International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Joint Runet++: A Joint Region-Based Unet++-Based Optic Disc and Cup Segmentation with Ensemble Generalization Loss for Glaucoma Disease Prediction

[*1]Jincy C Mathew, [2]V. Ilango, [3]Dr V Asha

Submitted: 08/12/2023 **Revised:** 19/01/2024 **Accepted:** 29/01/2024

Abstract: Glaucoma is a chronic eye disease that leads to irreversible vision loss. The cup-to-disc ratio (CDR) plays an important role in diagnosing and screening glaucoma. Early detection of glaucoma is key to preventing vision loss, but there is a lack of recognizable early symptoms. In this paper, the proposed joint region-based U net++ (Joint R U net++) is used to segment the optic disc (OD) and the optic cup (OC) at the same time. The proposed Joint RUnet++ contains an attention-driven serial Unet++-based feature extraction module, a disc detection network (DDN), and a cup detection network (CDN). This DDN and CDN network contains an ROI detection network that crops the input feature maps based on the coordinates of bounding boxes. The Deep Net Ensemble model consists of three CNNs: VGG19, DarkNet19, and EfficientNet-B1 trained on multiple projections of retinal fundus images. Integrate Hybrid Beetle Antenna Search and Genetic Algorithm (HBAS-GA) to detect optimized fusion weights from deep-net models using deep-ensemble generalization losses. The experiment was built on RIM-ONE DL, refuge and real datasets. A comparison is made between the experimental results and previous prediction models in terms of accuracy, specificity, sensitivity, IOU slice, IOU cup and ROC. The Rim One dl has an accuracy score of 0.981, the refuge dataset has an accuracy score of 0.986, and the real-time dataset has an accuracy score of 0.986.

Keywords: *Convolutional neural network (CNN), Joint region-based Unet++ (JointRUnet++), Respective regions of interest (ROI), Disc detection network (DDN), Cup detection network (CDN), Optic cup (OC), Optic disc (OD).*

1. Introduction

Glaucoma is a hereditary disease that affects the optic nerve (ON). Glaucoma also affects the OC and OD and causes blindness. Glaucoma is generally related to intraocular pressure (IOP). The main causes of this disease are migraines, obesity, family history, ethnicity, and high blood pressure. The nerve fibers are affected, leading to a deterioration of the retinal layer and increasing the OD and cup-to-disc ratio. High IOP results are damage to the optic nerve. ON deterioration occurs in the OD region known as the optic nerve head (ONH). Glaucoma causes the disc's boundary to enlarge and change colour from pink to pale [1, 3]. It's an irreversible eye disease that should be treated and detected early [2].

Glaucoma disease is an ocular that affects ON eyeball and the brain. The fluid pressure inside the eye is known as IOP, which is high, while the ON is weak. If the blood sugar level is doubled, there is a chance of glaucoma, leading to vision loss. The pressure pushes against ON and affects the nerve fibers. Affected nerve fibers result in damage to the retinal nerve fiber layer (RNFL) and lead to enlargement of the cup-to-disc ratio (CDR) and optic disc ratio (ONH), and OD. It also caused the retinal thinning and pigment epithelium retinal around the ON, called peripapillary atrophy (PPA). These are reduced by early identification of glaucoma [14, 15].

It is still challenging for computer-aided diagnostic (CADx) systems to detect glaucoma diseases [4]. Glaucoma detection in an early stage is challenging, but it can be cured [5]. Late detection causes permanent blindness, so glaucoma must be detected at an early stage. As per the previous assessment, a simple checkup should not detect glaucoma symptoms. A regular eye checkup may reveal glaucoma symptoms, which should be followed by further treatment [7]. The progression of glaucoma detection with existing functional and structural matric used in the hospital, like standard automated perimetry (SAP), and optical coherence tomography (OCT), is a challenge due to the lack of a true reference standard [6].

Several tests are performed to detect glaucoma, a function-based visual field test, including IOP measurement and ONH. The ONH assessment has proven to be the most appropriate and widely used among glaucoma experts. This calculation is usually done manually based on fundus images, and it is a very tedious procedure that requires

[1, 3]Department of MCA, New Horizon College of Engineering, Bengaluru India.
[2] Department of Computer Application, CMR Institute of Technology, Bengaluru, India.
[1, 2, 3]Visvesveraya Technological University, Belagavi -590018, India.
*[*1]jincycm26@gmail.com , [2]ilango.v@cmrit.ac.in, [3]asha.gurudath @gmail.com*

skilled supervision and is time-consuming. So the design of the automated glaucoma identification technique is more attention while making faster and more reliable decisions [8]. The automatic algorithm segmentation focuses on OD, and subsequent algorithms propose to increase screen accuracy while focusing on OC segmentation [9]. With the help of the fundus camera, fundus images are recorded in which retinal parts such as OD, artery, veins and macula are displayed. Glaucoma can be identified quickly and easily using fundus images [13].

The understanding disc images are integrally slanted since the broad range of the normal ON presence joins by pathological findings. Later, blindness is caused, glaucoma is detected, and ophthalmologists identify the clinical conditions using a variety of instruments and methods. Glaucoma detection, classification, and identification are very difficult; existing methods are costly and lengthy. To overcome these disadvantages, deep learning algorithms are used [10-12]. Nowadays, deep learning improves performance on several ophthalmic diseases finding issues from fundus images, and ophthalmologists use this method in decision-making [22]. Recently, deep learning methods have been increasingly used for OD and OD segmentation.

Several performance models are based on deep learning approaches [16]. The deep learning system elevated specificity and sensitivity for detecting diseases. Fundus vessel segmentation and deep learning methods work well. Deep learning for OD and OC segmentation from fundus images [17].

Nowadays, high-precision deep learning methods establish multiple images of medical detection tasks, such as diabetic macular edema, age-related macular degeneration and possible glaucoma. For diabetic retinopathy, several papers show that deep learning methods can be leveraged to achieve a highly identifiable image assessment for the retinal background [18]. For applications in medical ophthalmology, deep learning methods are applied to retinal fundus images. The computer-aided image segmentation approaches for fundus images [23].

This work mainly aims to improve the performance of the glaucoma disease prediction model using the ensemble learning method. The main objectives of the research works are listed as follows:

- Pre-processing techniques such as cropping, channel separation, and data augmentation are used on the input fundus images.
- After pre-processing, a JointRUnet++ is employed to simultaneously segment OD and OC based on the ROI.
- Three pre-trained CNNs: VGG19, DarkNet19 and EfficientNet-B1, are combined to form an ensemble Deep Net model with multiple projections of retinal fundus images.
- The proposed ensemble model is integrated with HBAS-GA to detect the optimized fusion weights of the ensemble DeepNet model based on the deep ensemble generalization loss.

The structure of this paper is as follows: Section 2 covers extensive past research on glaucoma detection. Section 3 proposed methodologies. Section 4 discusses the results. Section 5 discussed experimental evaluations. Section 6 concludes the paper.

2. Literature Survey

Taj et al. [17] proposed the glaucoma classification in two stages depending on Deep CNN architectures. In the first phase, 4 images net pre-trained deep CNN architectures such as NasNet- large, AlexNet, InceptionResNetV2, and Inception V3, are used and observed, with the NasNet-large architecture giving better performance based on accuracy, sensitivity, an area under receiver operating characteristics curve metrics, and specificity. The result shows that a NasNet-Largearchitecture is a possible option based on performance as a single classifier that will increase the performance of automatic glaucoma classification.

Phasuk et al. [18] proposed a glaucoma automated screening approach with retinal fundus images by fusing the results of different classification networks and feeding them as input to a simple ANN to get the final result. Public datasets such as DRISHTI-GS, RIM-ONE R3, and ORIGA-650 are used. This result showed that this network performed other glaucoma screening algorithms with an AUC of 0.94. Xu et al. [19] proposed extracting glaucoma features from fundus images and the report optical coherence tomography (OCT). The CDR in fundus images and the thickness of the optic nerve fiber layer (RNFLT) in OCT reports, the proposed method to extract is to extract common features of the shape and scale of the target curve. The shared feature and the prior information-based features comprehensively characterized the target curve. Next, based on the feature space of the CDR and RNFLT curve, an ensemble classification algorithm for the identification of glaucoma is proposed.

Bengani et al. [20] proposed a deep learning approach to automatically segment OD in fundus retinal images, using transfer learning and semi-supervised learning concepts. A convolutional autoencoder (CAE) is trained from Kaggle's diabetic retinopathy dataset to learn features from many unlabeled fundus images automatically. The results show that transfer and semi-supervised learning overcome barriers leading to a large labelled dataset. This

segmentation approach is an automatic retinal image in processing systems to identify eye diseases.

Latif et al. [21] proposed a two-phase novel network for glaucoma diagnosis (ODGNet) and OD localization. Initially, a visual projection map combined with a shallow CNN is used to localize the effective OD from fundus images. Furthermore, pre-trained approaches based on transfer learning are used to identify glaucoma. Several transfer approaches learning-based such as ResNet, VGGNet, and AlexNet combined saliency maps estimated using 5 public retinal datasets such as HRF, ORIGA, RIM-ONE, DR-HAGIS and DRIONS-DB distinguish among glaucomatous, and normal images. The results show that the ODGNet estimated on ORIGA for glaucoma identification is a better predictive model with good accuracy, AUC, sensitivity and specificity. These results show that the given approach to OD localization based on shallow CNN and projection map is robust and cheaper with high accuracy.

An ensemble- CNN classifier based on a spatially elliptic curve model (SBEFCM) was proposed by David et al. [22] for glaucoma detection and corresponding OC and OD limits. This study used the CNN ensemble classification to classify diabetic retinopathy and glaucoma. The boundary between OD and OC is detected by the new spatially weighted ellipse fitting model SBEFCM. As well as SBEFCM extends and improves multi-ellipse fitting. The CNN ensemble classifier is applied to local retinal glaucoma LAG and RIM-ONE datasets. To avoid blending with the surrounding tissue and blood vessels, the input image of the fundus is pre-processed. The CNN summary classification method identified various performance metrics for glaucoma detection, such as AUC, specificity, accuracy, and sensitivity values. SBEFCM accurately detected the prediction ratio of OC and OD boundaries. In comparison, the proposed CNN ensemble outperforms existing glaucoma detection methods.

Tulsani et al. [23] suggested that the evaluation of OD and OC is one of the most important steps in glaucoma diagnosis. This work develops a new method for glaucoma identification using a segmentation approach for OD and OC. An improvement on a small dataset was achieved using the Dhristi dataset. Hyper-parameter tuning is used in addition to the user-defined loss function to create UNE++ model segmentation. The small size optic nerve target and developed loss function solve the class imbalance. Liu et al. [24] proposed many existing techniques that do not achieve satisfactory segmentation performance due to the lack of pixel-level annotated data during training. This paper proposes a joint segmentation method for ODs and cups based on partially supervised conditional generative adversarial networks (GANs) to overcome this limitation. Segmentation networks, generators, and discriminators are built into our architecture to learn mappings between images and segments.

Li, Liu et al. [25] proposed that AGCNN is an attention-based CNN for glaucoma detection. This study comprises 11,760 fundus images categorized as negative or positive based on attention-based glaucoma (LAG). Through a simulated eye-tracking experiment, attention maps for 5,824 of the 11,760 fundus images are derived from ophthalmologists. Ultimately, an AG-CNN structure prediction with attention subnets, area pathological localization subnets, and classification glaucoma subnets was designed. The attention maps and attention prediction subnet are predicted to highlight glaucoma-relevant regions using weakly supervised training. To increase the performance of the AGCNN framework, in contrast to other attention-based CNN methods, features of localized pathological regions are also visualized. As a final result of testing our LAG database and another publicly available glaucoma database, the proposed AG-CNN approach significantly advances the state-of-the-art in glaucoma detection. Ganesh et al. [26] propose a GD-YNet deep-learning framework for glaucoma segmentation and detection. Combining the power of aggregate transformations with the YNet architecture enables binary contextual classification and OD segmentation for glaucoma detection.

Training time for the above glaucoma detection methods was very high, cup segmentation results were lower, overall system performance was lower, and time complexity was high. The proposed new automated ensemble model for glaucoma eye disease detection and classification based on deep feature learning using retinal fundus images is proposed to solve these problems. This work develops an ensemble-deep network for automated early-stage glaucoma prediction using retinal fundus images.

Limitations: These approaches do not consider the connection between the optic disc and the cup. The models assign the same weight for every pre-trained deep learning model to do the aggregation and small amount of available training data. The previous method was not designed with an end-to-end approach. Visual field tests are not available in all hospitals due to equipment limitations.

3. Proposed Methodologies

This work proposed a new glaucoma eye disease detection and classification model based on automated Ensemble Deep Feature Learning in retinal fundus images. The steps involved in the proposed model are data acquisition, pre-processing, OD and OC classification, and segmentation. Initially, retinal images are collected from a public data

source and a real dataset. Then, pre-processing techniques such as cropping, channel separation, and data augmentation are applied to the input fundus images. The original fundus image is too large to transfer directly to OD segmentation. Therefore, the dimensions reduce the original image; the relevant disc part is cropped. The optic disc portion in the red channel is observed most clearly in the RGB image. Therefore, the image is split into Red, Blue and Green channels, and only the red channel is used for disc segmentation, whereas for segmentation of the cup, the best results are obtained when all three RGB channels are used. Flipping and rotation are used as data expansion techniques to overcome the data imbalance problem. The images are first flipped horizontally to double the size of the training set. After that, all the images are flipped vertically, increasing the volume of the available data. Figure 1 shows a flowchart of U net++-based optical disk segmentation.

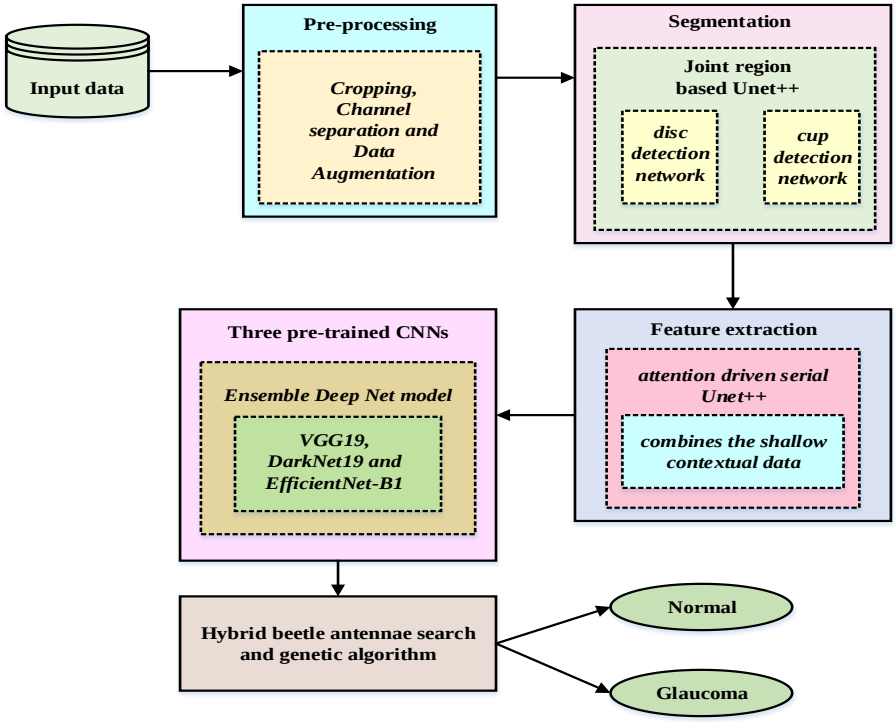

Fig 1: Flow diagram for a joint region-based U net++-based OD and cup segmentation.

After pre-processing, a joint-region-based Unet++ (JointRUnet++) simultaneously segments the optic disc and ocular cup based on the region of interest (ROI). The proposed JointRUnet++ includes a Unet++-based serial feature extraction module, a Disk Detection Network (DDN), and a Cup Detection Network (CDN). The proposed attention-based serial feature extraction module of Unet++ combines shallow context data into deep features and includes an attention unit to assign different weights to different spaces and channels. It is used to direct different levels of attention to disc/cup regions of different attention. Then the proposed DDN and CDN are segmented into OD and OC, respectively. These DDNs and CDNs contain an ROI detection network that crops the input feature maps based on bounding box coordinates. Finally, these trimmed feature maps are provided as input to various pre-trained deep learning models (VGG19, DarkNet19, and EfficientNet-B1) that discriminate between healthy and glaucomatous images. In this study, 3 pre-trained CNNs: VGG19, DarkNet19, and EfficientNet-B1, are combined into an ensemble deep-net model with multiple projections of retinal fundus images. The proposed ensemble model is integrated with HBAS-GA to detect optimized DeepNet model fusion weights based on deep ensemble generalization losses.

3.1 Pre-processing

Pre-processing techniques such as cropping, channel separation, and data enhancement are used to input fundus images. A picture is the removal of external parts by cropping to improve encircling, change the proportions of an angle, or emphasize a subject and resizing image results in a change in its physical size. A tile is divided into the portion of data that is lost due to cropping. The most noticeable slice was visible in the red channel of the RGB image. The disc was most clearly visible in the red channel of the RGB image. The image was split into red, blue and green channels, and for the cup segmentation, all three RGB channels were used to get the best results. The original dataset contained only a limited number of images that could be used as training data. A neural network could not be applied and trained with such an amount of data.

Flipping and data rotation has been used as augmentation techniques to overcome this problem. First, the images were flipped horizontally to double the size of the training set. Images are then flipped vertically, an additional increasing amount of data available. Flipped images were rotated from 0° to 180° in 20° increments after flipping. As a result, the model was effectively trained with a larger data set.

3.2. Joint region-based U net++ for segmentation

OC and OD segmentation task models for building custom UNET++ for optimal performance describe changes to the UNET++ model. In addition, the user-defined loss function was developed to account for the class imbalance that occurs due to the relatively small size of OD and OC and is compared to the entire picture. UNET++ [23] improves the segmentation process between the encoder and decoder by providing convolutional layers. DenseNet inspires the improvement of the U-NET architecture. Unlike the U-NET model, UNET++ has three additions: redesigned hop paths, dense hop connections, and deep monitoring. Figure 2 shows the redesigned jump path in green, reducing the semantic gap between encoders and decoders. Due to the connection hopping used in U-NET, semantically different functions between encoders and decoders are combined. By combining the output of the previous convolution layer with the appropriate dense block resampling output layer, similar feature maps enhance the semantic optimization performed by UNET++. In Figure 2, denser hop connections are inspired by dense blocks to improve DenseNet gradient flow. A dense skip connection ensures that all feature maps are collected and reach the correct node. Deep Surveillance is shown in red because its architecture includes U-NETs of varying depths and is not subject to network depth choices.

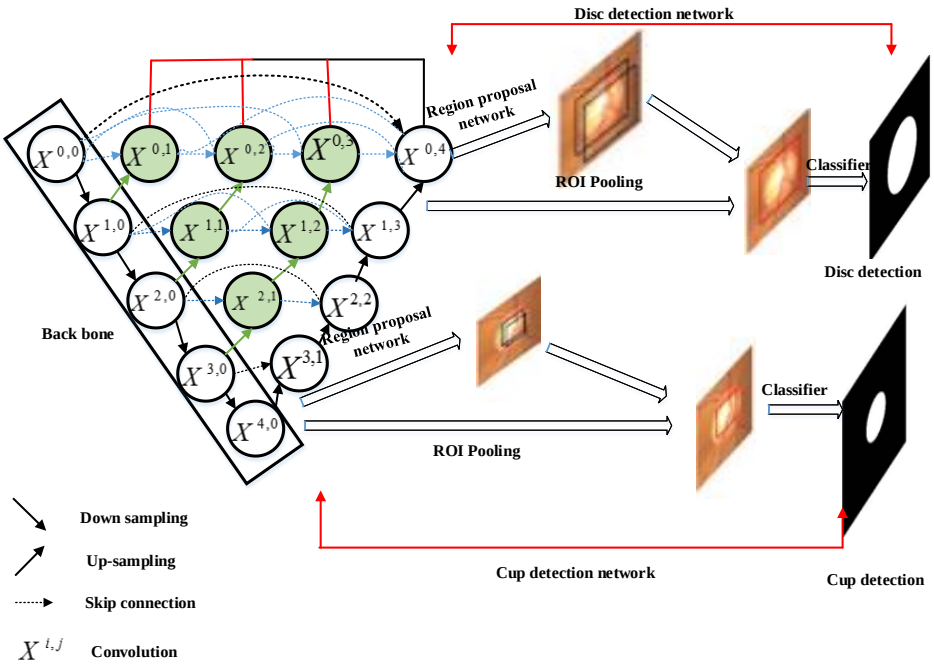

Fig 2: Illustrate U net++ with CDN and DDN methods.

Within the UNET++ architecture, all U-NETS decoders are partially connected to the encoder. In UNET++, due to the use of Deep Supervision, the benefit of the joint representation of the image is synchronously trained using its constituent U-NETs. UNET++ uses extensive monitoring to trim and adjust model complexity while maintaining speed and efficiency. Class imbalance is observed when the segmented pixels are a very small percentage of the total pixels. After pre-processing, the optical disc and optical coverage are segmented based on the ROI using Joint R U Net++ simultaneously.

3.2.1. Joint segmentation

There is a significant difference between OD segmentation and OC, as OC is located at the centre of the disc. An attention mechanism used a sectioned portion of the outer diameter to take advantage of this important preliminary information to guide OC localization. Complex training for both segmentation tasks within a single framework. Therefore, the results of OD and OC can influence each other positively. The paper aims to recognize the minimum outer diameter and the cup bounding boxes that segment them directly. Outside diameters and cup, boundaries are estimated using ellipses inscribed in the boxes. CDN and DDN were developed to segment optical discs and cups. In our method, a method based on object recognition is used.

Then, by directly segmenting the OD and OC, this study suggests detecting the minimum bounding box of the OD and OC. The computation of bounding box ellipses can be used to segment OD and OC. The framework called CDN and DDN is proposed, which involves sharing feature maps for different tasks (OD segmentation and OC segmentation) and uses attention mechanisms to improve performance. It produces the output of both OD and OC segmentation in one frame, and the results interact.

The feature extraction network is an input image. A neural network extracted signals to classify features. The image has certain structural elements, such as points, edges or objects. DDNs and CDNs contain ROI detection networks that clip input feature maps based on bounding box coordinates. Unet++'s attention-driven serial feature extraction includes a unit for assigning weights to different spaces and channels based on shallow context data.

3.2.2. Attention-driven serial Unet++ (ADSU) for feature extraction

In Figure 3, two Unet++ blocks are connected in series to form a pipeline. The fundus images are the input of this network. When the output of the shallow feature is reached, the first Unet++ block is passed to the second Unet++ block. It is also passed back to the chain with the output of the deep features for the second block. The second block to be concatenated with the output elements (the depth function) is returned. The residual connection and shallow contextual information can be used again. The serial Unet++ structure allows the original information of the shallow layers to be transmitted directly to subsequent deeper layers. The deeper layer can focus on residual learning and avoid model degradation.

Merge the concatenated features, apply an attention module with different concentrations to the fundus images, and reduce the images to three channels by applying two convolution layers. Finally, add the original eye background images to the extracted feature channels to get the images without the eye background [27].

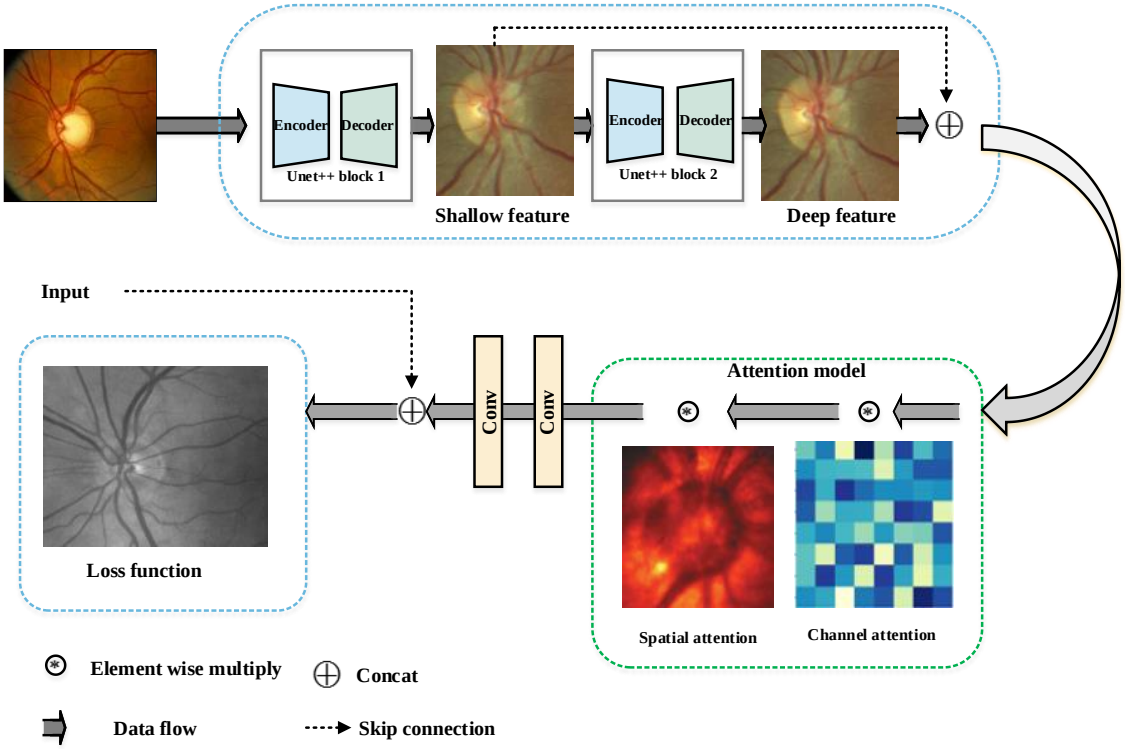

Fig 3: The whole structure of the ADSU net, including two U net++ blocks.

The feature extractor must fully use an image's information so that the fundus image can be removed and the image is as clear as possible due to several previous networks using encoder-decoder structures. The feature extractors successfully achieved excellent performance and built a serial Unet++ block using encoder-decoder structures as feature extractors. Specifically, use the Unet++ model, which adds more short links and skips routes to encourage contacting and merging information. Specifically, since the input patches were resized to 256 × 256 pixels, keep three layers of Unet++ to downscale the resolution to 1/8. Therefore, the definition of the r th feature at the q th layer is as follows:

$$x^{q,r} = g(x^{q,0} \oplus x^{q,1} \oplus ... \oplus x^{q,r-1} \oplus up(x^{q+1,r-1})) \quad (1)$$

Here $g(.)$ denotes the convolution layer, $up(.)$ denotes upsampling operation, and $\oplus$ denotes the concatenate operation.

3.2.3. Disc detection network (DDN) and cup detection network (CDN)

The suggested method contains two main parts: DDN and CDN. DDN is for OD segmentation, while CDN is for OC segmentation. OD and optical trays are generated by DDN and CDN, which have a similar design. The attention mechanism between the DDN and the CDN uses the disk attention model. It informs the CDN where the optical disk is located according to the output of the DDN. Fundus images detect OD and OC using the proposed DDN and CDN. To leverage the knowledge that the OC resides in the OD, design an end-to-end framework that recognizes the OD and the OC simultaneously, uses DDN and CDN, and integrates the prior knowledge with the disc.

3.2.4. Region of interest (ROI) poolin:

Unet++ generated several possible bounding boxes. ROI technique uses coordinates of candidate bounding boxes to crop small portions of feature maps. An attention mechanism is used in ROI, so there is no difference in any cropped object map size. ROI pooling includes the following procedures:

Image-level coordinates are used to determine candidate bounding boxes. So they have to be transformed into maps of objects by dividing them in a certain ratio.

The area to combine is separated into $k \times k$ equal blocks with k output size marked.

Each small block is subjected to the maximum pooling technique. As a result, pooled feature maps have a size of $k \times k$.

3.3. Ensemble Deep Net model

Using deep ensemble learning models, better generalization has a definitive model because deep learning and ensemble learning combine advantages. Compared to the performance of a single model, an ensemble can make more accurate predictions.

3.3.1. VGG-19

VGG-19 has 19 layers, 3 are fully connected, and 16 are convolutional layers. The parameters in deep networks can be reduced by using filters in convolutional layers. This architecture uses a 3x3 pixel filter. A maximum pooling layer is used, and there are two steps. There are approximately 138 million calculation parameters in VGG-19.

3.3.2. DarkNet19

Convolutional layers 19 and pooling layers 5 make up DarkNet-19 architecture. Activation maps are generated by circulating the input data through selected filters and convolutional layers. Filters are used to obtain features in activation maps. In the pooling layer, combine the output of neurons from 1 layer into a single neuron output to transfer the output to the next layer, reducing the data size. A similar configuration to the VGG models is used in the DarkNet-19 model, and its channels are duplicated. Another benefit of the DarkNet-19 model is its batch normalization, which speeds up convergence, stabilizes the training process, and makes model operations more efficient.

The DarkNet-19 model is selected for this study due to its state-of-the-art architecture and real-time object detection. Because of this, computer-based systems are playing a more active role in real-world disease detection. The second reason is that the dataset contains a resolution with images of 768x768 pixels, and the DarkNet-19 model accepts inputs of 256x256 pixels. Thus, DarkNet-19 can use the resolution of each image without losing any information.

3.3.3. EfficientNet-B1

Currently, the model includes an image of the fundus of a normal eye and an eye disease associated with glaucoma in the input layer. EfficientNet-B1 models are contained in the hidden layer. ImageNet databases are used to train EfficientNet models, one of the largest datasets for benchmarking image classification models. People have annotated 1,000 classes in this database. Pre-trained models can classify images into appropriate categories based on their rich features. This study used pre-trained EfficientNet models as transfer learning models that improve the classification of glaucomatous eye diseases. This type of learning approach for glaucoma classification using fundus images can reduce training time, increase convergence rates, and achieve optimal performance in detecting patient data samples of fundus images as either normal or glaucoma. Figure 4 shows the block diagram of an ensemble learning model by VGG19, dark net 19, and efficient net-B1.

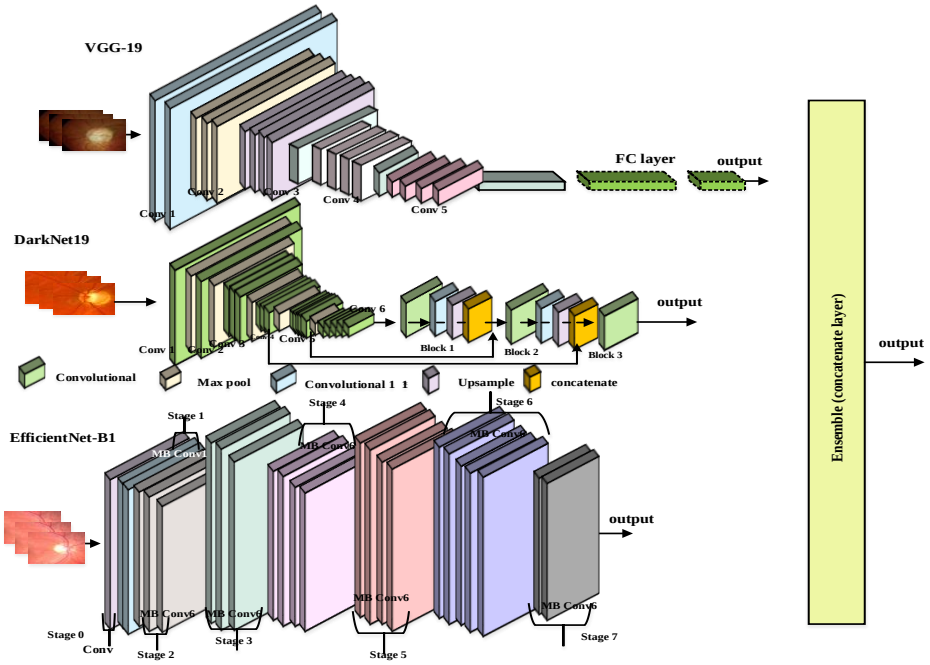

Fig 4: Block diagram of an ensemble learning model by VGG19, dark net 19, and efficient net-B1.

At last, these cropped feature maps are given as input to different deep learning models (VGG19, DarkNet19, and EfficientNet-B1) between healthy and glaucomatous images by differentiation. In this study, 3 pre-trained CNNs: VGG19, Dark Net19 and EfficientNet-B1, are combined to form an ensemble Deep Net model with multiple projections of retinal fundus images.

The output of the file classifier for given fundus image input data was defined as:

$$fitness = L_k(c_i \left| d_i \right.) = \sum_{k}^{N} w_k L_k(c_i \left| d_i^p \right.) \qquad (2)$$

Here, w_k denotes each weight assigned, $w_k \geq 0$,and $\sum_{k=1}^{N} w_k = 1$. Moreover, $L_k(c_i \left| d_i^p \right.)$ indicates on the class label posterior probability c_i for a given d_i^p , a particular projection p was input for the k -th ensemble.

The proposed ensemble model is integrated with the Hybrid Beetle Antenna Search and Genetic Algorithm (HBAS-GA) to detect the optimized fusion weights of the ensemble DeepNet model based on deep ensemble generalization loss. The proposed ensemble model offers several compelling advantages, feature repetition, improved feature propagation, and a significant reduction in the number of parameters. A BAS accelerates the convergence speed of a GA.

3.4. HBAS-GA

Inpredation, the BAS algorithm uses the characteristics of beetles to create a new intelligent algorithm. When a beetle finds food, it uses its antenna to receive information about food. In a left-handed environment, high concentrations will cause the beetle to move left, and in a right-handed environment, high concentrations will cause the beetle to move right. Below is a description of the standard BAS process.

The weight of the ensemble deep net model, which is normalized by the following formula for the complex variable vector,

$$\vec{w} = \frac{rands(f,1)}{\left\| rands(f,1) \right\|} \qquad (3)$$

Here $rands()$ denotes a random function and f denotes an optimization problem dimension. According to the above derivation, the position vector can be calculated as follows,

$$\begin{cases} b_{t1} = b - x * \vec{w}/2 \\ b_{tr} = b - x * \vec{w}/2 \end{cases} \qquad (4)$$

When b_{t1} b_{tr} indicating the position vector. x is a detection area can be covered if the perceived length is chosen such that the contraction range initially jumps out of the local minimum and then decays. To optimization problem, the fitness can be obtained as,

$$b = b - \zeta * \vec{w} * sign(I(b_{tr}) - I(b_{t1})) \qquad (5)$$

Here ζ denotes the step size of weight and $sign()$ indicates the sign function. A formula is shown below for length x , and the following formula shows some specific parameters, antenna length x and step size ζ .

$$x^t = x^{t-1} eta + 0.01 \qquad (6)$$

$$\delta^t = \delta^{t-1} eat \quad (7)$$

$$eat = 1 - 0.01 \quad (8)$$

$eta \in (0,1)$ denotes the coefficient attenuation when t is the iteration of the evolution process. The parameters can sometimes be set to constants.

3.4.1 Genetic algorithm calculation flow

It is important to note that genetic algorithms are very good at quickly finding solutions in the solution space without getting trapped in locally optimal solutions. The main calculation flow of the genetic algorithms is as follows:

Generic algorithms use a specific amount of binary code used in a specific range of parameters. Depending on the problem, there can be different forms of coding; for example, the code to solve the travelling salesman problem is a real type.

Detection and evaluation of fitness value: Individuals of a genetic generation can be determined by advantages and disadvantages according to various evaluation criteria, and adaptability and environment will continue to be iterative.

Select: Create new groups by identifying individuals who are well-adapted to the environment. Different methods can be used depending on the problem, and there are many ways to choose the methods.

Cross: The ability of the global search algorithm can be increased by choosing randomly obtained offspring that replace one or several with a certain probability.

The variation involves selecting one or more positions from the generated binary data to make changes, such as changing the "0" to a "1".

3.4.2 A Hybrid Optimization Method (BAS-GA)

There is only one beetle defined for optimization in the basic BAS algorithm. Using information, this beetle develops the best solution.BAS cannot guarantee the accuracy of the search when optimizing high-dimensional problems as it tends to fall into the local optimum. Beetle cut and mutation operations are defined in the base GA. Comparing fitness scores after mutations and crossover operations will determine which individual is optimal. There is a difference between the learning modes for BAS and GA [28]. GA emphasizes global search while focusing on local exploration BAS. BAS-GA, a new hybrid optimization method, combines the advantages of these two basic algorithms. As a result of the learning mode, welding robots tend to find the best solutions through spot weld sequencing. Figure 5 shows the flowchart for the proposed algorithm.

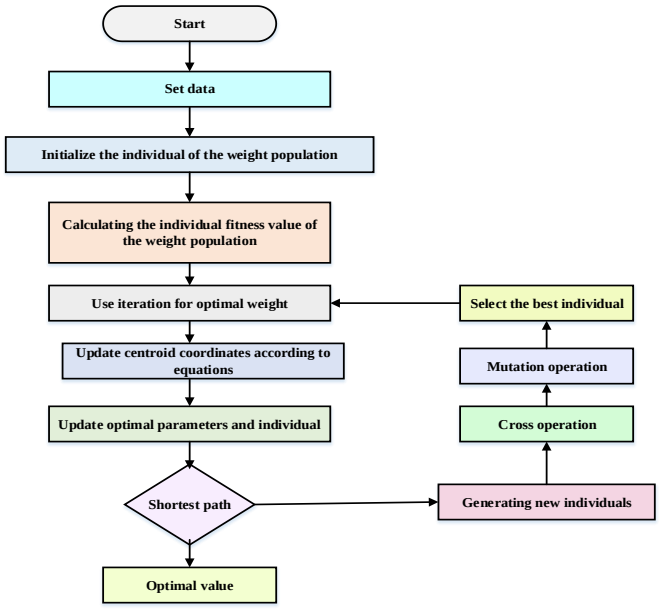

Fig 5: flow chart for BAS-GA algorithm.

4. Result and Discussion

Experiments, measurements, and evaluation datasets are described in this section. A real dataset, refuge dataset, and RIM-ONE DL were used in the experiment. The proposed model will be implemented using the python programming language. Also, the effectiveness of this approach will be validated by comparing it with existing methods in terms of intersection over union (IOU), dice coefficient, sensitivity, specificity, and accuracy.

4.1 Dataset description

RIM-ONE DL: In the RIM-ONE database, healthy and glaucomatous samples are represented in a balanced proportion. The model parameters and the loss function are validated using the RIM-ONE data set. The proposed method has better results in OD and OC on the second public database. The total number of images in the RIM-ONE DL dataset is 970: 776 for training and 195 for testing. Here the image size is 728*728, and the cropped image size is 256*256.

Refuge dataset: A dataset of retina fundus images was obtained as part of the MICCAI 2018 Retinal Fundus Glaucoma Challenge (REFUGE). The OD and OC have manually annotated 7independent ophthalmologists from Sun Yat-sen University's Zhongshan Ophthalmic Center. The suggested total number of images is 359; for training, 286; and for testing, 72. Here the actual image size is 2056*2124 after cropped image size is 256*256.

Real dataset: In this context, real data was collected from hospitals, such as various patients' eye images. Real data would be when a historical backup of a system contains real information about people. The total number of the image is displayed. Real-time images are 400, training

images are 320, and testing images are 80. Here initial image size is 800*800, and after cropping image size is 256*256**.**

An Intel(R) Core(TM) i5-3570 processor with four cores and four logical processors running at 3.20 GHz is used for the tests. The 64-bit operating system is x64-based and has built-in physical memory (RAM) of 8 GB.

4.2 Graphical representation of rim-one-dl, refuge, andrealtime datasets

As part of the proposed model, data will be collected, pre-processed, segmented, and classified into OD and OC categories. The fundus images are first collected from public data sources and real datasets. To segment OD and OC concurrently, Joint RUnet++ is applied after pre-processing. Figure 6 shows the graphical representation of the input image, pre-possessed images like cropping, channel separation, and data augmentation, and the segmented image for a rim-one-dl dataset, refuge dataset, and real-time dataset.

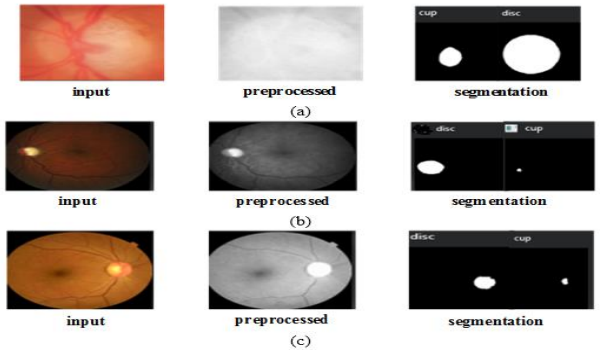

Fig 6:Graphical representation of the input, pre-processed, and segmentation images from (a) rim-one-dl, (b) refuge dataset, and (c) real-time dataset.

5. Performance Evaluation

The rim-one-dl dataset, the real-time dataset, and the refuge dataset were evaluated to find the accuracy, sensitivity, and specificity. Rim-one-dl dataset compared to several methods like Deep CNN (deep convolutional neural network), attention-based CNN (AG-CNN), and ensemble CNN. The refuge dataset was compared to several other methods, including ensemble SVM, multi-task CNN, and GD-Y net.

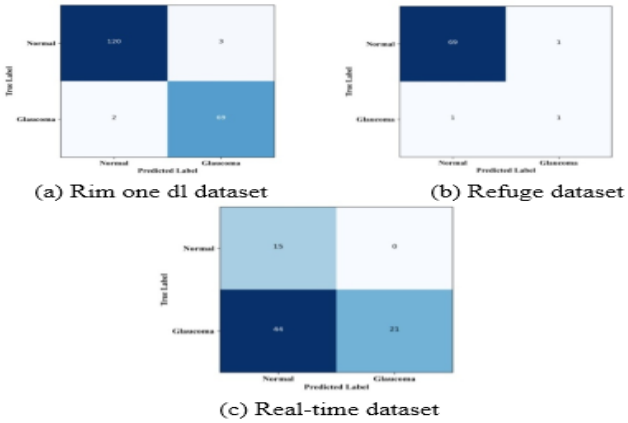

Fig 7: Confusion matrix for rim one dl dataset, refuge dataset, and real-time dataset.

Confusion matrices are tables used to measure the performance of classification algorithms, which visualize and summarize a classification algorithm's performance. A confusion matrix is shown in figure 7 for RIM-ONE DL, refuge, and real-time dataset.

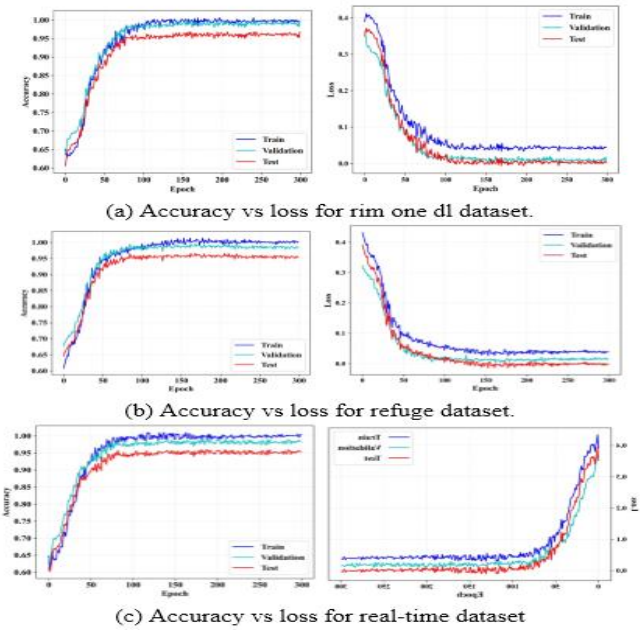

Fig 8: Accuracy vs loss for rim one dl, refuge, and real-time datasets.

The accuracy and loss values are compared between the training and testing data on rim-one-dl and the refuge dataset to illustrate the best-performing model. During training or validation, this is the sum of the errors made for each example. Each optimization iteration is measured against the loss value, which indicates how well or poorly the model performs. The accuracy metric measures the performance of the algorithm in an interpretable way. Figure 8 shows the training and testing time accuracy and loss function of the Rim-One DL and Refuge dataset.

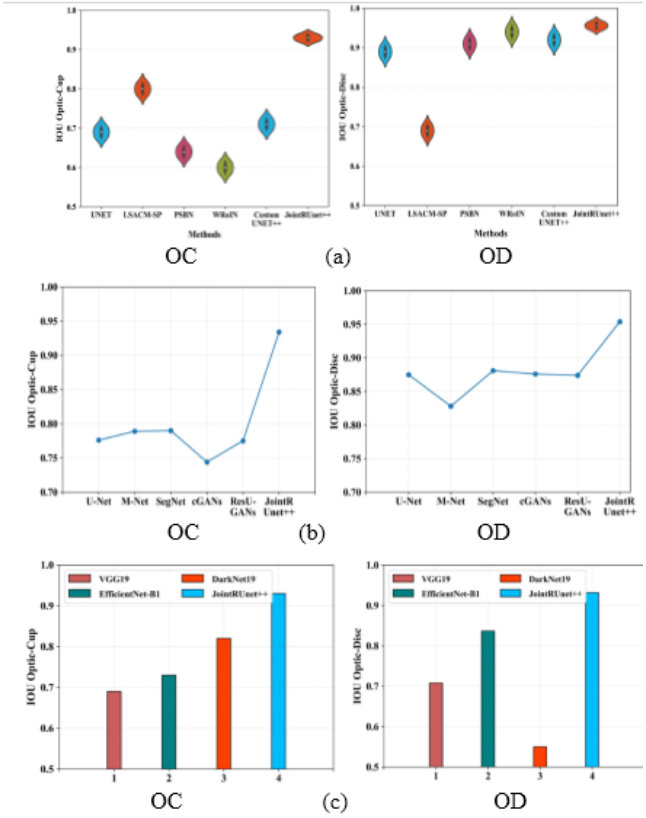

Fig 9: OC and OD IOU for (a) rim-one-dl dataset, (b) refuge dataset, (c) real-time dataset.

IOU calculates the point of intersection. A number between 0 and 1 indicates the overlap between the predicted and bounding box ground truth. Two boxes that completely overlap would have an intersection area equal to their union area; therefore, the IOU would be 1. Figure 9 above shows the IOU for the rim one dl data set, the refuge data set, and the real-time data set. Rim One dl record is compared to existing methods like U-net, LSACM-SP, PSBN, WRoIN and custom UNET++ [23]. The Refuge dataset is compared to U-net, M-net, segment, cGANs and ResU-GANS [24]. Due to a lack of annotated pixel-level fundus images, U-Net has the worst performance of all. Unlike the M-Net, which consists of a multi-scale input layer, a U-shaped convolutional network, a side output layer, and a multi-label loss function, the M-Net has a more flexible architecture. The OCIOU value is 0.93, the OD value is 0.956, the refuge OC value is 0.934, and the OD value is 0.954. The OC value of the real-time data set is 0.93, and the OD value is 0.932.

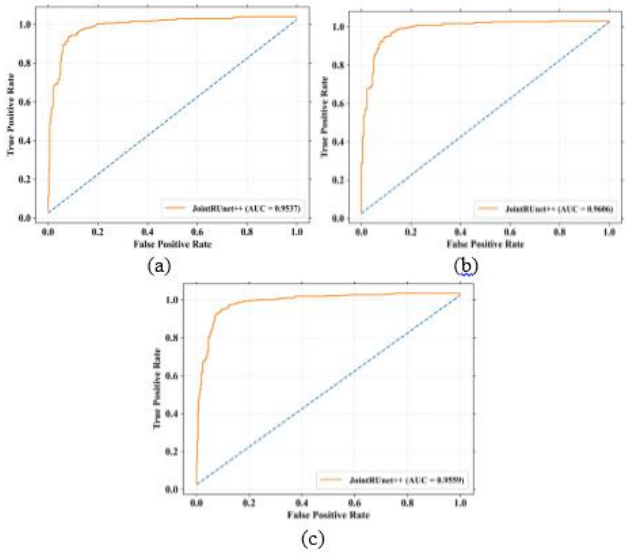

Fig 10: ROC graph for (a) rim-one-dl dataset, (b) refuge dataset, (c) real-time dataset

The ROC curve provides a performance measure of classification problems at various threshold settings. In calculating AUC, ROC is a separability measure, and ROC is a probability curve. This parameter measures the model's ability to distinguish between classes. Figure 10 shows the ROC plot for the rim-one-dl dataset, the refuge dataset, and the real-time dataset. The Rim-one-dl dataset AUC value is 0.9537, the refuse dataset AUC value is 0.9606, and the real-time dataset AUC value is 0.9559. Tables (1), (2) and (3) show the comparison with existing methods.

Table 1: Accuracy, specificity, and sensitivity comparison with existing methods.

Datasets	Methods	Accuracy	Specificity	Sensitivity
Rim one dl	JointRUnet++	0.981	0.977	0.981
	Deep CNN	0.8	0.87	0.696
	AG-CNN	0.852	0.855	0.848
	Ensembled CNN	0.944	0.913	0.944
Refuge	JointRUnet++	0.986	0.991	0.989
	Ensembled SVM	0.957	0.96	0.89
	Multitask-CNN	0.852	0.88	0.91
	GD-Y Net	0.975	0.96	0.971
	JointRUnet++	0.986	0.991	0.989

Real time	VGG19	0.833	0.875	0.875
	EfficientNet-B1	0.912	0.9375	0.93
	DarkNet19	0.75	0.687	0.687

Table 2: OC and disc IOU comparison with existing methods.

Datasets	Methods	IOU-disc	IOU-cup
Rim-one-dl	JointRUnet++	0.956	0.93
	UNET	0.89	0.69
	LSACM-SP	0.69	0.8
	PSBN	0.91	0.64
	WRoIN	0.94	0.6
	CustomUNET++	0.92	0.71
Refuge	JointRUnet++	0.954	0.934
	UNET	0.875	0.776
	M-Net	0.828	0.789
	SegNet	0.881	0.79
	cGANs	0.876	0.744
	ResU-GANs	0.874	0.775
Real time	JointRUnet++	0.932	0.93
	VGG19	0.708	0.69
	EfficientNet-B1	0.837	0.81
	DarkNet19	0.55	0.73

Table 3: Dice disc and dice cup comparison with the existing method.

Dataset	Methods	Dice disc	Dice cup
Rim one dl	JointRUnet++	0.981	0.932
	FCN-DenseNet	0.9	0.69
	Stack-UNet	0.95	0.82
Refuge	JointRUnet++	0.987	0.976
	FCN-GoogleNet	0.956	0.92
Real time	JointRUnet++	0.943	0.92
	VGG19	0.828	0.76
	EfficientNet-B1	0.911	0.91
	DarkNet19	0.69	0.71

6. Conclusion

The work proposes a deep net framework called CDN and DDN for OD and OC segmentation. This work proposes a new glaucoma eye disease detection and classification model based on automated Ensemble Deep Feature Learning using retinal fundus images. The steps involved in the proposed model are data collection, pre-processing, OD and OC segmentation and classification. This work aims mainly to improve the performance of the glaucoma disease prediction model using an ensemble learning model. A comparison will be made between the experimental results and earlier prediction models in terms of accuracy, specificity, sensitivity, IOU disc, IOU cup, and ROC. Rim one dl, refuge and real datasets will be used to evaluate the effectiveness of the proposed system. The rim one dl dataset has an accuracy value of 0.981, the refuge dataset has an accuracy value of 0.986, and the real-time dataset has an accuracy value of 0.986. Additional layers can be applied to account for these details to diagnose glaucoma eye disease. A patient's OD and OC can also be studied and monitored at various stages to improve diagnosis and decision-making.

Funding: No funding is provided for the preparation of manuscript.

Conflict of Interest: Authors declare that they have no conflict of interest.

Consent to participate: All the authors involved have agreed to participate in this submitted article.

Consent to Publish: All the authors involved in this manuscript give full consent for publication of this submitted article.

Authors Contributions: All authors read and approved the final manuscript.

Data Availability Statement: Data sharing is not applicable to this article.

Reference

[1] F. Abdullah, R. Imtiaz, H. A. Madni, H. A. Khan, T. M. Khan, M. A. Khan, and S. S. Naqvi, "A review on glaucoma disease detection using computerized techniques," *IEEE Access*. vol. 9, pp. 37311-33, 2021.

[2] S. Serte, and A. Serener, "A generalized deep learning model for glaucoma detection," *In 2019 3rd International symposium on multidisciplinary studies and innovative technologies (ISMSIT)*. IEEE, pp. 1-5, 2019.

[3] T. Nazir, A. Irtaza, A. Javed, H. Malik, D. Hussain, and R. A. Naqvi, "Retinal image analysis for diabetes-based eye disease detection using deep

learning," *Applied Sciences*. vol. 10, no. 18, p. 6185, 2020.

[4] Q. Abbas, "Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning," *International Journal of Advanced Computer Science and Applications*. vol. 8, no. 6, 2017.

[5] M. Aamir, M. Irfan, T. Ali, G. Ali, A. Shaf, A. Al-Beshri, T. Alasbali, and M. H. Mahnashi, "An adoptive threshold-based multi-level deep convolutional neural network for glaucoma eye disease detection and classification," *Diagnostics*. vol. 10, no. 8, pp. 602, 2020.

[6] D. Mirzania, A. C. Thompson, and K. W. Muir, "Applications of deep learning in detection of glaucoma: a systematic review," *European Journal of Ophthalmology*. vol. 31, no. 4, pp. 1618-42, 2021.

[7] T. Afroze, S. Akther, M. A. Chowdhury, E. Hossain, M. S. Hossain, and K. Andersson, "Glaucoma detection using inception convolutional neural network v3," *In Applied Intelligence and Informatics: First International Conference*, AII 2021, Nottingham, UK, July 30–31, 2021, Proceedings 1, Springer International Publishing pp. 17-28, 2021.

[8] D. R Nayak, D. Das, B. Majhi, S. V. Bhandary, and U. R. "Acharya, ECNet: An evolutionary convolutional network for automated glaucoma detection using fundus images," Biomedical Signal Processing and Control. vol. 67, p. 102559, 2021.

[9] S. Serte, and A. Serener, "Graph-based saliency and ensembles of convolutional neural networks for glaucoma detection," *IET Image Processing*. vol. 15, no. 3, pp. 797-804, 2021.

A. Smitha, and P. Jidesh, Classification of multiple retinal disorders from enhanced fundus images using semi-supervised GAN. *SN Computer Science*. vol. 3, pp. 1-1, 2022.

[10] H. Cho, and Y. H. Hwang, J. K. Chung, K. B Lee, J. S Park, H. G Kim, J. H. Jeong, "Deep learning ensemble method for classifying glaucoma stages using fundus photographs and convolutional neural networks," *Current eye research*. vol. 46, no. 10, pp. 1516-24, 2021.

[11] Y. Hagiwara, J. E. Koh, J. H. Tan, S. V. Bhandary, A. Laude, E. J. Ciaccio, L. Tong, U. and R. Acharya, "Computer-aided diagnosis of glaucoma using fundus images: A review," *Computer methods and programs in biomedicine*. vol. 165, pp. 1-2, 2018.

[12] H. N. Veena, A. Muruganandham, and T. S. Kumaran, "A novel optic disc and optic cup segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images," *Journal of King Saud University-Computer and Information Sciences*. vol. 34, no. 8, pp. 6187-98, 2022.

[13] S. Yu, D. Xiao, S. Frost, and Y. Kanagasingam, "Robust optic disc and cup segmentation with deep learning for glaucoma detection," *Computerized Medical Imaging and Graphics*. vol. 74, pp. 61-71, 2019.

[14] H. Fu, J. Cheng, Y. Xu, D. W Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE transactions on medical imaging*. vol. 37, no. 7, pp. 1597-605, 2018.

[15] F. Li, Y. Wang, T. Xu, L. Dong, L. Yan, M. Jiang, X. Zhang, H. Jiang, Z. Wu, and H. Zou, "Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs," *Eye*. vol. 36, no. 7, pp. 1433-41, 2022.

[16] A. Taj, M. Sajid, and K. S. Karimov, "An ensemble framework based on Deep CNNs architecture for glaucoma classification using fundus photography," *Mathematical Biosciences and Engineering*. vol. 18, no. 5, pp. 5321-47, 2021.

[17] S. Phasuk, P. Poopresert, A. Yaemsuk, P. Suvannachart, R. Itthipanichpong, S. Chansangpetch, A. Manassakorn, V. Tantisevi, P. Rojanapongpun, and C. Tantibundhit, "Automated glaucoma screening from retinal fundus image using deep learning," *In 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)* pp. 904-907, 2019. IEEE.

[18] Y. Xu, S. Ke, Y. Yang, and M. Hu, "Shared Feature Learning Based on Prior Information for Glaucoma Diagnosis," *Journal of Medical Imaging and Health Informatics*. vol. 9, no.7, pp. 1453-7, 2019.

[19] S. Bengani, "Automatic segmentation of optic disc in retinal fundus images using semi-supervised deep learning," *Multimedia Tools and Applications*. vol. 80, pp. 3443-68, 2021.

[20] Latif, S. Tu, C. Xiao, S. Ur Rehman, A. Imran, and Y. Latif, "ODGNet: a deep learning model for automated optic disc localization and glaucoma classification using fundus images," *SN Applied Sciences*. vol. 4, no. 4, p. 98, 2022.

[21] DS. David, "Enhanced glaucoma detection using ensemble based CNN and spatially based ellipse fitting curve model," *Journal of Ambient Intelligence and Humanized Computing*. pp. 1-2, 2021.

A. Tulsani, P. Kumar, and S. Pathan, "Automated segmentation of optic disc and optic cup for glaucoma assessment using improved UNET++ architecture," *Biocybernetics and Biomedical Engineering*. vol. 41, no. 2, pp. 819-32, 2021.

[22] S. Liu, J. Hong, X. Lu, X. Jia, Z. Lin, Y. Zhou, Y. Liu, and H. Zhang, "Joint optic disc and cup segmentation using semi-supervised conditional GANs," *Computers in biology and medicine*. vol. 115, p. 103485, 2019.

[23] Li, M. Xu, H. Liu, Y. Li, X. Wang, L. Jiang, Z. Wang, X. Fan, and N. Wang, "A large-scale database and a CNN model for attention-based glaucoma detection," *IEEE transactions on medical imaging*. vol. 39, no. 2, pp. 413-24, 2019.

[24] S. S. Ganesh, G. Kannayeram, A. Karthick, and M. Muhibbullah, "A novel context aware joint segmentation and classification framework for glaucoma detection," *Computational and Mathematical Methods in Medicine*. 2021.

[25] W. Zhao, Y. Zhao, L. Feng, and J. Tang, "Attention enhanced serial Unet++ network for removing unevenly distributed haze," *Electronics*. vol. 10, no. 22, p. 2868, 2021.

[26] B. C. Zhang, C. Wu, Z. X. Pang, Y. Li, and R. K. Wang, "Hybrid global optimum beetle antennae search-genetic algorithm based welding robot path planning," *In2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE pp. 1520-1524, 2019.