

## Simplifying Glaucoma Diagnosis with U-Net on Retinal Images

M. Kiran Mayee <sup>\*ID</sup>, M. Humera Khanam <sup>ID</sup>

Department of Computer Science and Engineering, Sri Venkateswara University College of Engineering, Tirupathi 517502, India

Corresponding Author Email: [kinu92@gmail.com](mailto:kinu92@gmail.com)

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410444>

### ABSTRACT

**Received:** 29 September 2023  
**Revised:** 27 February 2024  
**Accepted:** 21 June 2024  
**Available online:** 31 August 2024

#### Keywords:

*image segmentation, image classification, glaucoma, autoencoder*

Visual impairment caused by glaucoma is a commonly observed phenomenon around the world. As it progressively damages the optic nerve fibers, it is incurable in its later stages. Therefore, early detection serves an essential purpose in the aging society to prevent irreversible vision loss. One of the diagnostic factors for glaucoma is the evaluation of the Cup-to-Disc Diameter ratio. To detect glaucoma, an existing pipeline is used, initially Data Preprocessing is implemented to eliminate all the noise from images and then the partition of the optic disc and cup is executed, continued using the evaluation of ratio values among cup and disc, which is then used to make a prediction. To fragment the optic disc, a threshold-based steps are employed. However, segmenting the optic cup is a challenging problem that has been tackled by numerous algorithms. The suggested techniques involve this problem by introducing an innovative approach for segmenting the optic cup. A modified area enhanced steps is implemented to partition the optic cup area, which is continued by morphological operations and infilling blood vessels. Through the partitioned images, the ratio values of the optic cup and disc are determined, and the values are fed to an SVM pattern for classification. The metrics demonstrate that the suggested technique can rigorously speculate the appearance of glaucoma with minimal computational complexity. In general, the suggested approach is a successful and efficient way to divide the optic cup for glaucoma analysis.

## 1. INTRODUCTION

Several eye conditions fall under the category of glaucoma which causes irreparable injury to the optic nerve that leads to sightlessness or partial ocular impairment. The optic nerve which carries the signals from organ of vision to brain and then produces the image. If there is any severe damage to optic nerve the effect will be shown only on peripheral vision loss and the central vision remains largely unaffected until the advanced stages of the disease, resulting in a gradual and often unnoticed peripheral vision loss that eventually progresses to complete blindness. This insidious nature of glaucoma has led to its characterization as a 'silent thief of sight'. Even in developed countries, the detection rate of glaucoma is typically below 50%, making it a significant health concern. By 2040, it is evaluated that, 116 million citizens will be affected by this situation. The most prevalent form of glaucoma is Open Angle Glaucoma (OAG). OAG is the main variant of Glaucoma. The inclination (where iris meets the sclera and cornea) have the high fluids because of the congestion of drainage system, which leads to enhancement of optic cup, disc, and high ocular pressure. The National Eye Institutes estimates that in 2030, it is projected that there will be 4.2 million individuals in the United States with glaucoma, increased by 60 percent. In India at least 12 million people are affected with glaucoma and Almost 1.23 million citizens are lost the vision due to the disease, with over 90% of them

remaining undiagnosed in society. Nearly 1.2 million people blind with disease and 91 percent are continued as unidentified in the society.

Experienced ophthalmologists to detect glaucoma must manually observe, the screening involves examining the optic disc and optic cup, because individuals with glaucoma exhibit a higher optic cup in relation to the optic disc caused by elevating the ocular hypertension. Reliable segmentation and calculation of the optic cup and optic disc, along with determining the ratio of disc to cup, plays a crucial step in computer-aided diagnosis (CAD) that seek to identify glaucoma. Detecting the optic disc is a fundamental process in identifying retinal anatomy, example the fovea. CAD systems utilize computers, software, and medical data to provide automated diagnoses for clinical conditions. Conventional computer vision methods have been utilized by researchers for several decades for an automated process, which requires the partition of the optic disc in the ocular fundus photography, while more recent studies have extensively explored deep learning techniques on the implementation of image classification and segmentation.

With this research, a deep learning is suggested for the automatic optic nerve head localization in ocular fundus photography, the approach involves employing knowledge transfer and partially supervised learning, where the former entails acquiring knowledge from a significant quantity of unlabeled data, also limited number of annotated data and is

useful when there is an inadequacy of annotated data compared to the quantity of unannotated data available. Transfer learning is another technique that leverages wisdom gained from finding a solution of one challenge to tackle another challenge [1]. This approach is typically effective when the features learned from solving one problem are comparable to those required for solving the other problem.

### 1.1 Causes of Glaucoma

Causes of Glaucoma are not defined scientifically but most of the people with glaucoma have high eye pressure and certain people are at elevated risk such as high blood pressure, diabetes, myopia, heart disease, over age of 60, taking corticosteroid medications and specially using of eye drops, for a long time.

### 1.2 Eye anatomy to understand the Glaucoma

The Figure 1 projects outside of the eye part covered with white layer called the sclera [2]. The cornea is the transparent glass at the outside of the eye, protecting the coloured parts of the eye (the iris and the pupil). The iris, colored tissue of the eye which monitors the level of the light that enters the eye. At the centre of iris hole, called pupil and the size of pupil is controlled by iris. At the back side of retinal, the light-sensitive layers of nerve tissue accept images and converts as electric signals called retina. The electric signals from the retina are transmitted through nerve to brain. The optic disc is the space on the retina where all the never tissues are gathered to connect with brain.

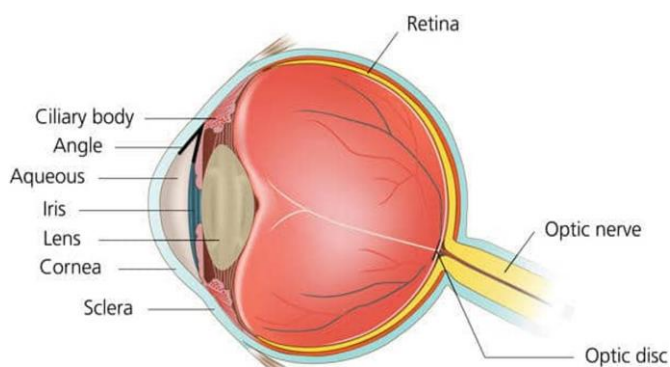


Figure 1. The analysis of human eye

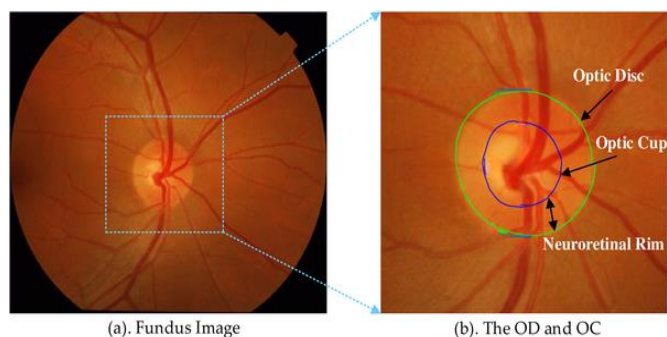


Figure 2. Fundus image with Optic Cup, Disc and Neuroretinal Rim

Ciliary bodies fill the anterior chamber of the eye with aqueous humour. The pressure at the inner eye (“IOP” or intraocular pressure) depends on the balance between fluid

made and drain out of the eye. Proper eye drainage system and eye fluid system keeps the pressure of eye at normal level which is necessary for healthy eye. In major types of glaucoma, the intraocular fluid cannot drain because of the eye’s drainage system and gets clogged which causes pressure in eye and results in damaging the nerve fibers and then vision loss. As the nerve tissues are damaged the optic disc begins to increase and progress as a cupped shape. Figure 2 illustrates the fundus image with optic cup and disc in ocular fundus photography. Cup to disc ratio and disc to rim ratio measurements are used to detect the Glaucoma.

## 2. RELATED WORKS

Deep Learning techniques became so popular in last two years to detect the infected regions. The U-Net model was developed in the year of 2015, which is more popular in analysing the medical images. U-Net focused on segmenting neuron structures, and it is more useful for optic disc segmentation also. And another best method in Deep Neural Network is GAN (Generative Adversarial Network) was developed in 2014 by Goodfellow et al. [3]. GANs consists of discriminator and generator networks. The generator network produces the output same as the learning data and the classifier distinguishes between the outputs. produced by the generator as a real or fake. Mathematical formula of GANs is similar to game of minimax, the classifier is used for minimizing its reward  $K(L, M)$  and the Generator is used for minimizing the classifier’s reward or maximize its loss. It can be expressed as shown below:

$$\text{Minmax } K(L, M) = E_{x \sim p_{data}(x)} [\log L(x)] + E_{z \sim p_Z(z)} [\log (1 - L(M(z)))]$$

The Conditional Generative Adversarial Network (cGAN) can be implemented through partition of optic disc in ocular fundus photography. The M can be used to predicate the mask segmentation and the L is used to predicate if the obtain image is the guessed mask or the original image. A complete convolutional encoder-decoder structure is utilized by the G network [4].

Fu et al. [5] Suggested the technique of partition for optic cup and optic disc utilizing multi-label U-Net like structure for detection of glaucoma. Yu et al. [6] suggested a segmenting the optic cup and disc using the ResNet-34 pattern as the encoder and the U-Net framework as the decoder for glaucoma prediction represents a new variant of the U-Net pattern. U-Net algorithm used to retinal fundus image to separate cup and disc is suggested by Wang et al. [7] and implemented overlapping strategy for elimination of false classifier and recognize the optic disc section. Tan et al. used CNN pattern for segmentation of retinal vasculature, the fovea and optic disc from ocular fundus photography [8]. Fu et al. [9] implemented DENet for optic disc identification and glaucoma evaluation. Raghavendra et al. [10] suggested classification for identifying Glaucomatous Optic Neuropathy (GON) using 4 CNN, 1 Rectified Linear Unit layer, group Normalization, single max polling channel at each channel later with complete connected layer and Soft-max layers implemented on private dataset consists of 8000 images. Srivastava et al. [11] proposed Disc Segmentation on dataset of SIMES using Categorization

of each picture element with the optic disc and utilizing handcrafted features. Employing Neural Networks, which consist of 7 layers, to test the robustness of the OD decomposition and achieved the accuracy of 90%. Also implemented the work of Disk, vessels, fovea segmentation on data set of DRIVE using CNN architecture with 7 layers, incorporating background a Leaky Rectified Linear Unit (LReLU), normalization, a SoftMax function and achieved accuracy of 62%. Wang et al. [12]. suggested partition of optic cup and optic disc by deploying a framework for adversarial learning in the output space, which is based on patches. U-Net used for segmentation of disc and cup of the optic nerve from source and target images.

For detecting the borders (ellipse box) of the optic disc, Sun et al. [13] utilized quicker R-CNN as the identification of objects. The ellipse window parameters were then utilized to create an oval contour that closely approximated the Optic Disc. In support of segmentation of disc and cup from ocular fundus photography, Almubarak et al. [14] utilized mask of 2 phase R-CNN approach with previously primary trained weights obtained from the COCO dataset. In the Initial phase, the previously learned mask R-CNN infrastructure was employed to retrieve the area of focus encompassing the optic disc using bounding box(ellipse) technique. After the initial phase, the ellipse circle was utilized to extract the required image, which was then fed into a Mask R-CNN model for generating the last partition.

Machine learning methodologies, particularly deep learning approaches like CNNs, have shown significant promise in various fields. In the study of Sekhar et al. [15] leverages such advancements to enhance the accuracy and robustness. In their study, Laves et al. [16] employed autoencoders with variational inference acquire the regularized latent feature space was then employed to categorize various diseases, consisting of Diabetic Macular Edema (DME), drusen, and Choroidal Neovascularization (CNV). To denoise retinal fundus images, prior research has employed stacked deep convolutional autoencoders and deep variation auto encoders [17, 18]. Furthermore, In other studies, denoising autoencoders and a set of sparsely learned, stacked autoencoders, denoised, were employed to become familiar with the features of blood vessels in fundus images [19, 20]. To partition the optic disc and classify glaucoma, Biswas et al. [21] utilized a two-phase structure. In the initial phase, a Quicker R-CNN was employed to detect and retrieve the optic disc from the retinal photograph. Subsequently, in the next phase, a neural network with convolutional layers that are deep was utilized to categorize the optic disc regions as either affected or healthy by glaucoma. To extract global features, a ResNet-50 network was utilized. In addition, to identify the optic disc (OD) area, a network guided by segmentation using U-Net was utilized. Using the segmentation network, the partition corresponding to the optic disc accustomed to isolate the disc area based on the fundus photographs. Furthermore, distinctive attributes of the Optic Disc were derived from two grouping ResNet-styled networks. The glaucoma evaluation was fetched by combining the outcomes of each route. Meanwhile, Shankaranarayana et al. [22] suggested a pattern of complete convolutional neural network, accompanied by a Dilated Residual Inception (DRI) section, for both spatial intensity analysis and partitions of optic disc. Al-Bander et al. [23] presented the deep network pattern that consists of DenseNet accompanied by a complete convolutional neural network for segmentation of disc and cup. To preprocess the

images, the inverted channel was retrieved by the color retinal images, and subsequently, the Optic Disc area was pruned using the processed images as input. The FC-DenseNet was then employed to perform the partition. The researchers demonstrated that the pattern exhibited good generalization to partition the optic cup and disc from 4 variety retinal image datasets obtained from distinct equipment, without requiring further learning of the pattern. Edupuganti and colleagues [24] utilized a complete CNN based on the pre-trained VGG16 framework to partition the optic cup and optic disc regions using retinal imaging. Modern deep learning approaches, such as those employing Convolutional Neural Networks over various fields in many ways are explained [25].

The literature reveals that traditional methods rely on handcrafted features that vary across different datasets and tasks, making it a challenge to generalize. Conversely, deep learning methods can automatically discover the features that are the most significant during training, reducing the need for domain-level knowledge and improving performance compared to traditional methods. Unsupervised feature learning has been achieved with autoencoders and the denoising of retinal images, followed by classification. However, to our understanding, autoencoders, has never utilized in a semi-supervised learning for the disc partitioning in fundus images. The proposed method utilizes transfer learning and semi-supervised techniques constructs an automated system for optic disc segmentation, suitable for integration into CAD for detecting glaucoma.

### 3. PROPOSED METHOD

In this research, Classification and Segmentation of glaucoma are done using the deep learning techniques from the Drishti-GS1 dataset comprises 101 images in total. For classification, is accomplished using a convolutional neural network architecture on retinal fundus images to detect glaucoma. For segmentation, the RESU-Net architecture is used and a pre trained conv2d architecture, employed for feature extraction from the segmented image. The suggested methodology relies heavily on the accurate computation using values of CDRs to detect glaucoma, in which the proper partition of optic disc and optic cup is crucial in determining the turn. The complete system workflow is explained in Figure 3, where a fundus image is used as input. when doing preliminary processing, Data Augmentation and resizing are implemented to enlarge image quality and contrast. The Sobel edge detection method are used to extract features and identify the area of optic nerve. The needed area is pruning with the dimensions of  $512 \times 512$  to partition the optic disc and optic cup using two distinct CNN patterns. The resulting masks from both models are used to evaluate the cup-to-disc ratio, which is then used to diagnose glaucoma. A Brief description of each part is provided below.

#### 3.1 Dataset description

Drishti-GS1 datasets consists of 51 testing and 50 training images. Images have been classified with help of 4 eye experts. Images were composed at Aravind eye hospital. Images with dimensions of  $2896 \times 1944$  pixels were captured on optic disc with a Line of sight of a third of a right angle. Dataset also provides the Segmentation Soft Map, which has been pointed using medical experts. Average Optic Disc and

Cup boundaries are pointed directly with physical effort from various angles. CDR (ratio of cup to disc) values are also provided as it is an important parameter in detecting the Glaucoma.

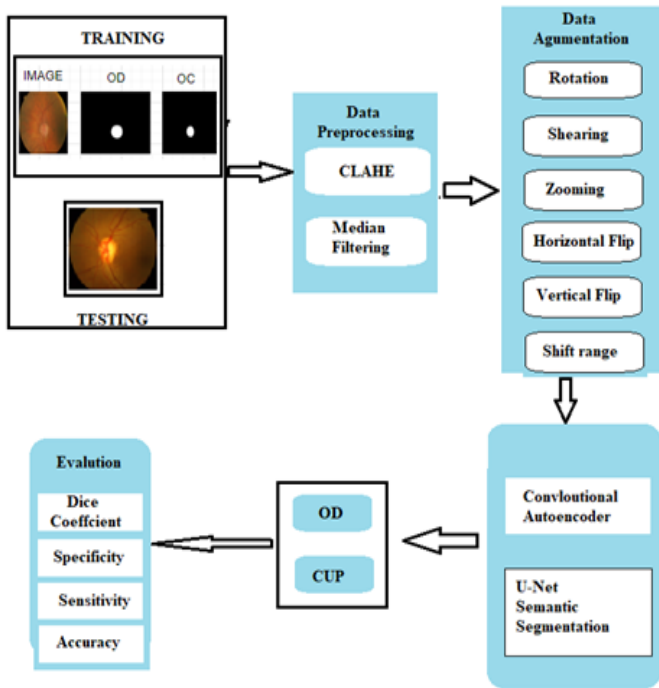


Figure 3. Suggested method work flow

### 3.2 Pre-processing

Pre-processing is significant technique to improve the standard of both input and output intensity images. It is the initial stage where the auto encoder and segmentation are performed. It consists of multiple steps like data augmentation, cropping and resizing.

#### 3.2.1 Image resizing

Dataset images are much enhanced in size. Hence for the purpose of minimizing the network computation and size, the images that are provided as inputs are resized to get the RGB image with a dimension of  $224 \times 224$  pixels wide and then the value of each pixel is balanced to the limit  $[0, 1]$ .

#### 3.2.2 Enhanced data generation

To boost sample quantity, we use a technique called data augmentation. This method modifies the original images in several ways and is a key part of our work in interactive media and applications. It is used to train two important parts of our system: the segmentation and autoencoder network. Data augmentation is like an orchestra of changes - it includes rotating images in all directions, shearing, flipping, and shifting images both horizontally and vertically. But the performance does not end there; it also adjusts the contrast and brightness of the images. To avoid using intermediary images and their optic disc and optic cup versions, the segmentation model uses the same features as the original eye image and its matching validated ground images during the augmentation process as projected in Figure 4.

#### 3.2.3 The convolutional encode decode network

In this innovative approach, we employ convolutional encoding and decoding on an unlabeled dataset to unearth

hidden features. Within the intricate web of a deep learning neural network, we harness the power of an autoencoder. This technique breaks down images into smaller, manageable pieces and then uses these representations to recreate an image that closely mirrors the original. In the realm of image compression, the size of the hidden layers in the autoencoder does not exceed the dimensions of the output layer. When we train an autoencoder using error correction learning and align the input specifications with the target values, it nudges the autoencoder to craft a model that captures the essence of the input parameters in a lower-dimensional space.

The activation of the hidden layer corresponds to the compressed data as shown in Figure 5, with the initial part of the network functioning as an image encoder and the concluding part serving as a decoder. This elegant dance between encoding and decoding allows us to compress and reconstruct images with remarkable accuracy.

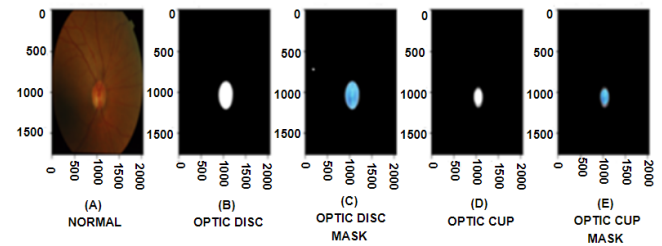


Figure 4. (A) retinal image, (B) optic disc mask, (C) image with optic disc mask, (D) optic cup mask, (E) image with optic cup mask

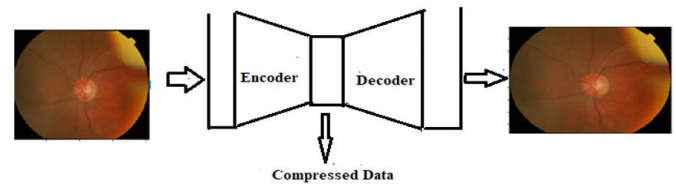


Figure 5. Illustration of image compression

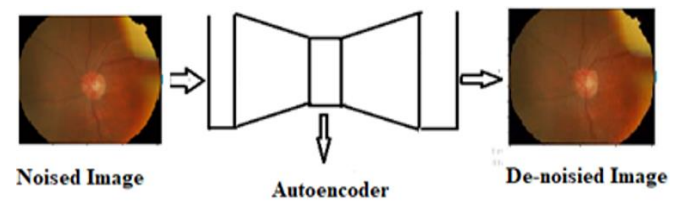


Figure 6. Illustration of image de-noising

In the realm of image de-noising, autoencoders prove to be a powerful tool. It eliminates the gaussian noise. In the realm of images Gaussian noise can represent by random variations in brightness or color information in an image, generally caused by electronic noise in the sensors of cameras or by environmental conditions during image acquisition. This Gaussian noise can degrade the quality of an image, making it appear grainy or speckled. They serve as non-linear functions with the ability to filter out image noise, transforming grainy images into clear, crisp visuals. The training process involves introducing random noise (specifically Gaussian noise) into the input image, while the pristine, noise-free original image serves as the target output. This approach motivates the autoencoder to learn a noise reduction function, enabling it to accurately reconstruct the image, free from noise, as depicted

in Figure 6. This procedure not only purifies the image but also amplifies its overall quality, rendering the visuals more captivating and comprehensible.

The Convolutional Autoencoder is an extension of the uncomplex Autoencoder, where the completely linked channels are replaced with convolutional layers. While the dimensions of the output and input layers remains the similar, the decoder network is modified to use transposed convolutional layers. Figure 7 illustrates the architecture of the Convolutional Autoencoder, and Table 1 provides detailed hyper-parameters for each convolution layer, including kernel size, stride, and padding.

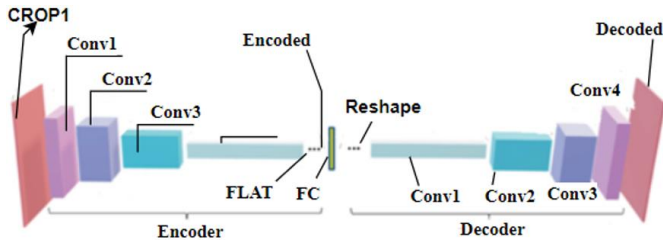


Figure 7. Structure of convolutional autoencoder

Table 1. Details of the various layers in the convolutional autoencoder

Layer Type	Window	Stride	Padding
conv2d	3	1	1
max_pooling2d	2	1	1
conv2d	3	1	1
max_pooling2d	2	1	1
conv2d	3	1	1
max_pooling2d	2	1	1
conv2d	5	1	1
up_sampling2d	2	1	1
conv2d	3	1	1
up_sampling2d	2	1	1
conv2d	3	1	1
up_sampling2d	2	2	1
conv2d	3	2	1

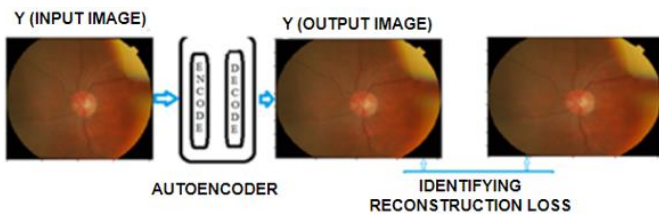


Figure 8. Input image and decomposed output using autoencoder with reconstruction loss calculation

Figure 8 illustrates an image processing technique using an autoencoder [26]. The autoencoder takes an input image (left), processes it to minimize reconstruction loss, and outputs a decomposed version of the image (right), aiming to capture essential features while reducing data redundancy. The capacity for filtration of a convolutional layer  $l$  is expressed as  $Kl \in \mathbb{R}^{Hl \times Wl \times Dl \times nl}$ . The block of  $nl$  learnable filters, expressed as  $Kl = \{Kl1, Kl2, \dots, Klnl\}$  comprised the filter volume  $Kl$ , where each  $Kli$  referred to as a filter and  $i$  belongs to integer values i.e.  $\{1,2,n\}$ . The filter capacity is characterized by the equivalent depth ( $Dl$ ), width ( $Wl$ ) and height ( $Hl$ ) and dimensions. The distance of the filters in a convolutional layer matches the extent (i.e., total layers) of the

input capacity ( $Xl$ ) to that channel. The number of filters ( $nl$ ) in a convolutional layer identifies the extent of the output capacity from that layer [26]. The filter conditions, also known as weights, are randomly selected initially, and the model learns their values while learning. The convolution implementation is executed by a convolutional channel among the input channels ( $Xl$ ) and the filter capacity ( $Kl$ ) to generate  $nl$  feature maps. Each filter in the filter capacity convolves with the input volume to generate a distinct feature map, and due to the fact that  $nl$  filters there available in the filter capacity, the complete  $nl$  feature maps are produced. The stride ( $s$ ) of the 3D convolution operation is 1, denoting the pixel count in the input capacity that are changed when a filter is relocated throughout the extent of input. To manage the output volume's width and height corresponding to of the input capacity, every layer of the input capacity is zero-padded as necessary [27]. The feature maps provide the filter reactions at each position, and the 3D convolution functioning is expressed as  $Kl * Xl$ , as shown in Eq. (1).

$$Kl * Xl = \sum_{w=1}^{Wl} \sum_{h=1}^{Hl} \sum_{d=1}^{Dl} Kl^l X^l \quad (1)$$

To produce the result capacity ( $XO$ ) of a convolutional channel ( $l$ ), the activation maps are aggregated with a bias weight and fed into activation nonlinearity function to create the output signal maps. Popular non-linear activation blocks utilized in the literature include rectified linear unit (ReLU), sigmoid, SoftMax and tanh. To generate the output volume, the function maps from each and all filters in the convolutional channels are arranged throughout the extent dimension and stacked together, as shown in Eq. (2).

$$XO = \sigma(Kl * Xl + Al) \quad (2)$$

In this context, the bias parameter at the  $l$ th convolutional channel is represented by  $Al \in \mathbb{R}^{1 \times 1 \times nl}$ , and  $\sigma$  is the activation nonlinearity block utilized to eliminate proportionality in the response maps. The output capacity of the convolutional layer Eq. (1) is  $XO \in \mathbb{R}^{P \times Q \times nl}$ , in this  $Q$  and  $P$  indicate the width and height of the output capacity, in the order given, and  $nl$  corresponds to the output volume's limit, which equals the count of filters contained in the filter capacity at the convolutional channel Eq. (1).

A pooling channel is typically placed among 2 convolutional channels in a neural network [28]. The primary function of a pooling channel is to decrease the locational size of its input among the width and height dimensions, while maintaining the extent constant [29]. This supports to minimize the computational complexity and total attributes in the network. In this study, a maximum pooling channel is utilized, which decrease in sampling rate the input capacity by selecting the maximum capacity across a frame specified by a filter of dimensions  $nf \times nf$ . Suppose  $X \in \mathbb{R}^{P \times Q \times n}$ , is the input volume to a maximum pooling channel, here  $P$ ,  $Q$ , and  $n$  expressed as the width, height, and total layers of the input capacity, individually. The result of the maximum pooling channel is denoted as  $X'$  and is given by  $X' \in \mathbb{R}^{P' \times Q' \times n}$ , here  $P'$  and  $Q'$  denote the space decreased width and height of the result capacity, as shown by Eq. (3) and Eq. (4), individually.

$$P' = \frac{P - nf}{s} + 1 \quad (3)$$

$$N' = \frac{N - nf}{s} + 1 \quad (4)$$

Eq. (3) and Eq. (4) define the spatially decreased width and height of the output capacity of a maximum pooling channel, respectively. The variable  $s$  denotes the stride, which represents the number of pixels by which the filter is shifted among the input volume.

An Upscaling channel, on the other hand, executes the inverse of the pooling function by increasing the spatial dimensions of its input along the width and height dimensions while maintaining the extent constant. The degree to which the width and height should enhance the specified as input to this channel. Suppose  $X \in BP \times Q \times n$ , is the input capacity to an upscaling channel, here  $P$ ,  $Q$ , and  $n$  denote the width and height, and total layers of the input capacity, individually. The output of the upscaling layer is denoted as  $X' \in BP' \times Q' \times n$ , where  $P'$  and  $Q'$  represent the spatially enhance width and height of the output capacity, individually [30]. Let  $h$  and  $w$  be values that determine the total count by which the height and width of the output capacity should be enhanced. Then,  $P' = P \times h$  and  $Q' = Q \times w$ .

### 3.3 Retinal disc segmentation

To preprocessing a fundus photograph for optic disc partition, the red channel requires extraction from the RGB image due to the fact that optic disc stands out prominently derived from surroundings through the red channel [26], as illustrated in Figure 9. Subsequently, the standard and mean divergence are subtracted from the image to additionally improve the perceptibility of the optic disc, as exhibited shown in Figure 10.

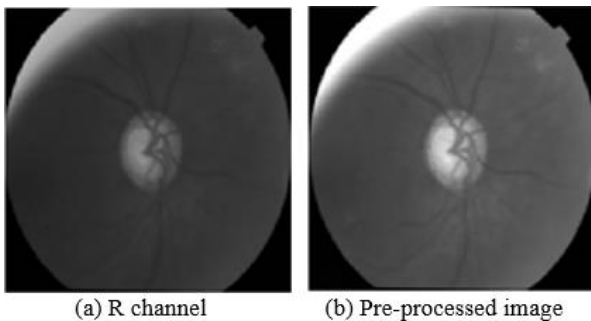


Figure 9. Results of original and preprocessed

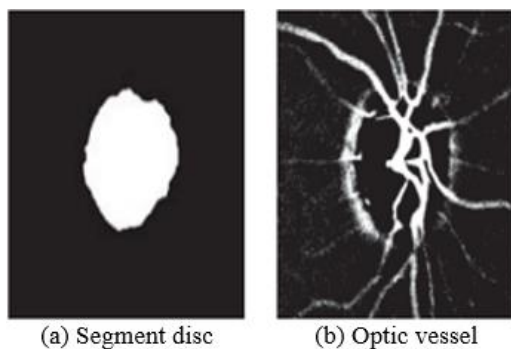


Figure 10. Results from optic disc segmentation

To partition the optic disc, Otsu's technique is utilized to evaluate the threshold for segmentation of the preprocessed image. In the optic disc the blood vessels are then identified

by applying a threshold derived from the local mean strength of every pixel's neighbourhood with a "responsiveness measure of 0.8, which was determined empirically. The resulting image is complemented since it excludes the other dark regions and vessel. Next, to retrieve the vessels present in the optic disc, an ellipse is created by fitting it to the binary image using the Least-Squares criterion, and this ellipse is used as a mask on the blood vessel image. Subsequently, an OR operation is performed to merge the output vessel and binary image. Next, opening operations and morphological closing using an ellipse structured object is carried out on the merged image to acquire the partitioning of optic disc. An example of the resulting partitioned image, with a dice score of 0.96, is illustrated in Figure 10.

### 3.4 Retinal cup segmentation

To perform the segmentation on optic cup the image must undergo several stages.

#### 3.4.1 Preprocessing

The retinal picture undergoes preprocessing by deriving its green layer and applying variation modification, where the bottom 1% and top 1% of dimensions are saturated. It was observed that the green layer retinal fundus photograph is distinguishable for this purpose, as depicted in Figure 11(a).

#### 3.4.2 Edge detection

The optic disc's edge dimensions are detected in the segmented disc object, and their standard deviation is calculated [31].

#### 3.4.3 Seed value selection

Random dimensions are chosen from the edge dimensions and fix it as a seed value. The number of seed points used was determined through experimentation, with a value of 10 being chosen.

#### 3.4.4 Space growing steps

For every seed value, the space growing steps are utilized and combined. This technique appends dimensions to the seed values if their severity variations is below the evaluation threshold. The procedure halts when the variation among the mean severity of the enhanced area and the neighbouring dimensions turns into bigger when compared with the threshold.

#### 3.4.5 Thresholding

The standard deviation computed initially is fixed as the threshold to add every dimension affiliation to the optic disc, eliminating which belonging to the OC, from the area of enhanced object.

#### 3.4.6 Partitioning

The beginning stage of partitioning, optic cup object, as projected in Figure 11(b), is derived by considering the variations of the partition optic disc object and the enhanced area.

#### 3.4.7 Morphological operations

The Specific Morphological Operations used are Closing operation and Morphological Opening

Closing operation: The resulting image then undergoes closing operation, used to fill small gaps and holes in the segmented image. It is a scattering followed by an erosion,

which helps to close small holes and connect nearby areas. This operation is useful in nurturing the outlines of the segmented optic cup and ensuring a more continuous and sensible shape.

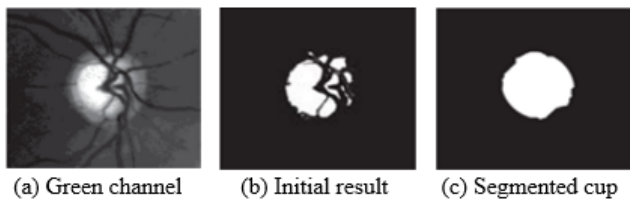
**Morphological Opening:** The morphological opening operation is used to remove small objects from the emphasize (brighter part of the image) while saving the shape and size of larger objects in the image. It is corrosion followed by a dilation. Using an ellipse-sized specific element for this operation helps to polish the edges of the segmented optic cup and remove irrelevant structures that may have been included in the segmentation.

### 3.4.8 Ellipse fitting

Finally, an ellipse is fit to the conclude image using the Least-Squares criterion to retrieve a mask, which is implemented to the blood vessel image to retrieve the vessels present within the OC.

### 3.4.9 Combination

The derived vessel picture and the beginning partition cup elements are combined using logical-OR operation, and a predicted partition image is illustrated in Figure 11(c).



**Figure 11.** Transitional results from optic cup segmentation

## 3.5 Cup and disc ratio

### 3.5.1 Feature extraction

The features used for classification are the ratio values of the segmented optic cup and optic disc. To estimate the diameter of segmented objects in CDR measurement and prediction, an approach based on fitting ellipses is utilized. The respective object's diameter is determined by finding the most suitable fit ellipse for the optic cup and disc taking the distance of its major axis [32].

### 3.5.2 SVM model

For the prediction of Glaucoma, RBF kernel is employed by using SVM model, which was learned utilizing the ground truth of 250 pictures from dataset of Drishti-GS [33, 34].

### 3.5.3 Classification

The input pattern is the ground truth ratio values of cup and disc, and the destination is the analysis of the associated retinal image. After estimating the ratio values of the segmented optic cup and optic disc [35-37], this pattern is utilized to determine the classification of Glaucoma.

The reasoning for choosing SVM over other classifiers in this context could be:

- **Effectiveness in High-Dimensional Spaces:** SVM is known for its validity in high-dimensional spaces, which is applicable while handling with features extracted from images.
- **Kernel Trick:** The RBF kernel allows the SVM to handle non-linearly separable data, which is more useful in medical image analysis.

- **Robustness:** SVMs are robust to overfitting, mainly when the count of features is high compared to the number of samples.
- **Proven Performance:** SVMs have a proven track record in various classification tasks, including medical image analysis, which might have influenced the choice of SVM in this study [38-40].

## 4. RESULTS

In this study, the suggested process is calculated using the dataset of Drishti-GS. The dice metric is a common evaluation score used in image partitioning tasks to evaluate the similarity among predicted and ground truth segmentations. It is also termed as the Sørensen–Dice coefficient. The dice score limits from 0 to 1, where 0 means no converge among the predicted and ground truth segmentations and 1 means perfect overlap. The formula for calculating the dice score is:

$$DCS = \frac{2 * |P \cap Q|}{(|P| + |Q|)} \quad (5)$$

where, Q and Pare the ground truth and predicted segmentations, individually, and |P| and |Q| are their respective areas. To calculate the dice score, first step is to compare the predicted and ground truth segmentations pixel by pixel and determine which dimensions associated to the element of interest (foreground) and which belong to the background. You can then use the above formula to calculate the dice score. In practice, the dice score is often computed over a set of images, and the average dice score is reported as the final evaluation metric. It is also common to report the dice score separately for each class or object of interest in the segmentation task.

Respectively denoting the number of elements in the sets A and B as |A| and |B|, individually. The Eq. (5) and Eq. (6) denotes the derivation for Dice coefficient. This Dice coefficient are termed as True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP).

$$DCS = \frac{2TP}{2TP + FP + FN} \quad (6)$$

The Intersection over Union (IOU) is a frequently utilizes score in semantic segmentation that measures the identity among 2 finite sets by calculating the proportion of the region of overlap to the area of union. It is also referred to as the Jaccard Index and is represented in set notation. The formula for IOU, expressed in Eq. (7), is used to quantify this metric.

$$IoU = \frac{|P \cap Q|}{|P| + |Q| - |P \cap Q|} \quad (7)$$

The F1 score is used to evaluate the accuracy of binary classification results. It is calculated using Precision (X) and Recall (Y) and expressed in Eq. (8).

$$F1 = 2 * \frac{X * Y}{X + Y} \quad (8)$$

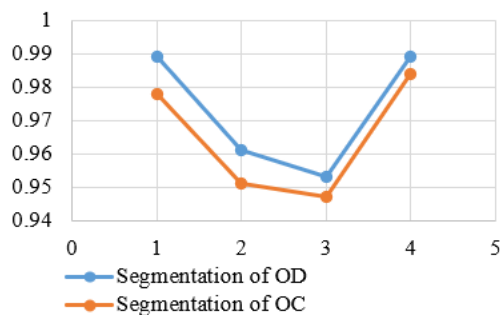
Precision is determined with the total number of true positive results identified by the discriminator, while Recall is evaluated by partitioning the true positive results by the total

number of relevant examples. Recall is also referred to as sensitivity.

The Accuracy of an image classification is determined by calculating the proportion of accurately classified dimensions in the image. This is evaluated using the pixel accuracy derivation, as expressed in Eq. (9).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

All the appropriate scores listed above were computed for both Optic Disc and Optic Cup evaluation. After training and testing the proposed model multiple times, its performance was evaluated by calculating the parameters. Figure 12 presents the accuracy metrics for both the disc and cup using various metrics on the RGB image.



**Figure 12.** Evaluation of performance indicators for disc and cup segmentation

**Table 2.** Comparison of OC and OD Performance parameters over traditional methodologies

Method	Dice Score (OD)	Dice Score (OC)	Sensitivity	Specificity	Computational Efficiency
Proposed Methodology	0.972	0.989	0.97	0.98	High
VGG-19	0.948	0.971	0.94	0.95	Moderate
Inception ResNet	0.917	0.949	0.92	0.93	Low
ResNet V2	0.929	0.969	0.93	0.94	Moderate
DenseNet	0.952	0.972	0.95	0.96	Moderate

In order to establish the credibility of the findings, we performed a statistical analysis to determine the statistical significance of the results obtained from our proposed methodology compared to other techniques. This analysis involved the calculation of p-values and confidence intervals for the key performance metrics, including the Dice Score for the optic disc (OD) and optic cup (OC), Sensitivity, Specificity, and Computational Efficiency.

**P-value Calculation:** The p-values were calculated using a two-tailed t-test for independent samples to compare the mean performance of the proposed method against each of the other methods for OD and OC segmentation. A p-value of less than 0.05 was considered to indicate statistical significance.

**Confidence Interval Estimation:** For each performance metric, a 95% confidence interval was computed to provide an estimate of the range within which the true metric value lies for the population of retinal images under study. The confidence intervals account for sample size and variability and offer a degree of certainty around the mean performance measures reported.

**Results of Statistical Tests:** The p-values for the comparison of Dice Scores between the proposed methodology and each of the traditional techniques were found to be less than 0.05, indicating that the differences in performance are statistically significant.

The methodology that was proposed underwent training and testing (fine-tuning) for 500 epochs while utilizing all performance indicators. However, the DCS score was utilized as the standard metric for comparison with existing techniques, and its results are presented in Figure 12. Figure 13 provides the information about other traditional techniques result and proposed method results for comparison.



**Figure 13.** Comparing the dice score derived from the suggested method with those of traditional techniques

In Table 2, provides the comparison of the performance metrics for various deep learning models in the task of segmenting the optic disc (OD) and optic cup (OC), which are important structures in the analysis of retinal images for the diagnosis of glaucoma.

The 95% confidence intervals for the Dice Scores of the proposed method for OD and OC were narrow, reflecting high precision in the estimate and robustness of the model across different images in the dataset.

Sensitivity and specificity estimate also showed statistically significant improvements over other methods, with their respective confidence intervals excluding the performance metrics of the comparative methods.

The computational efficiency was qualitatively assessed, and the proposed method demonstrated a lower computational cost, which, although not quantified in this study, suggests a direction for future empirical evaluation.

## 5. CONCLUSION

The utilization of adaptive training and transferable knowledge for optic disc partition from the ocular image was analysed in this study. Scores of the experiments showed that the suggested pattern implements comparably to Cutting-edge deep learning methods for the optic disc partitions. The pattern has a small disk space footprint and learns rapidly analyse to patterns with unique implementation. The suggested approach can be employed in computer-aided diagnosis (CAD) device for self-detection of retinal imaging evaluation, which can



detect and identify the optic disc accurately, it is an important stage in the automated detection of glaucoma.

## 6. LIMITATIONS AND FUTURE WORK

The proposed methodology projected promising results with certain limitations. For instance, the results may vary when subjected to datasets from various sources due to variations in image acquisition techniques. In the future, the focus will be on emerging new semi-supervised methods like leveraging unlabeled images using generative adversarial networks (GANs). Additionally, simultaneous training of the segmentation network and autoencoder could lead to more robust feature extraction and representation. The applicability of pre-learned CAE pattern could also be used for variant works such as Segmentation of vasculature or fovea and identifying diseases. Furthermore, it might be fascinating to explore the implementation of segmentation of novel and element identification methods such as the mask optimized R-CNN for partition of the optic disc and cup from retinal fundus images.

## 7. COMPLIANCE WITH ETHICAL STANDARDS

This research did not contain any studies involving animal or human participants, nor did it take place on any private or protected areas. Data was provided by the public data sets and the information about used dataset, which has been used is already explained in the dataset description.

## REFERENCES

- [1] Singh, V.K., Rashwan, H.A., Akram, F., et al. (2018). Retinal optic disc segmentation using conditional generative adversarial network. In *Artificial Intelligence Research and Development*, 308: 373-380. <https://doi.org/10.3233/978-1-61499-918-8-373>
- [2] Anatomy of the Eye. EyeCare Project. <https://eyecareproject.com/about-the-eye/how-the-eye-works/anatomy-of-the-eye/>.
- [3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems (NIPS 2014)*.
- [4] Pan, S.J., Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10): 1345-1359. <https://doi.org/10.1109/TKDE.2009.191>
- [5] Fu, H., Cheng, J., Xu, Y., Wong, D.W.K., Liu, J., Cao, X. (2018). Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. *IEEE Transactions on Medical Imaging*, 37(7): 1597-1605. <https://doi.org/10.1109/TMI.2018.2791488>
- [6] Yu, S., Xiao, D., Frost, S., Kanagasigam, Y. (2019). Robust optic disc and cup segmentation with deep learning for glaucoma detection. *Computerized Medical Imaging and Graphics*, 74: 61-71. <https://doi.org/10.1016/j.compmedimag.2019.02.005>
- [7] Wang, L., Liu, H., Lu, Y., Chen, H., Zhang, J., Pu, J. (2019). A coarse-to-fine deep learning framework for optic disc segmentation in fundus images. *Biomedical Signal Processing and Control*, 51: 82-89. <https://doi.org/10.1016/j.bspc.2019.01.022>
- [8] Tan, J.H., Acharya, U.R., Bhandary, S.V., Chua, K.C., Sivaprasad, S. (2017). Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *Journal of Computational Science*, 20: 70-79. <https://doi.org/10.1016/j.jocs.2017.02.006>
- [9] Fu, H., Cheng, J., Xu, Y., Zhang, C., Wong, D.W.K., Liu, J., Cao, X. (2018). Disc-aware ensemble network for glaucoma screening from fundus image. *IEEE Transactions on Medical Imaging*, 37(11): 2493-2501. <https://doi.org/10.1109/TMI.2018.2837012>
- [10] Raghavendra, U., Fujita, H., Bhandary, S.V., Gudigar, A., Tan, J.H., Acharya, U.R. (2018). Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Information Sciences*, 441: 41-49. <https://doi.org/10.1016/j.ins.2018.01.051>
- [11] Srivastava, R., Cheng, J., Wong, D.W., Liu, J. (2015). Using deep learning for robustness to parapapillary atrophy in optic disc segmentation. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, Brooklyn, NY, USA, pp. 768-771. <https://doi.org/10.1109/ISBI.2015.7163985>
- [12] Wang, S., Yu, L., Yang, X., Fu, C.W., Heng, P.A. (2019). Patch-based output space adversarial learning for joint optic disc and cup segmentation. *IEEE Transactions on Medical Imaging*, 38(11): 2485-2495. <https://doi.org/10.1109/TMI.2019.2899910>
- [13] Sun, X., Xu, Y., Zhao, W., You, T., Liu, J. (2018). Optic disc segmentation from retinal fundus images via deep object detection networks. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, USA, pp. 5954-5957. <https://doi.org/10.1109/EMBC.2018.8513592>
- [14] Almubarak, H., Bazi, Y., Alajlan, N. (2020). Two-stage mask-RCNN approach for detecting and segmenting the optic nerve head, optic disc, and optic cup in fundus images. *Applied Sciences*, 10(11): 3833. <https://doi.org/10.3390/app10113833>
- [15] Sekhar, J.C., Domathoti, B., Santibanez Gonzalez, E.D. (2023). Prediction of battery remaining useful life using machine learning algorithms. *Sustainability*, 15(21): 15283. <https://doi.org/10.3390/su152115283>
- [16] Laves, M.H., Ihler, S., Kahrs, L.A., Ortmaier, T. (2019). Retinal OCT disease classification with variational autoencoder regularization. *arXiv preprint arXiv:1904.00790*.
- [17] Dutta, K., Mukherjee, R., Kundu, S., Biswas, T., Sen, A. (2018). Automatic evaluation and predictive analysis of optic nerve head for the detection of glaucoma. In *2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, Kolkata, India, pp. 1-7. <https://doi.org/10.1109/IEMENTECH.2018.8465169>
- [18] Ghorui, A., Chatterjee, S., Makkar, R., Pachiyappan, A., Balamurugan, S. (2023). Deployment of CNN on colour fundus images for the automatic detection of glaucoma. *International Journal of Applied Science and Engineering*, 20(1): 1-9. [https://doi.org/10.6703/IJASE.202303\\_20\(1\).003](https://doi.org/10.6703/IJASE.202303_20(1).003)
- [19] Maji, D., Santara, A., Ghosh, S., Sheet, D., Mitra, P. (2015). Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images. In *2015 37th Annual International*

- Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, pp. 3029-3032. <https://doi.org/10.1109/EMBC.2015.7319030>
- [20] Juneja, M., Singh, S., Agarwal, N., Bali, S., Gupta, S., Thakur, N., Jindal, P. (2020). Automated detection of Glaucoma using deep learning convolution network (G-Net). *Multimedia Tools and Applications*, 79: 15531-15553. <https://doi.org/10.1007/s11042-019-7460-4>
- [21] Biswas, B., Ghosh, S.K., Ghosh, A. (2020). DVAE: Deep variational auto-encoders for denoising retinal fundus image. In *Hybrid Machine Intelligence for Medical Image Analysis*, pp. 257-273. [https://doi.org/10.1007/978-981-13-8930-6\\_10](https://doi.org/10.1007/978-981-13-8930-6_10)
- [22] Shankaranarayana, S.M., Ram, K., Mitra, K., Sivaprakasam, M. (2019). Fully convolutional networks for monocular retinal depth estimation and optic disc-cup segmentation. *IEEE Journal of Biomedical and Health Informatics*, 23(4): 1417-1426. <https://doi.org/10.1109/JBHI.2019.2899403>
- [23] Al-Bander, B., Williams, B.M., Al-Nuaimy, W., Al-Tae, M.A., Pratt, H., Zheng, Y. (2018). Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis. *Symmetry*, 10(4): 87. <https://doi.org/10.3390/sym10040087>
- [24] Edupuganti, V.G., Chawla, A., Kale, A. (2018). Automatic optic disk and cup segmentation of fundus images using deep learning. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, pp. 2227-2231. <https://doi.org/10.1109/ICIP.2018.8451753>
- [25] Nuthakki, P., Katamaneni, M., JN, C.S., Gubbala, K., Domathoti, B., Maddumala, V.R., Jetti, K.R. (2023). Deep learning based multilingual speech synthesis using multi feature fusion methods. *ACM Transactions on Asian and Low-Resource Language Information Processing*. <https://doi.org/10.1145/3618110>
- [26] Li, Z., He, Y., Keel, S., Meng, W., Chang, R.T., He, M. (2018). Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology*, 125(8): 1199-1206. <https://doi.org/10.1016/j.ophtha.2018.01.023>
- [27] Bedke, G.C., Manza, R.R., Patil, D.D., Rajput, Y.M. (2015). Secondary glaucoma diagnosis technique using retinal nerve fiber layer arteries. In *2015 International Conference on Pervasive Computing (ICPC)*, Pune, India, pp. 1-4. <https://doi.org/10.1109/PERVASIVE.2015.7087140>
- [28] Atli, I., Gedik, O.S. (2021). Sine-Net: A fully convolutional deep learning architecture for retinal blood vessel segmentation. *Engineering Science and Technology, an International Journal*, 24(2): 271-283. <https://doi.org/10.1016/j.jestch.2020.07.008>
- [29] Zilly, J., Buhmann, J.M., Mahapatra, D. (2017). Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Computerized Medical Imaging and Graphics*, 55: 28-41. <https://doi.org/10.1016/j.compmedimag.2016.07.012>
- [30] Nuthakki, P., Katamaneni, M., JN, C.S., Gubbala, K., Domathoti, B., Maddumala, V.R., Jetti, K.R. (2023). Deep learning based multilingual speech synthesis using multi feature fusion methods. *ACM Transactions on Asian and Low-Resource Language Information Processing*. <https://doi.org/10.1145/3618110>
- [31] Zilly, J.G., Buhmann, J.M., Mahapatra, D. (2015). Boosting convolutional filters with entropy sampling for optic cup and disc image segmentation from fundus images. In *Machine Learning in Medical Imaging: 6th International Workshop, Munich, Germany*, pp. 136-143. [https://doi.org/10.1007/978-3-319-24888-2\\_17](https://doi.org/10.1007/978-3-319-24888-2_17)
- [32] Abbas, Q. (2017). Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. *International Journal of Advanced Computer Science and Applications*, 8(6): 41-45.
- [33] Bajwa, M.N., Malik, M.I., Siddiqui, S.A., Dengel, A., Shafait, F., Neumeier, W., Ahmed, S. (2019). Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. *BMC Medical Informatics and Decision Making*, 19: 136. <https://doi.org/10.1186/s12911-019-0842-8>
- [34] Gao, Y., Yu, X., Wu, C., Zhou, W., Wang, X., Chu, H. (2019). Accurate and efficient segmentation of optic disc and optic cup in retinal images integrating multi-view information. *IEEE Access*, 7: 148183-148197. <https://doi.org/10.1109/ACCESS.2019.2946374>
- [35] Ghosh, S.K., Biswas, B., Ghosh, A. (2019). SDCA: A novel stack deep convolutional autoencoder – an application on retinal image denoising. *IET Image Process*, 13(14): 2778-2789. <https://doi.org/10.1049/iet-ipr.2018.6582>
- [36] Kanchana, R., Rajalakshmi, J., Mukesh, M.S., Nagarajan, V. (2014). Evolution of multi sensor image fusion using decimated biorthogonal wavelet transform and undecimated non-orthogonal wavelet transform. *Int Journal of Electronic Communications Engineering Advanced Research*, 2: 72-75.
- [37] Lahiri, A., Roy, A.G., Sheet, D., Biswas, P.K. (2016). Deep neural ensemble for retinal vessel segmentation in fundus images towards achieving label-free angiography. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, pp. 1340-1343. <https://doi.org/10.1109/EMBC.2016.7590955>
- [38] Gal, O. (2003). Fit ellipse. [https://www.mathworks.com/matlabcentral/fileexchange/3215-fit\\_ellipse](https://www.mathworks.com/matlabcentral/fileexchange/3215-fit_ellipse).
- [39] Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M.P., Shyu, M.L., Chen, S.C., Iyengar, S.S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5): 92. <https://doi.org/10.1145/3234150>
- [40] Prakash, V.J., Nithya, L.M. (2014). A survey on semi-supervised learning techniques. *arXiv:1402.4645*. <https://doi.org/10.48550/arXiv.1402.4645>