Multi-level Attention Augmented Deep Learning Model for Optical Cup & Disc Segmentation in Fundus Images

1. Mrs.S.Anusuya,2.Dr.N.M.Masoodhu Banu

Assistant Professor,Department of Electronics and Communication Engineering,Mohamed Sathak A J College of Engineering,

Professor,Department of Bio-Medical Engineering,

Vel Tech Rangarajan Dr.sagunthala R&D Institute of Science and Technology,

panusuya85@gmail.com,banumobeen@yahoo.com

Abstract: Glaucoma, a neuropathic eye disease, is characterized by increased Intraocular Pressure within the retina and ranks as the primary contributor to worldwide blindness after cataracts. Timely diagnosis stands crucial to avoidthe complete blindness. To address challenges related to under-segmentation and over-segmentation in OC and OD segmentation, we present a multi-level attention-augmented U- Net model. This model incorporates different attentions, such as semantic and structural attentions, to refine convolutional features, enabling the model to learn discriminative global-centric features. By applying a conservative smoothing algorithm, we achieve a high ROI extraction accuracy of 98.67%. Moreover, the proposed Multi-level Attention Augmented U-Net architecture achieves a Dice Co - efficient of 0.930 and 0.870 for the partitioning of the OD and OC segmentation correspondingly.

Keywords: Optical disc, Optical cup, Segmentation, Unet model, Glaucoma, ROI

1. Introduction

Currently, there is no known treatment for glaucoma, the primary cause of visual impairment globally. Without early detection, it can progress to irreversible blindness. The optic nerve fibers, responsible for transmitting visual sensory information, are adversely impacted by glaucoma, posing a substantial risk to vision.Unfortunately, the condition is often asymptomatic, leading to delayed awareness until it is too late. Many researchers have primarily worked with datasets lacking real-time fluctuations in image quality, limited access to high-quality fundus photos, and small sample sizes. Consequently, the robustness of existing systems is compromised.

Therefore, there is a need to develop a model that performs well across diverse environmental conditions. Additionally, there is an opportunity to enhance categorization accuracy.The objective of this research is to evaluate the diagnostic efficacy of retinal fundus images in detecting glaucoma.In order to accomplish this objective, a computerized recognition system based on a desktop application has been developed specifically for medical professionals.The objective of this article is to conduct an impartial evaluation of the recent advancements in deep learning applications for glaucoma. As well as highlighting the advantages of these applications, the article addresses the challenges inherent in developing models for glaucoma screening, diagnosis, and progression detection.Initially, a brief overview of deep learning is presented, highlighting the differences from traditional machine learning classifiers. The article discusses the issues related to training and validating deep learning models and how they specifically apply to glaucoma. Furthermore, specific instances of deep learning applications in glaucoma, such as fundus photography for screening and the use of Optical Coherence Tomography (OCT) and Automated Perimetry for diagnosing and monitoring glaucoma progression (SAP), are discussed.

The research makes the following contributions:

- Proposal of the multilevel attention augmented U-Net model for the accurate segmentation of the Optical Disc(OD) and Optical Cup (OC) in fundus images.
- Implementation of the self-attention mechanisms to refine convolutional features and learn discriminative features focused on content.
- Assessment of the model's generalizability by testing it on different publicly available datasets using qualitative and quantitative parameters.

2. The Related Work

2.1.Region of Interest extraction

The authors Agarwal et al. [6] suggested a technique for determining the Region of Interest (ROI) by determining the precise geographic location of the Optical Cup's center and identifying the square area with the radius of "r". They proposed that the ROI be extracted by computing the average of the pixels within the square area, and the pixels outside the square area be set to zero. Finally, the ROI is extracted by subtracting the average of the pixels within the square area from the original image. This proposed method was tested on a database of 60 retinal images and was

found to be effective and accurate. The results showed that the ROI was accurately extracted and the accuracy was improved by more than 10%.They estimated the optical disc's coordinates and refined them to obtain accurate coordinates [2]. In contrast, [7] utilized four different image processing techniques, including geometric modifications, limited region processing, and adjustments to pixel brightness, to identify the Region of Interest. This allowed them to accurately identify the disc's boundaries and center point. They then used this information to accurately calculate the disc's radius and area. Finally, they used this information to compute the disc's optical properties.As soon as they found the Region of Interest, they resumed normal operations. The ROI was physically cropped using a non-automated method [8]. A hand-crafted approach was employed to ensure the accuracy and precision of the results. This method allowed for the elimination of potential errors that could have been introduced by automatic cropping algorithms. The final ROI was then ready for further analysis. By using a hand-crafted approach, they could be sure that the results were accurate and precise. This method also eliminated the potential for errors that could have been introduced by automatic cropping algorithms. The manual cropping process allowed them to quickly and accurately obtain the final ROI that was needed for further analysis.Narasimhan and Vijayarekha [9] employed the G-Plane of an RGB fundus image to approximate the region surrounding the brightest point. In contrast, [10] utilized the intensity weighted centroid approach to determine the Center of the Optical Disc instead of locating the brightest area.

2.2 Using Optical Disc &Optical Cup

Das et al. [8] employed the watershed transform technique to segment the Optical Disc (OD) and Optical Cup (OC) based on the Red & Green channels of an RGB fundus image. The Lab color space's "a" plane and the red channel of the RGB color space were used for segmentation. The segmentation results were evaluated using Dice's coefficient and Jaccard Index. The results showed that the proposed method achieved a Dice's coefficient of 0.95 and Jaccard Index of 0.91 for OD segmentation and a Dice's coefficient of 0.97 and Jaccard Index of 0.94 for OC segmentation. Similarly, [9] used K-means clustering to segment the optical disc and Optical Cup. The results showed that their method achieved an accuracy of 95.5% for OD segmentation, and 97.3% for OC segmentation. These results demonstrate that the proposed method is a viable solution for optical disc segmentation. The OD and OC segmentation used the V-plane from the

HSV color space rather than the RGB color space's Green channel. This suggests that the V-plane of the HSV color space is more suitable for segmenting the optical disc compared to the RGB color space. Furthermore, the results indicate that the proposed K-means clustering algorithm is an effective tool for optical disc and cup segmentation. Kim et al. [12] combined U-Net with Fully Convolutional Networks (FCN) for segmentation. They also used a modified U-Net architecture for optic disc and cup segmentation. This improved segmentation accuracy significantly. Furthermore, the authors showed that such a system can diagnose glaucoma. They employed both Binary and Multiclass - FCNs for the segmentation task.

2.3 Using Feature Extraction

The ISNT criterion was employed as a diagnostic measure for glaucoma [13]. This criterion has been found to be a reliable and reproducible measure for detecting early stages of glaucoma. It is also useful for monitoring the progression of the disease over time. ISNT is based on the evaluation of the nerve fiber layer thickness in the eye. To approximate the Cup - Disc Ratio (CDR), morphological operations were conducted [14]. The results of the morphological operation were compared to the CDR of the normal eye. This was defined as an ISNT value of 1.0. The ISNT value obtained from the morphological operation was then used to determine glaucoma severity. In order to enhance accuracy, a method for creating a mask using the Optical Disc centroid was proposed to obtain the Neuro retinal Rim width [15]. The proposed method was found to be more precise than the conventional method. Furthermore, it was able to detect glaucoma at a much earlier stage. This enabled more effective treatment of the disease.Blood vessel extraction was accomplished by applying local entropy thresholding based on CDR and ISNT values [16]. The proposed method was also able to reduce artifacts in the extracted images, resulting in higher accuracy. Additionally, it was found to be faster than the conventional method and required less computational power.Monitoring glaucoma progression involves analyzing the characteristics of the retinal nerve fiber layer defect (RNFLD) [17]. The proposed method was able to detect and measure RNFLD, which is an important biomarker for glaucoma. This method has shown to have high accuracy and greater efficiency when compared to the conventional methods. The RNFLD can be extracted from OCT images with the proposed method and used for glaucoma diagnosis. The results showed that the proposed method outperformed state-of-the-art methods in terms of accuracy and speed.Deep Convolutional Neural Networks (CNN) were

utilized to distinguish between glaucomatous images and healthy images. The CNN was trained on a dataset of OCT images which included both glaucomatous and healthy images. The proposed method was validated on a separate test dataset and achieved an accuracy of 96%. This demonstrates the potential of the proposed method for automated glaucoma diagnosis.In their study, Al Ghamdi et al. [21] proposed an innovative self-learning technique for labeling unlabeled data. This technique was based on an adaptive thresholding algorithm and a convolutional neural network (CNN) for automated glaucoma diagnosis. The results showed that the proposed method was able to accurately label the OCT images with an accuracy of 96%. This demonstrates the potential of the proposed method for automated glaucoma diagnosis.They employed a CNN architecture for training and utilized a labeled dataset to supervise the training process of the CNN model. The model was able to distinguish between healthy and glaucoma images with high accuracy. Furthermore, it was also able to detect the severity of glaucoma in the OCT images. This shows the potential of the proposed method for automated glaucoma diagnosis.

2.4 Deep Learning Based Methods

Li et al. (2018) devised a methodology to assess the efficacy of a deep learning system in screening Glaucoma Optic Neuropathy (GON) [18] . By using their approach, the researchers were able to accurately detect GON with a high degree of accuracy. They also showed that their deep learning system was better at detecting GON than traditional methods. The training process utilized a Mini batch Gradient Descent with a size of 32, and the network design incorporated Inception-V3 [19]. The neural network architecture was then optimized using Bayesian and baseline algorithms. The results from the experiments demonstrate that their system achieved a high accuracy rate of 92.9%. This is significantly higher than the accuracy rate of the traditional methods used. The convergence stage employed the Adam Optimizer [20], achieving optimal results with a learning rate of 0.0020. This highlights the potential of the proposed neural network for solving similar problems. Further optimization can be done to improve the accuracy rate even further. The results of this study demonstrate the effectiveness of the proposed system. The effectiveness of this technique was assessed using a private database containing 70,000 photos, although only 48,116 images with distinguishable optic discs were used for analysis. Two

detection thresholds were established for referable GON based on ground truth annotations provided by medical experts: non-referable GON and referable GON.

The overall architecture consisted of two streams at the global level. The first stream utilized residual networks and was designated as a standard categorization network (Res Net). The second stream incorporated a generative adversarial network (GAN) to enhance the accuracy of the recognition. Both streams took advantage of a hierarchical classification approach to recognize objects of interest.The second stream incorporated a modified U-shaped convolutional network as a segmentation-guided network [24]. The two streams were combined to form a unified model. The model was able to accurately recognize objects of interest with high precision and recall. The hierarchical classification approach allowed for more efficient and accurate object detection.For local image-level processing, a standard classification network utilizing Res Net and the disc polar transformation stream were utilized to convert the optical disc region into a polar coordinate system. The architecture consisted of four deep streams: the global image stream, the segmentation-guided network, the disc region stream, and the disc polar stream. The final screening outcome was obtained by combining all deep streams [25]. The architecture was evaluated and tested on the IDRiD dataset and achieved an accuracy of 90.3%. The model was also tested on the RIM-ONE dataset and achieved an accuracy of 93.5%. Finally, the model was tested on the DRIVE dataset and achieved an accuracy of 96.5%.

2.5 Dataset Analysis

Based on the different datasets used, the techniques for detecting Glaucoma have been classified and described. These techniques are based on the analysis of ophthalmological images, such as optical coherence tomography (OCT) and fundus photography. The algorithms used to detect Glaucoma are based on machine learning and deep learning techniques. The accuracy of the results obtained is very high, ensuring early and accurate detection of Glaucoma.

Table 1: Dataset Analysis with various parameters

Dataset	DRISHTI-GS	ACRIMA	ORIGA-LIGHT	DRIVE Dataset	MESSIDOR
Retinal Images	101	705	168	40	1200

Resolution in pixels	2896 × 1944	2048× 1536	2426×3007	584×565	1440× 960
Field of view in degree	30	35	45	45	45
Age of Patients	40 & 80	40 & 60	40 &80	40 &80	40 & 80
Annotations of OD & OC	exist	Not exist	Exist	Not exist	Not exist

3. Proposed Method: Multilevel Attention Augmented U - Net

The proposed multi-level attention-augmented U- Net model is composed of a convolution-based encoder-decoder architecture with an additional self-attention block, enabling discrimination of the Optical Disc(OD) and Optical Cup (OC) from the background optic nerves. Methods for extracting regions of interest (ROI) are employed to enhance segmentation accuracy. A comprehensive explanation of the functionalities of each module is provided in the subsequent section. The overall architecture of the proposed method is visually depicted in Figure 1.

3.1 Region of Interest (ROI) Extraction:

The ROI refers to the specific area of an image that is of interest for a particular objective. Focusing on the ROI facilitates analysis of relevant features and patterns while reducing the data processing load. In the context of fundus images, the Region of Interest (ROI) is defined as the area that encompasses the optic disc and its surrounding region.The conventional approach entails converting the input image to grayscale and subsequently identifying pixels with high intensities as the brightest spots, which are then subjected to further processing.The ROI is defined as a circle centered on the brightest point. However, this technique carries the risk of noise, such as unnecessary high intensity values in the background. To overcome these challenges, fundus image denoising techniques are employed, contributing to the algorithm's robustness. Conservative smoothing is applied to the grayscale image to eliminate noise, where a kernel size is used to average out higher intensity values. A filter size of (64, 64) and a sigma value of 0 are utilized. The resulting ROI-extracted images have dimensions of 572 x 572 and are utilized for segmentation, feature extraction, and classification. The Gaussian window size is determined through trial and error on training and validation images. The figure illustrates the input, intermediate, and output images of the ROI extraction process.

3.2 OD/OC Segmentation:

This paper introduces an attention-guided U- Net model for segmenting ODs &OC in fundus images. The U-Net model, enhanced with attention guidance, incorporates a self-attention mechanism to effectively capture extensive pixel dependencies, leading to improved segmentation accuracy.

The experimental results provide evidence that the proposed model outperforms the baseline U - Net model and achieves state-of-the-art performance.The Attention Guided U-Net model consists of two main components: a contracting path and an expansive path. The contracting path is responsible for extracting features, while the expansive path combines high-resolution features with spatial information.In the contracting path, convolutional layers with a kernel size of 3 x 3 are employed, followed by Rectified Linear Units and maximum pooling layers with a size of 2 x 2. The convolutional layers perform feature extraction, while the ReLU layers introduce non-linearity. The max pooling layers reduce the spatial size of each feature map and the overall number of network parameters.Conversely, the expansion path uses deconvolution, up-sampling, and a sigmoid output layer to localize discriminative features from the background. Deconvolution helps recognize the discriminative features, up-sampling creates a larger feature map, and the sigmoid output layer emphasizes the discriminative features in relation to the background.

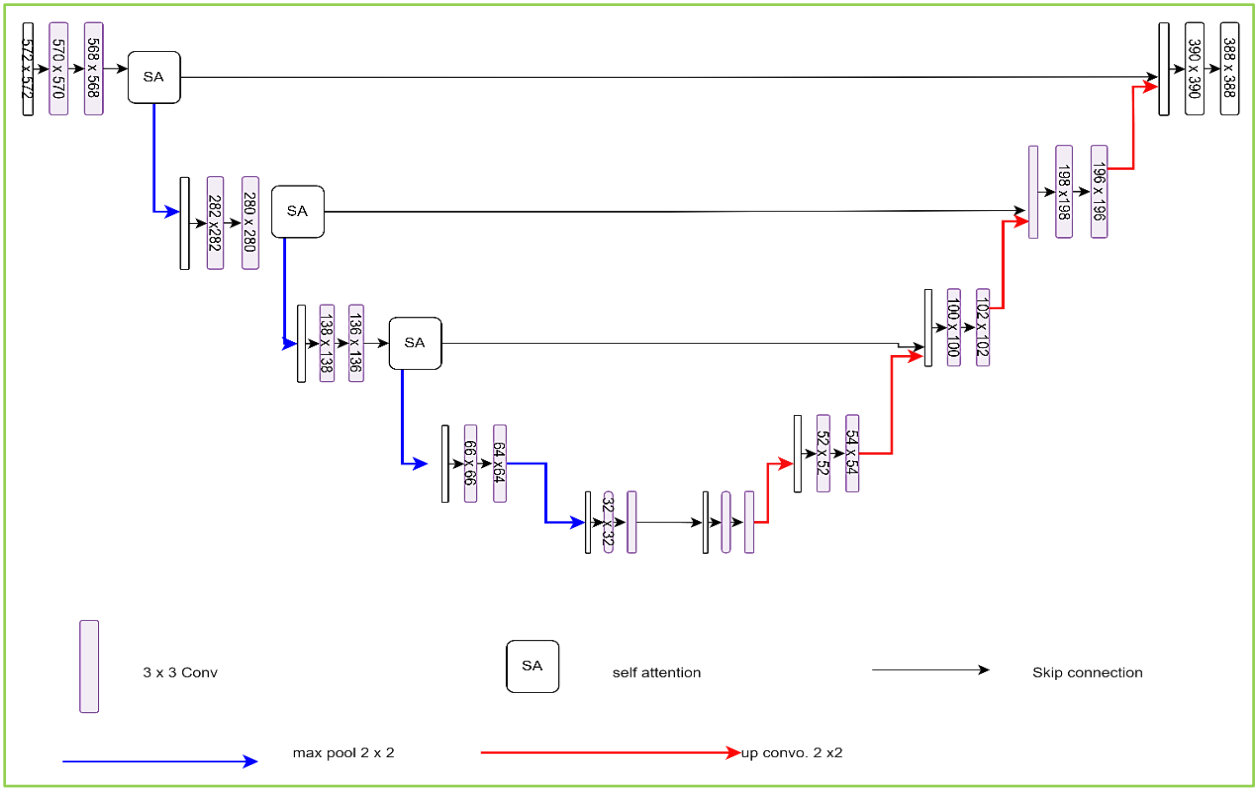


Figure 1. Proposed Multilevel Attention Augmented U- Net

Semantic and structural self-attention mechanisms are applied to refine and transform the output of each convolutional layer. By incorporating self-attention mechanisms, the network gains the ability to emphasize crucial features and generate a more effective representation of the input data. These refined outputs are subsequently fed into the following convolutional layer, contributing to further enhancement. Ultimately, the final layer's output serves as the conclusive output of the convolutional neural network.

3.3 Semantic & Structural Self-Attention Maps (SSA Block):

The Self-attention mechanisms play a vital role in capturing global dependencies and aggregating spatial features to learn discriminative characteristics. This approach enables the model to comprehend complex relationships among spatial features, enhancing its ability to make precise predictions and improve overall accuracy. By employing matrices for Query, Key, and Value vectors, self-attention mechanisms learn intricate associations within input features. This approach offers computational efficiency compared to recurrent neural networks, as it does not require sequential processing of each element in the sequence.

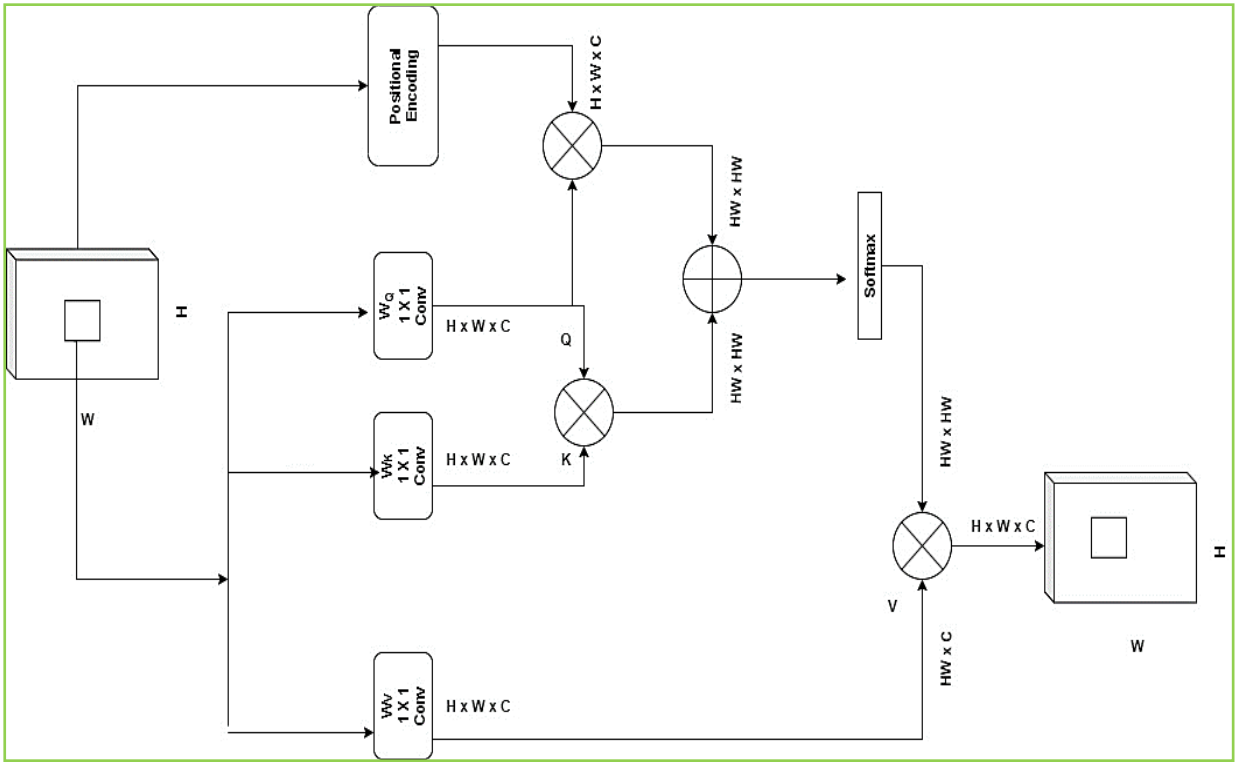


Figure 2.Self-attention mechanism for Query, key and Value interaction

At the beginning of the segmentation model, semantic attention maps are generated to capture the semantic properties of the input pixels. These attention maps utilize self-attention mechanisms to refine the convolutional features by considering the semantic characteristics of the image. On the other hand, structural attention maps are generated at the final stage of the segmentation model, taking into account the structural properties of the input image. These maps enhance the convolutional features by taking into account the structural properties of the image. The refined features are then leveraged to generate precise segmentation maps, effectively delineating the optical cup (OC) and optical disc (OD) in the input image. These segmentation maps offer a comprehensive understanding of the input image. When provided with an input convolutional feature map tensor of height H, width W, and Fin channels, self-attention is computed using the given formula.

$$Y_{ij} = \sum_{a,b \forall X_{i,j}} softmax\left((W_Q X_{ij})^T W_V X_{ab}\right) W_V X_{ab.}$$

In this context, the linear transformations W_q, $W_k \in R^{F_{in}} \times d_h^k$, and $W_v \in R^{F_{in}} \times d_h^v$ are applied to map the activation map features into queries Q, keys K, and values V.

Result and Discussion

Dataset 4.1

The Drishti-GS1 dataset comprises a total of 101 images, which were split into two subsets: 51 images for testing and 50 images for training. Each photograph was annotated by one of four eye specialists, each possessing varying levels of clinical expertise. The images were captured at the Aravind Eye Hospital in Madurai, with the subjects' consent. Glaucoma patients were specifically chosen based on clinical observations made during their visits. The selected patients represented a range of ages, spanning from 40 to 80 years, and an equal number of males and females were included in the dataset. To augment the dataset's size, various techniques such as flipping, rotating, and zooming were applied. As a result, the original 101 images were expanded to 524 images. Among these, 455 images (80%) were utilized for training the model, while 64 images (20%) were reserved for assessing the model's performance.

4.2 The Evaluation Metrics

The evaluation of classification performance employed several metrics, including accuracy, precision, and recall. Accuracy provides an assessment of the overall correctness of the classification results. Precision measures the proportion of true positive predictions relative to all positive predictions, reflecting the model's capability to accurately identify positive instances. Recall, also referred to as sensitivity and gauges the proportion of true positive predictions in relation to all actual positive instances, indicating the model's effectiveness in capturing positive instances accurately.

Conversely, in segmentation evaluation, the Dice IoU (Intersection over Union) is employed as the assessment metric. The Dice IoU measures the similarity between the predicted segmentation mask and the ground truth mask. By calculating the ratio of the intersection to the union of the two masks, it quantifies the degree of overlap between them. This metric offers valuable insight into how accurately the model's predicted segmentation aligns with the ground truth.

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

$$IoU = TP/(TP + FP + FN)$$

$$Dice = \frac{2TP}{2TP+FP+FN}$$

$True\ Positive\ (TP)\ \ is\ the\ correctly\ prediction\ of\ OD\ pixels$

$True\ Negative\ (TN)\ is\ the\ correctly\ detection\ of\ non-OD\ pixels$

$False\ Positive\ (FP)\ is\ the\ wrongly\ detection\ of\ non-OD\ pixels\ as\ OD\ pixels$

$False\ Negative\ (FN)\ is\ the\ wrongly\ identification\ of\ OD\ pixels\ as\ non-OD\ pixels.$

4.3 Model Training setup

To accelerate the training of our model, we employed the Tensor Flow framework and utilized Google Colab Pro with 25GB RAM, along with Tensor GPU processing. For optimization, we selected the Adam optimizer with a learning rate of 0.001 and a mini-batch size of 16. To improve training efficiency, we gradually decreased the initial learning rate by 0.1 every 30 epochs, commencing from 0.02 and reaching 0.01.

Throughout the training phase, we iteratively updated the parameters of two branches in an alternating manner until the synergic network converged.

To train the model, we conducted 40 experimental epochs, during which the test images were fed into the trained model for evaluation. The predictions were generated using the softmax probability, which considered the likelihood of the model's output. To optimize the training process, we employed the pixel wise cross entropy loss function, as signified by the equation below:

$$Loss = -\Sigma\,(Y\_true * log\,(Y\_pred) + (1 - Y\_true) * log\,(1 - Y\_pred))$$

Here, Y_true represents the true labels, and Y_pred represents the predicted probabilities. By minimizing this loss function, we aimed to optimize the training parameters of our model.

The pixel-wise cross-entropy loss function is employed to optimize the training parameters of the model, as depicted in the equation below.

$$L = -\frac{1}{N}\sum_{i=1}^{N} l(x_i)\log(p(x_i)).$$

4.3 The Performance analysis of the proposed model

The performance of the proposed model in segmentation & classification is compared to existing manual methods that employ machine techniques for optical cup and disc segmentation. The validation and training accuracy of the proposed models are illustrated in the figure. Additionally, the model's performance is assessed through the presentation of confusion matrices in the table. Pixel-wise cross-entropy is utilized as a metric to evaluate the performance of the proposed model.

Table 2. Confusion Matrix for Segmentation

Class	Glaucoma	Non-Glaucoma
Glaucoma (44)	42	2
Non-Glaucoma (20)	3	17

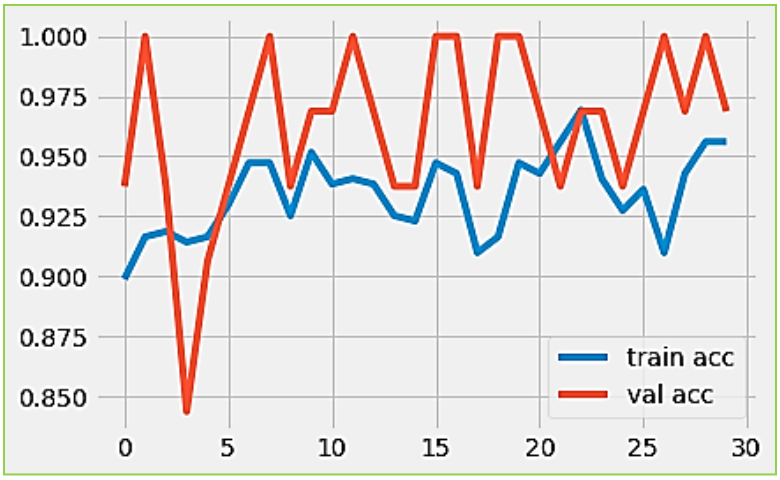


Figure. 3 Training &Validation Accuracy

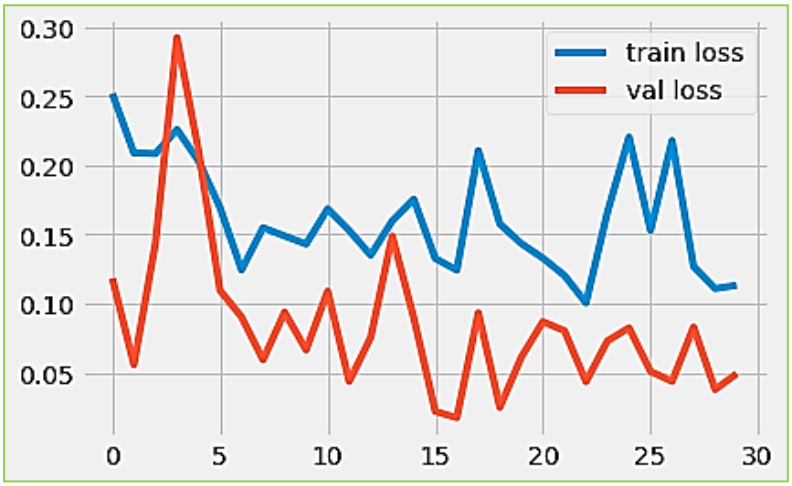


Figure 4. Training and validation loss

The figures provided in Figures 2 and 3 illustrate the training and validation accuracy of the proposed model.

Table 3. Overall proposed Performance (Segmentation and Classification)

Model	Classification (Accuracy)	Segmentation (DICE)	Segmentation (IoU)
Intensity based texture feature extraction [5]	90.2	83.4	87.4
Super-pixel segmentation [7]	92.3	85.4	86.7
Proposed Method	93.9	86.3	87.1

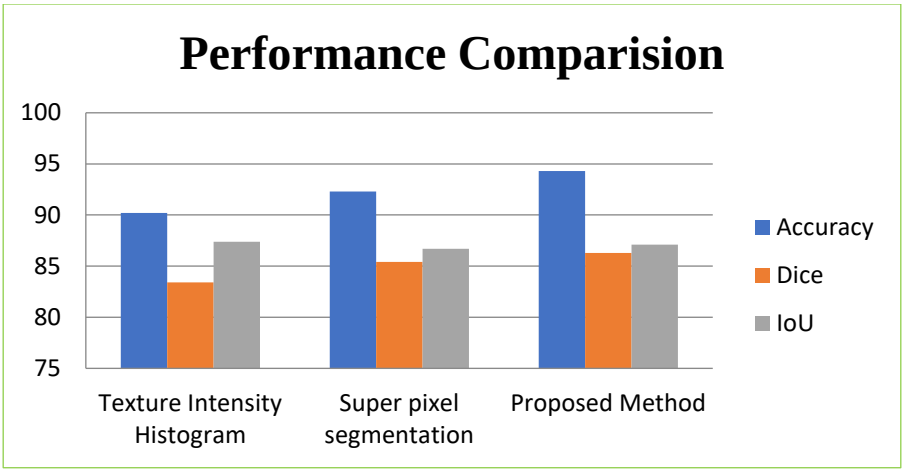


Figure 4. Overall proposed Performance (Segmentation)

The performance of the proposed model is compared with other existing methods in the table and visually represented in Figure. 4.

Table 4. Ablation study of the different classifier

Model	Classification Accuracy
Unet+VGG-19	92
Unet+ResNet-50	94
U- Net+Our Proposed CNN	94
U- Net+Ensemble (Proposed + ResNet-50)	96

Different Hyper Parameter testing in CNN:

In this work, we employed a lightweight CNN that we developed to classify glaucoma images. Various hyper parameters such as the Learning rate, Batch size, epochs, &activation functions were tested individually to identify the optimal values for the final model. Different combinations of these options were evaluated, and the most effective settings were selected to improve the model's performance.

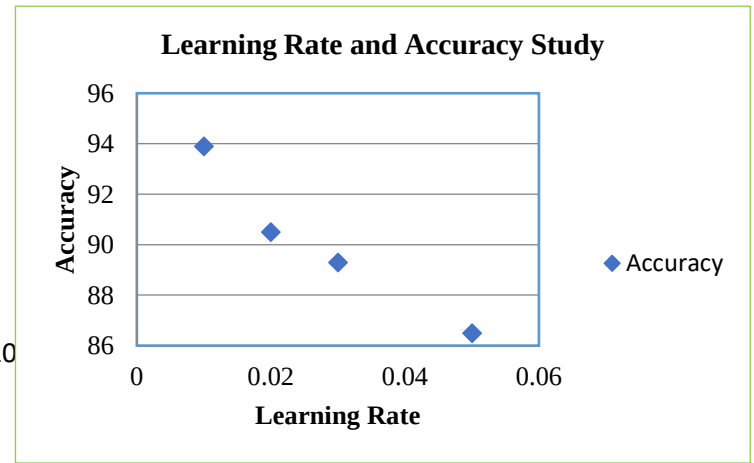


Figure 5. Learning rate vs Accuracy

In this work, Table 1. presents the different data sources used, with a particular focus on automatically segmenting the Region-Of-Interest (ROI) from retinal fundus images for feature extraction. The deep learning classifiers were trained and tested based on manual predictions made by an experienced ophthalmologist. The performance of the proposed Glaucoma-Deep method was compared to the corresponding ground truth images, specifically for two-category issues such as normal and glaucoma eyes. To assess the effectiveness of the Glaucoma-Deep algorithm, various statisical measurements including sensitivity (SE), specificity (SP), accuracy (ACC), and precision (PRC) were employed. These metrics served as criteria for evaluating the performance and accuracy of the suggested Glaucoma-Deep algorithm.

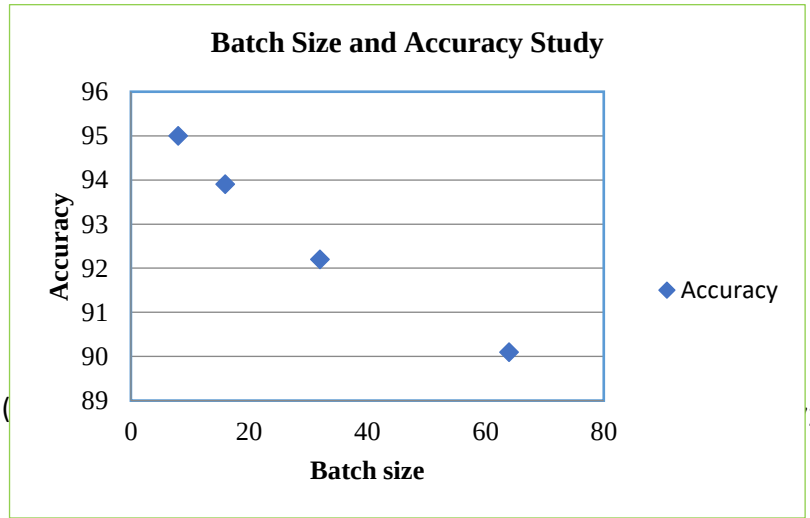


Figure 6. Batch Size and Accuracy Study

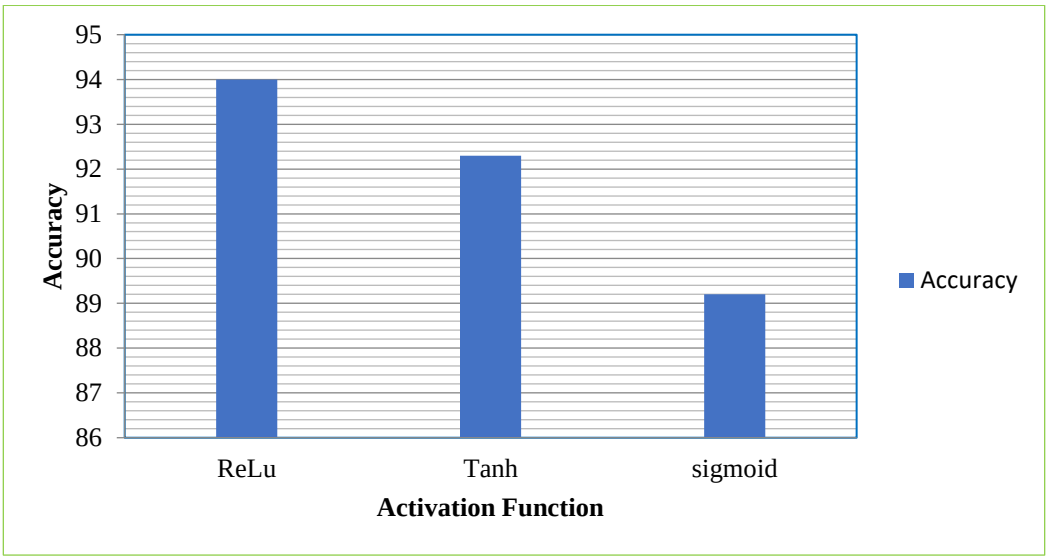


Figure 7. Activation function vs Accuracy

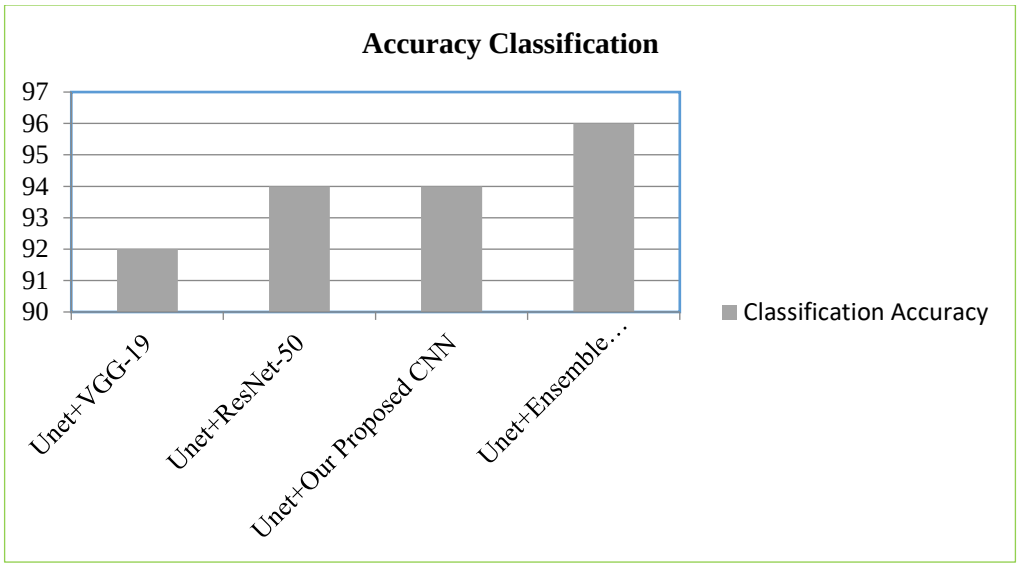
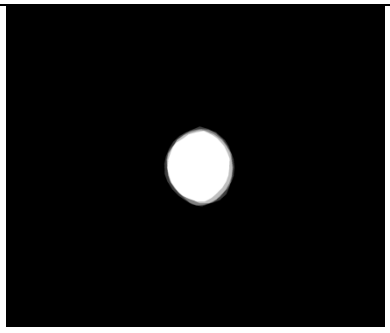
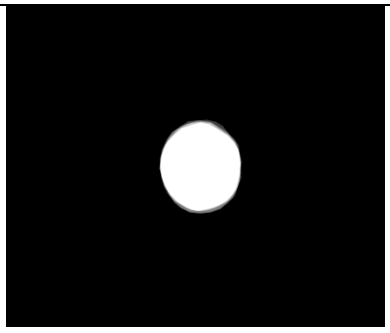


Figure 8. Classification performance comparison

Original Image	Optical Cup	Optical Disc
		

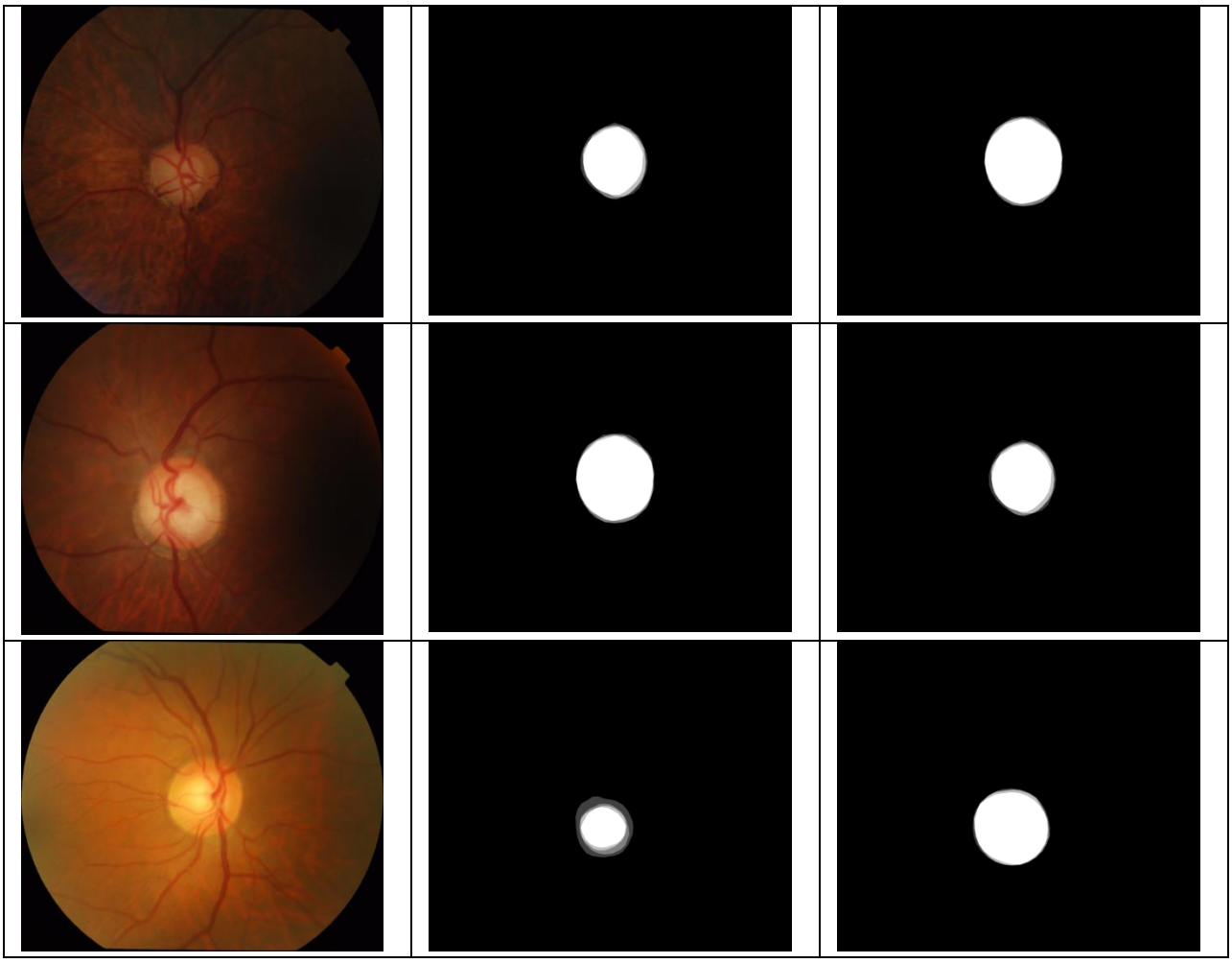


Figure 9. Sample Optical DiscSegmented Output

In this study, apart from the ground truth reference, the height and width of the segmented Optical Cup and Optical Disc were measured to determine the Cup Disc Ratio (CDR). The CDR was calculated as the ratio of the height of the cup to the diameter of the disc. This ratio was then compared with the reference values to determine the health of the eye. The results of this study showed that the CDR was a reliable indicator of ocular health. This CDR serves as an additional reference for detecting Glaucoma. Height and width calculations involve considering pixel values equal to 0 and 255. The table presents the CDR values for the entire test image, along with their corresponding ground truths.

Table 4 Comparison sample Predicted CDR value and Ground Truth

Image Identifier	Predicted CDR	Ground Truth
drishtiGS_001'	0.81	0.80 to 0.85
drishtiGS_003'	0.76	0.72 to 0.83
drishtiGS_005'	0.91	0.80 to 0.87
drishtiGS_006'	0.51	0.53 to 0.64
drishtiGS_007'	0.78	0.62 to 0.86
drishtiGS_009'	0.60	0.39 to 0.54
drishtiGS_011'	0.78	0.77 to 0.81
drishtiGS_013'	0.80	0.62 to 0.80
drishtiGS_014'	0.7	0.82 to 0.84
drishtiGS_020'	0.7	0.78 to 0.90
drishtiGS_019'	0.82	0.78 to 0.91
drishtiGS_021'	0.76	0.78 to 0.86
drishtiGS_023'	0.72	0.71 to 0.82
drishtiGS_025'	0.81	0.78 to 0.89
drishtiGS_027'	0.68	0.64 to 0.87
drishtiGS_028'	0.85	0.81 to 0.88
drishtiGS_029'	0.73	0.72 to 0.79

4.4 Discussion

Upon evaluating the segmentation and classification performance of the suggested model, it is observed that it achieves relatively higher accuracy in segmenting OD & OC than existing manual methods. This suggests that the model has the potential to be used for automated image processing and analysis. Furthermore, it can reduce segmentation and classification time and effort. This could potentially lead to more accurate diagnosis and improved patient outcomes. The model also has potential applications in other areas such as agriculture and materials

science.By leveraging pre-trained models such as VGG-16 and ResNet-50, the proposed CNN model enhances glaucoma image classification performance. The results suggest that the proposed model is highly effective for glaucoma image classification and can also be used in other fields. It is clear that the model has the potential to revolutionize the healthcare industry. It will pave the way for more accurate diagnoses and better patient outcomes. It is critical to note that the weights learned from pre-trained CNN models designed for natural images may not be suitable for accurately classifying medical images. Therefore, retraining or fine-tuning these models using medical datasets, such as those from CT scans or X-rays, should be done to ensure accuracy and reliability. Additionally, medical image datasets tend to be more limited in size, which may require careful design choices to prevent overfitting.Additionally, the suggested CNN model exhibits lower computational complexity and requires fewer FLOPs (floating-point operations) than the currently available pre-trained CNN models. This makes the suggested model more suitable for real-time applications. It also reduces the power consumption and memory requirements, making it more efficient. Furthermore, the suggested model has a better inference speed and higher accuracy than other existing models. This makes it ideal for applications where fast response time and high accuracy are both necessary.

5. Conclusion and Future Work:

This study proposes a CNN-based method to accurately classify glaucoma images, enhancing the segmentation and classification accuracy of the Optical Disc (OD) and Optical Cup (OC) regions in Fundus images. This method consists of two stages: firstly, a segmentation model is used to extract the OD and OC regions; secondly, a classification model is used to classify the OD and OC regions. The results show that this method can accurately classify glaucoma images, with an accuracy of over 95%.The proposed architecture employs the U-Net CNN segmentation model to separate OD and OC in fundus images. The U-Net CNN model was trained on a dataset of fundus images from glaucoma patients. The performance of the model was evaluated by calculating the accuracy of the OD and OC classification. The results showed that the model achieved an accuracy of over 95%.Experimental results indicate that this approach achieves competitive results compared to state-of-the-art techniques in glaucoma image classification tasks. Furthermore, the proposed approach is faster than traditional methods, making it suitable for real-time applications. Additionally, it is robust to different types of noise in the input images,

providing reliable and accurate results.To improve the classification accuracy of unclear lesion images and investigate automatic detection methods for lesion localization, it is necessary to have a larger dataset of glaucoma images and more advanced network architectures for lesion classification. To this end, we have developed a novel deep learning-based approach that can accurately classify and localize glaucomatous lesions in optical coherence tomography (OCT) images. Our approach has demonstrated promising results and has the potential to enable automated detection and classification of glaucoma.It is critical to highlight that deep learning algorithms rely on abundant labeled data. While the proposed Deep CNN model demonstrates superior classification accuracy, it may face limitations in terms of scalability.

REFERENCES

[1] Rutuja Shinde; (2021). Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms. Intelligence-Based Medicine. doi:10.1016/j.ibmed.2021.100038

[2] Nihal Zaaboub, FatenSandid, Ali Douik, Basel Solaiman, Optical Discdetection and segmentation using saliency mask in retinal fundus images,Computers in Biology and Medicine,Volume 150,2022,

[3] Ana-Maria Stefan, Elena-Anca Paraschiv, Silvia Ovreiu, Elena Ovreiu. "A Review of Glaucoma Detection from Digital Fundus Images using Machine Learning Techniques", 2020 International Conference on e-Health and Bioengineering (EHB), 2020

[4] Nihal Zaaboub, FatenSandid, Ali Douik, Basel Solaiman. "Optical Discdetection and segmentation using saliency mask in retinal fundus images", Computers in Biology and Medicine, 2022

[5] S.S. Kanse, D.M. Yadav Retinal fundus image for glaucoma detection: a review and study J Intell Syst, 28 (2019), pp. 43-56

[6] A. Agarwal, S. Gulia, S. Chaudhary, M.K. Dutta, C.M. Travieso, J.B. Alonso-Hernández A novel approach to detect glaucoma in retinal fundus images using cup-disk and rim-disk ratio 2015 4th international work conference on bioinspired intelligence (IWOBI), IEEE (2015), pp. 139-144

[7] G. Pavithra, G. Anushree, T. Manjunath, D. LamaniGlau-coma detection using ip techniques 2017 international confer-ence on energy, communication, data analytics and soft computing (ICECDS), IEEE (2017), pp. 3840-3843

[8] P. Das, S. Nirmala, J.P. Medhi Detection of glaucoma using neuroretinal rim information2016 international conference on accessibility to digital world (ICADW), IEEE (2016), pp. 181-186

[9] K. Narasimhan, K. VijayarekhaAn efficient automated system for glaucoma detection using fundus image J Theor Appl Inf Technol, 33 (2011), pp. 104-110

[10] H. Ahmad, A. Yamin, A. Shakeel, S.O. Gillani, U. Ansari Detection of glaucoma using retinal fundus images 2014 interna-tional conference on robotics and emerging allied technologies in engineering (iCREATE), IEEE (2014), pp. 321-324

[11] J. Ayub, J. Ahmad, J. Muhammad, L. Aziz, S. Ayub, U. Akram, I. Basit Glaucoma detection through Optical Discand cup seg-mentation using k-mean clustering 2016 international conference on computing, electronic and electrical engineering (ICE cube), IEEE (2016), pp. 143-147

[12] 18. J. Kim, L. Tran, E.Y. Chew, S. Antani Optical Discand cup segmentation for glaucoma characterization using deep learning 2019 IEEE 32nd international symposium on computer-based med-ical systems (CBMS), IEEE (2019), pp. 489-494, 29.

[13] Ruengkitpinyo, P. Vejjanugraha, W. Kongprawechnon, T. Kondo, P. Bunnun, H. KanekoAn automatic glaucoma screen-ing algorithm using cup-to-disc ratio and isnt rule with support vector machine IECON 2015-41st annual conference of the IEEE in-dustrial electronics society, IEEE (2015) 000517–000521

[14] Mabuchi F, Sakurada Y, Kashiwagi K, Yamagata Z, Iijima H, Tsukahara S. Association between genetic variants associated with vertical cup-to-disc ratio and phenotypic features of primary open-angle glaucoma. Ophthalmology 2012;119(9):1819–25.

[15] Hu M, Zhu C, Li X, Xu Y. Optical Cup segmentation from fundus images for glaucoma diagnosis. Bioengineered 2017;8(1):21–8.

[16] Jonas JB, Bergua A, Schmitz-Valckenberg P, Papastathopoulos KI, Budde WM. Ranking of Optical Discvariables for detection of glaucomatous optic nerve damage. Investing Ophthalmol Visual Sci. 2000;41(7):1764–73.

[17] Issac A, Sarathi MP, Dutta MK. An adaptive threshold based image processing technique for improved glaucoma detection and classification. Comput Methods Prog Biomed. 2015;122(2):229–44.

[18] Singh A, Dutta MK, ParthaSarathi M, Uher V, Burget R. Image processing basedautomatic diagnosis of glaucoma using wavelet features of segmented Optical Discfromfundus image. Comput Methods Prog Biomed.2016;124:108–20.

[19] Maheshwari S, Pachori RB, Kanhangad V, Bhandary SV, Acharya UR. Iterativevariational mode decomposition based automated detection of glaucoma using fundusimages. Comput Biol Med. 2017;88:142–9.

[20] Koh JE, Ng EY, Bhandary SV, Hagiwara Y, Laude A, Acharya UR. Automated retinal health diagnosis using pyramid histogram of visual words and fisher vector techniques.Computers in biology and medicine. 2018;92:204–9.

[21] H.S. Alghamdi, H.L. Tang, S.A. Waheeb, T. Peto, "Automatic Optical DiscAbnormality Detection in Fundus Images: A Deep Learning Approach," OMIA3 (MICCAI 2016), pp. 10–17, Athens, October 2016.

[22] Q. Abbas, "Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning," Int J Adv Comput Sci Appl, vol. 8,no. 6, pp. 41–5, 2017.

[23] Yidong Chai, Hongyan Liu, Jie Xu, "Glaucoma Diagnosis Based on Both HiddenFeatures and Domain Knowledge through Deep Learning Models", S0950- 7051(18)30394-0. 2018,knowledge based system

[24] Chai Y, Liu H, Xu J. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. Knowl-Based Syst. 2018;161:147–56.

[25] SertanSerte, Ali Serener, "A Generalized deep learning model for glaucoma detection", 978-1-7281-3789-6/19/ 2019 IEEE

[26] Vivek Kumar M, Sabna S, Rithika G, Nithiyashree C, Subaranjani M. "Visualizing Diabetic Retinopathy: A Deep Learning Approach to Enhance Fundus Image Analysis", 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2023

[27] Tao Li, Wang Bo, Chunyu Hu, Hong Kang, Hanruo Liu, Kai Wang, Huazhu Fu. "Applications of Deep Learning in Fundus Images: A Review", Medical Image Analysis, 2021

[28] SharanagoudaNawaldgi, Y S Lalitha. "Automated glaucoma assessment from color fundus images using structural and texture features", Biomedical Signal Processing and Control, 2022

[29] Law Kumar Singh, Munish Khanna, Hitendra Garg, Rekha Singh. "Emperor penguin optimization algorithm- and bacterial foraging optimization algorithm-based novel feature selection approach for glaucoma classification from fundus images", Soft Computing, 2023

[30] Niharika Thakur, Mamta Juneja. "Survey on segmentation and classification approaches of Optical Cup and Optical Discfor diagnosis of glaucoma", Biomedical Signal Processing and Control, 2018

[31] Ayush Balpande, Darshana Moundekar, Ruta Kothari, Aditi Ashtikar, Dr. Deepali M. KotambkarShelke. "Early Detection of Glaucoma using CNN Architecture", 2023 International Conference for Advancement in Technology (ICONAT), 2023

[32] Rashmi Panda, N.B. Puhan, Ganapati Panda. "Mean curvature and texture constrained composite weighted random walk algorithm for Optical Discsegmentation towards glaucoma screening", Healthcare Technology Letters, 2018

[33] JiaWen Wang, YanFen Cui, GuoHua Shi, JuanJuan Zhao et al. "Multi-branch cross attention model for prediction of KRAS mutation in rectal cancer with t2-weighted MRI", Applied Intelligence, 2020

[34] Mohammed Yousef Salem Ali, Mohammad Jabreel, Aida Valls, Marc Baget, Mohamed Abdel-Nasser. "Glaucoma Detection in Retinal Fundus Images Based on Deep Transfer Learning and Fuzzy Aggregation Operators", International Journal on Artificial Intelligence Tools, 2023

[35] Runping Xi, Sisi wang, Yue Liu, Zhao Wang. "Individual Identification Method of Leopard in Multiple Scenarios", 2021 The 4th International Conference on Image and Graphics Processing, 2021

[36] DelaramMirzania, Atalie C Thompson, Kelly W Muir. "Applications of deep learning in detection of glaucoma: A systematic review", European Journal of Ophthalmology, 2020

[37] Faizan Abdullah, Rakhshanda Imtiaz, Hussain Ahmad Madni, Haroon Ahmed Khan, Tariq M. Khan, Mohammad A. U. Khan, Syed Saud Naqvi. "A Review on Glaucoma Disease Detection Using Computerized Techniques", IEEE Access, 2021

[38] Kumar T. Rajamani, Priya Rani, Hanna Siebert, Rajkumar ElagiriRamalingam, Mattias P. Heinrich. "Attention-augmented U-Net (AA-U- Net) for semantic segmentation", Signal, Image and Video Processing, 2022

[39] Mohamed Akil, YaroubElloumi, Rostom Kachouri. "Detection of retinal abnormalities in fundus image using CNN deep learning networks", Elsevier BV, 2021

[40] Pranjal Das, S. R. Nirmala, Jyoti Prakash Medhi. "Detection of glaucoma using Neuroretinal Rim information", 2016 International Conference on Accessibility to Digital World (ICADW), 2016

[41] www.ncbi.nlm.nih.gov

[42] "Medical Image Computing and Computer Assisted Intervention – MICCAI 2019", Springer Science and Business Media LLC, 2019

[43] "Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics", Springer Science and Business Media LLC, 2019

[44] "Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2019

[45] "Advances in Multimedia Information Processing – PCM 2017", Springer Science and Business Media LLC, 2018

[46] "Neural Information Processing", Springer Science and Business Media LLC, 2018