Journal of *Imaging*

MDPI

Article

EffUnet-SpaGen: An Efficient and Spatial Generative Approach to Glaucoma Detection

Krishna Adithya Venkatesh [1], Bryan M. Williams [2], Silvester Czanner [3], Srinivasan Kavitha [1], David Friedman [4], Colin E Willoughby [5], Rengaraj Venkatesh [1] and Gabriela Czanner [3,*]

1 Aravind Eye Care System, 1 Anna Nagar, Madurai, Tamil Nadu, 625 020, India
2 School of Computing and Communications, Lancaster University, Bailrigg, Lancaster LA1 4WA, UK
3 School of Computer Science and Mathematics, Liverpool John Moores University, Liverpool L3 3AF, UK
4 Glaucoma Center of Excellence, Harvard Medical School, Boston, MA 02115, USA
5 Biomedical Research Institute, Ulster University, Coleraine, Co. Londonderry BT52 1SA, UK
* Correspondence: g.czanner@ljmu.ac.uk

Abstract: Current research in automated disease detection focuses on making algorithms "slimmer," reducing the need for large training datasets and accelerating recalibration for new data while achieving high accuracy. The development of slimmer models has become a hot research topic in medical imaging. In this work, we develop a two-phase model for glaucoma detection, identifying and exploiting a redundancy in fundus image data relating particularly to the geometry. We propose a novel algorithm for cup and disc segmentation "EffUnet" with an efficient convolution block and combine this with an extended spatial generative approach for geometry modelling and classification, termed "SpaGen." We demonstrate the high accuracy achievable by EffUnet in detecting the optic disc and cup boundaries, and show how our algorithm can be quickly trained with new data, by recalibrating the EffUnet layer only. Our resulting glaucoma detection algorithm "EffUnet-SpaGen" is optimized to significantly reduce the computational burden while at the same time surpassing current state-of-art in glaucoma detection algorithms with AUROC 0.997 and 0.969 in the benchmark online datasets ORIGA and DRISHTI respectively. Our algorithm also allows deformed areas of optic rim to be displayed and investigated, providing explainability, which is crucial to successful adoption and implementation in clinical settings.

Keywords: Glaucoma; Diagnosis; Generative model; Machine learning; Classification

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *J. Imaging* **2021**, *7*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

1. Introduction

Glaucoma is a neurodegenerative disease resulting in progressive optic nerve damage with a characteristic pattern of optic nerve damage and visual field loss. Late diagnosis is a major risk factor for permanent visual loss [1] and early glaucoma detection is key to preventing avoidable blindness. Detection of structural changes to the optic nerve using imaging or clinical examination is central to diagnosis but challenging even for highly skilled specialists. Patients can be misclassified which is a significant challenge, especially in low resource settings, where access to clinical expertise and specialist diagnostic equipment is limited. A low-cost and accurate automated method of quantifying glaucomatous structural changes would help meet this need [2].

A significant challenge of developing automated glaucoma detection algorithms is that a vast number of labeled color fundus images is required for training (Figure 1). Current algorithms are very promising and show high accuracy; however, they are computationally very complex, which requires strong computing infrastructure as well as large datasets for training, for example 30 thousand images to achieve an AUROC of 0.996 [3]. ~~The~~Such computationally complex algorithms ~~may~~are ~~be~~ challenging to implement on mobile devices for community and particularly rural disease screening, necessitating the investigation of further solutions. The access to a large amount of good quality ~~of~~ annotated data ~~in glaucoma~~ for training is a persistent challenge, due in part to the complexity of the diagnosis. Therefore, an automated detection system that is computationally flexible to require less computing power and that also requires fewer training images is a fundamental requirement.

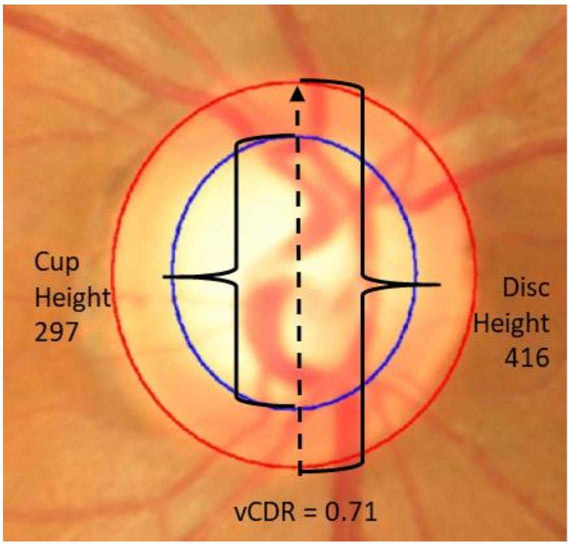

Figure 1. Colour fundus photograph of optic disc with two features: disc (red), and cup (blue).

In our paper, we present a new machine learning and generative model-based method that is able to discriminate between glaucomatous and healthy patients from standard fundus images of the optic nerve head. The proposed method revisits the convolution layers [4] and improves the generative statistical model [5]. The contribution of our work is as follows: (1) we propose a novel two-step algorithm for glaucoma detection, which traces the boundaries of the optic cup and disc efficiently, facilitating the extraction of the whole cup-to-disc profile and allowing presentation of this to the clinician for further inspection if desired, and provides an accurate glaucoma diagnosis; (2) we propose EffUnet, an efficient U-shaped convolutional neural network for efficient segmentation of the cup and disc; (3) to detect glaucoma, we propose a refined and extended spatial statistical generative model SpaGen, which takes into account the extracted profile and the cup to disc area ratio to improve detection; (4) we demonstrate the performance of our algorithm on two large publicly available datasets and show how it can be quickly recalibrated for independent data, by recalibrating the EffUnet layer only.

1.1. Background

Glaucoma is still diagnosed manually in clinical practice. Research into automated glaucoma diagnosis from fundus photographs is showing promising results. There are two main approaches to automated glaucoma detection from fundus photographs [6]. One approach involves initially automatically detecting the boundaries of the cup and disc using automated segmentation [7], which allows for the cup and disc boundaries to be used for glaucoma classification. See [8], [9], [10] for reviews and a recent approach in [5]. The alternative artificial intelligence (AI) approach to automated glaucoma diagnosis uses direct Deep Learning (DL) [3]~~, with a~~. While this has clear benefits of achieving good results while obviating the necessity for explicit automated cup and disc segmentation~~.~~ ~~With such approaches~~, ~~the AI is trained~~such approaches are trained to use all information in fundus images to differentiate glaucoma patients from those without glaucoma (see review in [11]), much of which may be~~is~~ redundant. These approaches require large numbers of expert-labelled images, ~~are~~ can be more difficult to translate to new devices and are typically not explainable. The large number of the expert-labelled images is a still a problem in glaucoma due the complexity of the gold standard definition of glaucoma. To remedy the problem of large number of images, there are other approaches like transfer learning. To solve the lack of inherent explainability there is a current research that investigates computational approaches to bring explainability to the algorithms.

A current focus is to make AI glaucoma detection algorithms "slim" in order to allow for wider use (including in low-resource settings) while also requiring fewer labelled images for training. One approach to achieve this is in realizing the redundancy in retinal fundus images for disease recognition and using this knowledge to develop lean algorithms. For example, attention maps from simple eye tracking experiments from glaucoma grading have been successfully used to improve automated glaucoma detection via an attention-based convolutional network (AG-CNN) approach [4]. However, this method requires additional data on attention maps.

Another approach to redundancy is in recognizing that the boundaries of the cup and disc in healthy eyes are similar to ellipses and hence a deviation from the ellipse can be utilized for discrimination [5]. Using this approach, the fundus image is reduced to a cup-to-disc profile vector of 24 numbers and a generative model is used for classification. However, this approach uses a computationally complex DL algorithm for cup and disc segmentation. One AI approach using slimmer algorithms is to create models that are easy to calibrate on new datasets. One such approach has been used in detecting diabetic retinopathy [12]; the researchers used a two-step architecture. The first step was an automated segmentation and the second step was a disease discrimination algorithm. Using this approach, the authors showed that, for new datasets, one needs to recalibrate the segmentation algorithm while the discrimination algorithm does not change, making the computation slimmer. This approach however still requires a computationally intensive DL method for discrimination.

1.1.1. Existing Segmentation Methods

U-Net is a U-shaped convolutional network which was originally developed for biomedical image segmentation [13]. It is composed of a down-sampling encoder layer and up-sampling decoder layer. The encoder consists of repeated groups of two convolution layers followed by a ReLU activation function and max pooling to produce a set of encoder feature maps. The decoder path also consists of convolution layers to output decoder feature maps. Skip connections transfer the corresponding feature maps from the encoder path and concatenate them with them to the upsampled decoder path.

Recently, there have been various adaptations of Unet. Mnet [14] is a convolution neural network with a multi-scale input layer and a multi-scale output layer. TernausNet [15] uses a pretrained VGG model as an encoder section of Unet. LinkNet [16] exploits ResNet-18 as an encoder and also used residual blocks instead of concatenation. In [7], a pretrained ResNet-34 is used as an encoder. However, most of these models are heavy and computationally expensive. There have also been several recent attempts to segment the optic cup and disc using deep learning-based approaches, including Unet [17] and a modified Mnet with bidirectional convolutional LSTM [18]. Some methods have also aimed to deliver models with lower memory requirements. Other methods [19] proposed a modified Unet with a novel augmentation based on contrast variations and [20] proposed CDED-Net, a computationally less expensive encoder-decoder approach with feature re-use, allowing a shallower structure to be employed.

1.1.2. Generative spatial generative model

Generative models are commonly used in statistics and also known as predictive models. The idea is to fit a model and to use the model for prediction or interpolation. This is a common paradigm in statistics for longitudinal data [21], [22].

In computer vision, statistical generative models are less frequently used, though their value is now being studied. For example, one group introduced a probabilistic generative layer to their convolutional neural network, and on standard benchmarks, they required 300-fold less training data, while achieving similar accuracy [23].

In glaucoma detection, one group published an algorithm that uses a generative model layer for classification after a DL algorithm is used for the segmentation of the cup and disc [5]. This approach required a dataset 100-times smaller for training and achieved similar accuracy of 0.996 in internal validation. The algorithm is however computationally expensive due to requiring a large DL network.

2. Materials and Methods

Our automated supervised classification of glaucoma from fundus images aims to be computationally lean to allow wide-spread use, and to allow simple calibration on new datasets. In this section, our methods are described.

2.1. Our framework

We propose a generative AI algorithm in a two-stage architecture (Figure 2). Firstly, automated segmentation of the optic cup and disc via EffUnet is done to extract the boundaries of the cup and disc (see Output 1, Figure 2). Then SpaGen algorithm [5] is updated by having two parameters for variance of noise (rather than one), and by introducing cup-to-disc area ratio (CDAR). The two variance parameter reflect the fact that variability in glaucoma images is larger than those of normal images. The CDAR is added to reflect the observation of clinicians. The boundaries of the cup and disc are then used to calculate the cup-to-disc ratio (CDR) values in 24 directions at 15-degree intervals (0, 15, 30…360 degrees), (see Output 2 in Figure 2). These 24 CDR values, as well as the ~~cup-to-disc area ratio (~~CDAR~~)~~, are then input to a spatial generative model, SpaGen. Finally, classification is carried out for each eye and output as a probability of glaucoma (see Output 3, Figure 2).

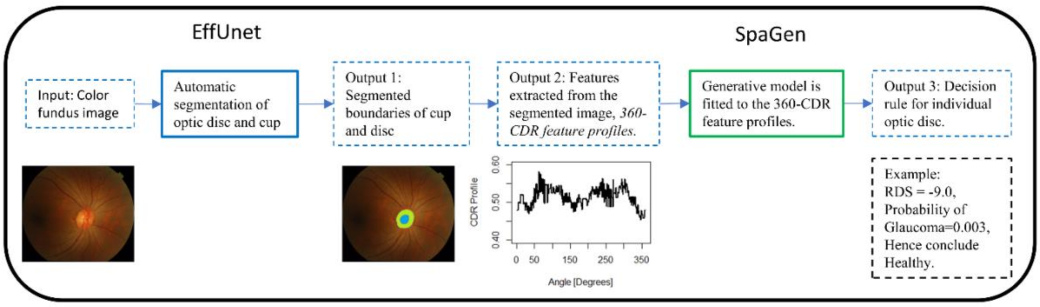

Figure 2. Framework of our EffUnet-SpaGen network. EffUnet is explained in Figure 3 and 4, SpaGen is explained in Figure 5.

2.2. Segmentation of cup and disc via EffUnet

We developed EffUnet as a U-shaped convolution network with a pre-trained efficient net-B1 [24] as the encoder. This is a modification of U-Net as the main body in our deep network (Figure 3 and 4).

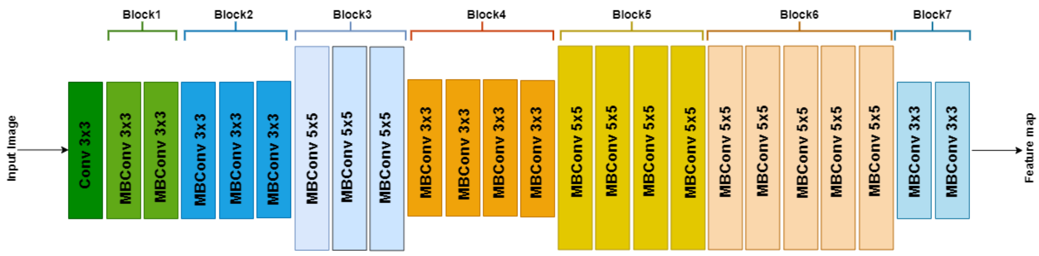

Figure 3: Architecture of EfficientNetB1 with MBConv as basic building blocks. The overall architecture can be divided into seven blocks as shown. Each MBConvX block is shown with the corresponding filter size.

In our modified U-Net architecture, we employ the EfficientNet-B1 as the down sampling encoder section of the U-Net architecture, while the decoder section is similar to the original U-Net architecture. EfficientNet's main building block is a mobile inverted bottleneck MBConv [24], [25], to which squeeze-and-excitation optimization [26] is also added.

To use EfficientNet-B1, the upsampling network has decoder blocks and each decoder block is composed of 2 × 2 upsampling 2D convolution of the previous layer output with stride of 2, concatenated corresponding feature maps from the encoder section. The concatenated tensor is then passed through two convolution layers with ReLU activation and batch normalized before passing to the next decoder block. The final layer of the architecture is convolution with softmax with channel number the same as the target classes and output image size the same as the input image.

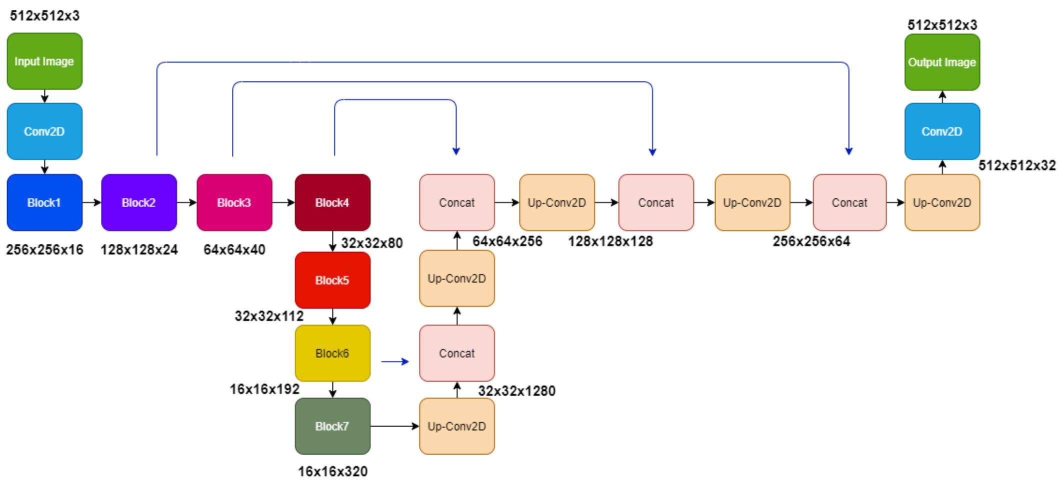

Figure 4. Framework of our EffUnet model. The Details of Block 1-7 are shown in Figure 3. The Output image (green rectangle on the right) is the Output 1 in the whole architecture shown in Figure 2.

Most existing segmentation models for cup and disc segmentation use a two-step process; disc segmentation to crop the region of interest and then multi-label segmentation to segment both cup and disk. Our model is applied on the entire image with just the black boundaries removed and resized to 512 x 512. Our EffUnet model is computationally less expensive with 12.6 M parameters hence 1.9x less parameters than ResNet34-Unet [7] which has 24.4M parameters. Our model converges a lot faster than the other models compared in Table 2.

2.3. Classification of images via SpaGen

We present here an improved generative spatial algorithm (Figure 54) for disease discrimination from the shape of the cup and disc of [5]. The key novelty is in allowing for different noise modelling in disease groups, and the incorporation of the cup-to-disc-area ratio (CDAR) (Figure 54), which is a significant factor in detecting glaucoma [27], not previously used in an automated model. This is accomplished by including two additional parameters: one for the noise component (σ_G^2) and one for the fixed component (see β_{CDAR}). Then the final improved spatial model is a hierarchical model

$$
\begin{gathered}
Y_{i,d} = \beta_0 + \beta_{G,0} I_G + \beta_{CDAR} CDAR \\
+\beta_{G,1} \sin(2\pi d/24) I_{G,d} + \beta_{G,1} \cos(2\pi d/24) I_{G,d} \\
+\beta_{G,3} \sin(4\pi d/24) I_{G,d} + \beta_{G,3} \cos(4\pi d/24) I_{G,d} \\
+\beta_{H,1} \sin(2\pi d/24) I_{H,d} + \beta_{H,1} \cos(2\pi d/24) I_{H,d} \\
+\beta_{H,3} \sin(4\pi d/24) I_{H,d} + \beta_{H,3} \cos(4\pi d/24) I_{H,d} \\
+z_i + e_{i,d}
\end{gathered} \tag{1}
$$

where $Y_{i,d}$ is CDR value of ith eye in dth direction ($d = 1, \ldots, 24$); I_G and I_H are the indicator functions for glaucoma and healthy; $I_{G,d}$ and $I_{H,d}$ are interaction terms. The term z_i is a random effect for of ith eye allowing to account for differences between eyes, $e_{i,d}$ is the random term accounting for random variations within eye. The joint probability distribution of random effect and random terms is

$$
\begin{bmatrix} z_i \\ e_i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_z^2 & 0 \\ 0 & V_e \end{bmatrix} \right), \tag{2}
$$

where V_e is a 24×24 variance-covariance matrix of error term. We allow this matrix to be different for glaucomatous and healthy groups:

$$
\begin{gathered}
V_e = \sigma_G^2 I_{24\times 24} \text{ in glaucomatous eye} \\
V_e = \sigma_H^2 I_{24\times 24} \text{ in healthy eye.}
\end{gathered} \tag{3}
$$

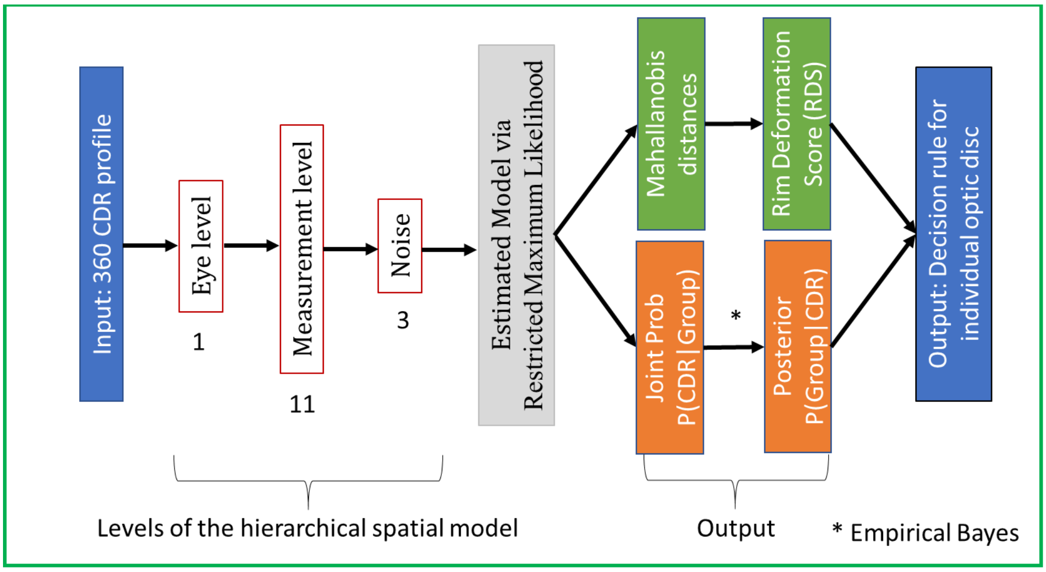

Figure 5~~2~~3. Framework of our SpaGen Model, with 15 parameters (1+11+3). This constitutes the second stage of the whole architecture (see Figure 2).

Then, assuming the prior probabilities of the diagnostic groups glaucomatous and healthy, p_G and p_H, and applying Bayes theorem, the posterior probability that a new eye with the observed profile vector Y_{new} of 24 values of CDR (pCDR) is glaucomatous:

$$
p_{new,G} = \frac{p_G f_G(Y_{new} \mid \beta, V)}{p_G f_G(Y_{new} \mid \beta, V) + p_H f_H(Y_{new} \mid \beta, V)}, \tag{4~~1~~}
$$

The posterior probability in equation (4~~1~~) can be used to propose a glaucoma detection rule. The simplest detection rule is to compare this posterior probability with a predefined probability threshold, p_{th}:

if $p_{new,G} \geq p_{th}$, conclude that the eye is glaucomatous

$$\text{if } p_{new,G} < p_{th}, \text{ conclude that the eye is healthy.} \quad (5)$$

The probabilities have the following property

$$\log\left(\frac{p_{new,G}}{1-p_{new,G}}\frac{1-p_G}{p_G}\right) = \frac{1}{2}\left[d_M(Y_{new}, \mu_H) - d_M(Y_{new}, \mu_G)\right] \quad (6)$$

where $d_M(Y_i, \mu_H)$ and $d_M(Y_i, \mu_G)$ is the Mahalanobis distance [28] of the observed data of patient i from the Healthy and Glaucomatous groups, respectively.

We then define the Rim Deformation Score (RDS) as

$$RDS = \frac{1}{2}\left[d_M(Y_{new}, \mu_H) - d_M(Y_{new}, \mu_G)\right] \quad (7\cancel{2})$$

and this can be compared to a predefined threshold, RDS_{th} to yield an equivalent decision rule

$$\text{if } RDS_{new,G} \geq RDS_{th}, \text{ conclude that the eye is glaucomatous}$$
$$\text{if } RDS_{new,G} < RDS_{th}, \text{ conclude that the eye is healthy.} \quad (8)$$

2.4. Experiments

We carried out internal validation of the performance of our EffUnet-SpaGen method in glaucoma detection on the ORIGA and DRISHTI datasets.

The ORIGA dataset is a subset of the data from the Singapore Malay Eye Study (SiMES), collected from 2004 to 2007 by the Singapore Eye Research Institute and funded by the National Medical Research Council. All images were anonymised before release. The ORIGA dataset comprises 482 healthy and 168 glaucoma images from Malay adults aged 40-80. The 650 images with manually labelled optic masks are divided into 325 training images (including 72 glaucoma cases), called ORIGA-A; and 325 testing images (including 95 glaucoma cases), called ORIGA-B [29]. The images were manually annotated, by an ophthalmologist clicking on several locations of the image to indicate the optic disc and optic rim, then a best-fitting ellipse was calculated automatically. We refer to this segmentation as the ground truth. Four graders also graded the image, and a fifth grader was used for consensus.

The DRISHTI dataset [30], called DRISTHI-GS1 by the authors and referred to here as DRISHTI) is a dataset collected and annotated by Aravind Eye Hospital, Madurai, India. All 101 images are provided with segmentation ground truth. Altogether, the set contains 70 Asian glaucomatous eyes. Selected patients were 40-80 years old. DRISHTI is split into 50 training images, called DRISHTI-A; and 51 testing images, called DRISHTI-B.

For the glaucoma classification threshold, we choose a ~~so-called~~ mathematically optimal threshold, which is the one that gives the closest point in receiver operating characteristic curve (ROC) to the top left corner, where the ROC is derived from the training dataset. We used the following criteria for accuracy: area under receiver operating characteristic curve (AUROC), sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV). We used a division of the 650 images of ORIGA into two sets, A and B, as recommended [29].

All experiments were run on a desktop computer with intel i7,16 GB RAM and a Nvidia RTX 2080 GPU, which was used to train the CNN. We trained the segmentation for 200 epochs, and use the best training result for the evaluation. Training time for segmentation is provided in Table1. We trained the SpaGen model by maximising the likelihood, which has global maximum due normal distribution of errors, the training time was 7 seconds.

3. Results

3.1. Segmentation model: computational complexity and accuracy

We used ORIGA's training and testing datasets (325 images, see Experiments). For each image, black boundaries were removed and the images were resized to 512 x 512. The performance of the proposed method EffUnet for segmenting the optic disc and optic cup was compared to the ground truth and evaluated using several standard metrics: IOU (Overlap), Dice coefficient (F-Measurement), Accuracy (Acc), Number of parameters and Number of Epochs needed:

$$\text{Dice: } DC = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (10)$$

$$\text{Jaccard: } JC = \frac{TP}{TP + FP + FN} \quad (11)$$

$$\text{Accuracy: } Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

where TP, TN, FP and FN are true positive, true negative, false positive and false negative, respectively.

Our EffUnet method is computationally less complex than the ResNet algorithm (see Number of parameters and Number of Epochs, Table 2). The ResNet algorithm requires 1.134 and 1.93 times more parameters to be tuned (see Ratio, Table 2). EffUnet is also more accurate for detecting boundaries of cup and disc (see IOU, Dice and Accuracy, in Table 2) than ResNet. ResNet-18.

Table 1. Computational efficiency and accuracy of segmentation of cup and disc jointly via EffUnet and ResNet-Unet. The training dataset is ORIGA-A, the test set is ORIGA-B. Ratio of parameters is the ratio of number of parameters in a method divided by the number of parameters in EffUnet method.

Methods	JC	DC	Acc	Number of Parameters	Ratio of Parameters	Training time (min)
ResNet34-Unet [7]	0.845	0.910	0.9966	24456444	1.93	55
ResNet18-Unet	0.846	0.911	0.9967	14340860	1.134	49
EffUnet (our method)	**0.854**	**0.916**	**0.9968**	**12641459**	**1**	**42**

The EffUnet algorithm achieves high accuracy in detecting the boundaries of the optic disc when compared to 18 published algorithms (Table 3). It achieves the highest DC of 0.9991 and the highest JC of 0.9983. Its Accuracy is very high at Acc=0.9985 which is only 0.0004 smaller than that of the fully convolutional DenseNet, which used the same ORIGA dataset and same train-test split. The rest of the 15 algorithms used other datasets.

Table ~~3~~2. Comparison of segmentation methods for optic disc. Note: [31], [32] and [33] did segmentations of both cup and disc.

Author	Method	Optic Disc			Dataset
		DC	JC	Acc	
Wong et al [34]	Support vector machine based classification mechanism	-	0.940	0.990	SiMES
Yu et al. [30]	Directional matched filtering and level sets	-	0.844	-	MESSIDOR
Mookiah et al. [35]	Attanassov intuitionistic fuzzy histon (A-IFSH) based method	0.920	-	0.934	Private
Giachetti et al. [36]	Iteratively refined model based on contour search constrained by vessel density	-	0.861	-	MESSIDOR
Dashtbozorg et al. [37]	Sliding band filter	-	0.890	-	MESSIDOR
		-	0.850	-	INSPIRE-AVR
Basit and Fraz [38]	Morphological operations, smoothing filters, 3*and the marker controlled watershed transform	-	0.710	-	Shifa
		-	0.456	-	3*CHASE-DB1
		-	0.547	-	3*DIARETDB1
		-	0.619	-	DRIVE
Wang et al. [39]	Level set method	-	0.882	-	DRIVE
		-	0.882	-	DIARETDB1
		-	0.891	-	DIARETDB0
Hamednejad et al. [40]	DBSCAN clustering algorithm	-	-	0.782	DRIVE
Roychowdhury et al. [41]	Region-based features and supervised classification	-	0.807	0.991	DRIVE
		-	0.802	0.996	DIARETDB1
		-	0.776	0.996	DIARETDB0
		-	0.808	0.991	CHASE-DB1
		-	0.837	0.996	MESSIDOR
		-	0.729	0.985	STARE
Girard et al. [42]	Local K-means clustering	-	0.900	-	MESSIDOR
Akyol et al. [43]	Keypoint detection, texture analysis, and visual dictionary	-	-	0.944	DIARETDB1
		-	-	0.950	DRIVE
		-	-	0.900	ROC
Abdullah et al. [44]	Circular Hough transform and grow-cut algorithm	-	0.786	-	DRIVE
		-	0.851	-	DIARETDB1
		-	0.832	-	CHASE-DB1
		-	0.879	-	MESSIDOR
		-	0.861	-	Private
Tan et al. [45]	7-Layer CNN	-	-	-	DRIVE
Zahoor et al. [46]	Polar transform	-	0.874	-	DIARETDB1
		-	0.844	-	MESSIDOR
		-	0.756	-	DRIVE
Sigut et al. [47]	Contrast based circular approximation	-	0.890	-	MESSIDOR
Noor et al. [31]	Colour multi-thresholding segmentation	0.590	-	0.709	DRIVE
Khalid et al. [32]	Fuzzy c-Means (FCM) and morphological operations	-	-	0.937	DRIVE
Yin et al. [48]	Statistical model	-	0.920	-	ORIGA
Fu et al. [14]	Multi-label deep learning and Polar transformation (DL)	-	0.929	-	ORIGA
Al-Bander et al. [33]	Fully convolutional DenseNet	0.965	0.933	**0.999**	ORIGA
Proposed method	**EffUnet**	**0.999**	**0.998**	**0.999**	ORIGA

The EffUnet algorithm achieved high accuracy in detecting the boundaries of the optic cup when compared to 5 published algorithms (Table 4). It achieved DC 0.8706, JC 0.7815 and Acc 0.9983. The values of DC and JC are higher than those of DenseNet and value of Acc was similar to that derived from DenseNet, which also used the ORIGA dataset with the same split to train and test subsets.

Table 3~~4~~. Comparison of segmentation methods for optic cup.

Author	Method	Optic Cup			Dataset
		DC	JC	Acc	
Hatanaka et al. [49]	Detection of blood vessel bends and features determined from the density gradient	-	-	-	Private
Almazroa et al. [50]	Thresholding using type-II Fuzzy method	- - -	- - -	0.761 0.724 0.815	BinRushed Magrabi MESSIDOR
Noor et al. [31]	Colour multi-thresholding segmentation	0.510	-	0.673	DRIVE
Khalid et al. [32]	Fuzzy c-Means (FCM) and morphological operations	-	-	0.903	DRIVE
Yin et al. [51]	Sector based and intensity with shape constraints	0.830	-	-	ORIGA
Yin et al. [48]	Statistical model	0.810	-	-	ORIGA
Xu et al. [52]	Low-rank superpixel representation	-	0.744	-	ORIGA
Tan et al. [53]	Multi-scale superpixel classification	-	0.752	-	ORIGA
Fu et al. [14]	Multi-label deep learning and Polar transformation	-	0.770	-	ORIGA
Al-Bander et al. [33]	Fully convolutional DenseNet	0.866	0.769	**0.999**	ORIGA
Proposed method	**EffUnet**	**0.870**	**0.782**	0.998	ORIGA

The EffUnet algorithm, when trained on ORIGA and fine-tuned on DRISHTI-A, achieves high accuracy in detecting the optic cup and optic disc in DRISHTI-B compared 4 published algorithms (Table 5). The model achieves a cup DC 0.9229, cup JC 0.8612, disc DC 0.9991 and disc JC 0.9983, which is the state-of-the-art performance on the DRISHTI-B set.

Table 4~~5~~. Comparison of segmentation methods for optic cup and disc. The model was finetuned on DRISHTI-A (n=50 images) and evaluated on DRISHTI-B set (n=51 images).

Author	Optic Disc		Optic Cup	
	DC	JC	DC	JC
Sevastopolsky [54]	-	-	0.850	0.750
Zilly et al. [55]	0.973	0.914	0.871	0.850
Al-Bander et al. [33]	0.949	0.904	0.828	0.711
Shuang et al. [7]	0.974	0.949	0.888	0.804
Proposed method	**0.999**	**0.998**	**0.923**	**0.861**

3.2. Segmentation model: reliability of vertical CDR

The segmentation model has very good reliability for determining the vertical CDR (vCDR, Figure ~~5~~6). After EffUnet segmented the cup and disc, the vertical heights of the cup and disc were calculated (in pixels) and the vertical cup-to-disc ratio was calculated (see vCDR_EffUnet in Figure ~~5~~6). This was then compared to the values from the manual annotation of the images where an ophthalmologist clicks several pixels of cup and disc (see vCDR_Manual in Figure ~~5~~6, which is the same as vCDR in Figure 1). For this reliability analysis, we used Bland-Altman analysis (Figure ~~5A~~6A).

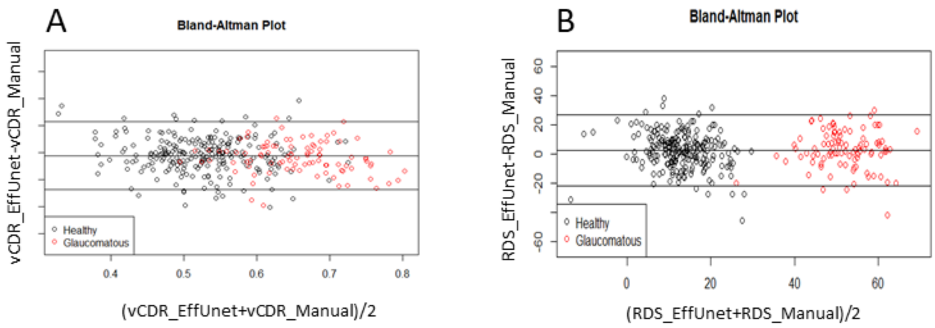

Figure 6. Reliability analysis of vertical cup-to-disc ratio (CDR) and rim deformation score (RDS) via Bland-Altman plot. Data used: segmentation trained on ORIGA-A, test set is ORIGA-B.

3.3. EffUnet-SpaGen: reliability of RDS

The segmentation model has very high reliability in terms of the Rim Deformation Score (RDS, equation (7~~2~~)) (Figure ~~5B~~6B). The RDS values calculated from EffUnet (see RDS_EffUnet, Figure ~~5B~~6B) are in good agreement with those calculated using the manually segmented cup and disc (see RDS_Manual in Figure ~~5B~~6B).

3.4. EffUnet-SpaGen: internal validation for glaucoma detection in ORIGA and DRISHTI datasets

The accuracy of EffUnet-SpaGen is high in internal validation. We trained both stages of EffUnet-SpaGen on the ORIGA-A data, and achieved 0.997 AUROC (Table 6). The CDAR alone gives 0.844 and 0.856 accuracy, for ORIGA and DRISHTI, respectively. CDAR improves the accuracy from 0.939 to 0.994 for ORIGA, and 0.879 to 0.923 for DRISHTI, if 1 variance parameter used. CDAR improves the accuracy from 0.965 to 0.997 for ORIGA, and 0.923 to 0.969 for DRISHTI, if 2 variance parameters are used. So, in summary, it improves the accuracy by 3.7 to 5.5%.

Table ~~5~~6. Ablation study of accuracy of EffUnet-SpaGen in internal validation on ORIGA and on DRISHTI. For ORIGA: train set for segmentation and glaucoma detection is ORIGA-A (n=325) (253:72 of healthy: glaucomatous), test set is ORIGA-B (n=325) (229:96 of healthy:glaucomatous). For DRISHTI: train set for segmentation is whole ORIGA and DRISHTI-A, train set for glaucoma detection is ORIGA and test is DRISTHI-B. CDAR is the Cup/Disc Area Ratio.

Segmentation Model	Generative model (n of parameters)	Results for ORIGA (top), DRISHTI (bottom)				
		AUROC	Sen	Spe	PPV	NPV
EffUnet	~~Cup/Disc Area Ratio~~C-DAR (2)	0.844	0.847	0.726	0.882	0.663
		0.856	0.737	**0.923**	0.966	0.545

EffUnet	CDR profile of 24 values & 1 variance parameter (13)	0.939	0.842	0.921	0.816	0.934
		0.879	0.789	**0.923**	0.968	0.600
EffUnet	CDR profile of 24 values & 2 variance parameters (14)	0.965	0.863	0.961	0.901	0.944
		0.933	0.895	**0.923**	0.971	0.750
EffUnet	CDR profile of 24 values & 1 variance parameters & ~~Cup/Disc Area Ratio~~CDAR (14)	0.994	0.979	0.961	0.912	0.991
		0.923	0.842	**0.923**	0.970	0.667
EffUnet	**CDR profile of 24 values & 2 variance parameters & ~~Cup/Disc Area Ratio~~CDAR (15)**	**0.997**	**0.989**	**0.974**	**0.940**	**0.996**
		0.969	**0.947**	**0.923**	**0.973**	**0.857**

3.5. Comparison results of our method for ORIGA dataset

Our approach EffUnet-SpaGen on the ORIGA dataset has the best performance published to date (AUROC=0.997) when compared to state-of-art architectures (Table 3). Gabor [56] and Wavelet [57] methods use manual features with Support Vector Machine (SVM) classifiers to get the diagnostic results. GRI [58] is a probabilistic two-stage classification method to extract the Glaucoma Risk Index. The Superpixel [59] method segments the optic disc and optic cup using superpixel classification for glaucoma screening. Chen et al. [60] and Zhao et al. [61] proposed two convolutional neural network (CNN) methods, both of which achieved good accuracy. MacCormick et al. [5] used dense fully convolutional deep learning (DL) models for segmentation and a spatial model for Disc Deformation Index (DDI) and classification had high accuracy (0.996 AUROC) but this process was highly computationally intensive (Table ~~7~~6).

Table ~~7~~6. Detection of glaucoma in ORIGA. The training set is ORIGA-A and the test set is ORIGA-B.

Author	Method of Glaucoma Detection	AUROC
Dua et al. [57]	Wavelet	0.660
Acharya et al. [56]	Gabor	0.660
Cheng et al. [59]	Superpixel	0.830
Bock et al. [58]	GRI	0.810
Chen et al. [60]	CNN	0.830
Zhao et al. [61]	CNN	0.869
Liao et al. [62]	EAMNet	0.880
MacCormick et al. [5]	DL + DDI	0.996
Proposed method	**EffUnet-SpaGen**	**0.997**

The visual results of our segmentation show good results on challenging images (Figure 7).

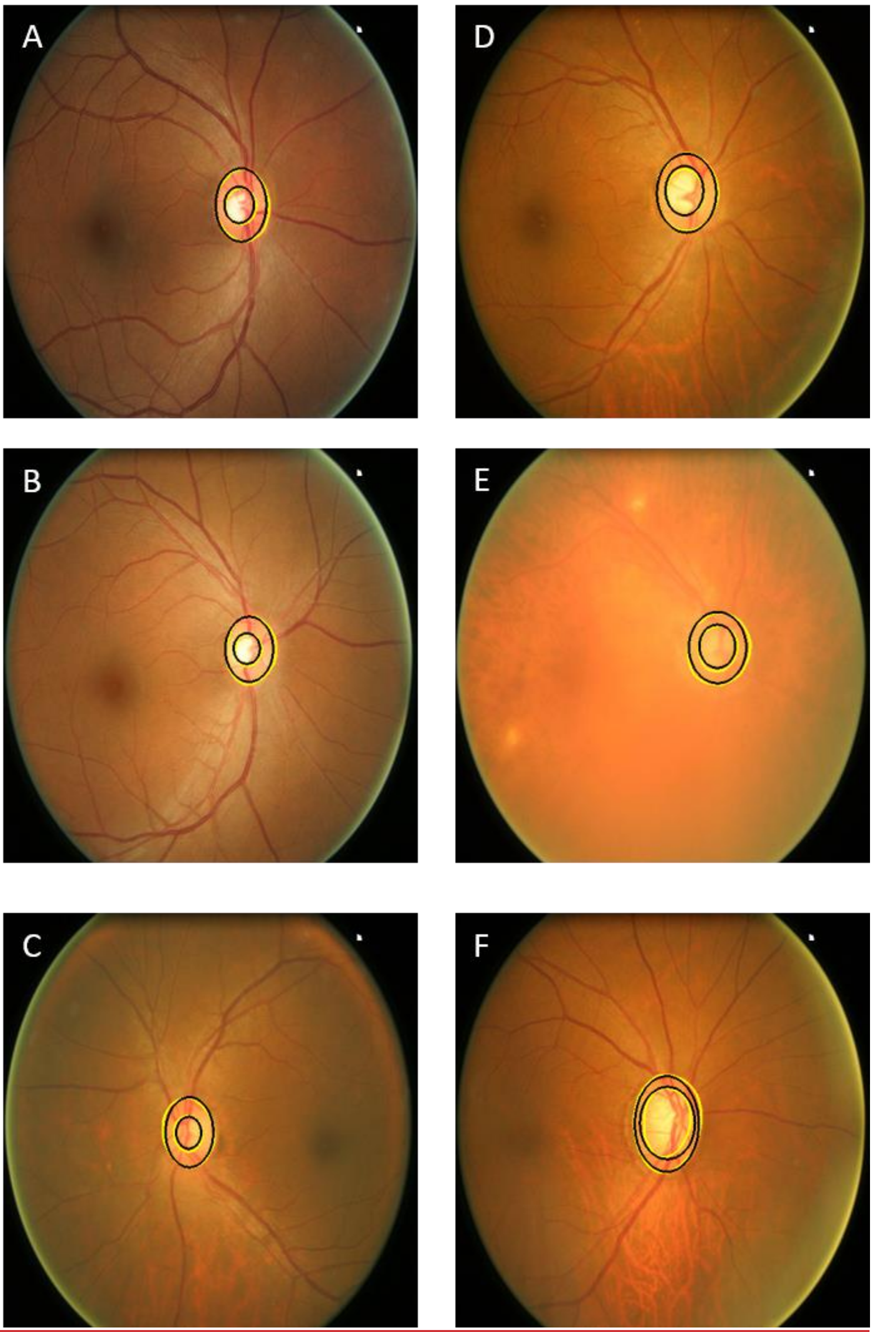

Figure 7. Visual results of several images, of normal eyes (A-C) and glaucomatous eyes (D-F). The challenging images are E, C and F.

4. Discussion

We present a new interpretable approach to glaucoma diagnosis, which combines a computationally-lean cup and disc segmentation algorithm (EffUnet) with an improved generative spatial algorithm (SpaGen). This hybrid approach is an important improvement over existing machine learning algorithms, allowing for an interpretable explanation of the findings by providing visualization measurements of the cup and disc, on which the diagnosis is based. As well as allowing us to present these areas and the key points of interest, such as rim thinning, this approach provides us with a point at which errors can be detected and mitigated, which direct deep learning approaches cannot currently do. Our approach allows lean computation, excellent results with less data, and the incorporation of additional information.

The EffUnet-SpaGen algorithm for the automated grading of optic nerve head images from fundus photographs achieves excellent performance in identifying eyes with glaucoma and distinguishing them from eyes without glaucoma. We have also demonstrated the generalisability of our work to two distinct populations by updating our method for and evaluating it on the DRISHTI dataset. As with all projects in medical imaging, it would be beneficial to demonstrate that these improved results persist in additional datasets and particularly on additional populations. It has been demonstrated already that deep learning models for glaucoma, as well as other diseases, experience a drop in performance when evaluated on new populations, even though the imaging may appear to be similar [63]. While we have tested on multiple populations in this work, it is important to continue to evaluate on the widest possible demographic.~~, highlighting~~ This highlights the need for the development of more publicly available datasets with glaucoma ground truth. To address this issue, we are currently developing segmentation masks for the LAG [64] dataset with Aravind Eye Hospital, Pondicherry, India, in an attempt to alleviate this problem.

In the task of accurately diagnosing glaucoma, we achieved an AUROC of 0.997 on the ORIGA dataset and 0.969 on DRISHTI, performing similarly or better than competing approaches, including [5] (0.996) and [62] (0.88). This represents an almost perfect result for internal validation and is the best performance reported to date for AI algorithms targeted at the diagnosis of glaucoma, compared with results that are publicly available and tested on curated datasets. Furthermore, our AUROC improves on that of a recent deep learning algorithm, which achieved 0.986 [3]. We have also demonstrated that our cup and disc segmentation technique achieves excellent performance compared with previous work.

Both EffUnet and SpaGen are computationally lean, with EffUnet requiring almost half the number of parameters of ResNet34. This allows it to estimate the glaucoma score in less than a second, making our computational speed comparable with Deep Learning approaches, while achieving similar results. Furthermore, the interpretation of the results is intuitive: the deformation of the rim is calculated along the whole cup and disc as a deviation from the normal ellipsoid-like shape, meaning that the exact deformation can be easily visualised by a clinician. Our approach also allows us to intuitively factor in additional information such as the cup to disc size and area ratio which, as we have demonstrated, allows for more accurate results.

5. Conclusions

We have presented a supervised hybrid machine and statistical learning classification framework for glaucoma detection from fundus images that is computationally flexible for wide clinical use. We achieved this by introducing a two-step framework consisting of computationally lean automated segmentation (EffUnet) and statistical learning spatial generative algorithm (SpaGen).

The segmentation produced by our proposed AI acts as a device-independent representation of the shape of the cup and disc, up to changes in field of view and aspect ratio, which our SpaGen algorithm can accommodate. This means that, while we may need to update the segmentation training with new data, we do not need to retrain the glaucoma classification rule.

On the standard benchmark dataset, EffUnet-SpaGen outperformed state-of-art deep-learning methods (0.997 AUROC) while requiring smaller datasets (n=325) for training the segmentation and classification approaches.

EffUnet is computationally less demanding (using 1.9x fewer parameters than other machine learning approaches) and SpaGen is a generative model that efficiently models the noise in data, requiring only 15 parameters. The 15-parameter model is a probabilistic generative model, that efficiently models the ellipsoid shape of the optic nerve head. It shows that there is large data redundancy in the fundus image, with most of the necessary information appearing to lie in the boundaries of the optic nerve head. Combined, this allows EffUnet-SpaGen to be trained efficiently on a n=325 dataset, which is consistent with a 300-fold decrease in training data compared to [23].

Our work removes the barriers to wider clinical use without requiring a prohibitive amount of training data in a real-world setting. Given tested in real clinical settings, this AI will translate to improvements in the management of eye care and help with the prevention of blindness from glaucoma.

Author Contributions: Conceptualization, all authors; methodology, K.A.V., B.V. and G.C.; software, K.A.V., B.W. and G.C.; validation, K.A.V. and G.C.; formal analysis, K.A.V. and G.C.; investigation, K.A.V., B.W. and G.C.; data curation, K.A.V. and G.C.; writing—original draft preparation, K.A.V., B.W. and G.C.; writing—review and editing, all authors; supervision, B.W. and G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: GC & SC are thankful for their support from GCRF fund that enabled the travels for this collaboration.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] L. Rossetti *et al.*, ‘Blindness and glaucoma: a multicenter data review from 7 academic eye clinics’, *PloS One*, vol. 10, no. 8, p. e0136632, 2015.

[2] L. Balyen and T. Peto, ‘Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology’, *Asia-Pac. J. Ophthalmol.*, vol. 8, no. 3, pp. 264–272, 2019.

[3] Z. Li, Y. He, S. Keel, W. Meng, R. T. Chang, and M. He, ‘Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs’, *Ophthalmology*, vol. 125, no. 8, pp. 1199–1206, 2018.

[4] L. Li *et al.*, ‘A large-scale database and a CNN model for attention-based glaucoma detection’, *IEEE Trans. Med. Imaging*, vol. 39, no. 2, pp. 413–424, 2019.

[5] I. J. MacCormick *et al.*, ‘Accurate, fast, data efficient and interpretable glaucoma diagnosis with automated spatial analysis of the whole cup to disc profile’, *PloS One*, vol. 14, no. 1, p. e0209409, 2019.

[6] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, ‘Artificial intelligence in retina’, *Prog. Retin. Eye Res.*, vol. 67, pp. 1–29, 2018.

[7] S. Yu, D. Xiao, S. Frost, and Y. Kanagasingam, ‘Robust optic disc and cup segmentation with deep learning for glaucoma detection’, *Comput. Med. Imaging Graph.*, vol. 74, pp. 61–71, 2019.

[8] A. Almazroa, R. Burman, K. Raahemifar, and V. Lakshminarayanan, ‘Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey’, *J. Ophthalmol.*, vol. 2015, 2015.

[9] M. S. Haleem, L. Han, J. Van Hemert, and B. Li, ‘Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: a review’, *Comput. Med. Imaging Graph.*, vol. 37, no. 7–8, pp. 581–596, 2013.

[10] F. Abdullah *et al.*, ‘A review on glaucoma disease detection using computerized techniques’, *IEEE Access*, vol. 9, pp. 37311–37333, 2021.

[11] D. S. Ting *et al.*, ‘Deep learning in ophthalmology: the technical and clinical considerations’, *Prog. Retin. Eye Res.*, vol. 72, p. 100759, 2019.

[12] J. De Fauw *et al.*, 'Clinically applicable deep learning for diagnosis and referral in retinal disease', *Nat. Med.*, vol. 24, no. 9, pp. 1342–1350, 2018.
[13] O. Ronneberger, P. Fischer, and T. Brox, 'U-net: Convolutional networks for biomedical image segmentation', 2015, pp. 234–241.
[14] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, 'Joint optic disc and cup segmentation based on multi-label deep network and polar transformation', *IEEE Trans. Med. Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018.
[15] V. Iglovikov and A. Shvets, 'Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation', *ArXiv Prepr. ArXiv180105746*, 2018.
[16] A. Chaurasia and E. Culurciello, 'Linknet: Exploiting encoder representations for efficient semantic segmentation', 2017, pp. 1–4.
[17] E. S. Kumar and C. S. Bindu, 'Two-stage framework for optic disc segmentation and estimation of cup-to-disc ratio using deep learning technique', *J. Ambient Intell. Humaniz. Comput.*, pp. 1–13, 2021.
[18] M. K. Khan and S. M. Anwar, 'M-Net with Bidirectional ConvLSTM for Cup and Disc Segmentation in Fundus Images', in *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2021, pp. 472–476.
[19] R. Imtiaz, T. M. Khan, S. S. Naqvi, M. Arsalan, and S. J. Nawaz, 'Screening of Glaucoma disease from retinal vessel images using semantic segmentation', *Comput. Electr. Eng.*, vol. 91, p. 107036, 2021.
[20] M. Tabassum *et al.*, 'CDED-Net: Joint segmentation of optic disc and optic cup for glaucoma screening', *IEEE Access*, vol. 8, pp. 102733–102747, 2020.
[21] C. H. Morrell, L. J. Brant, S. Sheng, and E. J. Metter, 'Screening for prostate cancer using multivariate mixed-effects models', *J. Appl. Stat.*, vol. 39, no. 6, pp. 1151–1175, 2012.
[22] D. M. Hughes, A. Komárek, G. Czanner, and M. Garcia-Finana, 'Dynamic longitudinal discriminant analysis using multiple longitudinal markers of different types', *Stat. Methods Med. Res.*, vol. 27, no. 7, pp. 2060–2080, 2018.
[23] D. George *et al.*, 'A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs', *Science*, vol. 358, no. 6368, 2017.
[24] M. Tan and E. Le QV, 'Rethinking model scaling for convolutional neural networks.', *arXiv*, 2019.
[25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, 'Mobilenetv2: Inverted residuals and linear bottlenecks', 2018, pp. 4510–4520.
[26] J. Hu, L. Shen, and G. Sun, 'Squeeze-and-excitation networks', 2018, pp. 7132–7141.
[27] G. Wollstein, D. F. Garway-Heath, and R. A. Hitchings, 'Identification of early glaucoma cases with the scanning laser ophthalmoscope', *Ophthalmology*, vol. 105, no. 8, pp. 1557–1563, 1998.
[28] P. C. Mahalanobis, 'Analysis of race-mixture in Bengal', 1925.
[29] Z. Zhang *et al.*, 'Origa-light: An online retinal fundus image database for glaucoma analysis and research', 2010, pp. 3065–3068.
[30] H. Yu *et al.*, 'Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets', *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 644–657, 2012.
[31] N. M. Noor, N. E. A. Khalid, and N. M. Ariff, 'Optic cup and disc color channel multi-thresholding segmentation', 2013, pp. 530–534.
[32] N. E. A. Khalid, N. M. Noor, and N. M. Ariff, 'Fuzzy c-means (FCM) for optic cup and disc segmentation with morphological operation', *Procedia Comput. Sci.*, vol. 42, pp. 255–262, 2014.
[33] B. Al-Bander, B. M. Williams, W. Al-Nuaimy, M. A. Al-Taee, H. Pratt, and Y. Zheng, 'Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis', *Symmetry*, vol. 10, no. 4, p. 87, 2018.

[34] D. W. K. Wong, J. Liu, N. M. Tan, F. Yin, B.-H. Lee, and T. Y. Wong, 'Learning-based approach for the automatic detection of the optic disc in digital retinal fundus photographs', 2010, pp. 5355–5358.

[35] M. R. K. Mookiah *et al.*, 'Automated detection of optic disk in retinal fundus images using intuitionistic fuzzy histon segmentation', *Proc. Inst. Mech. Eng. [H]*, vol. 227, no. 1, pp. 37–49, 2013.

[36] A. Giachetti, L. Ballerini, and E. Trucco, 'Accurate and reliable segmentation of the optic disc in digital fundus images', *J. Med. Imaging*, vol. 1, no. 2, p. 024001, 2014.

[37] B. Dashtbozorg, A. M. Mendonça, and A. Campilho, 'Optic disc segmentation using the sliding band filter', *Comput. Biol. Med.*, vol. 56, pp. 1–12, 2015.

[38] A. Basit and M. M. Fraz, 'Optic disc detection and boundary extraction in retinal images', *Appl. Opt.*, vol. 54, no. 11, pp. 3440–3447, 2015.

[39] C. Wang and D. Kaba, 'Level set segmentation of optic discs from retinal images', *J Med Bioeng*, vol. 4, pp. 213–220, 2015.

[40] G. Hamednejad and H. Pourghassem, 'Retinal optic disk segmentation and analysis in fundus images using DBSCAN clustering algorithm', 2016, pp. 122–127.

[41] S. Roychowdhury, D. D. Koozekanani, S. N. Kuchinka, and K. K. Parhi, 'Optic disc boundary and vessel origin segmentation of fundus images', *IEEE J. Biomed. Health Inform.*, vol. 20, no. 6, pp. 1562–1574, 2015.

[42] F. Girard, C. Kavalec, S. Grenier, H. B. Tahar, and F. Cheriet, 'Simultaneous macula detection and optic disc boundary segmentation in retinal fundus images', 2016, vol. 9784, p. 97841F.

[43] K. Akyol, B. Şen, and Ş. Bayır, 'Automatic detection of optic disc in retinal image by using keypoint detection, texture analysis, and visual dictionary techniques', *Comput. Math. Methods Med.*, vol. 2016, 2016.

[44] M. Abdullah, M. M. Fraz, and S. A. Barman, 'Localization and segmentation of optic disc in retinal images using circular Hough transform and grow-cut algorithm', *PeerJ*, vol. 4, p. e2003, 2016.

[45] J. H. Tan, U. R. Acharya, S. V. Bhandary, K. C. Chua, and S. Sivaprasad, 'Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network', *J. Comput. Sci.*, vol. 20, pp. 70–79, 2017.

[46] M. N. Zahoor and M. M. Fraz, 'Fast optic disc segmentation in retina using polar transform', *IEEE Access*, vol. 5, pp. 12293–12300, 2017.

[47] J. Sigut, O. Nunez, F. Fumero, M. Gonzalez, and R. Arnay, 'Contrast based circular approximation for accurate and robust optic disc segmentation in retinal images', *PeerJ*, vol. 5, p. e3763, 2017.

[48] F. Yin *et al.*, 'Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis', 2012, pp. 1–6.

[49] Y. Hatanaka *et al.*, 'Improved automated optic cup segmentation based on detection of blood vessel bends in retinal fundus images', 2014, pp. 126–129.

[50] A. Almazroa, S. Alodhayb, K. Raahemifar, and V. Lakshminarayanan, 'Optic cup segmentation: type-II fuzzy thresholding approach and blood vessel extraction', *Clin. Ophthalmol. Auckl. NZ*, vol. 11, p. 841, 2017.

[51] F. Yin *et al.*, 'Sector-based optic cup segmentation with intensity and blood vessel priors', in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 1454–1457.

[52] Y. Xu *et al.*, 'Optic cup segmentation for glaucoma detection using low-rank superpixel representation', 2014, pp. 788–795.

[53] N.-M. Tan, Y. Xu, W. B. Goh, and J. Liu, 'Robust multi-scale superpixel classification for optic cup localization', *Comput. Med. Imaging Graph.*, vol. 40, pp. 182–193, 2015.

[54] A. Sevastopolsky, 'Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network', *Pattern Recognit. Image Anal.*, vol. 27, no. 3, pp. 618–624, 2017.

[55] J. Zilly, J. M. Buhmann, and D. Mahapatra, 'Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation', *Comput. Med. Imaging Graph.*, vol. 55, pp. 28–41, 2017.

[56] U. R. Acharya *et al.*, 'Decision support system for the glaucoma using Gabor transformation', *Biomed. Signal Process. Control*, vol. 15, pp. 18–26, 2015.

[57] S. Dua, U. R. Acharya, P. Chowriappa, and S. V. Sree, 'Wavelet-based energy features for glaucomatous image classification', *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 80–87, 2011.

[58] R. Bock, J. Meier, L. G. Nyúl, J. Hornegger, and G. Michelson, 'Glaucoma risk index: automated glaucoma detection from color fundus images', *Med. Image Anal.*, vol. 14, no. 3, pp. 471–481, 2010.

[59] J. Cheng *et al.*, 'Superpixel classification based optic disc and optic cup segmentation for glaucoma screening', *IEEE Trans. Med. Imaging*, vol. 32, no. 6, pp. 1019–1032, 2013.

[60] X. Chen, Y. Xu, D. W. K. Wong, T. Y. Wong, and J. Liu, 'Glaucoma detection based on deep convolutional neural network', 2015, pp. 715–718.

[61] R. Zhao, Z. Chen, and X. Duan, 'Automatic detection of glaucoma based on aggregated multi-channel features', *J Comput-Aided Comput Graph*, vol. 29, pp. 998–1006, 2017.

[62] W. Liao, B. Zou, R. Zhao, Y. Chen, Z. He, and M. Zhou, 'Clinical interpretable deep learning model for glaucoma diagnosis', *IEEE J. Biomed. Health Inform.*, vol. 24, no. 5, pp. 1405–1412, 2019.

[63] H. Liu *et al.*, 'Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs', *JAMA Ophthalmol.*, vol. 137, no. 12, pp. 1353–1360, 2019.

[64] L. Li, M. Xu, X. Wang, L. Jiang, and H. Liu, 'Attention based glaucoma detection: A large-scale database and cnn model', 2019, pp. 10571–10580.