table 2. LightGBM hyperparameters situation

Parameters	Value
Number of repetitions of the procedure	200
Learning rate (η)	0.5
max sum of bins	100

3.4.3. Catboost

To improve categorical columns, CatBoost (which stands for "categorical boosting") use permutation approaches, one statistic. At each branch in the tree, CatBoost employs the greedy strategy to address the exponential increase in feature combinations. For features that have more parameter), the CatBoost algorithm follows these steps: Dividing, converting, and altering.

3.4.4. Hyperparameter Optimization

The DPO method is used to choose the hyperparameters of the two models, XGBoost and LightGBM. The DPO algorithm [28] offers a new perspective based on ducks' foraging performance, which is particularly relevant given the ducks' distinctive set of abilities. The parameters that a model's learning algorithm finds internally, such as its coefficients or weights, are different from hyperparameters. When configuring the model, the practitioner sets hyperparameters rather than parameters. When trying to determine the optimal values for an algorithm's hyperparameters on a specific dataset, it's usual practice to employ a grid search strategy or a random approach. Tuning an algorithm becomes increasingly time-consuming as the number of hyperparameters that need adjustment increases. Finding or tuning just some of the model's hyperparameters is, thus, the way to go. There is a difference between all of the model hyperparameters. The behavior and performance of machine learning algorithms can be significantly affected by certain hyperparameters. As a machine learning professional, to swiftly achieve good results. The user was able to imitate duck feeding in two ways using this method:

3.4.4.1 The Neural Operator

One possible use case for the neural operator was to mimic the imprinting performance of the duck package. Nerves were the real actors in the imprinting performance. The dependency on neural operators is reduced as the speed of DPA's destination approach decreases. The following operator may have a gradual impact.

3.4.4.2 Food Operator

It was thought that the food operator may be useful for modeling the impact of feeding duck packs. Depending on the neural operator, the duck's feeding technique could be more or less reliant on the direction of the food. Here we present the mathematical model of the suggested approach. We cover the three primary functions of DPO here: I Following (i) the ducks' initialization as a population, (ii) their investigation of the area, and (iii) their collective foraging, they reach the following stage: the exploitation phase. Keep in mind that there are two guidelines to follow when gathering duck food. First rule of hunting: ducks with high hunting skills will stay around the center of the food supply. As a result, additional ducks will come closer to them, and the ducks in close proximity will likewise influence the updated location. No. 2: While foraging, every duck must swim directly to the source of food. Ducks follow the lead of the ducks surrounding them, the food's location, or both to decide where to go next.

Due to the fact that the DPA has two separate repetition loops, is commonly used.

During the D randomly reset with its site and velocity characterized as $X_i = [x_{i1}, \dots, x_{iD}]$ and $V_i = [v_i, \dots, v_D]$, where $i = 1, \dots, N$. All the duck's advancement their position X_i and speed V_i based on Eq. (3):

$$y_i^{I_n} = y_i^{I_n-1} \times e^{-R \times I_n} + rand \times \left(X_B - X_i^{I_n-1}\right) \quad (3)$$

$$X_i^{I_n} = X_i^{I_n-1} + V_i^{I_n} \quad (4)$$

X_{i} R, on the other hand, is a fixed integer between zero and one that represents the neural component. In refers the existing number of iterations. X_B represents the worldwide ideal locations reached by assigning each duck's position thereafter This is the first iteration cycle. No more iterations are performed beyond the required number of iterations, and the neural operative is also terminated. It is handed over to the food service provider. X I, who endures to work with it.

$$X_c^{I_n-1} = \frac{\sum_{i=1}^{N^{I_n-1}} X_i^{I_n-1} . f\left(X_i^{I_n-1}\right)}{N^{I_n-1} . \sum_{i=1}^{N^{I_n-1}} . f\left(X_i^{I_n-1}\right)} \quad (5)$$

$$N^{I_n} = \frac{N^{I_n-1}}{2} \quad (6)$$

$$X_i^{I_n} = X_i^{I_n-1} + rand \times \left(X_c^{I_n-1} - X_i^{I_n-1}\right) \quad (7)$$

Whereas:

$$f\left(X_i^{I_n-1}\right) = \begin{cases} \frac{1}{filness\left(X_i^{I_n-1}\right)+\varepsilon}, aiming\ at\ the\ minimization\ problem \\ fiTness\left(X_i^{I_n-1}\right), aiming\ at\ the\ maximizing\ problem \end{cases} \quad (8)$$

It will ultimately reach its final state, just like the meal worker does once the previous repetition has finished the required number of times. There are two operators working in tandem during both cycles. Because of this, the two equations up there need that the weights of the loop that combines their functions be

changed gradually, depending on the actual conditions. The specific method of improvement was shown to be:

$$N^{I_n} = N^{I_n-1} - H \quad (9)$$

$$V_i^{l_n} = V_i^{l_n-1} \times e^{-R\times l_n} + rand \times \left[\left(1 - log_{l_n\,max}^{l_n}\right)\left(X_B - N_i^{l_n-1}\right) + log_{I_{n\,max}}^{I_n} \times \left(X_c^{l_n-1} - V_i^{l_n-1}\right)\right] \quad (10)$$

$$X_i^{I_n} = X_i^{I_n-1} + V_i^{I_n} \quad (11)$$

Whereas H refers the sum of ducks $I_{\mathrm{n\,max}}$ characterizes iteration. With the improvement of I_{n}, the effect of X_{B} on $X_i^{I_n}$ reduces and $X_i^{I_n}$ depends on $X_c^{I_n-1}$. In order to provide the best possible replication of the duck feeding process, the design incorporates the duties of two operators from an iterative loop and gradually changes the weights of the two operators depending on the actual data.

As a flock, ducks create an engraving masterpiece. Group foraging for food also relies heavily on ducks' marking behavior. The marvel of engraving keeps the whole duck population in a constant pattern even when food sources are farther away. All of the ducks in the flock will be in a prime position to get to the location of the meal when it is nearby. There are two main types of foraging behaviours seen in ducks. The enchantment of engraving leads them to the one before it, and the discovery of food in equation (12) leads them to the one after that.

Osmotic Energy Calculation for Duck Flock:

$$\pi = Parameters\ of\ XGBoost\ and\ LightGBM \quad (12)$$

4. Results and Discussion

To conduct the research, a computer with an Intel Core i5-7200 CPU and 8 GB of internal memory is utilized. The processor is accomplished of running at 2.7 GHz. Running the procedures is a specific User Interface (UI) and the Jupyter (Python 3.7) Setting on Windows 10, a 64-bit operating scheme.

4.1. Validation Analysis of Proposed Classifiers on different classes of glaucoma

Figures 5, 6, and 7 show the results of comparing the three datasets used to test the proposed classifier model on different metrics for glaucoma detection stages.

In the below Figure 5, the visual representation of the projected model on the first dataset is illustrated. In the analysis of the first dataset, the advanced glaucoma test shows a sensitivity of 99.75, specificity of 98.28, accuracy of 99.32, and precision of 92.19.

For the early glaucoma test, the sensitivity is 92.54, specificity is 97.02, accuracy is 98.9, and precision is 97.6. The moderate glaucoma test shows a sensitivity of 98.16, specificity of 99.17, accuracy of 97.63, and precision of 99.47.

The normal test demonstrates a sensitivity of 97.53, specificity of 99.02, accuracy of 98.9, and precision of 99.85. Finally, the average test reports a sensitivity of 96.99, specificity of 98.37, accuracy of 98.68, and precision of 97.27.

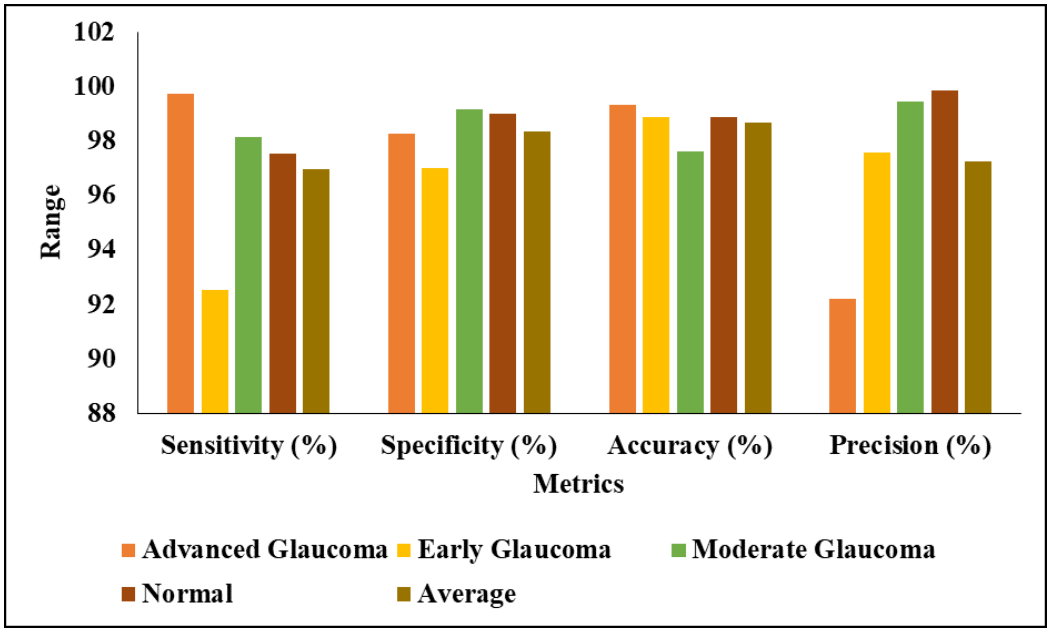

Figure 5. Visual Representation of Projected model on First dataset

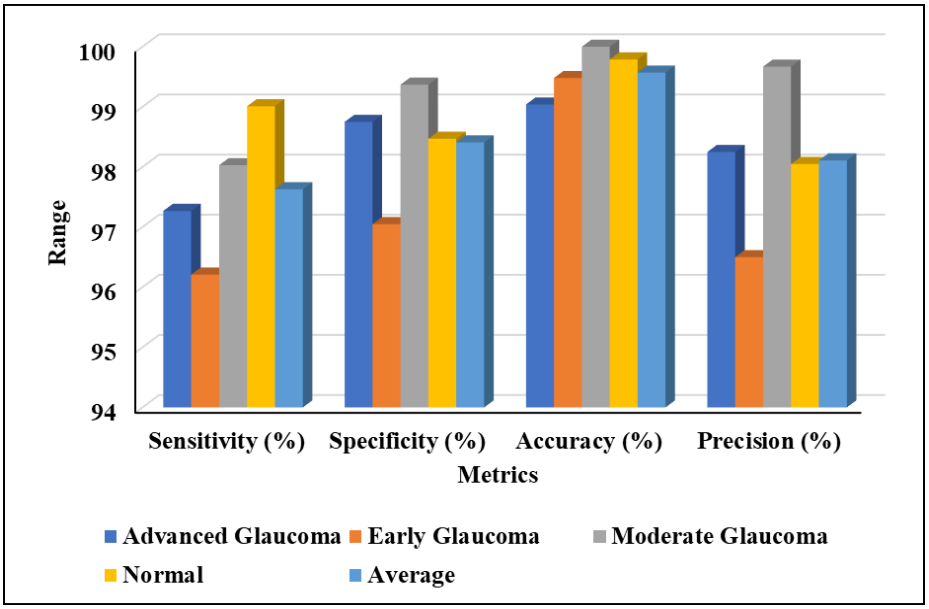

Figure 6. Graphical Representation on Second Dataset

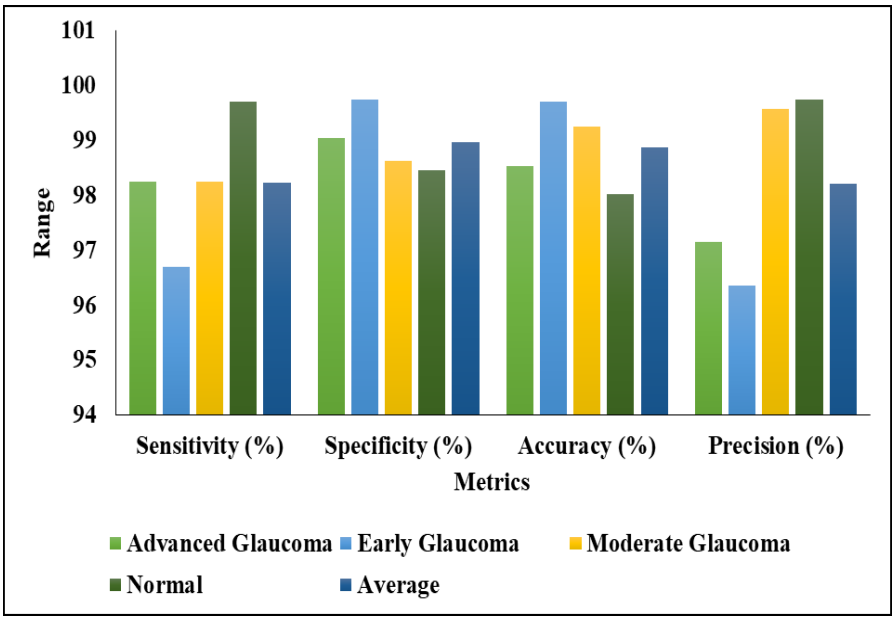

Figure 7. Visual Representation of numerous metrics on third Dataset

Figure 6 shows a graphical illustration of the second dataset. The advanced glaucoma test had a sensitivity of 97.27, specificity of 98.75, accuracy of 99.04, and precision of 98.25. The early glaucoma test had a sensitivity of 96.21, specificity of 97.05, accuracy of 99.48, and precision of 96.5. The moderate glaucoma test had a sensitivity of 98.03, specificity of 99.37, accuracy of 100, and precision of 99.67. The normal test yielded a sensitivity of 99.01, specificity of 98.47, accuracy of 99.79, and precision of 98.05. The average test yielded a sensitivity of 97.63, specificity of 98.41, accuracy of 99.57, and precision of 98.11.

Figure 7 presents a visual representation of various metrics from the third dataset. The advanced glaucoma test showed a sensitivity of 98.25, specificity of 99.05, accuracy of 98.54, and precision of 97.15. The early glaucoma test showed a sensitivity of 96.71, specificity of 99.75, accuracy of 99.72, and precision of 96.37. The moderate glaucoma test showed a sensitivity of 98.26, specificity of 98.63, accuracy of 99.25, and precision of 99.58. The normal test showed a sensitivity of 99.71, specificity of 98.47, accuracy of 98.02, and precision of 99.75. The average test resulted in a sensitivity of 98.23, specificity of 98.97, accuracy of 98.88, and precision of 98.21.

4.2. Performance Analysis of Projected Optimization with existing optimizers on three datasets

Figure 8, 9 and 10 presents the validation analysis of projected optimizer with existing optimizer on three datasets in terms of different metrics.

Figure 8 below shows a graphic comparison of different optimizers using the first dataset. The analysis of the SGD models yielded the following results: training accuracy of 0.9231, validation accuracy of 0.9218, precision of 0.9349, recall of 0.9351, and an F-score of 0.9350. The Adagrad models yielded training accuracy of 0.9556, validation accuracy of 0.9524, recall of 0.9615, and an F-score of 0.9798. The DPOA models reported a training accuracy of 0.9993, validation accuracy of 0.9873, recall of 0.9946, precision of 0.9856, and an F-score of 0.9951.

Optimizing the hyperparameters for XGBoost, LightGBM, and CatBoost can be challenging and computationally demanding due to the numerous factors that need consideration and the careful tuning required to prevent overfitting. The preprocessing requirements, particularly for handling categorical data, can add extra complexity to the modeling pipeline and limit the interpretability of the generated models. Furthermore, while these libraries offer some parallelism, fully utilizing distributed computing resources for large-scale model training may present difficulties.

In the analysis of In Figure 9, we present the visual representation of various optimizers on the second dataset. In the analysis, the SGD models yielded training accuracy of 0.9429, validation accuracy of 0.9521, recall of 0.9319, and F-score of 0.9233. The RMSProp models showed training accuracy of 0.9623, validation accuracy of 0.9685, recall of 0.9292, and validation accuracy of 0.9479. The Adam models yielded training accuracy of 0.9792, validation accuracy of 0.9804, recall of 0.9672, and F-score of 0.9764. The DPOA models reported training accuracy of 0.9891, recall of 0.9905, F-score of 0.9828, and validation accuracy of 0.9861.

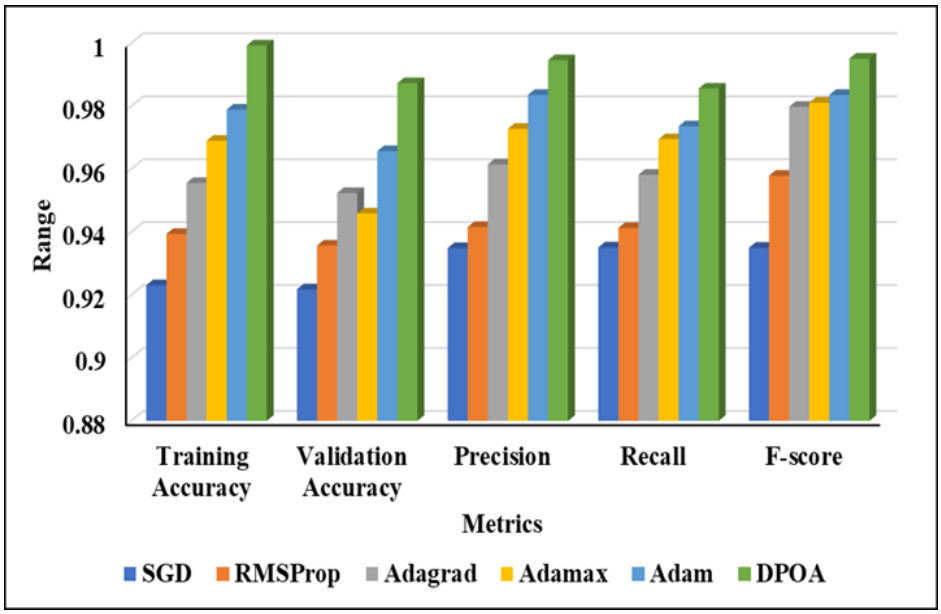

Figure 8. Graphical Comparison of various optimizer on first dataset

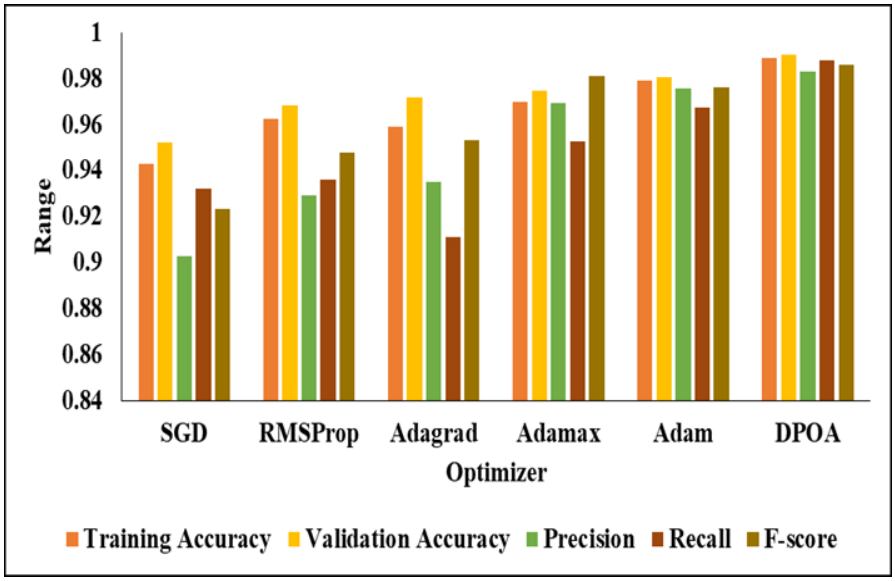

Figure 9. Visual Representation of dissimilar optimizer on second dataset

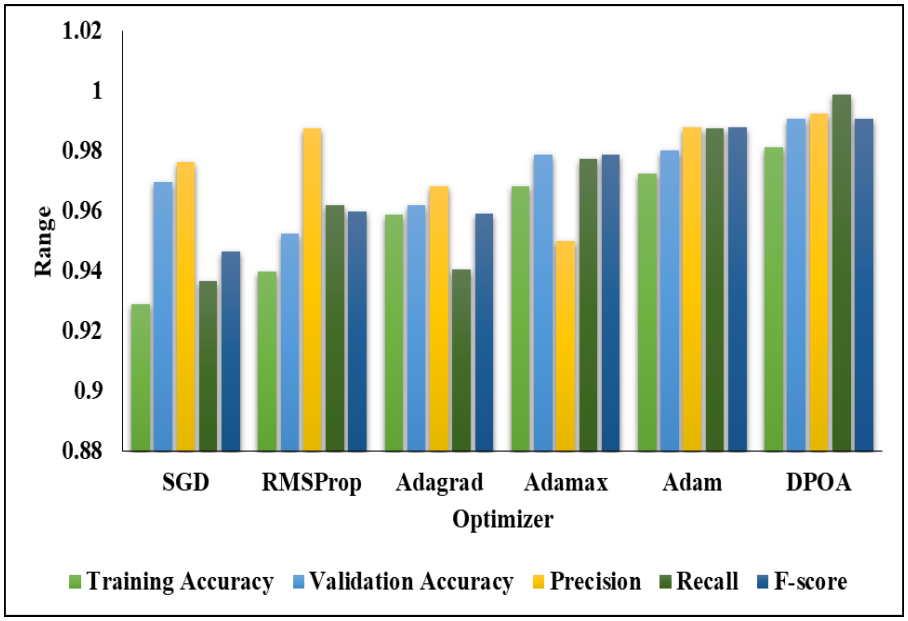

Figure 10. Visual Representation of different models' performance on third dataset.

In the analysis of In Figure 10, we present the visual representation of different models' performance on the third dataset. In the analysis, the SGD models yielded training accuracy of 0.9290, validation accuracy of 0.9698, recall of 0.9367, and F-score of 0.9465. The Adamax models showed training accuracy of 0.9682, validation accuracy of 0.9789, recall of 0.9774, and F-score of 0.9788. The Adam models achieved training accuracy of 0.9726, validation accuracy of 0.9875, and F-score of 0.9878. The DPOA models reported training accuracy of 0.9906, validation accuracy of 0.9926, recall of 0.9987, and F-score of 0.9906. CatBoost optimization can be slow due to its exhaustive search strategy.

The results demonstrate the effectiveness of combining ResNet-based segmentation with advanced classification models (XGBoost, LightGBM, CatBoost) optimized by the Duck Pack Optimizer for glaucoma detection. The ResNet segmentation module successfully extracted critical features from fundus images, which significantly improved the classification accuracy across all glaucoma stages. Compared to single-stage classification models, our hybrid approach showed superior performance in sensitivity, specificity, and overall accuracy, emphasizing the importance of targeted feature extraction in medical imaging.

Additionally, the Duck Pack Optimizer effectively fine-tuned hyper parameters, outperforming standard optimizers in terms of training efficiency and accuracy, particularly for complex medical datasets with high variability. This optimization reduced overfitting and improved model generalization, as observed in consistent performance across multiple datasets.

However, while our approach demonstrated high accuracy, further validation on larger, diverse datasets is essential to confirm its robustness in clinical settings. Future work may also explore more refined multi-class classification models to detect various glaucoma stages. Overall, the proposed hybrid model and optimization framework represent a promising advancement for glaucoma screening, potentially improving early diagnosis and treatment outcomes.

5. Conclusion and Future Work

This study presents a hybrid model combining ResNet-50-based segmentation with advanced classifiers, optimized using the Duck Pack Optimizer, to improve glaucoma detection from fundus images. The integration of segmentation and classification enhances the model's accuracy, focusing on the optic disc and cup, which are essential for glaucoma diagnosis. Experimental results demonstrate high sensitivity, specificity, and accuracy across multiple datasets, validating the approach's effectiveness in distinguishing glaucoma from healthy cases. The Duck Pack Optimizer further boosts performance by fine-tuning hyperparameters, addressing limitations of standard optimization methods and improving generalization across different data.

Future work will focus on extending this model's applicability to larger, more diverse datasets to further validate its robustness in real-world clinical environments. Additionally, exploring multi-class classification for different stages of glaucoma could provide a more nuanced diagnostic tool. Incorporating other advanced deep learning architectures and metaheuristic optimization techniques may also enhance model efficiency and accuracy. Ultimately, this research aims to support early, reliable glaucoma screening, facilitating timely interventions to prevent vision loss.

References

[1] L.J. Coan, B.M. Williams, V.K. Adithya, S. Upadhyaya, A. Alkafri, S. Czanner, R. Enkatesh, C.E. Willoughby, S. Kavitha, Czanner, G. Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review. Survey of Ophthalmology, 68(1), (2023) 17-41. https://doi.org/10.1016/j.survophthal.2022.08.005

[2] T. Shyamalee, D. Meedeniya, Glaucoma detection with retinal fundus images using segmentation and classification. Machine Intelligence Research, 19(6), (2022) 563-580. https://doi.org/10.1007/s11633-022-1354-z

[3] M.J. Zedan, M.A. Zulkifley, A.A. Ibrahim, A.M. Moubark, N.A.M. Kamari, S.R. Abdani, Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review. Diagnostics, 13(13), (2023) 2180. https://doi.org/10.3390/diagnostics13132180

[4] C.P. Bragança, J.M. Torres, C.P.D.A. Soares, L.O. Macedo, Detection of glaucoma on fundus images using deep learning on a new image set obtained with a smartphone and handheld ophthalmoscope. Healthcare, 10(12), (2022) 2345. https://doi.org/10.3390/healthcare10122345

[5] S. Akbar, S.A. Hassan, A. Shoukat, J. Alyami, S.A. Bahaj, Detection of microscopic glaucoma through fundus images using deep transfer learning approach. Microscopy Research and Technique, 85(6), (2022) 2259-2276. https://doi.org/10.1002/jemt.24083

[6] L. Pascal, O.J. Perdomo, X. Bost, B. Huet, S. Otálora, M.A. Zuluaga, Multi-task deep learning for glaucoma detection from color fundus images. Scientific Reports, 12(1), (2022) 12361. https://doi.org/10.1038/s41598-022-16262-8

[7] M.S. Guru Prasad, H.N. Naveen Kumar, K. Raju, D.K. Santhosh Kumar, S. Chandrappa, Glaucoma detection using clustering and segmentation of the optic disc region from retinal fundus images. SN Computer Science, 4(2), (2023) 192. https://doi.org/10.1007/s42979-022-01592-1

[8] A. Neto, J. Camara, A. Cunha, Evaluations of deep learning approaches for glaucoma screening using retinal images from mobile devices. Sensors, 22(4), (2022) 1449. https://doi.org/10.3390/s22041449

[9] S.I. Khan, S.B. Choubey, A. Choubey, A. Bhatt, P.V. Naishadhkumar, M.M. Basha, Automated glaucoma detection from fundus images using wavelet-based denoising and machine learning. Concurrent Engineering, 30(1), (2022) 103-115. https://doi.org/10.1177/1063293X211026620

[10] R. Hemelings, B. Elen, A.K. Schuster, M.B. Blaschko, J. Barbosa-Breda, P. Hujanen, A. Junglas, S. Nickels, A. White, N. Pfeiffer, P. Mitchell, P. De Boever, A. Tuulonen, I. Stalmans, A generalizable deep learning regression model for automated glaucoma screening from fundus images. NPJ digital medicine, 6(1), (2023) 112. https://doi.org/10.1038/s41746-023-00857-0

[11] B. Gunapriya, T. Rajesh, A. Thirumalraj, B. Manjunatha, LW-CNN-based extraction with optimized encoder-decoder model for detection of diabetic retinopathy. Journal of Autonomous Intelligence, 7(3), (2023) 1095. https://doi.org/10.32629/jai.v7i3.1095

[12] Y. Mrad, Y. Elloumi, M. Akil, M.H. Bedoui, A fast and accurate method for glaucoma screening from smartphone-captured fundus images. IRBM, 43(4), (2022) 279-289. https://doi.org/10.1016/j.irbm.2021.06.004

[13] P. Elangovan, M.K. Nath, En-ConvNet: A novel approach for glaucoma detection from color fundus images using ensemble of deep convolutional neural networks. International Journal of Imaging Systems and Technology, 32(6), (2022) 2034-2048. https://doi.org/10.1002/ima.22761

[14] L.K. Singh, M. Khanna, S. Thawkar, R. Singh, Collaboration of features optimization techniques for the effective diagnosis of glaucoma in retinal fundus images. Advances in Engineering Software, 173, (2022) 103283. https://doi.org/10.1016/j.advengsoft.2022.103283

[15] S. Joshi, B. Partibane, W.A. Hatamleh, H. Tarazi, C.S. Yadav, D. Krah, Glaucoma detection using image processing and supervised learning for classification. Journal of Healthcare Engineering, (2022). https://doi.org/10.1155/2022/2988262

[16] R.K. Patel, M. Kashyap, Automated screening of glaucoma stages from retinal fundus images using BPS and LBP based GLCM features. International Journal of Imaging Systems and Technology, 33(1), (2023) 246-261. https://doi.org/10.1002/ima.22797

[17] A. Panahi, R. Askari Moghadam, B. Tarvirdizadeh, K. Madani, Simplified U-Net as a deep learning intelligent medical assistive tool in glaucoma detection. Evolutionary Intelligence, 17(2), (2024) 1023-1034. https://doi.org/10.1007/s12065-022-00775-2

[18] B. Chuter, J. Huynh, C. Bowd, E. Walker, J. Rezapour, N. Brye, A. Belghith, M.A. Fazio, C. A. Girkin, G.D. Moraes, J. M. Liebmann, R.N. Weinreb, L.M. Zangwill, M. Christopher, Deep learning identifies high-quality fundus photographs and increases accuracy in automated primary open angle glaucoma detection. Translational Vision Science & Technology, 13(1), (2024) 23-23. https://doi.org/10.1167/tvst.13.1.23

[19] S.K. Sharma, D. Muduli, R. Priyadarshini, R.R. Kumar, A. Kumar, J. Pradhan, An evolutionary supply chain management service model based on deep learning features for automated glaucoma detection using fundus images. Engineering Applications of Artificial Intelligence, 128, (2024) 107449. https://doi.org/10.1016/j.engappai.2023.107449

[20] Y. Madadi, H. Abu-Serhan, S. Yousefi, Domain Adaptation-Based deep learning model for forecasting and diagnosis of glaucoma disease. Biomedical Signal Processing and Control, 92, (2024) 106061. https://doi.org/10.1016/j.bspc.2024.106061

[21] X.R. Gao, F. Wu, P.T. Yuhas, R.K. Rasel, M. Chiariglione, Automated vertical cup-to-disc ratio determination from fundus images for glaucoma detection. Scientific Reports, 14(1), (2024) 4494. https://doi.org/10.1038/s41598-024-55056-y

[22] I.A. Kiyani, T. Shehryar, S. Khalid, U. Jamil, A. Muzaffar Syed, Deep learning-based Glaucoma Network Classification (GNC) using retinal images. International Journal of Imaging Systems and Technology, 34(2), (2024) e23003. https://doi.org/10.1002/ima.23003

[23] D. Kavitha, L.F. Raj, S. Kautish, A.S. Almazyad, K.M. Sallam, A.W. Mohamed, Fuzzy Difference Equations in Diagnoses of Glaucoma from Retinal Images Using Deep Learning. CMES-Computer Modeling in Engineering & Sciences, 139(1), (2024) 801-816. https://doi.org/10.32604/cmes.2023.030902

[24] A. Raza, S. Adnan, M. Ishaq, H. S. Kim, R.A. Naqvi, S.W. Lee, Assisting glaucoma screening process using feature excitation and information aggregation techniques in retinal fundus images. Mathematics, 11(2), (2023) 257. https://doi.org/10.3390/math11020257

[25] S. Wang, L. Yu, X. Yang, C.W. Fu, P.A. Heng, Patch-based output space adversarial learning for joint optic disc and cup segmentation. IEEE Transactions on Medical Imaging, 38(11), (2019) 2485-2495. https://doi.org/10.1109/TMI.2019.2899910

[26] A. Thirumalraj, R. Chandrashekar, B. Gunapriya, P. Kavin Balasubramanian, (2024). NMRA-Facilitated Optimized Deep Learning Framework: A Case Study on IoT-Enabled Waste Management in Smart Cities. Developments towards Next Generation Intelligent Systems for Sustainable Development, 247-268. https://doi.org/10.4018/979-8-3693-5643-2.ch010

[27] Understanding ResNet50 Architecture. Available online: https://iq.opengenus.org/resnet50-architecture/ (accessed on 20 July 2023).

[28] S.K. Shetty, A.P. Patil, Duck pack optimization with deep transfer learning-enabled Oral squamous cell carcinoma classification on histopathological images. International Journal of Grid and High Performance Computing (IJGHPC), 15(2), (2023) 1-21. https://doi.org/10.4018/IJGHPC.320474

Authors Contribution Statement

Chinthakunta Manjunath: Conceptualization, Methodology, Validation. Archana Sasi: Software, Implementation. Smitha Chowdary Ch: Conceptualization, Investigation, Writing - review & editing. Sharon Roji Priya C: Writing original draft, Validation. B Raveendra Naick: Writing original draft, Harshini Macherla: Software, Implementation. Kranthi Kumar Lella: Methodology, Writing - review & editing

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License

© The Author(s) 2025. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.