Journal of Theoretical and Applied Information Technology
15th July 2024. Vol.102. No. 13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENHANCING GLAUCOMA DIAGNOSIS: DEEP LEARNING MODELS FOR AUTOMATED IDENTIFICATION AND EXPLAINABILITY USING FUNDUS IMAGES

MANEESHA VADDURI[1], KUPPUSAMY. P[2*]

[1,2*]School of Computer Science and Engineering, VIT-AP University,
Amaravathi, Andhra Pradesh 522237, India

E-mail: [1]maneesha.21phd7026@vitap.ac.in, [2*]drpkscse@gmail.com

ABSTRACT

Glaucoma is a serious eye condition that poses a significant threat to vision health, often resulting in permanent sight loss by damaging the optic nerves. Detecting glaucoma early is crucial for effective management, aiming to reduce intraocular pressure and inflammation. However, current detection methods are resource-intensive and prone to human error, failing to detect the disease in its early stages. Deep Learning (DL) offers promising avenues for automated diagnosis, yet concerns persist regarding model reliability. Addressing this, the Enhanced Deep Learning Approach for Glaucoma Diagnosis (EDAGD) is introduced. Leveraging SegNet and ResNet-50 architectures, EDAGD achieves exceptional segmentation accuracies of 98.58% for the Optic Disc (OD) and 96.52% for the Optic Cup (OC) on the RIM-ONE dataset, while also demonstrating robust performance on the ACRIMA and REFUGE datasets. Furthermore, EDAGD utilizes cutting-edge visualization techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) and Grad-CAM++ to generate interpretable heatmaps, aiding in pinpointing critical regions for diagnosis. By accurately classifying segmented images, EDAGD achieves impressive performance metrics of 97.97% accuracy, 98.41% sensitivity, and 96.58% specificity. The potential impact of automated glaucoma diagnosis on healthcare systems includes reducing the burden on ophthalmologists, increasing accessibility to diagnostic tools in remote areas, and potentially lowering healthcare costs. By integrating advanced Deep Learning techniques with explainable AI methods, our approach not only improves the accuracy of glaucoma diagnosis but also builds trust among clinicians. This fosters seamless integration into clinical practice, ultimately advancing patient care by enabling timely and accurate diagnosis of glaucoma.

Keywords: *Glaucoma, segmentation, classification, Fundus images, Explainability*

1. INTRODUCTION

Glaucoma is a serious ocular condition that occurs as a result of an abnormality in the fluid equilibrium inside the eye, leading to an elevation in the intraocular pressure that affects the nerve cells. Elevating the Intraocular Pressure (IOP) results in detrimental effects on the optic nerve, leading to impaired vision. Glaucoma's delayed progression may result in ocular discomfort, and other symptoms may include sudden eye pain, halos around lights, impaired vision, and headache, especially with extreme IOP levels. Vision impairment may be effectively treated and cured with the use of medicinal and surgical interventions [1]. A variety of surgeries and medicinal interventions are effective in treating early-stage glaucoma. Current methodologies need a longer duration to get the most satisfactory outcomes. Early identification of glaucoma is crucial for avoiding blindness [2].

The global occurrence of glaucoma in individuals between 40 and 80 years of age is approximately 3.54%. This translates to roughly one in every 200 persons aged 40 might be affected by glaucoma, a number that escalates to one in eight by the age of 80. Among the various factors contributing to the risk of glaucoma, elevated IOP stands out as a significant contributor to optic nerve and blood vessel damage [3]. If not treated, glaucoma can lead to total impairment of the optic nerves and irreversible vision loss. Given its insidious nature, glaucoma often progresses gradually, causing minimal or no symptoms, earning it the title "sneak thief of sight." This highlights the importance of regular eye examinations for timely identification and assistance to avert permanent visual impairment [4]. Glaucoma remains a major contributor to

irreversible vision impairment worldwide, second only to cataracts. It is responsible for about 12% of blindness cases each year. The number of people impacted by glaucoma, especially those aged between 40 and 80, is anticipated to reach 111.8 million by the year 2040. Furthermore, statistics indicate that 2.4% of the general population and 4.7% of individuals aged 70 and above face the risk of developing this condition [5].

Glaucoma is defined by the deterioration of the Retinal Ganglion Cells (RGCs), which can be triggered by various underlying disorders. This degeneration of RGCs poses significant health risks resulting in severe vision impairment if not treated.

(i) Alterations in the Optic Nerve Head (ONH) structure and the layer of nerve fibers.

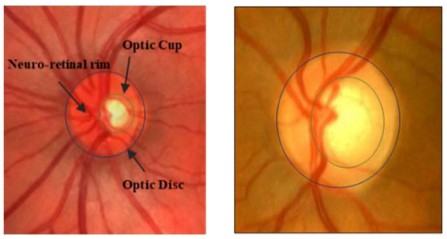
(ii) Concurrent limitations in visual field functionality.

Glaucoma poses a significant threat to vision, often resulting in peripheral vision deterioration and potential blindness if left untreated. Despite the importance of prompt diagnosis and treatment, there is no known solution for glaucoma. Therefore, the development of automated methods for early glaucoma detection is essential for effective disease management [6]. Retinal fundus images are crucial for evaluating the condition of different eye parts, such as the optic nerve, retina, vitreous humor, macula, and blood vessels. Ophthalmologists utilize fundus cameras to capture these images, which are then employed in diagnosing eye diseases such as glaucoma. Glaucoma can cause changes in the morphology of the cup area, the focal segment of the ONH. These changes serve as early indicators of glaucoma progression, as the ONH transports optical data from the retina to the brain.

The asymptomatic form of glaucoma is one of its problems in the early stages, leading to gradual optic nerve damage and eventual vision deterioration. Detecting glaucoma early is crucial to prevent irreversible visual impairment. Increased OC excavation, leading to changes in the cup-to-disc ratio (CDR), is a physiological indication of glaucoma progression. Ophthalmologists use CDR measurements to monitor glaucoma development [7].

Glaucoma presents in various forms, including angle-closure glaucoma, normal-tension glaucoma, primary congenital glaucoma, open-angle glaucoma, and others. Open-angle glaucoma is the most common form, characterized by gradual pressure rise due to partial drainage canal blockage. On the other hand, angle-closure glaucoma results from sudden drainage blockage, leading to exponential pressure increases. The damage caused by glaucoma is primarily due to elevated intraocular pressure resulting from blocked drainage canals. Without treatment, glaucoma can progress to complete blindness, underscoring the importance of early treatment.

The objective is to enhance the visibility of the decision-making process employed by DL methods, thereby bolstering their reliability and facilitating informed decision-making by medical professionals.



a. Anatomy of a Healthy OD b. Glaucoma-affected OD

Figure 1: Digital Fundus Images Focused on OD

Figure 1 depicts the fundus image of the retina serving as a crucial ocular record, capturing details such as the OD, the OC, and the blood vessels, providing insight toward assessing the state of the retina. Ophthalmologists and medical professionals commonly utilize these images to monitor the development of ocular disorders, particularly glaucoma. Within a fundus image, the OD, also referred to as the ONH, denotes the location where nerve fibers and blood vessels penetrate the retina. Positioned within the OD is the OC, the brightest region at the retina's center. The region between the OD and OC is called the neuroretinal rim, with its width ratios varying based on the OD structure [8]. The primary indication of glaucoma is often the enlargement of the OC, termed cupping, typically resulting from elevated pressure of the eye that deteriorates the optic nerve and leads to the loss of nerve cells. This damage subsequently increases cup size, manifesting as cupping.

Timely and accurate detection of glaucoma is essential to prevent irreversible vision loss. However, existing techniques often require prolonged durations to yield satisfactory outcomes, and manual detection, reliant on expert knowledge, poses challenges due to the scarcity of ophthalmologists in hospitals. Moreover, these methods are prone to human error and may fail to

identify the disease in its early stages. Hence, there is an urgent need to implement automatic techniques for detecting glaucoma with high efficiency and accuracy [9]. DL offers promising opportunities for automated diagnosis by processing large volumes of data and identifying patterns that human observers may overlook, thus providing a potential solution to the shortcomings of current detection methods [10].

The Discrete Wavelet Transform (DWT) is very efficient in detecting glaucoma. The Dynamic Time Warping (DTW) technique is constrained to a dyadic scale and lacks adaptability [11]. The drawbacks of dyadic DTW-based approaches include fixed time-frequency coverage, variance shift, signal independence, and limited frequency resolution. In addition, the Empirical Wavelet Transform (EWT) approaches have been used to identify glaucoma in its early stages. However, these techniques have some disadvantages, such as a lack of mathematical theory, which results in increased processing complexity, as well as issues with mode mixing and boundary distortion [12].

The Cup-to-Disc Ratio (CDR) is a measurement used to determine the vertical diameter of the OC and disc in the eye. It is often used in screening tests to identify glaucoma [9]. During the period of significant and gradual changes in color intensity between the OC and optic rim in the eye, there is a large concentration of blood vessels in the field of vision, making the process of splitting the OC and disc very challenging. Furthermore, the process of manually segmenting the OC and disc in the eye for identifying highly subjective areas is time-consuming [13]. Therefore, these challenges are resolved by the use of automated techniques for segmenting OC and OD. Glaucoma detection in the eye involves the use of active contour-based algorithms and thresholding algorithms to segment the OC and OD [14]. The thresholding-based approaches provide advantages in glaucoma detection by producing binary pictures from single-channel and multi-channel processed images. These techniques rely on the color intensity difference between the OC and OD [15]. Thresholding methods are often used for segmenting fundus images. However, segmenting low-contrast images has several obstacles to achieving high accuracy [16]. Contour-based approaches are used to detect the boundaries of the OD and OC using a given set of points. It minimizes an energy function. However, these segmentation strategies based on active contours are prone to being trapped in a local minimum, and the effectiveness of the detection process relies heavily on the initial configuration of the contour method.

Recently, DL methods have been used to segment the OC and OD with reduced computational time [17]. In this research, we employ a highly efficient DL model, the EDAGD, across three distinct public datasets to assess the model's generalizability. These datasets include RIM-ONE, ACRIMA, and REFUGE. computational framework involves utilizing segmentation of the OC and OD with the SegNet model, and classification with ResNet-50 to analyze the images.

Furthermore, the trustworthiness of the proposed EDAGD is enhanced by incorporating Grad-CAM and Grad-CAM++.which provide explanatory visualization for the model's outcomes. The generated heatmaps offer transparency by emphasizing the specific areas in the fundus image that impact the model's decision-making process. This transparency enables specialists within the field, such as eye surgeons, optometrists, and ophthalmologists to understand the rationale behind the predicted results.

Experimental research designs have been widely used in the field of medical imaging and AI. For instance, studies by Li et al. (2024) in China [18] and Chun et al. (2023) in South Korea [19] have demonstrated the effectiveness of DL models in detecting diabetic retinopathy using retinal images. Similarly, in the automotive industry, experimental designs have been used to test DL algorithms for autonomous driving (Fang et al., 2024) [20]. These studies from diverse regions and industries highlight the robustness and versatility of experimental designs in validating AI models.

Typically, DL classification models pose challenges in terms of explanation and comprehension, particularly for individuals outside the domain. New methods within explainable artificial intelligence (XAI) are striving to tackle this problem by offering perspectives into the decision-making mechanisms of DL models. XAI enhances the clarity, comprehension, and justification of DL model results using visual or textual methods. For instance, activation maps visually interpret outcomes by highlighting the area of focus in input images which influences classification outcomes. Similarly, Grad-CAM generates coarse localization maps using output gradients to offer insight into model decisions.

Grad CAM++ represents a refined version of Grad-CAM, addressing certain limitations by adjusting weights for improved object detection and

explanation of several instances of objects within a picture. These parameters, are pivotal in producing visual explanations, the result is generated by combining positive partial derivatives from the feature maps using weights of the last convolutional layers about particular class scores. These heatmaps emphasize important image areas crucial for prediction. For instance, if the model places greater emphasis on the center of the OD and correlates it with a larger OC with higher glaucoma likelihood, it aids ophthalmologists in understanding model outputs, thereby enhancing system reliability and trustworthiness. As a result, automated glaucoma detection streamlines the process, ensuring consistent and reliable decisions while also saving time and effort.

2. LITERATURE

Table 1 provides a comprehensive overview of the methodologies, datasets, metrics (including Accuracy and AUC), and limitations discussed in the literature review on Glaucoma detection.

A technique employing transfer learning utilizing VGG-16 and the AlexNet was employed for the classification of glaucoma, utilizing ONH pictures collected from multiple public datasets. They built a binary set of images: In the first one, the dataset was expanded using many data augmentation methods, such as rotation, flipping, cropping, and random scaling. while the other one involved constructing 3D topographical maps of the ONH were created using shading information extracted from 2D images (SHS method). Assessment of both datasets for the categorization of glaucoma revealed enhanced performance compared to standard Convolution Neural Network (CNN) classification methods [21].

A glaucoma diagnosis method was developed depending on two procedures, utilizing the RIM-ONE, ORIGA, and DRISHTI-GS datasets. The model incorporated three types of CNN architectures: ResNet-152, AlexNet, and ResNet-50, as the ensemble classifiers. A graph saliency region technique was employed to crop the OD. The model's performance was evaluated across three different methods: without a saliency map, using a feature map and a CNN model, and employing a feature map in combination with an ensembling approach. The ensemble approach yielded the most promising outcomes, an AUC of 94% and 88% accuracy [22].

Glaucoma Net was introduced to detect primary open-angle glaucoma (POAG) using images acquired from a variety of groups and places. Two CNNs make up the model, which aims to replicate the human-grade process. The first CNN concentrates on learning discriminative traits, whereas the second CNN combines them for grading purposes. By emulating the human-grade procedure and applying a collection of network designs, the system's diagnostic accuracy was greatly enhanced [23].

A technique using 3D CNN was suggested for the purpose of detecting glaucoma by analyzing fundus pictures obtained using the datasets RIM-ONE and the DRISHTI-GS. The algorithm transformed 2D to detect fundus pictures into 3D structures for both the RGB and gray channels. The training was conducted across all four channels. The greatest results were obtained from the gray channel, which had 83.2% AUC, 85.54% sensitivity, 66.45 Kappa, 80.95% specificity, and 83.23% accuracy [24].

A CNN model was created to identify glaucoma. It attained an F1 score of 96.2% when tested on 295 videos and 1811 fundus pictures [25].

A recent study evaluated the diagnostic precision, practicality, and comprehensibility of a vision transformer DL approach in identifying primary open-angle glaucoma. and recognizing significant regions in retinal images [26].

A DL-based approach was devised to identify early-stage glaucoma by using the grey channels of the fundus images and employing the data augmentation methods. The ResNet-50 architectural model has shown exceptional performance on many datasets such as DRISHTI-GS, G1020, ORIGA, and RIM-ONE. Using the G1020 dataset, it achieves 98.48% detection accuracy, 99.30% sensitivity, 96.52% specificity, 97% AUC, and 98% F1-score [27].

The process of glaucoma screening was examined by using classification and segmentation techniques on several datasets of fundus pictures. The research used Xception, Inception ResNet V2, and ResNet152 V2 architectures for the purpose of comparison. Activation maps were produced via the Grad-CAM methodology to improve comprehensibility. The segmentation task used a U-Net architecture that integrated Inception ResNet V2 and Inception V3 models. The Xception model attained a remarkable accuracy of 97% when used in the REFUGE dataset for categorization [28].

A new CNN architecture was proposed for classifying fundus images. For instance, KR-NET was introduced as an approach for the DRiFT2 database, achieving 96.4% accuracy, 95.0% sensitivity, and 98.5% specificity [29].

Table 1: Overview Of Literature Review

Methodology	Dataset	Metrics (Accuracy (%), AUC (%)		Limitations
VGG-16, AlexNet [21]	DRISONS-DB, HRF, RIMONE, Drishti-GS1	Accuracy 94.3 AUC 99.1		Insufficient exploration of data preprocessing biases may hinder Glaucoma-Net's clinical applicability and diagnostic accuracy.
Ensemble ResNet Framework [22]	RIM-ONE, DRISHTI-GS, and ORIGA	Accuracy 91.1, AUC 83.3		Insufficient exploration of multi-frequency scale.
CNN [23]	OHTS, LAG	Accuracy OHTS- 0.930 LAG- 0.969		The model's sensitivity is compromised due to dataset imbalance, notably the scarcity of POAG images.
3DCNN [24]	RIM-ONE, DRISHTI-GS	Accuracy 83.23, AUC 83.2,		The study focuses on one database, limiting applicability across diseases and imaging modalities.
CNN [25]	DRISHTI-GS	Accuracy 98 (OD), 97 (OC)		The study lacks comparison with existing methods, hindering the assessment of model effectiveness and performance.
CNN [26]	Customized from 5 public datasets	AUC 0.92		Imbalanced datasets, limited data, and cropping may affect model performance and generalizability.
ResNet-50 [27]	RIM-ONE ORIGA G1020 DRISHTI-GS1	Accuracy 96.15 92.59 98.48 97.03	AUC 94.20 93.00 97.00 96.00	The constrained availability of high-resolution images may affect the model's performance and specificity.
Xception [28]	Refuge	Accuracy 97		Class imbalance and noisy images affected the accuracy of the classification.
CNN [29]	DRISHTI-GS1 Refuge RIM-ONE Private ACRIMA	Accuracy 90.33 88.16 90.52 88.78 91.52	AUC 85.24 80.41 91.62 91.57 95.72	Dependency on pre-trained models limits the flexibility and generalizability of glaucoma classification.
ResNet-152 [30]	Private Dataset	Accuracy 93.5		Deep learning models reliant on small datasets risk bias and may lack generalizability.
ResNet50 [31]	NTUH	Accuracy 90.80 AUC 91.20		Deep learning models may require extensive training data for optimal performance across diverse glaucoma presentations.
CNN [32]	ORIGA	Accuracy 93.5 AUC 95.1		Inadequate investigation of false positives and negatives, and lack of comparability with existing approaches, hinders a complete evaluation.
CNN(CoG-NET) [33]	Drishti, RIM-ONE, Refuge, ACRIMA	Accuracy 95.3 AUC 99		Lack of comprehensive analysis of potential biases or confounding factors

Recent research made significant progress in the identification and localization of the glaucoma using fundus imaging. A comparison study was done utilizing the Inception-v4, VGG-16, and ResNet-152 architectures employing Grad-CAM for glaucomatous area localization. The ResNet-152-M model exhibited superior performance, achieving 96% accuracy. Additionally, a web-based application (Medinoid) was developed to offer decision-making support, diagnostic confidence scores, and suspected areas for input fundus images [30].

DL models trained on fundus image datasets have garnered attention in glaucoma diagnosis research. Specifically, research using ResNet50 architectures sought to discover distinct regions that are associated with changes in ganglion cell complexes (GCC) thickness and areas of central focus. By incorporating GCC thickness information, these models have shown improved accuracy in glaucoma determination. The focus primarily lies on

the ONH for the detection of glaucoma, with high prediction accuracy achieved even with cropped macular images. Moreover, the ability of these models to pinpoint detailed GCC impairment regions underscores their significant contribution to glaucoma diagnosis, enabling early detection and timely treatment interventions [31].

A glaucoma screening method was proposed using XAI and CNN, utilizing three datasets: HRF, Drishti-GS, and ORIGA-Light. The XAI approach of CAM was used to produce heatmaps for the interpretation of fundus images. The CNN design exhibited outstanding results using the ORIGA-Light data [32].

An advanced CNN model called CoG-Net was proposed to predict glaucoma. The study used four publicly accessible datasets: REFUGE, Drishti-GS, ACRIMA, and RIM-ONE. The use of activation maps as an explainable approach allowed for the identification of crucial areas in the fundus picture that are responsible for categorization. The CoG-Net model demonstrated superior performance, with an accuracy of 95.3% [33].

3. PROPOSED EDAGD MODEL

3.1 Dataset Description

To provide a comprehensive analysis, this study incorporated three distinct fundus image datasets: RIM-ONE [34], ACRIMA [35], and REFUGE [36]. Each dataset was divided into training, testing, and validation subsets in a 70:20:10 ratio. Various augmentation techniques were applied to maintain class balance between the glaucoma and healthy image classes within each subset. The ACRIMA dataset consists of 3244 images in total, with 2270 images for training, 649 for testing, and 325 for validation. The REFUGE dataset contains 1386 images, with 970 for training, 278 for testing, and 138 for validation. The RIM-ONE dataset includes 4520 images, with 3164 for training, 904 for testing, and 452 for validation and all these details are outlined in Table 2. This balanced distribution and comprehensive augmentation ensure robust training and evaluation across different datasets.

Table 2: Dataset Overview: Training, Testing, and Validation Images

Dataset	Training	Testing	Validation
RIM-ONE [34]	4520	1291	645
ACRIMA [35]	3244	927	463
REFUGE [36]	1386	396	198

3.2 Overall Process View

The proposed methodology, named EDAGD, utilizes segmentation and classification as its main components. The method integrates pre-processing and feature extraction approaches that are often used in the study of medical images. In order to enhance the comprehensibility of the system, XAI methods are also included. The methodology's high-level design is shown in Figure 2, while the suggested architecture is outlined in Algorithm 1.

Algorithm 1 Pseudo-code for the Entire Procedure

FUNCTION main():
Dataset: acquire *RIM-ONE*;
Preprocessed_data: preprocess (Dataset,2);
Augmented_data: augment (preprocessed_data,8);
For Each Segmentation_process **do**
A_x: initialize_model (SegNet);
A_y: train_model (A_x, augmented_data);
RETURN Segmented_images;
ForEach Classification_process **do**
B_x: initialize model (ResNet50, Segmented_images);
B_y: train_model (B_x, augmented_data);
B_z:interpret_XAI(B_y, 2);
RETURN class_prediction, Heatmap;
End FUNCTION

In this research, three distinct datasets were examined to enhance the applicability of the findings. Additionally, visualizations were created to provide clear explanations of the results.

3.3 Image Pre-Processing

The median filter and Contrast Limited Adaptive Histogram Equalization (CLAHE) were employed as the main pre-processing methods. CLAHE is renowned for enhancing contrast and picture quality [37], [38]. Meanwhile, the median filter technique was utilized to reduce noise and preserve edge features. Additionally, to tackle data imbalance, prevent overfitting, and augment the training dataset size, various augmentation techniques were employed [39], [40].

3.4 Process of Segmentation

SegNet and U-Net are both composed of an encoder network and a matching decoder network. These networks are used to segment images by categorizing each pixel. The architecture of SegNet is shown in Figure 3, where the VGG16 network is utilized as the encoder with 13 convolutional layers.

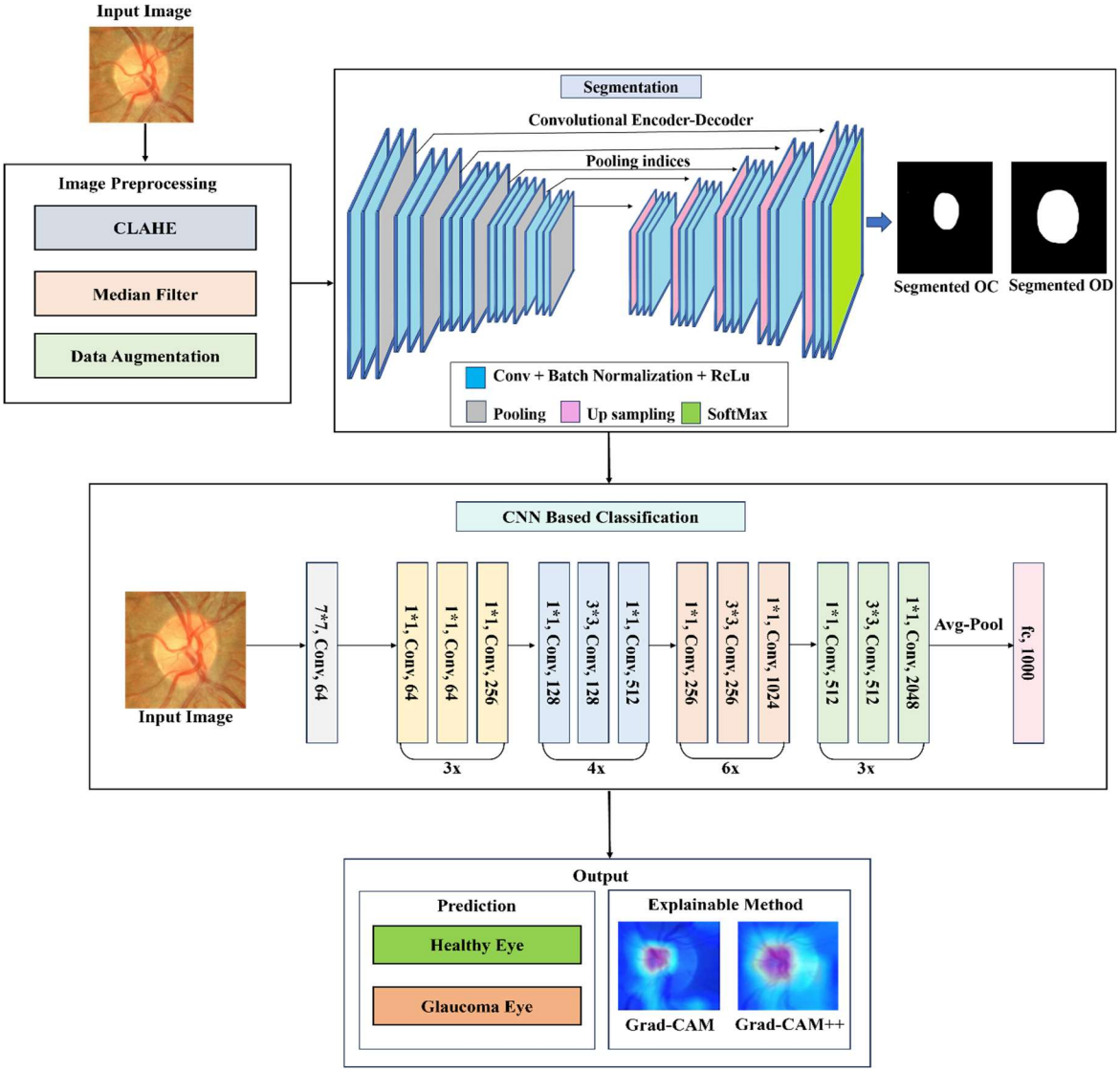


Figure 2: Methodology Outline for the Proposed EDAGD Approach

Each layer in the encoder has a matching decoder layer. To reduce training time, some images are resized and fully connected layers are removed to maintain maps with high-resolution features. Following decoding, the result is supplied to a classifier to perform pixel-wise categorization. SegNet is classified as a CNN due to the absence of fully connected layers.

Each layer in the encoder has a matching decoder layer. To reduce training time, some images are resized and fully connected layers are removed to maintain maps with high-resolution features. Following decoding, the result is supplied to a classifier to perform pixel-wise categorization. SegNet is classified as a CNN due to the absence of fully connected layers. The encoder produces transfer pool indices, which are used by the decoder to up-sample the input and create a sparse feature map. Employing convolution, the trainable filter bank is utilized to increase the feature map's density. In conclusion, the feature maps generated by the decoder are transmitted to a soft-max classifier to perform pixel-wise categorization.

SegNet's efficient encoder network allows for excellent results with time and memory efficiency. The use of max-pooling and sub-sampling in the encoder provides translation invariance and a larger spatial window for all pixels. However, this can result in lossy picture representation at the borders, which is not ideal for segmentation. To address this, boundary information is recorded prior to sub-sampling. Due to memory

Journal of Theoretical and Applied Information Technology
15th July 2024. Vol.102. No. 13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

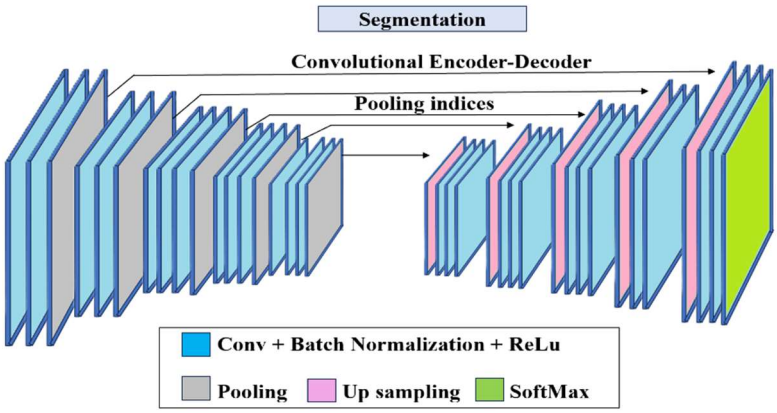


Figure 3: The SegNet architecture used for OC and OD Segmentation

constraints, only the indices of the maximal features within each pooling window are saved for each encoder feature map. SegNet also has a trainable decoder, which sets it apart from Fully decoder, Convolutional Networks (FCNs).

Once the max-pooling indices are generated from the encoders, the associated indices are utilized by the decoder network to up-sample the input feature maps. The image is shown in Figure 4. The values in the feature map are denoted by W, X, Y, and Z. This process involves enlarging the feature maps using the maximum pooling indices and then convolving them with a trainable decoder filter to produce a rich feature map. Unlike other networks that use R, G, and B channels, the SegNet decoder can generate multiple channels, enhancing its versatility.

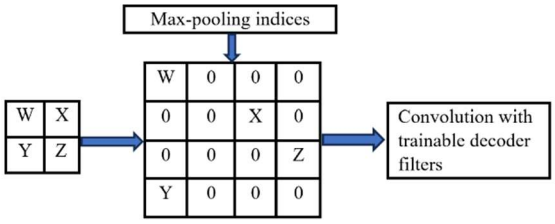


Figure 4: SegNet Decoder. W, X, Y, Z are the values in the feature map

3.5 Process of Classification

The primary objective of the classification procedure is to provide accurate results. Therefore, the pre-processed original images along with the segmented masks produced using the SegNet model were employed to provide input images for the process of classification. In the classification task, the ResNet50 model was used, which shown exceptional performance in prior research on detecting Glaucoma. The modified ResNet50 model replaces the top dense layers with three additional layers: global average pooling (GAP) to reduce parameters, a Softmax layer, and a 512-unit dense layer for a binary class classification (Glaucoma or Normal). It also includes a dropout layer with an average rate of 0.7 to address overfitting. In addition, the ResNet50 model's auxiliary classifier effectively resolved the issue of disappearing gradients. During the segmentation period, the SegNet model underwent fine-tuning to improve its efficiency. The model underwent training for 150 epochs, utilizing the Adam optimizer with a learning rate set at 0.001.

Throughout the process of training the model, several hyperparameters had a substantial impact on determining its performance. The learning rate, responsible for controlling the magnitude of optimization steps, was meticulously calibrated to strike a satisfactory equilibrium between convergence speed and stability. Similarly, the batch size, which dictates the quantity of data points handled in each iteration, and the number of epochs, which indicates the overall number of iterations across the dataset, were modified to successfully train the model while efficiently managing computational resources. The selection of the loss function, which is pivotal in guiding the optimization process, was decided deliberately based on the problem's nature, whether it involves regression or classification. The model parameters were updated during training using both Stochastic Gradient Descent (SGD) and Adam optimizers. A comparative analysis was undertaken to assess the influence of different optimizers on the model's convergence and overall performance. This analysis yielded valuable information on the intricacies of the optimization process.

The mathematical foundation behind the theoretical principles used may be described as follows. One hyperparameter that controls the size of the step made at each iteration towards minimizing a loss function throughout training is the learning rate. It affects the training process's convergence and

stability. Excessive learning rates might cause the optimization method to surpass the minimal value and prevent convergence. Conversely, if the learning rate is very low, the algorithm may experience prolonged convergence or get trapped in a poor solution. To calculate the amount of parameter updates, multiply the learning rate by the loss gradient with respect to the model parameters. The update rule for a parameter X at iteration i may be defined using the learning rate α as shown in equation (1), where $\nabla l(X^i)$ represents the gradient of the loss function with respect to the parameter (X^i).

$$X_{i+1} = X_i - \alpha \cdot \nabla l(X_i) \quad (1)$$

SGD is a basic optimization approach that iteratively updates the parameters of the model by randomly selecting a limited sample of the training data. The update strategy for the SGD optimizer is defined by the gradient of the loss function with respect to the model parameters, as specified in equation (2). In this equation, Θ_i represents the parameters of the model at iteration i, α denotes the learning rate which determines the magnitude of the optimization step, and $\nabla K_i(\Theta_i)$ represents the loss function's gradient K_i about the parameters of the model at time t.

$$\Theta_{i+1} = \Theta_i - \alpha \cdot \nabla K_i(\Theta_i) \quad (2)$$

The original output scores (logits) are transformed into a probability distribution throughout several classes using the SoftMax activation function. The definition of the variable is applicable to every class a in the output layer. It is calculated according to equation (3), where Y_a represents the raw score (logit) connected with class a, and N represents the entire number of classes.

$$Softmax(Y_a) = e^{Y_a} \Big/ \sum_{a=1}^{n} e^{Y_a} \quad (3)$$

The Global Average Pooling (GAP) layer is utilized in order to decrease the spatial dimensions. It substitutes the conventional fully connected layers in the last section of a CNN, while simultaneously decreasing the spatial dimensions (width and height) to a singular value for each feature map. This technique aids in decreasing the overall number of parameters inside the network, reduces the problem of overfitting, and improves the capacity to understand and comprehend the model. This calculates the average of each feature map over its entire spatial dimensions, as described in equation (4). The feature map f has dimensions $h \times w \times c$, where h represents the height, w represents the width, and c represents the number of channels. $f_{a,b,c}$ refers to the activation at position (a, b) in channel c of the feature map, and $GAP(f)_c$ represents the average value for channel c.

$$GAP(f)_c = (1 \,/\, h \times w) \sum_{a=1}^{h} \sum_{b=1}^{w} f_{a,b,c} \quad (4)$$

Dropout is a regularization method that mitigates the problem of overfitting. During training, a certain number of the neurons or units are randomly set to zero, which is referred to as "drop out". This technique helps the model acquire more resilient and generalized representations. This enhances the model's capacity to generalize to unfamiliar data by reducing its reliance on individual neurons. The dropout procedure is applied to each layer neuron separately during training. Equation (5) depicts this procedure, where P is the dropout rate and an is the input value (activation) to the dropout layer. The probability of a neuron being turned off during training is represented by the dropout rate.

$$Dropout(a) = \begin{cases} a * 1/(1-P), \text{with probability } 1_{-P} \\ 0, with\ probability\ P \end{cases} \quad (5)$$

3.6 Explainable strategy

A DL algorithm with medical interpretability was suggested to obtain a precise automatic diagnosis of the glaucoma while also improving transparency by identifying particular areas to aid in diagnosis [28]. Deep neural networks are often non-transparent, making their predictions unnoticeable to humans. Nevertheless, this problem may be reduced by including explainable and interpretable strategies in categorization models [29].

In order to improve the clarity and understandability of CNN-based models, both the Grad-CAM and the Grad-CAM++ were included. These methods produce rough localization maps by using gradients from the final convolutional layer to emphasize pivotal areas in the picture for the prediction concept. Figure 5 depicts the comprehensive procedure of Grad-CAM. Grad-CAM enhances the Class Activation Mapping (CAM) method by integrating gradient weights, which allows for the visualization of significant areas indicated by CNN models [29]. CAM exposes the activation patterns in the feature mappings of the final convolutional layer, illustrating the specific areas of interest that CNN prioritizes throughout its analysis of the input. By integrating Grad-CAM with

Journal of Theoretical and Applied Information Technology
15th July 2024. Vol.102. No. 13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

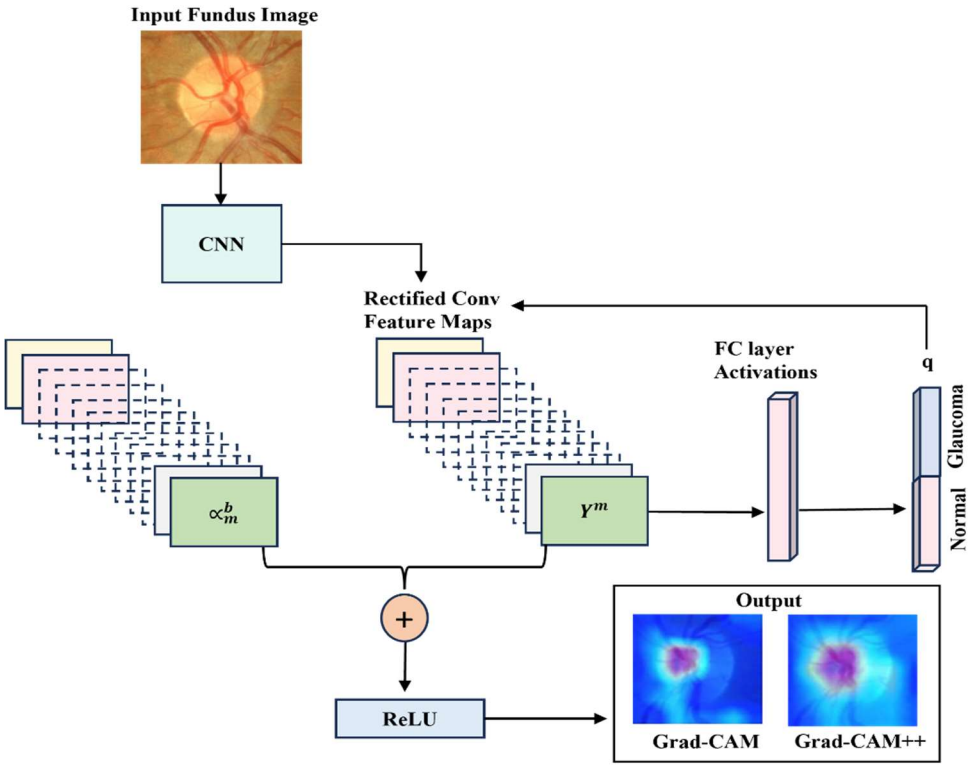


Figure 5: Summary of the Grad-CAM Procedure

smooth visualizations, it generates detailed visual representations that effectively distinguish between different classes. In addition, Grad-CAM tackles the problems associated with CAM, including the requirement for model re-training and the constraints imposed by the architecture. The Grad-CAM output may be superimposed into the input image.

To produce the ultimate localization map for glaucomatous areas ($L^c_{Grad-CAM}$), Initially, gradients are calculated for the class score 'c' (x^c) regarding the feature maps (B^n) of the last convolutional layer ($\frac{\partial x^c}{\partial B^n_{kl}}$). These feature maps have dimensions with width (w) and height (h), where n belongs to the set {1, 2, ..., n} and n represents the total number of feature maps. Subsequently, the computed gradients ($\frac{\partial x^c}{\partial B^n_{kl}}$) are subjected to a global-average pooling process across the width and height measurements (k and l).

This pooling process produces the weights (W^c_n) for the class of glaucoma, as outlined in (6) [29].

$$W^c_n = \frac{1}{w * h} \sum_k \sum_l \left(\frac{\partial x^c}{\partial B^n_{kl}} \right) \quad (6)$$

The Rectified Linear Unit (ReLU) functions by selecting just the non-negative activations once it receives an input formed by combining the weights and pertinent feature maps linearly. Subsequently, the last localization map of the glaucoma-affected regions is derived using (7):

$$(L^c_{Grad-CAM} = ReLU \left(\sum_n W^c_n \, B^n_{kl} \right) \quad (7)$$

4. RESULT ANALYSIS

4.1 Segmentation and Classification Results

A learning rate of 1 × 10−4 was utilized along with 150 epochs and a batch size of 8 using the ADAM optimizer. For the RIM-ONE and ACRIMA datasets, a dropout rate of 0.7 was incorporated alongside 512 dense layers. Meanwhile, for the REFUGE dataset, a dropout rate of 0.8 was opted for with 1024 dense layers to achieve superior performance following numerous experiments involving hyperparameter tuning.

To assess the efficacy of the fundus image segmentation process, various metrics were employed. These metrics include true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). Here's what each of these terms means:

TP: These are the examples of occurrences when the model accurately predicts. Glaucoma images as Glaucoma.

Table 3. Segmentation and classification results on different datasets

Dataset	Type	Acc (%)	Sen (%)	Spe (%)	DC (%)	JC (%)	AUC (%)
RIM-ONE	ODS	98.59	98.62	98.25	98.67	98.32	
	OCS	96.53	97.38	92.46	97.36	96.14	
	Classification	97.96	98.42	96.59			98%
ACRIMA	ODS	98.31	98.56	95.51	98.24	98.34	
	OCS	97.33	97.79	92.12	98.12	97.29	
	Classification	96.25	96.67	92.34			97%
REFUGE	ODS	93.01	95.24	82.22	95.21	92.08	
	OCS	93.34	95.31	85.12	95.32	92.26	
	Classification	90.12	92.64	83.52			92%

FP: These are examples of occurrences when the model incorrectly predicts Normal images as Glaucoma.
FN: These instances happen when Glaucoma images are wrongly predicted as Normal by the model.
TN: These are cases where the model accurately predicts Normal images as Normal.

SegNet was used for segmentation and ResNet50 for classifying using the RIM-ONE dataset. To illustrate the broad applicability of EDAGD approach, evaluated it on various datasets, including ACRIMA and REFUGE, as depicted in the Table 3. The experimental results indicate that EDAGD model that was trained using the RIM-ONE dataset worked as well as with the other datasets.

The EDAGD method, employing SegNet for segmentation and ResNet50 for classification, consistently yielded good results across all three datasets, showcasing its generalizability. Table 3 displays the evaluation criteria obtained for ODS, OCS, and classification across the three datasets.

Equations (8) through (12) represent the corresponding calculations. In these equations, sensitivity signifies the likelihood of correctly predicting glaucoma in individuals with the condition, while specificity indicates the probability of correctly predicting the absence of glaucoma in those with normal eyes. The Dice Coefficient (DC), also known as the F1-score, assesses the performance of segmentation by measuring the similarity between anticipated and actual segmentation. A number greater than 88% is regarded to indicate an excellent fit between the two. Divide the sum of the active pixel counts in both masks by the count of active pixels at the intersection of the two masks to obtain this coefficient. The Jaccard coefficient (JC), also known as Intersection-Over-Union (IoU), quantifies the degree of overlapping among the target mask and predicted output. Higher values of the coefficient indicate a higher level of similarity between the images. There is a positive correlation between DC and JC. The Area Under the Curve (AUC) is a metric that quantifies the classifier's capacity to differentiate between classes. It serves as a concise representation of the Receiver Operating Characteristic (ROC) curve. Table 3 demonstrates that higher AUC values indicate a more advanced level of diagnostic prediction ability for the model.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (9)$$

$$Specificity = \frac{TN}{TN + FP} \quad (10)$$

$$Dice\ Coefficient\ (DC) = \frac{2 \times Area\ of\ Overlap}{Total\ pixels\ combined} \quad (11)$$

$$IoU\ (or) Jaccard\ Index\ (JC) = \frac{Area\ of\ Overlap}{Total\ pixels\ combined} \quad (12)$$

Figure 6 presents both accuracy and loss curves obtained while classifying the three distinct datasets, namely: a. RIM-ONE, b. ACRIMA, and c. REFUGE datasets. Typically, in well-calibrated models, the validation loss slightly exceeds the training loss, yet both curves closely align. Initially, both training and validation losses are substantial, then rapidly decrease.

Among the datasets, the RIM-ONE dataset exhibits the lowest loss values, with corresponding training and validation losses of 0.0687 and 0.0732, respectively. Additionally, the accuracy curves for both training and validation exhibit similar trends across all datasets.

Journal of Theoretical and Applied Information Technology
15th July 2024. Vol.102. No. 13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

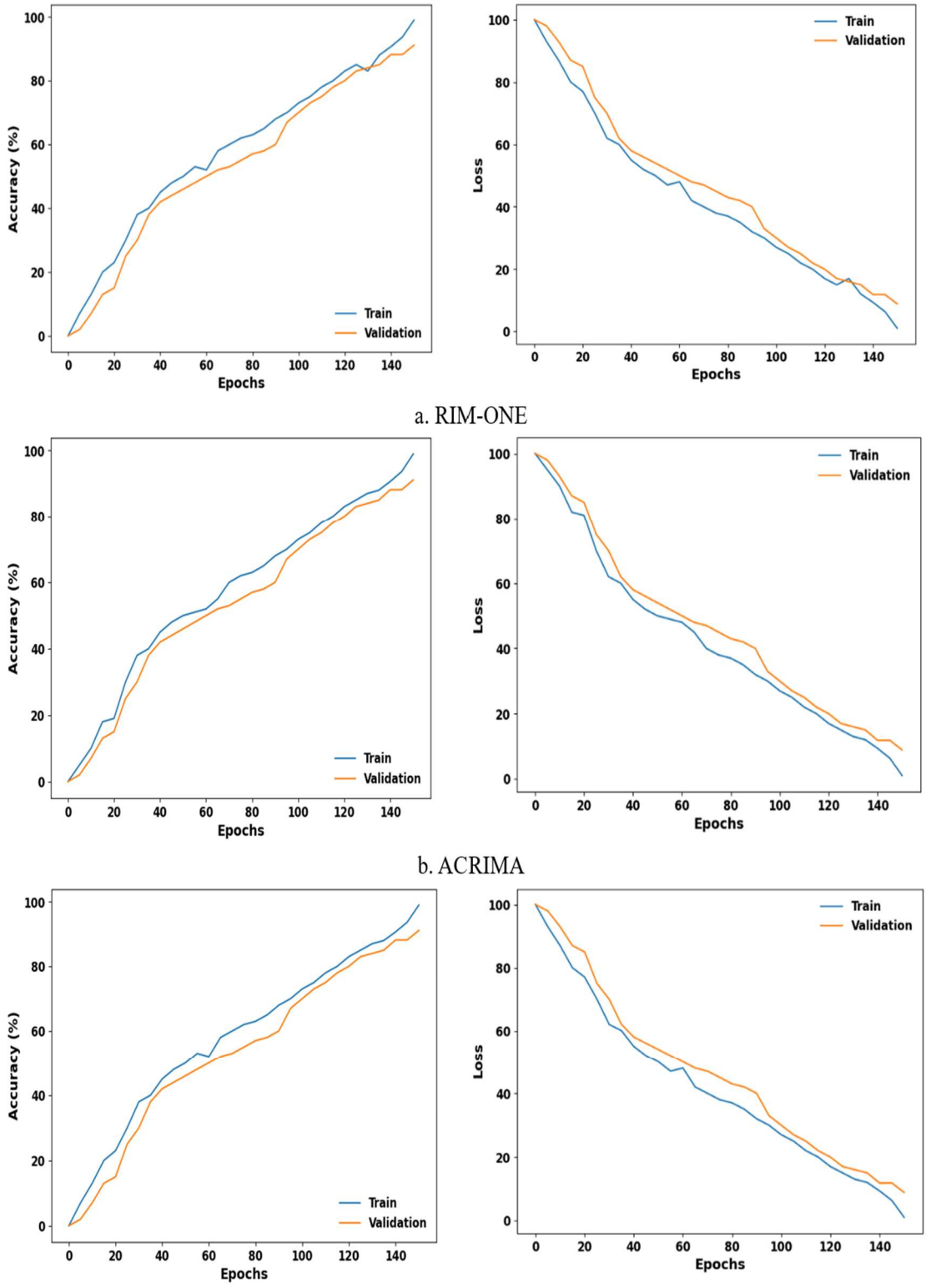


Figure 6. Accuracy and Loss curves for the a. RIM-ONE, b. ACRIMA and c. REFUGE Datasets

The RIM-ONE dataset demonstrates the highest level of accuracy, achieving 98.16% and 96.24% for training and validation accuracies, respectively.

The accuracy and loss curves displayed in Figure 6 illustrate no signs of underfitting or overfitting issues, demonstrating the effectiveness of the EDAGD method in handling the training and testing of image data and surpassing the existing approaches as outlined in Table 4 and depicted in Figure 7.

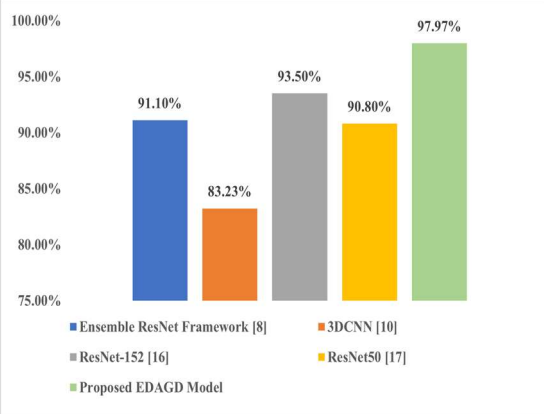


Figure 7: Performance Comparison of the proposed EDAGD model with existing approaches in terms of Accuracy.

The proposed EDAGD model outperformed the other state of art works with 97.97% Accuracy, and the comparisons are clearly outlined in Table 4.

Table 4: Comparison of the proposed EDAGD Model with existing approaches.

Model	Accuracy
Ensemble ResNet Framework [8]	91.1%
3DCNN [10]	83.23%
ResNet-152 [16]	93.5%
ResNet50 [17]	90.80%
Proposed EDAGD Model	97.97%

The confusion matrix for the predictions over the three datasets is shown in Figure 9. The matrix is divided into four regions: the top left represents true positives (TP), the top right represents false positives (FP), the bottom left represents false negatives (FN), and the bottom right represents true negatives (TN).

The algorithm accurately predicts glaucoma in images with the highest values across all datasets. Subsequently, the next greatest values are attributed to normal images that were accurately classified as normal. The lowest numbers reflect inaccurate forecasts for the corresponding classes. Therefore, the technique exhibits excellent performance on all three datasets.

Additionally, the ROC curve illustrates how well the classification model performs across various classification thresholds, providing insight into the balance between sensitivity and specificity. In Figure 8, the ROC curves for all three datasets are located in close proximity to the top-left corner, indicating better classifier performance with higher

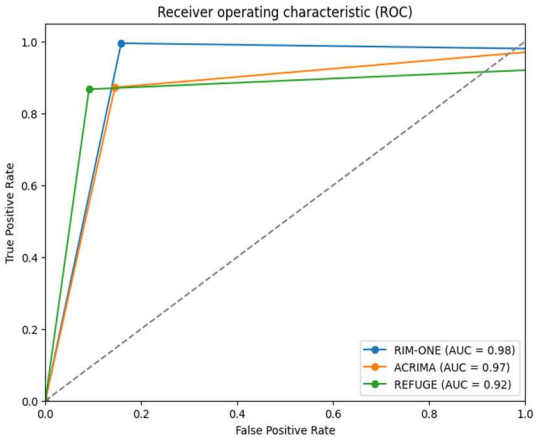


Figure 8: ROC Curve for three datasets

accuracy. Particularly, the classifiers for the RIM-ONE and ACRIMA datasets exhibit superior performance compared to the other datasets.

4.2. Explainable Strategy Results

In this research EDAGD model was implemented, which encompasses the segmentation, the classification, and the explainability aspects of glaucoma diagnosis. Within this framework, segmentation masks and the explainability of visual representations were encompassed. Applying the Grad-CAM and Grad-CAM++ methodologies, heatmaps were generated highlighting significant areas in the input images that influenced the predictions. In particular, the cupping region is located in the central region of the OD highly correlates with the cases of glaucoma. A high cup-to-disc ratio often raises suspicion and the likelihood of glaucoma in patients. Conversely, for healthy eyes without visible cupping issues or signs of glaucoma, there are no glaucomatous areas to highlight and visualize. On the other hand, if the eyes are healthy and do not show any indicators of glaucoma such as obvious cupping or glaucomatous regions, there is no need to emphasize or visualize any glaucomatous areas.

In glaucomatous eyes, the optic nerve, particularly the ONH, is prominently highlighted, indicating its significance in classification. The importance of regions is denoted by varying shades,

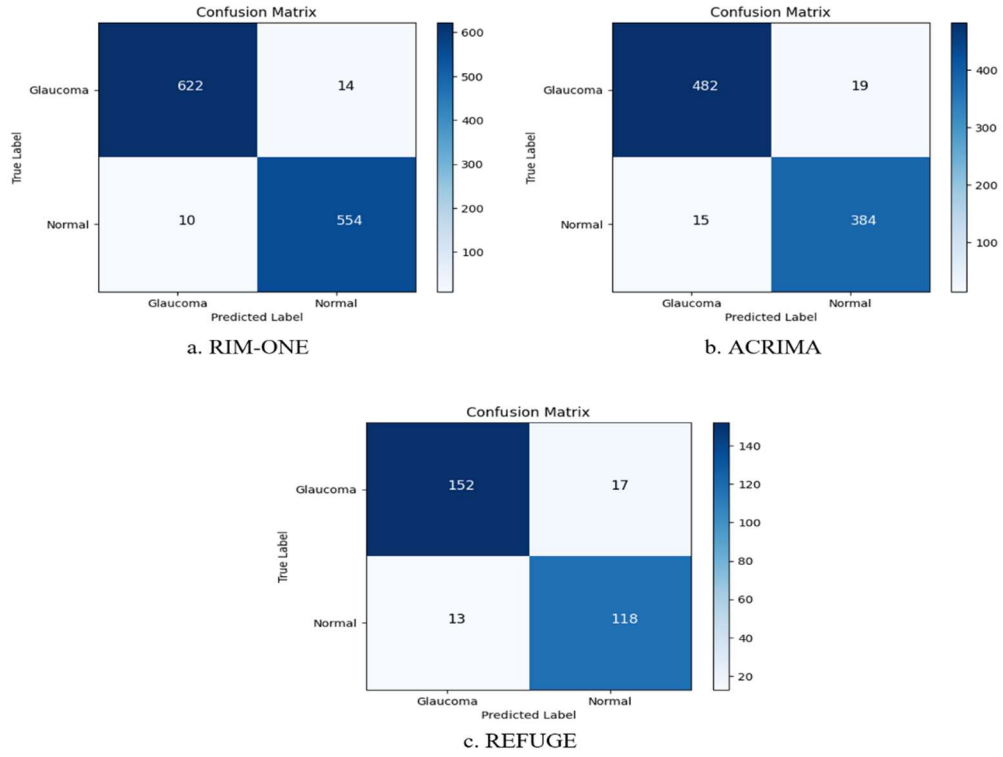


Figure 9: Confusion Matrix Comparison Across Diverse Datasets

with red being the most important and blue the least. Figure 10 depicts a comparison between the visualizations of explainability achieved via the use of Grad-CAM and Grad-CAM++.

Grad-CAM++ enhances its visual representations of CNN model predictions compared to Grad-CAM. It enhances the ability to identify the location of objects, clarifies the presence of many instances of objects in a single picture, and enables more effective information distillation via explanation, as compared to Grad-CAM.

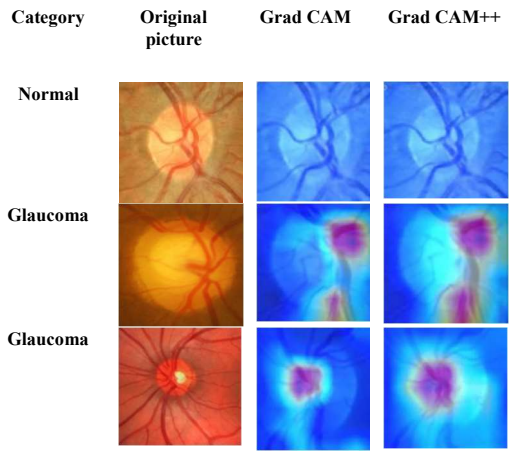


Figure 10: Collection of Input Images Paired with Heatmaps

5. DISCUSSIONS

In this work EDAGD framework was developed for identifying glaucoma that incorporates explainability, aiming to improve the reliability of the model. By combining segmentation and classification algorithms, the study demonstrates enhanced performance in identifying glaucoma using fundus pictures compared to using these techniques individually. The SegNet model, which utilizes DL algorithms, is employed for the segmentation procedure, effectively detecting glaucoma eye problems. This supplementary assessment could potentially improve clinical decision-making and automate the diagnostic process, reducing time-consuming elements and minimizing human error.

The algorithm developed in this research addresses the significance of treating different types of noise present in pictures. Through rigorous training, the algorithm provides accurate predictions by efficiently eliminating irrelevant data. The use of a median filter approach reduces noise in pictures while preserving their borders and intricate features. This technique, particularly effective in eliminating salt-and-pepper noise, replaces outlier pixel values generated by noise with more representative values from the surrounding area. As a result, the model receives a more refined and accurate input for

analysis, facilitating the extraction of relevant and consistent characteristics.

While the research focuses specifically on the identification of glaucoma, it does not extend to defining the progression of the disease. However, the EDAGD model outperforms previous research in terms of categorization and explainability approaches, achieving an accuracy of 97.97% on the RIM-ONE dataset and utilizing three separate fundus imaging datasets to showcase the applicability of the EDAGD approach across different scenarios. This surpasses previous studies with accuracies of 91.1% [22], 83.23% [24], 93.5% [30], and 90.80% [31].

6. CONCLUSION

This research presents a new and advanced EDAGD framework based on DL. The framework includes approaches for segmenting, classifying, and explaining glaucoma diseases using various datasets of fundus images. The goal is to effectively diagnose glaucoma disorders. The primary objective was to leverage explainable techniques to elucidate the rationale behind predictions made by the underlying classification model. Meticulously examining three distinct datasets, namely the RIM-ONE, the ACRIMA, and the REFUGE, ensured the robustness and generalizability of the framework.

Through a thorough and detailed comparison and examination of various CNN models, SegNet and ResNet-50 architectures were utilized, and the EDAGD framework achieved remarkable segmentation accuracies of 98.58% for the OD and 96.52% for the OC on the RIM-ONE dataset and maintained robust performance on the ACRIMA and REFUGE datasets. Additionally, EDAGD employs advanced visualization techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) and Grad-CAM++ to produce interpretable heatmaps, which help identify critical regions for diagnosis. The framework's classification of segmented images demonstrates impressive performance metrics, achieving 97.97% accuracy, 98.41% sensitivity, and 96.58% specificity.

The experimental results demonstrate that the proposed EDAGD model surpasses current approaches such as Ensemble ResNet Framework [22], 3DCNN [24], ResNet-152 [30], and ResNet50 [31] with an accuracy of 97.97%.

While the research focuses specifically on the identification of glaucoma, it does not extend to defining the progression of the disease. Future expansions of the study could involve incorporating advanced DL methods such as contrastive learning and vision transformers to improve glaucoma classification as well as the progression of the disease.

REFERENCES:

[1] Vadduri, Maneesha, and P. Kuppusamy. "Diabetic Eye Diseases Detection and Classification Using Deep Learning Techniques—A Survey." In International Conference on Information and Communication Technology for Competitive Strategies, pp. 443-454. Singapore: Springer Nature Singapore, 2022.

[2] Vadduri, Maneesha, and P. Kuppusamy. "Enhancing Ocular Healthcare: Deep Learning-Based multi-class Diabetic Eye Disease Segmentation and Classification." IEEE Access (2023).

[3] World Health Organization, ‘‘World report on vision,’’ 2019

[4] Coan, Lauren J., Bryan M. Williams, Venkatesh Krishna Adithya, Swati Upadhyaya, Ala Alkafri, Silvester Czanner, Rengaraj Venkatesh, Colin E. Willoughby, Srinivasan Kavitha, and Gabriela Czanner. "Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review." Survey of ophthalmology 68, no. 1 (2023): 17-41.

[5] Juneja, Mamta, Shaswat Singh, Naman Agarwal, Shivank Bali, Shubham Gupta, Niharika Thakur, and Prashant Jindal. "Automated detection of Glaucoma using deep learning convolution network (G-net)." Multimedia Tools and Applications 79 (2020): 15531-15553.

[6] Singh, Law Kumar, Pooja, Hitendra Garg, Munish Khanna, and Robin Singh Bhadoria. "An enhanced deep image model for glaucoma diagnosis using feature-based detection in retinal fundus." Medical & Biological Engineering & Computing 59 (2021): 333-353.

[7] Yousefi, Siamak, Michael H. Goldbaum, Madhusudhanan Balasubramanian, Felipe A. Medeiros, Linda M. Zangwill, Jeffrey M. Liebmann, Christopher A. Girkin, Robert N. Weinreb, and Christopher Bowd. "Learning from data: recognizing glaucomatous defect patterns and detecting progression from visual field measurements." IEEE Transactions on Biomedical Engineering 61, no. 7 (2014): 2112-2124.

[8] Kim, Paul Y., Khan M. Iftekharuddin, Pinakin G. Davey, Márta Tóth, Anita Garas, Gabor Holló, and Edward A. Essock. "Novel fractal feature-based multiclass glaucoma detection and progression prediction." ieee journal of biomedical and health informatics 17, no. 2 (2013): 269-276.

[9] Cheng, Jun, Fengshou Yin, Damon Wing Kee Wong, Dacheng Tao, and Jiang Liu. "Sparse dissimilarity-constrained coding for glaucoma screening." IEEE Transactions on Biomedical Engineering 62, no. 5 (2015): 1395-1403.

[10] Vermeer, Koen A., Frans M. Vos, Barrick Lo, Qienyuan Zhou, Hans G. Lemij, Albert M. Vossepoel, and Lucas J. Van Vliet. "Modeling of scanning laser polarimetry images of the human retina for progression detection of glaucoma." IEEE transactions on medical imaging 25, no. 5 (2006): 517-528.

[11] Niwas, Swamidoss Issac, Weisi Lin, Chee Keong Kwoh, C-C. Jay Kuo, Chelvin C. Sng, Maria Cecilia Aquino, and Paul TK Chew. "Cross-examination for angle-closure glaucoma feature detection." IEEE journal of biomedical and health informatics 20, no. 1 (2015): 343-354.

[12] George, Yasmeen, Bhavna J. Antony, Hiroshi Ishikawa, Gadi Wollstein, Joel S. Schuman, and Rahil Garnavi. "Attention-guided 3D-CNN framework for glaucoma detection and structural-functional association using volumetric images." IEEE journal of biomedical and health informatics 24, no. 12 (2020): 3421-3430.

[13] Mary, Viola Stella, Elijah Blessing Rajsingh, and Ganesh R. Naik. "Retinal fundus image analysis for diagnosis of glaucoma: a comprehensive survey." IEEE Access (2016).

[14] Acharya, U. Rajendra, Sumeet Dua, Xian Du, and Chua Kuang Chua. "Automated diagnosis of glaucoma using texture and higher order spectra features." IEEE Transactions on information technology in biomedicine 15, no. 3 (2011): 449-455.

[15] Cheng, Jun, Jiang Liu, Yanwu Xu, Fengshou Yin, Damon Wing Kee Wong, Ngan-Meng Tan, Dacheng Tao, Ching-Yu Cheng, Tin Aung, and Tien Yin Wong. "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening." IEEE transactions on medical imaging 32, no. 6 (2013): 1019-1032.

[16] Parashar, Deepak, and Dheeraj Kumar Agrawal. "Automated classification of glaucoma stages using flexible analytic wavelet transform from retinal fundus images." IEEE Sensors Journal 20, no. 21 (2020): 12885-12894.

[17] Maheswari, Chevula, Gurukumar Lokku, and K. Nagi Reddy. "Detection of Glaucoma Using HMM Segmentation and Random Forest Classification." In Inventive Systems and Control: Proceedings of ICISC 2022, pp. 585-597. Singapore: Springer Nature Singapore, 2022.

[18] Li, Zihan, Yuan Zheng, Dandan Shan, Shuzhou Yang, Qingde Li, Beizhan Wang, Yuanting Zhang, Qingqi Hong, and Dinggang Shen. "ScribFormer: Transformer Makes CNN Work Better for Scribble-based Medical Image Segmentation." IEEE Transactions on Medical Imaging (2024).

[19] Chun, Ji-Won, and Hun-Sung Kim. "The present and future of artificial intelligence-based medical image in diabetes mellitus: focus on analytical methods and limitations of clinical use." Journal of Korean Medical Science 38, no. 31 (2023).

[20] Fang, Jianwu, Fan Wang, Jianru Xue, and Tat-Seng Chua. "Behavioral intention prediction in driving scenes: A survey." IEEE Transactions on Intelligent Transportation Systems (2024).

[21] Wang, Peipei, Mingyuan Yuan, Yan He, and Jiuai Sun. "3D augmented fundus images for identifying glaucoma via transferred convolutional neural networks." International Ophthalmology 41 (2021): 2065-2072.

[22] Chaudhary, Pradeep Kumar, and Ram Bilas Pachori. "Automatic diagnosis of glaucoma using two-dimensional Fourier-Bessel series expansion based empirical wavelet transform." Biomedical Signal Processing and Control 64 (2021): 102237.

[23] Lin, Mingquan, Bojian Hou, Lei Liu, Mae Gordon, Michael Kass, Fei Wang, Sarah H. Van Tassel, and Yifan Peng. "Automated diagnosing primary open-angle glaucoma from fundus image by simulating human's grading with deep learning." Scientific reports 12, no. 1 (2022): 14080.

[24] de Sales Carvalho, Nonato Rodrigues, Maria da Conceicao Leal Carvalho Rodrigues, Antonio Oseas de Carvalho Filho, and Mano Joseph Mathew. "Automatic method for glaucoma diagnosis using a three-dimensional convoluted neural network." Neurocomputing 438 (2021): 72-83.

[25] Veena, H. N., A. Muruganandham, and T. Senthil Kumaran. "A novel optic disc and optic

cup segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images." Journal of King Saud University-Computer and Information Sciences 34, no. 8 (2022): 6187-6198.

[26] Fan, Rui, Kamran Alipour, Christopher Bowd, Mark Christopher, Nicole Brye, James A. Proudfoot, Michael H. Goldbaum et al. "Detecting glaucoma from fundus photographs using deep learning without convolutions: Transformer for improved generalization." Ophthalmology science 3, no. 1 (2023): 100233.

[27] Shoukat, Ayesha, Shahzad Akbar, Syed Ale Hassan, Sajid Iqbal, Abid Mehmood, and Qazi Mudassar Ilyas. "Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach." Diagnostics 13, no. 10 (2023): 1738.

[28] Neto, Alexandre, José Camara, and António Cunha. "Evaluations of deep learning approaches for glaucoma screening using retinal images from mobile device." Sensors 22, no. 4 (2022): 1449.

[29] Sonti, Kamesh, and Ravindra Dhuli. "A new convolution neural network model “KR-NET” for retinal fundus glaucoma classification." Optik 283 (2023): 170861.

[30] Kim, Mijung, Jong Chul Han, Seung Hyup Hyun, Olivier Janssens, Sofie Van Hoecke, Changwon Kee, and Wesley De Neve. "Medinoid: computer-aided diagnosis and localization of glaucoma using deep learning." Applied Sciences 9, no. 15 (2019): 3064.

[31] Guo, Jing-Ming, Yu-Ting Hsiao, Wei-Wen Hsu, Sankarasrinivasan Seshathiri, Jiann-Der Lee, Yan-Min Luo, and Peizhong Liu. "A Study of the Interpretability of Fundus Analysis with Deep Learning-Based Approaches for Glaucoma Assessment." Electronics 12, no. 9 (2023): 2013.

[32] Deperlioglu, Omer, Utku Kose, Deepak Gupta, Ashish Khanna, Fabio Giampaolo, and Giancarlo Fortino. "Explainable framework for Glaucoma diagnosis by image processing and convolutional neural network synergy: analysis with doctor evaluation." Future Generation Computer Systems 129 (2022): 152-169.

[33] Juneja, Mamta, Sarthak Thakur, Archit Uniyal, Anuj Wani, Niharika Thakur, and Prashant Jindal. "Deep learning-based classification network for glaucoma in retinal images." Computers and Electrical Engineering 101 (2022): 108009.

[34] Fumero, Francisco, Silvia Alayón, José L. Sanchez, Jose Sigut, and M. Gonzalez-Hernandez. "RIM-ONE: An open retinal image database for optic nerve evaluation." In 2011 24th international symposium on computer-based medical systems (CBMS), pp. 1-6. IEEE, 2011.

[35] Diaz-Pinto, Andres, Sandra Morales, Valery Naranjo, Thomas Köhler, Jose M. Mossi, and Amparo Navea. "CNNs for automatic glaucoma assessment using fundus images: an extensive validation." Biomedical engineering online 18 (2019): 1-19.

[36] Zhang, Zhuo, Feng Shou Yin, Jiang Liu, Wing Kee Wong, Ngan Meng Tan, Beng Hai Lee, Jun Cheng, and Tien Yin Wong. "Origa-light: An online retinal fundus image database for glaucoma analysis and research." In 2010 Annual international conference of the IEEE engineering in medicine and biology, pp. 3065-3068. IEEE, 2010.

[37] Vijayalakshmi, D., and Malaya Kumar Nath. "A novel multilevel framework based contrast enhancement for uniform and non-uniform background images using a suitable histogram equalization." Digital Signal Processing 127 (2022): 103532.

[38] Vijayalakshmi, D., and Malaya Kumar Nath. "A systematic approach for enhancement of homogeneous background images using structural information." Graphical Models 130 (2023): 101206.

[39] Thakoor, Kaveri A., Xinhui Li, Emmanouil Tsamis, Paul Sajda, and Donald C. Hood. "Enhancing the accuracy of glaucoma detection from OCT probability maps using convolutional neural networks." In 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp. 2036-2040. IEEE, 2019.

[40] Meedeniya, Dulani. Deep Learning: A Beginners' Guide. CRC Press, 2023.

[41] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, pp. 234-241. Springer International Publishing, 2015.

[42] Ko, Yu-Chieh, Shih-Yu Wey, Wei-Ta Chen, Yu-Fan Chang, Mei-Ju Chen, Shih-Hwa Chiou, Catherine Jui-Ling Liu, and Chen-Yi Lee. "Deep learning assisted detection of

glaucomatous optic neuropathy and potential designs for a generalizable model." PLoS One 15, no. 5 (2020): e0233079.

[43] Selvaraju, Ramprasaath R., Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. "Grad-cam: Visual explanations from deep networks via gradient-based localization." In Proceedings of the IEEE international conference on computer vision, pp. 618-626. 2017.