IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 5, 2021, accepted March 15, 2021, date of publication March 23, 2021, date of current version April 1, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3068204

The Use of U-Net Lite and Extreme Gradient Boost (XGB) for Glaucoma Detection

OLUWATOBI JOSHUA AFOLABI[1], GUGULETHU P. MABUZA-HOCQUET[2], FULUFHELO V. NELWAMONDO[3,4], (Senior Member, IEEE), AND BABU SENA PAUL[4], (Member, IEEE)

[1]Department of Electrical/Engineering Science, University of Johannesburg, Johannesburg 2006, South Africa
[2]Defence and Security, Council for Scientific and Industrial Research, Pretoria 0001, South Africa
[3]Nextgen Enterprises and Institutions, Council for Scientific and Industrial Research, Pretoria 0001, South Africa
[4]Institute for Intelligent Systems, University of Johannesburg, Johannesburg 2006, South Africa

Corresponding author: Oluwatobi Joshua Afolabi (joshua.aoj@gmail.com)

ABSTRACT Glaucoma has been credited to be the foremost cause of preventable loss of sight in the world second only to cataract. Its effect on the eye is usually irreversible and can only be prevented by early detection. In this paper, we developed a glaucoma detection technique. This technique includes a modified U-Net model called 'U-Net lite' and an extreme gradient boost (XGB) algorithm. The novel U-Net lite model is designed to have fewer parameters than the original U-Net model. The U-Net lite's parameters are 40 times fewer than the original U-Net model which makes the proposed model faster and cheaper to train. The proposed model is utilized to segment both the optic cup and the optic disc from the fundus images. The extreme gradient boost algorithm is utilized to analyze extracted features from segmented optic cups and discs and hence detect glaucoma. The proposed U-Net lite model was both trained and tested on the DRIONS, DRISHTI-GS, RIM-ONE V2 and the RIM-ONE V3 databases. When tested for optic disc segmentation on the four databases, the model achieved the following average dice-scores: 0.96 on RIM-ONE V3, 0.97 on RIM-ONE V2, 0.96 on DRIONS, and 0.97 on DRISHTI-GS. The XGB algorithm achieved an accuracy of 88.6% and an AUC-ROC of 93.6 % in detecting glaucoma from the RIM-ONE V3 and DRISHTI-GS database. The proposed glaucoma detection technique achieves a state-of-the-art accuracy and is useful for observing structural changes in an optic cup and optic disc.

INDEX TERMS Fundus image, glaucoma, segmentation, U-Net.

I. INTRODUCTION

Glaucoma is an eye ailment characterized by a growing deterioration of the optic nerve head as well as ganglion cells in the retina [1], [2]. It is a foremost source of preventable loss of sight with no clear symptom at its preliminary stages. About 50% of its victims are not usually aware of its presence [3]–[5]. Glaucoma develops because of an obstruction to the flow of the aqueous humour in the eye canal. The obstruction results to a continuous rise in the eye pressure and consequently increasing the size of the optic cup as seen in Fig.1. The enlarged optic cup causes a continuous loss of fibres located at the optic nerve and this is perceived as a gradual loss of vision in its victims. In the preliminary stage of the disease, victims have no symptom or sign but as the disease advances, victims notice a narrowing of the visual field beginning from the peripheral [2], [6] [7]. The damaging action of the disease cannot be reversed and if left unchecked may lead to a permanent loss of sight [8]. Therefore, a procedure that allows for speedy detection of the disease is significant.

The diagnosis of glaucoma is typically conducted by assessing variation in the structure of the optic nerve head [9], [10]. One of the methods that have been utilized to identify the presence of glaucoma is the optic Cup-to-Disc Ratio (CDR).

The CDR is the ratio of the longitudinal diameter of the optic cup to the longitudinal diameter of the optic disc [11]. The CDR method depends on the accurate segmentation of the optic disc as well as the optic cup. Several techniques have been utilized to segment the optic disc and the optic cup. The mostly utilized technique involves (i) the pre-processing

The associate editor coordinating the review of this manuscript and approving it for publication was Junhua Li.

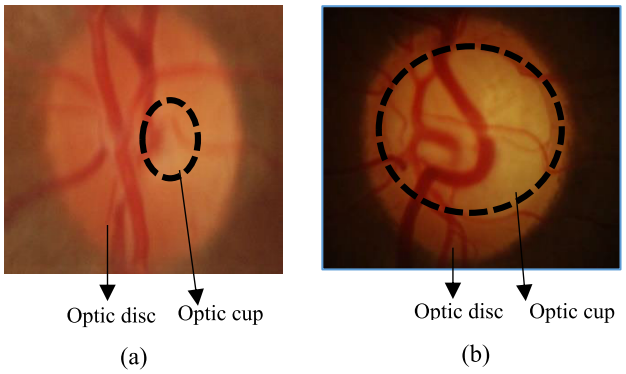


FIGURE 1. **(a) A fundus image with a non-glaucomatous optic cup. (b) A fundus image with an enlarged glaucomatous cup.**

of fundus image (ii) determining the regions in the fundus image which are of interest (iii) localizing the optic disc (OD) and finally localizing the optic cup (OC). This technique has been utilized in many studies [12]–[17] with little variation in its implementation. However, the technique is computationally intensive especially when it is tested on large batches of fundus images. This is because the technique must be applied to each of the fundus images individually. Moreover, the accuracy of the technique is substantially influenced by the differing pixel intensities of the fundus images across the databases. Therefore, the above-explained technique is not robust to noise and presence of pathologies in the fundus images.

The following are the contributions of this paper. (1) A proposed segmentation model called U-Net lite. The novel model has 40 times fewer parameters than the original U-Net model which makes it faster and cheaper to train (1) A segmentation algorithm that yields a high dice-score OC and OD segmentation. (3) A glaucoma classification algorithm based on the extreme gradient boost (XGB). The XGB classifier was trained with carefully engineered features from the fundus images. This novel approach eliminates the challenges of varying CDR threshold values when using the CDR method for glaucoma detection. Therefore, there is no need to set a threshold value when using the model.

The rest of this paper is arranged as follows: related work is discussed in section 2, the proposed experimental approach is discussed in section 3, section 4 presents the achieved experimental results, section 5 presents the discussion and analysis of the achieved results, section 6 discusses the limitation of the study, section 7 presents the conclusion, and the last section presents the future work.

II. RELATED WORK

Modern advances in object recognition and image processing have brought about the application of deep learning models and systematic algorithms for medical images segmentation. In modern computing, U-Net has been referred to as the gold standard for many biomedical segmentation exercises and the reason is not far fetched as the architecture has achieved high scores in many segmentation exercises [18]–[37]. However, a major downside of the U-Net architecture which is also true of many deep learning architectures is the high cost of computation and training. Alexander *et al.* [18] compared the performance of U-Net models with E-net [19] and Box E-net models [18]. The E-net model has been used for real time semantic segmentation and was designed to have fair segmentation performance but efficient processing performance. The Box E-net is an improvement on the E-net. The improvement was achieved by replacing some convolution layers with box-convolution layers. Alexander *et al.* concluded that although both the Box E-net and the E-net architectures are 15 times faster than the U-Net model, they are still about 2% less accurate than the U-Net model. Hence, there is a need for a model that combines both speed and accuracy in a segmentation process. Our proposed model is designed to be 40 times lighter than the original U-Net model while achieving a comparable accuracy performance with the original U-Net.

Luo *et al.* [20] proposed the use of 'Attention-Dense-U-Net' model to segment blood vessels from fundus images. The proposed model incorporates a densely connected network as well as an attention mechanism to the original U-Net model. Although Luo *et al.* did not give information about the model speed or the size of the model, it can be inferred that the model would be about the same size and speed as the original U-Net model. The huge size of the model translates to a high computational cost and long training time. We used a variant of the proposed model architecture to segment blood vessels [21] and we achieved similar results to that of *Luo et al*. For instance, *Luo et al*. achieved an accuracy of 0.9663 and a sensitivity of 0.8075 on the DRIVE database and we achieved an accuracy of 0.9615 and a sensitivity of 0.8309 on the same database.

Zeng *et al.* [22] proposed the use of a network based on U-Net to segment nuclei from histology images. The performance of the U-Net model was compared with the performance of other models which include the CellProfiler (CP) model [23], Fiji [24] and CNN models [25]. The proposed U-Net model outperformed the other models. Again, this work validates the superiority of U-Net model in medical segmentations.

Yahyatabar *et al*. [26] proposed the use of densely connected U-Net models for lung segmentation. With layers of their proposed U-Net model densely connected, the authors still claim that the model is the lightest model for lung segmentation in CR images. Although the proposed model is very light, it achieved comparable results with other models. This reveals that a carefully trained light model can achieve comparable results with its heavier counterpart.

Luo *et al.* [27] proposed the use of Vessel-Net for retinopathy screening. The proposed Vessel-Net model is built on a U-Net model. The architecture of the proposed model embraces three tiers of fundus information which are the global stream (learnt by a Res-Net-50 model [28]), disc region stream (learnt by a pre-trained U-net model) and the

vessel-related stream (learnt by a Ladder-Net model [29]). The authors recorded an area under curve (AUC) score of 0.8464.

The U-Net model has also been used for a variety of segmentation tasks which include iris segmentation [30], left ventricle endocardial border segmentation [31], dental panoramic image segmentation [32], stroke lesion segmentation [33] , liver and spleen segmentation [34] and even non-biomedical segmentation tasks like road crack detection [35], detection of salt domes [36] and defect segmentation [37].

The following methods have mostly been used to segment the optic disc (OD) and the optic cup (OC) from fundus images.

Maninis *et al*. [38] proposed a method that utilized transfer learning technique to train convolutional neural networks (CNN) [39]. CNN was built on a VGG-16 architecture [40]. The proposed method was utilized to segment the OD and OC from fundus images. Maninis *et al*. recorded a dice score of 0.96. Although the method proposed achieved a high dice score, the drawback of this approach is that the size of the architecture is large. There are about 1.85×10^7 parameters to be trained which introduce a lot of computational complexities. Our proposed method has only 7.8×10^5 parameters. The dice score, as well as the Intersection-over-Union (IoU) score, are metrics utilized to measure the goodness of a segmentation process. A good segmentation process will have high dice and IoU scores.

Zilly *et al.* [16] utilized a technique that included the use of boosted CNN, filtered entropy [41], normalized contrast and standardized patches to segment the optic cups from fundus images. The AdaBoost algorithm [42] was utilized for the boosting operation. The proposed method was assessed using the DRISHTI-GS database [43], [44] and the RIM-ONE database [45]. The method achieved a dice and IoU score of 0.85 and 0.87 respectively.

Improving on what was done by Zilly *et al*., Buhmann *et al.* [46] proposed a new method which does not require the cropping of the disc location before segmenting the optic cup. This method includes a process that picks points holding salient information on the fundus image by using entropy sampling. This method eliminates the computational complexity involved in the method proposed by Zilly *et al.* The method achieved higher dice-score than Zilly *et al*. [16] but a lower IoU score. The drawbacks of the methods proposed by both Zilly *et al*. and Buhmann *et al.* are that the methods require a lot of pre-processing and post-processing. This is because the segmentation process involves randomly selecting points of interest on the entropy maps and thereby increasing computational cost. Our proposed method has very little pre-processing and no post-processing.

Tabassum *et al.* [47] proposed a method that jointly segmented the optic disc and the optic cup. The segmentation process was treated as a semantic pixel-wise labelling problem. The method achieved a high dice score on the DRISHTI-GS and RIM-ONE dataset. The method achieved a dice score of 0.92 and an IoU score of 0.86 on the DRISHTI-GS database. However, the drawback of the proposed method is that it has a high number of parameters needed to be trained which incurs a high training cost. For instance, the authors reported that it takes 5.5 hours to train the RIM-ONE dataset using the Intel(R) Xeon(R) W-2133 CPU 3.60GHz processor, 32GB RAM, and Nvidia 2080TI GPU. Our proposed method uses a model that trains the same dataset for 32.5 minutes using Kaggle's 2 CPU cores, 14 GB RAM, 1 NVIDIA Tesla K80 GPU.

Jiang *et al.* [48] proposed a joint segmentation of optic cup and optic disc by using an end-to-end region-based convolutional neural network. The proposed method assumes that the shape of the optic cup and disc is elliptical. After segmentation of optic cups and discs, the authors detected glaucoma using the vertical optic cup to disc ratio. When the proposed method was tested on the ORIGA dataset, the authors achieved an average overlapping error of 0.209 and 0.063 for the optic cup and optic disc respectively. The drawback is that the proposed method assumes an elliptical shaped optic cup and disc which is not always the case.

Qin *et al.* [49] proposed a method for optic cup and optic disc segmentation based on a modified fully convolutional network (FCN) combined with the inception building blocks as used in GoogleNet. The method was tested on the REFUGE dataset. A dice score of 0.92 and IoU score of 0.90 was recorded for the optic cup segmentation process. However, the drawback of the proposed method is that the pre-processing may need adjustment on the parameters of the Hough circle transformation algorithm to achieve optimum results.

Shah *et al.* [50] proposed a parameter shared branched network and a weak region of interest network for the accurate segmentation of the optic cup and optic disc. The networks employ the use of dynamic cropping and are trained using a single neural network. The proposed networks were then tested on the DRISHTI-GS database and a dice score of 0.96 was achieved for optic disc segmentation. The drawback is that the proposed method involves the use of two networks, and this adds to the overall computational cost.

Thakur *et al.* [51] proposed the use of an algorithm they called a 'Level Set Based Adaptively Regularized Kernel-Based Intuitionistic Fuzzy C Means (LARKIFCM)'. The algorithm involves the use of clustering to segment optic disc and optic cup. The method achieved a dice score of 0.92 when tested on the DRISHTI-GS database. The drawback of the proposed method is that it is time consuming and parameters of the algorithm depend on the dataset used as input. The proposed method may not be suitable for large datasets as tuning of parameters is needed for each dataset.

After a successful segmentation process, the CDR is usually utilized to detect glaucoma. The CDR method of detecting glaucoma was employed by Patel *et al.* [52]. In their work, a CDR threshold value of 0.5 was utilized to classify fundus image to either the glaucomatous or the non-glaucomatous class i.e. fundus images with CDR values

less than or equal to 0.5 were considered non-glaucomatous and CDR values higher than 0.5 were considered glaucomatous. A total of 100 fundus images were used and an accuracy of 0.78 was recorded.

Zhao *et. al*. [53] estimated the CDR value of optic nerve heads by using a semi-supervised learning model. The proposed method comprises two phases: a supervised learning phase using a random forest regressor and a convolutional neural network phase. The method was tested on 421 fundus images and achieved a CDR error that is lower than 0.0563 and an area-under-curve (AUC) of 0.905.

In a study done by Virk *et al.* [54], 50 fundus images were classified into either glaucoma or non-glaucoma class. Virk *et al.* concluded that fundus images with CDR values between 0.3 and 0.5 should be classified as non-glaucomatous while those of above 0.5 should be classified as glaucomatous. Virk *et al.* recorded an accuracy of 80% when these threshold values were utilized to detect glaucoma.

In another study done by Mohamed *et al*. [55], fundus images from the RIM-ONE database were tested for glaucoma. The testing algorithm included the CDR method and a CDR threshold of 0.6 was utilized. Mohamed *et al*. concluded that CDR values for non-glaucomatous fundus images fall between 0.4 and 0.6 and those of glaucomatous fundus images are higher than 0.6. It should be noted that the work was carried out on only the RIM-ONE database and the use of CDR threshold of 0.6 may only be suitable for this database.

After segmenting the optic cups and optic discs from fundus images, Mvoulana *et al.* [56] employed the CDR method to detect glaucoma. In their study, a CDR threshold value of 0.63 was employed to classify fundus images to either glaucomatous or non-glaucomatous. The threshold value was computed by evaluating the CDR mean and standard deviation of every fundus images in glaucoma and non-glaucoma classes. Fundus images with CDR value greater than 0.63 were classified glaucomatous.

Murthi *et al.* [57] used the least square fitting algorithm to segment the optic cups and the optic discs from the fundus images. After the segmentation process, the ellipse fitting algorithm was utilized to smoothen the boundaries of the disc and cup. Murthi *et al.* further utilized the CDR method to detect glaucoma. A CDR threshold value of 0.65 was utilized. For example, fundus images which have CDR values of 0.68 were considered glaucomatous.

Khan *et al.* [58] separated the optic cups and the optic discs from fundus images by utilizing the mean threshold morphological technique. Together with several attributes, a CDR threshold value of 0.5 was used to recognize glaucomatous fundus images.

Lotankar *et al.* [59] suggested a technique for detecting glaucoma by extracting several attributes from the optic nerve head. Attributes extracted from the optic nerve head included the rim to disc area ratio, the cup to disc area ratio and the cup to disc ratio. Lotankar *et al.* proposed that CDR values for non- glaucomatous fundus images range from 0.2 to 0.4 and 0.5 to 1 for glaucomatous fundus images.

Roslin *et al*. [60] segmented the blood vessels in the optic discs by using an edge detection algorithm which was based on the Prewitt operators. In their method, the CDR of each fundus image was measured. The authors proposed that the phases of glaucoma development can be studied from the CDR values of the fundus images. To classify fundus images into glaucomatous or non-glaucomatous, a CDR threshold value of 0.3 was utilized. Fundus images that have CDR threshold values of 0.3 or less were labelled non-glaucomatous and fundus images which have CDR threshold values that are higher than 0.3 were labelled glaucomatous.

The major drawback in the use of CDR to detect glaucoma is that different CDR threshold values have been utilized by the authors who employed this method [61]. Also, the CDR threshold values utilized depended largely on the dataset and much more on the judgement of the authors. These factors have made the use of CDR threshold method in detecting glaucoma a subjective and less accurate approach especially when being used on fundus images from different databases. In our proposed method, we eliminate the challenges of varying CDR threshold values when using the CDR method for glaucoma detection.

III. PROPOSED EXPERIMENTAL APPROACH

A. IMAGE DATABASE

The experiment performed in this work makes use of four publicly available databases. The databases consist of fundus images and their corresponding segmented optic discs and optic cups for model training and testing. The databases are RIM-ONE v2 [45] , RIM-ONE v3 [45], DRIONS [62] and DRISHTI-GS [43], [44].

The RIM-ONE database was exclusively developed to focus on optic nerve head segmentation. The fundus images are of high resolution and were captured using a Nidek AFC-210 fundus camera. The camera has a body of Canon EOS 5D Mark II and has a resolution of 21.1 megapixels. The version 2 of the database (RIM-ONE v2) has 455 images including 318 training images and 137 testing images. However, the version has only segmented optic discs ground truths and no optic cups ground truths. The version 3 (RIM-ONE v3) has 159 images including 127 images for model training and 32 images for testing. The ground-truth images were provided by two ophthalmologists.

The DRIONS database consists of 110 fundus images. The fundus images belong to subjects with glaucoma and eye hypertension diseases. The images were selected from an eye database that belongs to the Ophthalmology Service at Miguel Servet Hospital, Spain.

The DRISHTI-GS database includes 50 fundus images. The images are of high- resolution with a dimension of 2896×1944. The ground-truth images were provided by 4 experts. The database consists of both the cup and disc ground-truths.

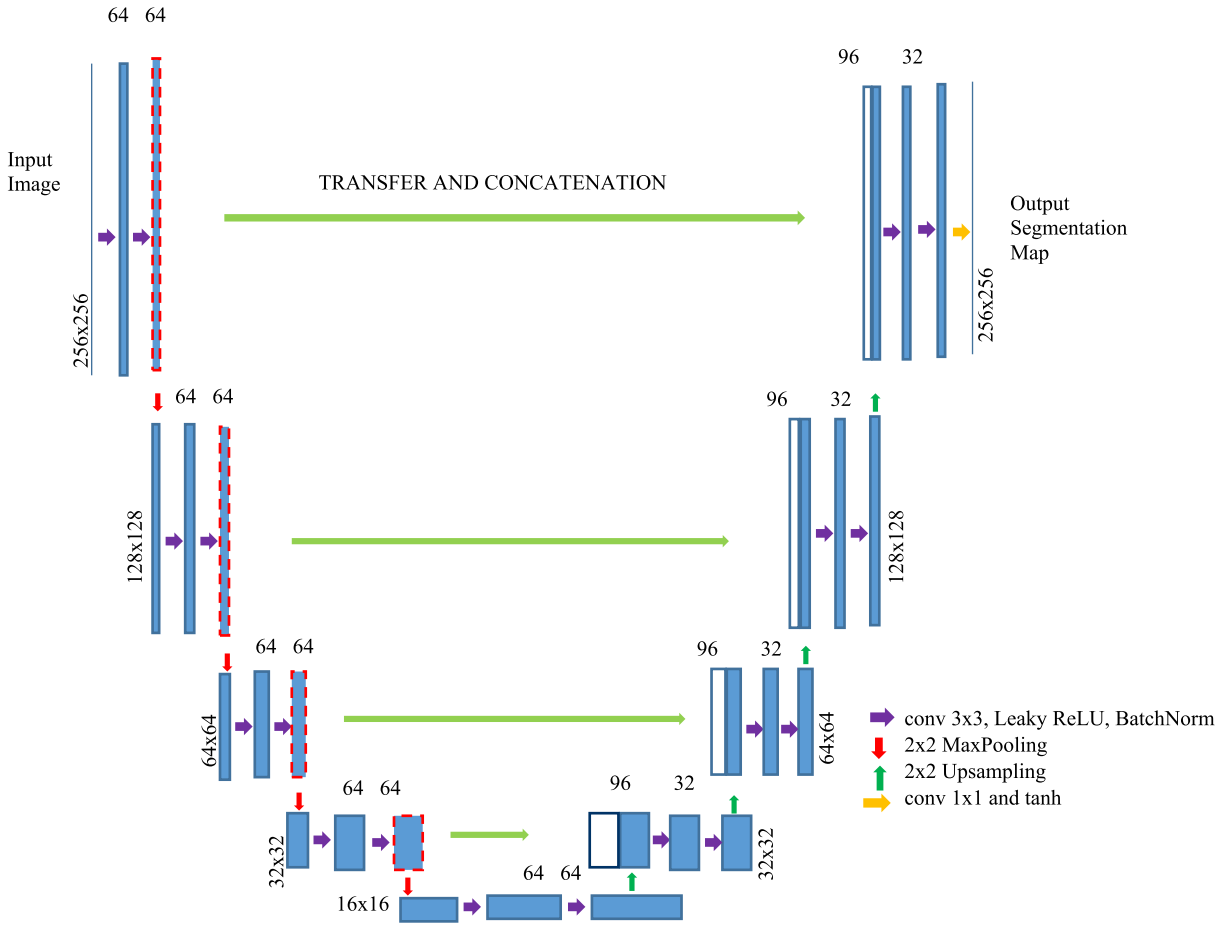


FIGURE 2. The architecture of the proposed U-Net lite model.

B. NETWORK ARCHITECTURE

The technique adopted in this research is a combination of two phases. The first phase consists of a segmentation process and the second phase is a detection process. The segmentation process is done using a U-Net lite model while the detection process is built using an extreme gradient boost (XGB).

The original U-Net model [63] is a convolutional network that has been widely utilized for biomedical image segmentation. It was conceived as an improvement over the Fully Convolutional Network [39]. The network has two layers: the down-sampling encoding layer and the up-sampling decoding layer. The encoding layer is made of two batches of 3×3 convolutional layers connected to an activation layer. The activation layer (rectified linear unit ReLU) is followed by a 2×2 max-pooling layer. This configuration is then repeated in successions. The decoding layer concatenates the up-sampled feature maps with the output of the encoding layer. The upsampling was done using 2×2 convolutional layers. Although the architecture has been widely utilized [64]–[67], it is a cumbersome model with lots of parameters to be trained.

In this work, the segmentation process was done using the proposed U-Net lite model. The model architecture is shown in Fig.2.

The proposed U-Net lite model has more convolutional layers and is designed to have the same size of filters (i.e. 3×3) in both the downsampling encoding layer and the upsampling decoding layer which is a major difference when compared to the original U-Net. The kernels are initialized to the 'glorot uniform' and the bias of the kernels are initialized to the 'he-normal'. The output layer of the proposed model has a filter size of 1×1. The architecture of the proposed model has 40 times fewer parameters than the original U-Net. The original U-Net model has about 3.1×107 parameters while the proposed model has 7.8×105 parameters. Our trial showed that models with huge training parameters tend to quickly over-fit. Each layer of the U-Net model is batch normalized and this helps to bring the average activation of the layers closer to zero [68].

The Leaky ReLU activation is utilized because it does not saturate quickly and helps the model to converge faster [69].

The output of the proposed model was connected to a 'tanh' activation layer.

The proposed U-Net model is different from other U-Net models in the configuration of its encoding and decoding layers. The widths of the convolutions are greatly reduced, and this process reduces over-fitting. The kernels were also carefully initialized, and this helps the training process to be faster. By using batch normalization with no drop-out, we improved the performance of the model greatly. Furthermore, the use of 'tanh' instead of the traditional 'sigmoid' at the output of the model improved the rate of convergence of the model.

The detection process includes an XGB classifier trained with extracted features from the segmented discs and cups. To the best of our knowledge, the proposed pipeline has never been used for a glaucoma detection process. The extracted features were normalized before feeding them into the classifier.

C. SYSTEM WORKING PROCEDURE

The proposed system pipeline is shown in Fig.3.

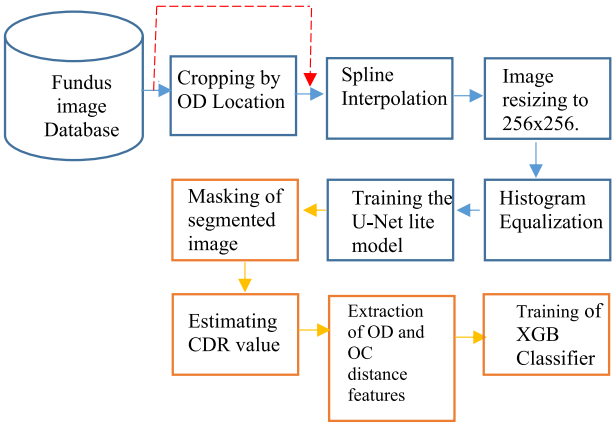


FIGURE 3. CDR value estimation and training of classifiers.

In pursuance of an accurate OC segmentation, the fundus images are cropped based on the location of the OD (the OD location was acquired from the OD segmentation process). This is done to accentuate the boundary of the OC. The cropped fundus images are scaled down using spline interpolation of the binomial order and resized to 256 x 256 pixels. The resizing is necessary to enhance the training speed and allow for more images per batch while training. Prior to passing the re-sized fundus images into the U-Net lite model, the contrast of the images is further refined by stretching out the most frequent intensity values in the images. This process enhances the training of the model and allows it to learn better. The scikit-image histogram-equalization is utilized for this process.

The OD segmentation process is like that of the OC except that there is no cropping of the fundus image (as shown by the dotted red jumper arrow in Fig.3). The proposed segmentation process is further described by the following algorithm.

1. Cropping of the fundus images based on the location of the OD. This procedure is needed only for the OC segmentation and not needed for the OD segmentation.
2. Applying spline interpolation to the RGB fundus images using the binomial order and nearest mode of filling.
3. Resizing the images to 256 × 256 pixels.
4. Applying histogram equalization to the images
5. Rescaling of images. All values of images are set to be between 1 and 0.
6. Training the proposed model with the scaled images.

The outputs of the segmentation process (i.e. segmented optic cups and discs) are further post-processed to detect glaucoma. The post-processing steps are described as:

Step 1: The segmented optic cups and optic discs are masked at 90°. These are the vertical cup and disc features.

Step 2: The maximum values of the non-zero cup and disc features are extracted.

Step 3: The vertical CDR values are acquired by dividing the vertical optic cup length by the vertical optic disc length as shown in equation 2

$$CDR = (Vertical\ cup\ length)/(vertical\ disc\ length \quad (1)$$

Step 4: Further extraction of the optic cup and disc features from the segmented cups and discs. The extracted features consist of the vertical separation between the cup and disc estimated at a minimum of 18° interval. This is done to catch the expansion in cup size and the minuscule loss of optic nerves along the optic cup fringe. A total of ten (10) vertical separations (labelled T0-T9) is acquired. This is displayed in Fig.4. The decision to use 10 vertical separations is taken after different numbers of vertical separations are tested. The numbers of vertical separations tested are 5,10,15 and 20. The T values of 5, 10,15 and 20 vertical separations for a fundus image are shown and discussed in the result section.

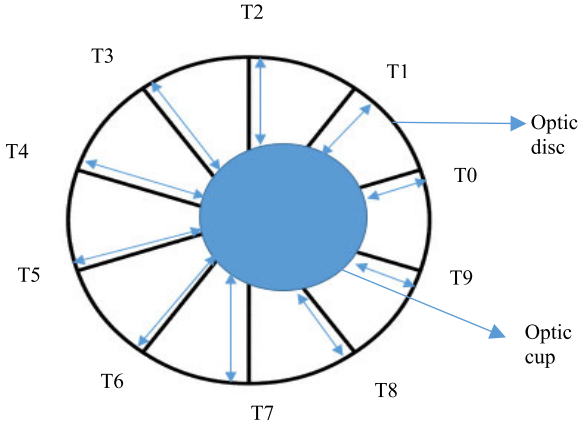


FIGURE 4. Distance between the optic cup and the optic disc evaluated along the optic cup's peripheral.

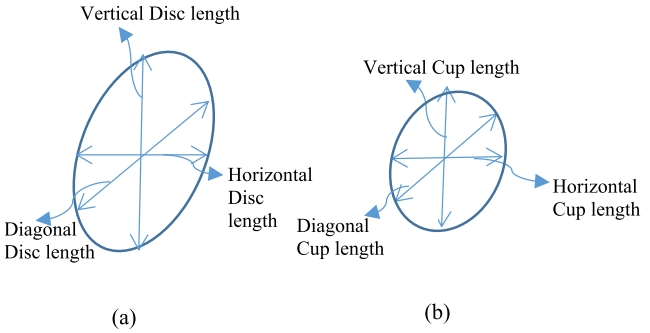


FIGURE 5. (a) The horizontal, diagonal and vertical opti disc lengths. (b) The horizontal, diagonal and vertical optic cup lengths.

Step 5: The vertical cup and its disc length, the horizontal cup, and its disc length, as well as the diagonal cup and its disc length, are acquired and utilized to train an XGB classifier.

The vertical, horizontal, and diagonal lengths are estimated as shown in Fig. 5 (a) and 5 (b) respectively.

D. MODEL TRAINING

The U-Net lite model was trained with the four databases discussed in section 2.1. After trying several gradient descent-based optimization algorithms [70], the stochastic gradient optimizer was used to compile the model. The model was compiled using a learning rate of $1e^{-2}$ for the optic disc segmentation and $1e^{-3}$ for the optic cup segmentation. The Nesterov was set to be true and momentum was set to be 0.95. The loss function utilized in (3) has the same value as the dice-score.

$$f(X, Y) = \frac{2\sum_{i,j}^{h,w} x_{i,j} y_{i,j}}{\sum_{i,j}^{h,w} x_{i,j}^2 + \sum_{i,j}^{h,w} y_{i,j}^2} \tag{2}$$

$$C(X, Y) = -log f(X, Y) \tag{3}$$

where the likelihood that the pixels predicted for the foreground is $X = (x_{i,j})$ and the given output is $Y = (y_{i,j})$, and h, w are the height and width respectively.

A comparative metric to dice-score is the IoU score. The IoU (5) is a metric utilized in many segmentation tasks to quantify the overlay that exists between the ground truth and the output of a model. As seen in (5), it quantifies the pixels present in both the ground truth and the model's output and divides the shared pixel by all the pixels in the ground truth and the model's output. Dice-score (4) is fundamentally the same as IoU, except that it awards more score to each correct pixel in the model's output by increasing the pixels shared by the ground truth and model's output by a factor of 2.

$$D(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|} \tag{4}$$

$$I(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \tag{5}$$

The model is trained over 65 epochs for both optic cup and optic disc segmentation. The model is trained using Kaggle's 2 CPU cores, 14 GB RAM, 1 NVIDIA Tesla K80 GPU. A batch size of 8 and an image size of 256 by 256 is utilized. No kind of data augmentation is used for the training process.

IV. RESULTS

This section evaluates the performance of the proposed segmentation model as well as the trained classifiers. Table 1 to Table 4 shows the average results of the proposed model when evaluated on the testing images in the databases. In Fig. 6 to Fig. 17, the fundus images acquired from the database are referred to as 'Database image', 'Model's segmentations' are the output of the proposed model and 'Ground truths' are the images available as ground-truths in the databases.

TABLE 1. Optic disc segmentation performance for RIM-ONE v2 database.

	RIM-ONE V2	
	Optic Disc	
	IoU Score	Dice Score
Proposed Method	0.91	0.97

TABLE 2. Optic disc and cup segmentation performance for RIM-ONE v3 database.

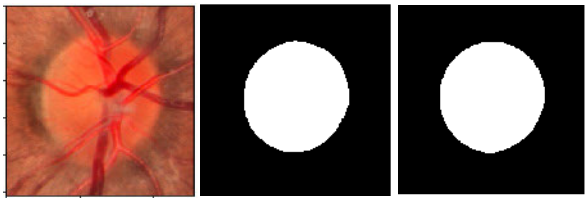
	RIM-ONE v3			
	Optic Disc		Optic cup	
	IoU score	*Dice-score*	*IoU score*	*Dice-score*
Sevastopolsky [71]	0.89	0.95	0.69	0.82
Zilly1 [46]	0.89	0.94	**0.80**	0.82
Maninis [38]	0.89	**0.96**	-	-
Al-Bander [72]	0.83	0.90	0.56	0.69
Proposed method	**0.90**	**0.96**	0.73	**0.86**

TABLE 3. Optic disc and cup segmentation performance for DRISHTI-GS database.

	DRISHTI-GS			
	Optic Disc		Optic cup	
	IoU score	*Dice-score*	*IoU score*	*Dice-score*
Sevastopolsky [71]	-	-	0.75	0.85
Zilly1 [46]	**0.91**	**0.97**	0.85	0.87
Zilly2 [16]	-	-	0.86	0.83
Al-Bander [72]	0.90	0.95	0.71	0.83
N.Thakur [51]	**0.91**	0.92	**0.87**	0.90
Proposed method	0.90	**0.97**	0.79	**0.95**

TABLE 4. Optic disc segmentation performance for DRIONS database.

	DRIONS	
	Optic Disc	
	IoU Score	Dice Score
Sevastopolsky [71]	0.89	0.94
Maninis [38]	0.88	**0.97**
Proposed method	**0.90**	0.96

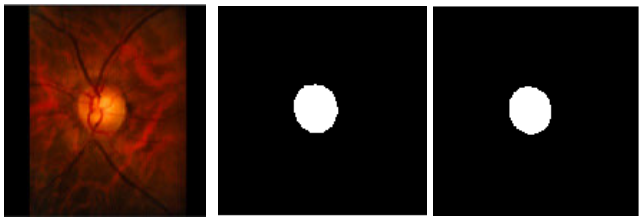


(i) (ii) (iii)

FIGURE 6. Model's best OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

(i) (ii) (iii)

FIGURE 7. Model's worst OD segmentation (i)Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.



(i) (ii) (iii)

FIGURE 8. Model's best OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

A. RIM-ONE DATABASE

The proposed model was tested on two versions of the RIM-ONE database: version 2 and version 3. However, RIM-ONE v2 database does not have ground-truths for optic cups. The average performance of the proposed model on RIM-ONE v2 database is shown in Table 1. The proposed model's best performance achieved a dice-score of 0.99 and an IoU score of 0.97. This is shown in Fig. 6. The worst performance achieved a dice-score of 0.81 and an IoU score of 0.65. The model's worst performance on the database is shown in Fig. 7.

(i) (ii) (iii)

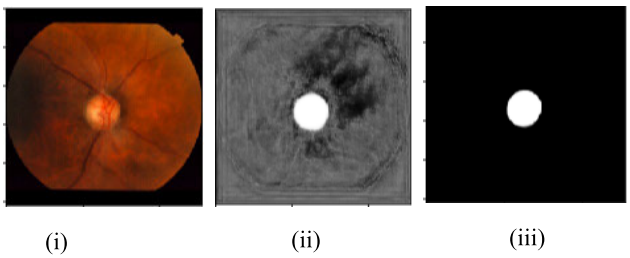
FIGURE 9. Model's worst OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

(i) (ii) (iii)

FIGURE 10. Model's best OC segmentation (i) Database image. (ii)Model's OC segmentation, (iii) Ground truth OC segmentation.

(i) (ii) (iii)

FIGURE 11. Model's worst OC segmentation (i) Database image. (ii)Model's OC segmentation, (iii) Ground truth OC segmentation.



(i) (ii) (iii)

FIGURE 12. Model's best OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

The performance of the proposed model on RIM-ONEv3 is shown in Table 2. We compared our result with that of Sevastopolsky [71], Zilly1 [46], Maninis [38] and Al-Bander [72] using the dice-score and IoU score as our assessment. The best and worst performance of the proposed model on the database is shown in Fig.8, Fig.9, Fig.10 and Fig.11 for both optic disc and optic cup segmentation.

The best optic disc segmentation as seen in Fig.8 has a dice-score of 0.99 and an IoU score of 0.96. The worst optic disc

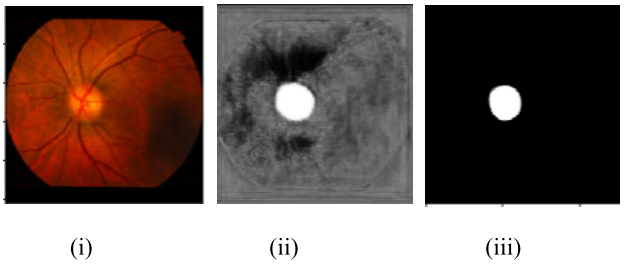


FIGURE 13. Model's worst OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

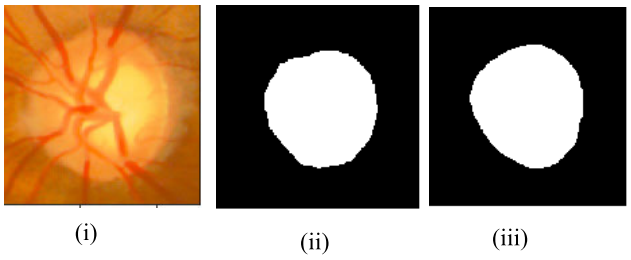


FIGURE 14. Model's best OC segmentation (i) Database image. (ii)Model's OC segmentation, (iii) Ground truth OC segmentation.

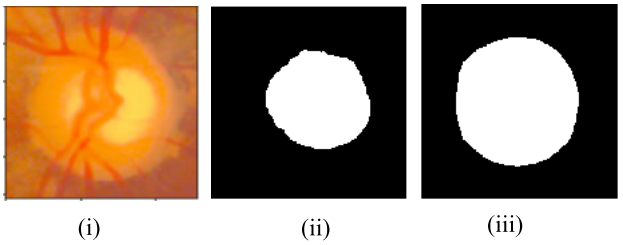


FIGURE 15. Model's worst OC segmentation (i) Database image. (ii)Model's OC segmentation, (iii) Ground truth OC segmentation.

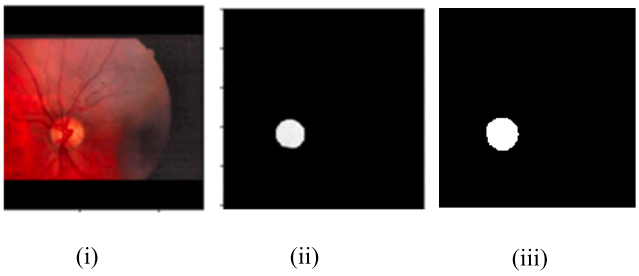


FIGURE 16. Model's best OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

segmentation as seen in Fig.9 has a dice-score of 0.89 and an IoU score of 0.77.

The best optic cup segmentation as seen in Fig.10 has a dice-score of 0.98 and an IoU score of 0.91. The worst optic cup segmentation as seen in Fig.11 has a dice-score of 0.30 and an IoU score of 0.15.

B. DRISHTI-GS DATABASE

The proposed model was tested on the DRISHTI-GS database. The performance of the proposed model in this database is shown in Table 3. For the optic disc segmentation, the best performance of the proposed model has a dice-score of 0.99 and an IoU score of 0.95. The best optic disc performance of the proposed model is shown in Fig. 12. The worst optic disc performance has a dice-score of 0.95 and an IoU score of 0.84. The worst optic disc performance is shown in Fig. 13.

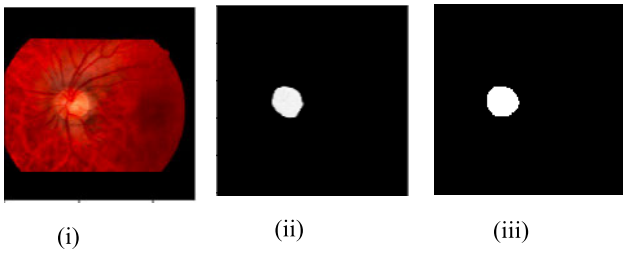


FIGURE 17. Model's worst OD segmentation (i) Database image. (ii)Model's OD segmentation, (iii) Ground truth OD segmentation.

In Table 3, it will be seen that the method used by Thakur *et al.* has high IoU scores. However, the major drawback of the method is that the parameters of the model must be set by the user and the values of those parameters vary across different databases. Hence, the parameters cannot be generalized for all databases.

For the optic cup segmentation, the best performance of the proposed model has a dice-score of 0.99 and an IoU score of 0.89. The best optic cup segmentation performance of the proposed model is shown in Fig. 14. The worst optic cup segmentation has a dice-score of 0.92 and an IoU score of 0.61. The worst optic cup performance is shown in Fig.15.

C. DRIONS DATABASE

The model was tested on the DRIONS database. The performance of the model in this database is shown in Table 4. The DRIONS database has ground-truths only for the optic discs. The best performance of the proposed model for the optic disc segmentation has a dice sore of 0.99 and an IoU score of 0.94. The best optic disc performance of the proposed model is shown in Fig. 16. The worst optic disc segmentation performance has a dice-score of 0.95 and an IoU score of 0.82. The worst optic disc performance of the proposed model is shown in Fig. 17.

It should be noted that the proposed model has a higher dice score than IoU score in all databases tested (Tables I-IV). This is because the dice metric gives more incentive to true positives detected while the IoU metric tend to penalize wrong classifications significantly. This means that the proposed model does generally well in the segmentation process but can be adversely affected by fundus images which have strong blood vessels occlusion as seen in Fig.7, Fig. 11, and Fig.15

D. CLASSIFIERS PERFORMANCE

The trained classifier is utilized to detect glaucoma from segmented optic discs and optic cups. A popular method of

detecting glaucoma from segmented optic disc and cup is the CDR (already discussed in section 1 and section 2.3). We compared the performance of CDR for glaucoma detection against some classifiers. The RIM-ONE v3 and the DRISHTI-GS databases are utilized for the glaucoma detection process because they both have the optic disc and optic cup ground-truths. Furthermore, out of the four databases in view, only the RIM-ONE v3 and the DRISHTI-GS databases have each fundus image labelled appropriately as 'glaucomatous' or 'non-glaucomatous'.

Table 5 shows the performance in detecting glaucoma of different CDR threshold values when tested on fundus images from the RIM-ONE v3 and the DRISHTI-GS databases. The metrics utilized for comparison are Precision, Accuracy, Recall, and Area under the Receiver Operating Characteristic Curve (AUC_ROC).

TABLE 5. Performance of different CDR threshold values on fundus images from the RIM-ONE v3 and DRISHTI-GS databases.

	CDR THRESHOLD VALUES				
	0.700	0.600	0.500	0.400	0.300
AUC_ROC	0.801	**0.874**	0.831	0.671	0.5471
ACCURACY	0.7976	**0.872**	0.832	0.6763	0.5547
RECALL	0.602	**0.795**	0.886	0.954	1.000
PRECISION	1.000	**0.945**	0.804	0.617	0.533

From Table 5, it is obvious that the goodness of the CDR method for glaucoma detection is substantially influenced by the CDR threshold value in use. If a low CDR threshold value like 0.300 is utilized, the detection process will have a high recall but a very low precision i.e. although all the glaucomatous samples are recognized, numerous non-glaucomatous samples are wrongly recognized as glaucomatous. The opposite is also valid for a very high CDR threshold value. This is seen in the CDR value of 0.700 (high precision, low recall). The AUC_ROC metric is a summary of the trade-off between recall and precision. Therefore, a better model will have a higher AUC_ROC value. The model proposed must have an AUC_ROC value that is higher than 0.874.

To train the classifiers, vertical separations are extracted from the optic disc and cup as discussed in section III. The effect of different vertical separations is tested on a fundus image. The numbers of vertical separations tested are 5,10, 15 and 20. The vertical separations are measured at an interval of 72°, 36°, 24° and 18° respectively. The result of this test is presented in Table VI.

From Table 6, taking just 5 vertical separations (measured at 72°) does not give an accurate picture of the varying distance between the optic cup and disc. Taking 15 and 20 vertical separations result in several duplications of the measured distance. This is because the intervals (24° and 18°) are small and no notable change occurred in the optic cup and disc. The optimum number of vertical separations is therefore 10. Although the result shown is just for a fundus image, the same phenomenon applies to all fundus images.

TABLE 6. The vertical distance between an optic cup and optic disc measured in pixels.

	NUMBER OF VERTICAL SEPARATIONS			
	5	10	15	20
VERTICAL DISTANCE BETWEEN OPTIC CUP AND OPTIC DISC MEASURED IN PIXELS	3.61	**4.00**	4.00	4.00
	5.39	**3.61**	4.00	4.00
	6.00	**3.16**	3.61	3.61
	7.28	**5.39**	3.62	3.61
	4.00	**5.00**	3.61	3.16
		6.00	5.39	3.16
		3.61	5.37	5.39
		7.28	5.39	5.39
		1.41	6.00	5.00
		4.00	6.00	5.00
			6.00	6.01
			7.28	6.00
			7.28	3.61
			7.27	3.61
			4.00	7.28
				7.28
				1.42
				1.41
				5.10
				4.00

An XGB classifier, a logistic regression, a support vector machine (SVM), a random forest classifier and a k-nearest neighbour (KNN) classifier were trained with all the obtained optic cup and optic disc attributes. This is done to assess the classifiers and afterwards pick the best performing classifier. The classifiers are evaluated using 5 folds cross-validation. The performance of each classifier is shown in Table 7.

Table 7 displays the results of the 5 classifiers tested on fundus images from both the RIM-ONEv3 and DRISHTI-GS databases. The XGB classifier has the highest AUC-ROC and precision average value and has an accuracy of up to 0.996 in one of the cross-validation sets. All the classifiers tested accomplish a higher AUC-ROC value than the CDR technique and the XGB classifier has the highest AUC-ROC. It can then be concluded that the XGB classifier has a superior classification capacity than the CDR technique.

We further compared our technique with other methods of detecting glaucoma. Other methods that have been used recently are:

a. GIST features extraction combined with the Support Vector Machine (SVM) classifier [73]
b. Combination of GIST and Pyramid Histogram of Oriented Gradients (PHOG). [73]
c. CDR combination with Super-Pixel identification technique. [74]

TABLE 7. Performance of the Classifiers on both the rim-one v3 and the drishti-gs database.

CLASSIFIER	CROSS-VALIDATION SET NUMBER	AUC-ROC	ACCURACY	PRECISION	RECALL
XGB	1	0.996	0.967	0.969	0.967
	2	0.971	0.931	0.931	0.931
	3	0.919	0.862	0.889	0.867
	4	0.964	0.893	0.895	0.893
	5	0.888	0.821	0.837	0.821
	6	0.878	0.821	0.837	0.821
	AVERAGE VALUE	**0.936**	**0.883**	**0.893**	0.883
LOGISTIC REGRESSION	1	0.927	0.867	0.867	0.867
	2	0.898	0.792	0.798	0.795
	3	0.938	0.930	0.938	0.933
	4	0.935	0.857	0.865	0.857
	5	0.912	0.893	0.912	0.892
	6	0.937	0.964	0.967	0.964
	AVERAGE VALUE	0.925	**0.883**	0.891	**0.885**
SVM	1	0.938	0.933	0.941	0.933
	2	0.890	0.793	0.850	0.800
	3	0.938	0.862	0.889	0.867
	4	0.908	0.786	0.850	0.786
	5	0.944	0.821	0.823	0.821
	6	0.959	0.929	0.938	0.929
	AVERAGE VALUE	0.930	0.854	0.881	0.856
RANDOM FOREST CLASSIFIER	1	0.993	0.967	0.969	0.967
	2	0.893	0.897	0.898	0.898
	3	0.879	0.793	0.798	0.795
	4	0.944	0.821	0.837	0.821
	5	0.781	0.750	0.762	0.750
	6	0.995	0.964	0.967	0.964
	AVERAGE VALUE	0.914	0.865	0.872	0.866
KNN	1	0.996	0.967	0.969	0.967
	2	0.888	0.828	0.841	0.831
	3	0.976	0.897	0.898	0.898
	4	0.875	0.893	0.912	0.893
	5	0.929	0.893	0.895	0.893
	6	0.747	0.750	0.762	0.750
	AVERAGE VALUE	0.902	0.871	0.879	0.872

d. Wavelet Energy features extraction and the application of z-score. [75]
e. GIST features combined with radon [76].

Table 8 shows the performance of the XGB classifier against the methods above when tested on the DRISHTI-GS database

V. DISCUSSIONS

This work proposes a technique for detecting glaucoma. The first phase of the technique includes a segmentation process. The segmentation process is done using U-Net lite model. The benefits of our model include the following: the proposed model architecture has fewer parameters to be trained. The original U-Net architecture has about 3.1×10^7 parameters while the proposed U-Net lite has about 7.8×10^5 which is about $40\times$ less the size of the original. The modified architecture, therefore, has fewer parameters to be trained.

TABLE 8. Modified unet +xgb compared with other methods when tested on the drishti-gs database.

	METHOD USED	AUC-ROC	ACCURACY	CLASSIFIER USED
N.Gour et.al. [73]	GIST	0.670	0.701	SVM
N.Gour et.al. [73]	PHOG	0.720	0.753	SVM
N.Gour et.al. [73]	GIST and PHOG	0.860	0.792	SVM
J.Cheng *et.al.* [74]	CDR and Super-pixel features	0.710	0.713	SVM
U. Raghavendra *et.al.* [75]	GIST and Radon	0.77	0.763	SVM
S.Dua *et.al.* [76]	Wavelet feature extraction	0.710	0.752	SVM
Orlando *et.al.* [77]	OverFeat + VGG-S	0.760	-	Logistic regression
Proposed method	Modified U-Net and Feature Extraction	**0.885**	**0.860**	XGB

The proposed model also requires less training epochs. For instance, the original U-Net model requires about ten (10) hours to train on a Nividia Titan GPU [40]. Also, Sevastopolsky [48] trained his model on the RIM-ONE v3 database for about 382 epochs (this is about 2.8 hours of training) using the platform provided by Amazon web service, Zilly *et al*. [46] trained their model for about 55 minutes and Maninis *et al.* [38] trained their model for 200 epochs and 3.1 hours. Our proposed model was trained on the same database for 65 epochs and 32.5 minutes using Kaggle's 2 CPU cores, 14 GB RAM, 1 NVIDIA Tesla K80 GPU. Our proposed model was trained for 100 epochs, 45 minutes on both the DRISHTI-GS and the DRIONS database. The much less training time of our proposed model translates to a cheaper cost of model training.

The second phase of the proposed pipeline deals with the glaucoma detection process. Features extracted from the fundus images were used to train an XGB classifier, SVM, logistic regression, KNN and a random forest classifier. The XGB classifier has a higher AUC when compared with the other classifiers. Furthermore, the number of features extracted from the fundus images was varied and the effect studied. It was found out that extracting 5 features (or vertical separations) did not give a full view of the changing geometry of the optic cup and disc and extracting more than ten (10) features only resulted into duplication of data. Literatures studied show that an XGB classifier has never been used for a glaucoma detection process. The use of a trained XGB classifier replaces the use of the traditional CDR method for glaucoma detection from the segmented optic disc and optic cup. As discussed earlier (section 1), the traditional CDR method is very subjective, and the threshold value utilized depends on the author. The CDR threshold value that has been chosen by different authors ranges from 0.3 to 0.6. The proposed framework achieved higher accuracy and AUC-ROC

when compared with other methods of glaucoma detection (other methods of glaucoma detection include methods that use the CDR threshold technique and methods that do not use the CDR threshold technique).

VI. LIMITATION OF STUDY

In as much as the model achieves state-of-the-art results in the segmentation process, it is still affected by the poor image quality. This is truer about the optic cup segmentation. In some cases, the optic cups are extremely difficult to identify (e.g. Fig.10) in the ground-truth images and this makes the model to output a very loose approximate of the optic cup location. Also, the presence of other ocular diseases such as diabetic retinopathy is not accounted for in the glaucoma detection process. Hence, optic discs and cups labelled as 'normal' might be influenced by other visual sicknesses asides glaucoma. The impact of which is not measured in this study.

VII. CONCLUSION

In this work, we developed a glaucoma detection model that includes two stages. The first stage includes a segmentation process, and the second stage includes a glaucoma detection process. The proposed method successfully achieved the following: (1) A segmentation model that consists of a modified U-Net model which has 40x less parameters than the original U-Net model. This makes the training of the proposed model to be fast and cost-effective. (1) For the optic disc segmentation, the model achieves an IoU score of 0.97 and a dice-score of 0.97 on the RIM-ONE v2 database, an IoU score of 0.90 and a dice-score of 0.96 on the RIM-ONE v3 database, an IoU score of 0.90 and a dice-score of 0.97 on the DRISHTI-GS database, an IoU score of 0.90 and a dice-score of 0.96 on the DRIONS database. (3). The proposed architecture achieves an AUC-ROC score of 0.936, an accuracy of 0.883, a precision of 0.893 and a recall of 0.883 when used to detect glaucoma on both the RIM-ONE v3 and the DRISHTI-GS database.

In summary, the proposed method offers very light architecture and achieves desired results in both the segmentation process and the glaucoma detection process.

FUTURE WORK

The study will be done by utilizing more openly accessible databases. This will help the model to train better. Furthermore, more glaucoma detection techniques such as the ISNT ratio and area covered by blood vessels will be utilized.

REFERENCES

[1] J. Moon, K. H. Park, D. M. Kim, and S. H. Kim, "Factors affecting ISNT rule satisfaction in normal and glaucomatous eyes," *Korean J. Ophthalmol.*, vol. 32, no. 1, pp. 38–44, Jan. 2018, doi: 10.3341/kjo.2017.0031.

[2] K. Qiu, G. Wang, X. Lu, R. Zhang, L. Sun, and M. Zhang, "Application of the ISNT rules on retinal nerve fibre layer thickness and neuroretinal rim area in healthy myopic eyes," *Acta Ophthalmol.*, vol. 96, no. 2, pp. 161–167, Mar. 2018.

[3] A. Giangiacomo and A. L. Coleman, "The epidemiology of glaucoma," in *Glaucoma*. Berlin, Germany: Springer, 2009, pp. 13–21.

[4] P. Mitchell, W. Smith, K. Attebo, and P. R. Healey, "Prevalence of open-angle glaucoma in Australia. The Blue Mountains eye study," *Ophthalmology*, vol. 103, no. 10, pp. 1661–1669, 1996. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/8874440

[5] G. Michelson, J. Hornegger, S. Wärntges, and B. Lausen, "The papilla as screening parameter for early diagnosis of glaucoma," *Deutsches Aerzteblatt Int.*, vol. 105, nos. 34-35, pp. 583–589, Aug. 2008. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19471619

[6] P. J. Foster, F. T. Oen, D. Machin, T. P. Ng, J. G. Devereux, G. J. Johnson, P. T. Khaw, and S. K. Seah, "The prevalence of glaucoma in Chinese residents of Singapore: A cross-sectional population survey of the Tanjong Pagar district," *Arch. Ophthalmol.*, vol. 118, no. 8, pp. 1105–1111, 2000, doi: 10.1001/archopht.118.8.1105.

[7] M. Dirani, J. G. Crowston, P. S. Taylor, P. T. Moore, S. Rogers, M. L. Pezzullo, J. E. Keeffe, and H. R. Taylor, "Economic impact of primary open-angle glaucoma in Australia," *Clin. Exp. Ophthalmol.*, vol. 39, no. 7, pp. 623–632, Sep. 2011. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1442-9071.2011.02530.x

[8] A. A. Almazroa, S. Alodhayb, E. Osman, E. Ramadan, M. Hummadi, M. Dlaim, M. Alkatee, K. Raahemifar, and V. Lakshminarayanan, "Retinal fundus images for glaucoma analysis: The RIGA dataset," *Proc. SPIE*, vol. 6, Mar. 2018, Art. no. 105790.

[9] J. B. Jonas and A. Dichtl, "Optic disc morphology in myopic primary open-angle glaucoma," *Graefe's Arch. Clin. Exp. Ophthalmol.*, vol. 235, no. 10, pp. 627–633, Oct. 1997. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/9349946

[10] J. B. Jonas, G. C. Gusek, and G. O. Naumann, "Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes," *Invest. Ophthalmol. Vis. Sci.*, vol. 29, no. 7, pp. 1151–1158, 1988. [Online]. Available: http://www.iovs.org/cgi/content/abstract/29/7/1151

[11] C. Burana-Anusorn, W. Knogprawechnon, T. Kondo, S. Sintuwong, and K. Tungpimolrut, "Image processing technique for glaucoma detection using the cup-to-disc ratio," *Thammasat Int. J. Sci. Technol.*, vol. 19, no. 1, pp. 22–34, Mar. 2013.

[12] A. Poshtyar, J. Shanbehzadeh, and H. Ahmadieh, "Automatic measurement of cup to disc ratio for diagnosis of glaucoma on retinal fundus images," in *Proc. 6th Int. Conf. Biomed. Eng. Informat.*, Dec. 2013, pp. 24–27. [Online]. Available: https://ieeexplore.ieee.org/document/6746900, doi: 10.1109/BMEI.2013.6746900.

[13] H. Ahmad, A. Yamin, A. Shakeel, S. O. Gillani, and U. Ansari, "Detection of glaucoma using retinal fundus images," in *Proc. Int. Conf. Robot. Emerg. Allied Technol. Eng.*, Apr. 2014, pp. 321–324.

[14] L. Shyam and G. S. Kumar, "Blood vessel segmentation in fundus images and detection of glaucoma," in *Proc. Int. Conf. Commun. Syst. Netw.*, Jul. 2016, pp. 34–38. [Online]. Available: https://ieeexplore.ieee.org/document/7823982, doi: 10.1109/CSN.2016.7823982.

[15] R. Panda, N. B. Puhan, A. Rao, D. Padhy, and G. Panda, "Recurrent neural network based retinal nerve fiber layer defect detection in early glaucoma," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 692–695.

[16] J. G. Zilly, J. M. Buhmann, and D. Mahapatra, "Boosting convolutional filters with entropy sampling for optic cup and disc image segmentation from fundus images," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, Oct. 2015, pp. 136–143.

[17] J. S. Park, H. S. Cho, and J. I. Cho, "Automated extraction of optic disc regions from fundus images for preperimetric glaucoma diagnosis," in *Proc. 17th Int. Conf. Control, Automat. Syst. (ICCAS)*, Oct. 2017, pp. 1107–1110.

[18] A. Karimov, A. Razumov, R. Manbatchurina, K. Simonova, I. Donets, A. Vlasova, Y. Khramtsova, and K. Ushenin, "Comparison of UNet, ENet, and BoxENet for segmentation of mast cells in scans of histological slices," in *Proc. Int. Multi-Conf. Eng., Comput. Inf. Sci. (SIBIRCON)*, Oct. 2019, pp. 544–547. [Online]. Available: https://ieeexplore.ieee.org/document/8958121

[19] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: A deep neural network architecture for real-time semantic segmentation," 2016, *arXiv:1606.02147*. [Online]. Available: http://arxiv.org/abs/1606.02147

[20] Z. Luo, Y. Zhang, L. Zhou, B. Zhang, J. Luo, and H. Wu, "Micro-vessel image segmentation based on the AD-UNet model," *IEEE Access*, vol. 7, pp. 143402–143411, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8859307

[21] A. O. Joshua, F. V. Nelwamondo, and G. Mabuza-Hocquet, "Blood vessel segmentation from fundus images using modified U-net convolutional neural network," *J. Image Graph.*, vol. 8, no. 1, pp. 21–25, 2020.

[22] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images," *IEEE Access*, vol. 7, pp. 21420–21428, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8632907

[23] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "CellProfiler: Image analysis software for identifying and quantifying cell phenotypes," *Genome Biol.*, vol. 7, no. 10, 2006, Art. no. R100. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/17076895

[24] F. Dong, H. Irshad, E.-Y. Oh, M. F. Lerwill, E. F. Brachtel, N. C. Jones, N. W. Knoblauch, L. Montaser-Kouhsari, N. B. Johnson, L. K. F. Rao, B. Faulkner-Jones, D. C. Wilbur, S. J. Schnitt, and A. H. Beck, "Computational pathology to discriminate benign from malignant intraductal proliferations of the breast," *PLoS ONE*, vol. 9, no. 12, Dec. 2014, Art. no. e114885. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25490766

[25] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, and A. Sethi, "A dataset and a technique for generalized nuclear segmentation for computational pathology," *IEEE Trans. Med. Imag.*, vol. 36, no. 7, pp. 1550–1560, Jul. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7872382

[26] M. Yahyatabar, P. Jouvet, and F. Cheriet, "Dense-Unet: A light model for lung fields segmentation in chest X-ray images," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 1242–1245.

[27] D. Luo and L. Shen, "Vessel-net: A vessel-aware ensemble network for retinopathy screening from fundus image," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 320–324.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[29] A. Rasmus, H. Valpola, M. Honkala, M. Berglund, and T. Raiko, "Semi-supervised learning with ladder networks," 2015, *arXiv:1507.02672*. [Online]. Available: http://arxiv.org/abs/1507.02672

[30] W. Zhang, X. Lu, Y. Gu, Y. Liu, X. Meng, and J. Li, "A robust iris segmentation scheme based on improved U-Net," *IEEE Access*, vol. 7, pp. 85082–85089, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8744291

[31] V. Zyuzin and T. Chumarnaya, "Comparison of Unet architectures for segmentation of the left ventricle endocardial border on two-dimensional ultrasound images," in *Proc. Ural Symp. Biomed. Eng., Radioelectron. Inf. Technol. (USBEREIT)*, Apr. 2019, pp. 110–113.

[32] S. Sivagami, P. Chitra, G. S. R. Kailash, and S. R. Muralidharan, "UNet architecture based dental panoramic image segmentation," in *Proc. Int. Conf. Wireless Commun. Signal Process. Netw. (WiSPNET)*, Aug. 2020, pp. 187–191.

[33] Y. Zhang, J. Wu, Y. Liu, Y. Chen, E. X. Wu, and X. Tang, "MI-UNet: Multi-inputs UNet incorporating brain parcellation for stroke lesion segmentation from T1-weighted magnetic resonance images," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 526–535, Feb. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9099078

[34] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: A full-scale connected UNet for medical image segmentation," 2020, *arXiv:2004.08790*. [Online]. Available: http://arxiv.org/abs/2004.08790

[35] L. Wang, X.-H. Ma, and Y. Ye, "Computer vision-based road crack detection using an improved I-UNet convolutional networks," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Aug. 2020, pp. 539–543.

[36] M. Alfarhan, M. Deriche, and A. Maalej, "Robust concurrent detection of salt domes and faults in seismic surveys using an improved UNet architecture," *IEEE Access*, early access, Dec. 10, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9290013, doi: 10.1109/ACCESS.2020.3043973.

[37] D. Lin, Y. Li, S. Prasad, T. L. Nwe, S. Dong, and Z. M. Oo, "CAM-UNET: Class activation MAP guided UNET with feedback refinement for defect segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 2131–2135.

[38] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," 2016, *arXiv:1609.01103*. [Online]. Available: http://arxiv.org/abs/1609.01103

[39] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7478072, doi: 10.1109/TPAMI.2016.2572683.

[40] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556

[41] R. C. Gonzalez, S. L. Eddins, and R. E. Woods, *Digital Image Processing using MATLAB*, 2nd ed. New Delhi, India: McGraw-Hill, 2012.

[42] H. Doğan and O. Akay, "Using AdaBoost classifiers in a hierarchical framework for classifying surface images of marble slabs," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8814–8821, Dec. 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417410005233, doi: 10.1016/j.eswa.2010.06.019.

[43] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-GS: Retinal image dataset for optic nerve head(ONH) segmentation," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 53–56.

[44] J. Sivaswamy, S. R. Krishnadas, A. Chakravarty, G. D. Joshi, and A. U. S. Tabish, "A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis," *JSM Biomed. Imag. Data Papers*, vol. 2, no. 1, p. 1004, 2015.

[45] F. Fumero, S. Alayon, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation," in *Proc. 24th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2011, pp. 1–6.

[46] J. Zilly, J. M. Buhmann, and D. Mahapatra, "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation," *Comput. Med. Imag. Graph.*, vol. 55, pp. 28–41, Jan. 2017. [Online]. Available: https://www.clinicalkey.es/playcontent/1-s2.0-S0895611116300775

[47] M. Tabassum, T. M. Khan, M. Arsalan, S. S. Naqvi, M. Ahmed, H. A. Madni, and J. Mirza, "CDED-Net: Joint segmentation of optic disc and optic cup for glaucoma screening," *IEEE Access*, vol. 8, pp. 102733–102747, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9103492.nb

[48] Y. Jiang, L. Duan, J. Cheng, Z. Gu, H. Xia, H. Fu, C. Li, and J. Liu, "JointRCNN: A region-based convolutional neural network for optic disc and cup segmentation," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 2, pp. 335–343, Feb. 2020. [Online]. Available: https://search.datacite.org/works/10.1109/tbme.2019.2913211

[49] P. Qin, L. Wang, and H. Lv, "Optic disc and cup segmentation based on deep learning," in *Proc. IEEE 3rd Inf. Technol., Netw., Electron. Automat. Control Conf.*, Mar. 2019, pp. 1835–1840.

[50] S. Shah, N. Kasukurthi, and H. Pande, "Dynamic region proposal networks for semantic segmentation in automated glaucoma screening," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 578–582.

[51] N. Thakur and M. Juneja, "Optic disc and optic cup segmentation from retinal images using hybrid approach," *Expert Syst. Appl.*, vol. 127, pp. 308–322, Aug. 2019. [Online]. Available: https://search.datacite.org/works/10.1016/j.eswa.2019.03.009

[52] S. C. Patel and M. I. Patel, "Analysis of CDR of fundus images for glaucoma detection," in *Proc. 2nd Int. Conf. Trends Electron. Informat. (ICOEI)*, May 2018, p. 1071.

[53] R. Zhao, X. Chen, X. Liu, Z. Chen, F. Guo, and S. Li, "Direct cup-to-disc ratio estimation for glaucoma screening via semi-supervised learning," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 4, pp. 1104–1113, Apr. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8794624

[54] J. K. Virk, M. Singh, and M. Singh, "Cup-to-disk ratio (CDR) determination for glaucoma screening," in *Proc. 1st Int. Conf. Next Gener. Comput. Technol. (NGCT)*, Sep. 2015, pp. 504–507.

[55] N. A. Mohamed, M. A. Zulkifley, W. M. D. W. Zaki, and A. Hussain, "An automated glaucoma screening system using cup-to-disc ratio via simple linear iterative clustering superpixel approach," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101454. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809419300035, doi: 10.1016/j.bspc.2019.01.003.

[56] A. Mvoulana, R. Kachouri, and M. Akil, "Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images," *Comput. Med. Imag. Graph.*, vol. 77, Oct. 2019, Art. no. 101643. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S089561111930062X, doi: 10.1016/j.compmedimag.2019.101643.

[57] A. Murthi and M. Madheswaran, "Enhancement of optic cup to disc ratio detection in glaucoma diagnosis," in *Proc. Int. Conf. Comput. Commun. Informat.*, Jan. 2012, pp. 1–5.

[58] F. Khan, S. A. Khan, U. U. Yasin, I. U. Haq, and U. Qamar, "Detection of glaucoma using retinal fundus images," in *Proc. 6th Biomed. Eng. Int. Conf.*, Oct. 2013, pp. 1–5.
[59] M. Lotankar, K. Noronha, and J. Koti, "Detection of optic disc and cup from color retinal images for automated diagnosis of glaucoma," in *Proc. IEEE UP Sect. Conf. Electr. Comput. Electron. (UPCON)*, Dec. 2015, pp. 1–6.
[60] M. Roslin and S. Sumathi, "Glaucoma screening by the detection of blood vessels and optic cup to disc ratio," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2016, pp. 2210–2215.
[61] A. O. Joshua, G. Mabuza-Hocquet, and F. V. Nelwamondo, "Assessment of the cup-to-disc ratio method for glaucoma detection," in *Proc. Int. SAUPEC/RobMech/PRASA Conf.*, Jan. 2020, pp. 1–5.
[62] E. J. Carmona, M. Rincón, J. García-Feijoó, and J. M. Martínez-de-la-Casa, "Identification of the optic nerve head with genetic algorithms," *Artif. Intell. Med.*, vol. 43, no. 3, pp. 243–259, Jul. 2008.
[63] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015, *arXiv:1505.04597*. [Online]. Available: http://arxiv.org/abs/1505.04597
[64] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1597–1605, Jul. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8252743
[65] V. Iglovikov and A. Shvets, "TernausNet: U-Net with VGG11 encoder pre-trained on ImageNet for image segmentation," 2018, *arXiv:1801.05746*. [Online]. Available: http://arxiv.org/abs/1801.05746
[66] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.
[67] A. O. Joshua, F. V. Nelwamondo, and G. Mabuza-Hocquet, "Segmentation of optic cup and disc for diagnosis of glaucoma on retinal fundus images," in *Proc. Southern Afr. Univ. Power Eng. Conf./Robot. Mechtron./Pattern Recognit. Assoc. South Afr. (SAUPEC/RobMech/PRASA)*, Jan. 2019, pp. 183–187.
[68] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167
[69] A. L. Maas, A. Y. Hannun, and Y. A. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. 30th Int. Conf. Mach. Learn.*, 2013, p. 3. [Online]. Available: https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf
[70] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa, "A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks," in *Proc. Int. Conf. Comput. Techn., Electron. Mech. Syst. (CTEMS)*, Dec. 2018, pp. 92–99.
[71] A. Sevastopolsky, "Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network," *Pattern Recognit. Image Anal.*, vol. 27, no. 3, pp. 618–624, Jul. 2017. [Online]. Available: https://search.proquest.com/docview/1938627054, doi: 10.1134/S1054661817030269.
[72] B. Al-Bander, B. Williams, W. Al-Nuaimy, M. Al-Taee, H. Pratt, and Y. Zheng, "Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis," *Symmetry*, vol. 10, no. 4, p. 87, Mar. 2018. [Online]. Available: https://search.proquest.com/docview/2040897837
[73] N. Gour and P. Khanna, "Automated glaucoma detection using GIST and pyramid histogram of oriented gradients (PHOG) descriptors," *Pattern Recognit. Lett.*, vol. 137, pp. 3–11, Sep. 2020. [Online]. Available: https://search.datacite.org/works/10.1016/j.patrec.2019.04.004
[74] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N.-M. Tan, D. Tao, C.-Y. Cheng, T. Aung, and T. Y. Wong, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1019–1032, Jun. 2013. [Online]. Available: https://search.datacite.org/works/10.1109/tmi.2013.2247770
[75] U. Raghavendra, S. V. Bhandary, A. Gudigar, and U. R. Acharya, "Novel expert system for glaucoma identification using non-parametric spatial envelope energy spectrum with fundus images," *Biocybern. Biomed. Eng.*, vol. 38, no. 1, pp. 170–180, 2018, doi: 10.1016/j.bbe.2017.11.002.
[76] S. Dua, U. R. Acharya, P. Chowriappa, and S. V. Sree, "Wavelet-based energy features for glaucomatous image classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 80–87, Jan. 2012. [Online]. Available: https://search.datacite.org/works/10.1109/titb.2011.2176540
[77] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "Convolutional neural network transfer for automated glaucoma identification," in *Proc. 12th Int. Symp. Med. Inf. Process. Anal.*, Jan. 2017, Art. no. 101600.

OLUWATOBI JOSHUA AFOLABI received the master's degree in electrical and electronic engineering from the University of Johannesburg, South Africa, where he is currently pursuing the Ph.D. degree. He has published and presented his research work both locally and internationally. His current research interest includes deep learning techniques to detect glaucoma from fundus images.

GUGULETHU P. MABUZA-HOCQUET received the Ph.D. degree in electrical and electronic engineering from the University of Johannesburg, South Africa. She is currently employed as a Senior Biometrics Researcher with the Council for Scientific and Industrial Research (CSIR). She is a Supervisor for students enrolled for Masters and Ph.D. degrees and a Mentor for permanent employees within the organization. She has published and presented her research work both locally and internationally. Her current research focuses on studying the science behind the human eye and the iris by using computer vision and machine learning techniques to extract features for the automated authentication of children and adults' identity, and the possible detection of certain diseases.

FULUFHELO V. NELWAMONDO (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in electrical engineering (computational intelligence) from the University of the Witwatersrand, Johannesburg, South Africa. He was a Postdoctoral Fellow with the Graduate School of Arts and Sciences, Harvard University. He was also the Executive Director for the CSIR Modeling and the Digital Science Unit. He is currently an Electrical Engineer by training. He is a Senior Member of the Association of Computing Machinery (ACM), a member of the South African Institute of Electrical Engineers, a Visiting Professor of Electrical Engineering with the Institute of Intelligent Systems, University of Johannesburg. He is a registered Professional Engineer and the Executive Manager for the CSIR Cluster: Next Generation Enterprises and Institutions.

BABU SENA PAUL (Member, IEEE) received the B.Tech. and M.Tech. degrees in radiophysics and electronics from the University of Calcutta, India, in 1999 and 2003, respectively, and the Ph.D. degree from the Department of Electronics and Communication Engineering, Indian Institute of Technology Guwahati, India. He was with Philips India Ltd., from 1999 to 2000. From 2000 to 2002, he served as a Lecturer with the Department of Electronics and Communication Engineering, SMIT, India. He joined the University of Johannesburg, in 2010. He has served as the Head of the Department with the Department of Electrical and Electronic Engineering Technology, University of Johannesburg, from April 2015 to March 2018. He is currently serving as the Director of the Institute for Intelligent Systems, University of Johannesburg. He has successfully supervised several postgraduate students and Postdoctoral Research fellows. He has attended and published over 60 research papers in international and national conferences, symposiums, and peer-reviewed journals. He is a Life Member of IETE. He was awarded the IETE Research Fellowship.