

Received December 1, 2021, accepted December 21, 2021, date of publication December 28, 2021, date of current version January 10, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3139160

Deep Learning-Based Glaucoma Detection With Cropped Optic Cup and Disc and Blood Vessel Segmentation

MIR TANVIR ISLAM¹, SHAFIN T. MASHFU¹, ABRAR FAISAL, SADMAN CHOWDHURY SIAM, INTISAR TAHMID NAHEEN¹, AND RIASAT KHAN¹, (Member, IEEE)

Department of Electrical and Computer Engineering, North South University, Dhaka 1229, Bangladesh

Corresponding author: Riasat Khan (riasat.khan@northsouth.edu)

This work was supported by North South University, Bangladesh, under Conference Travel and Research Grant CTRG-21-SEPS-25.

ABSTRACT Glaucoma is an irreversible neurodegenerative disease, where intraocular hypertension is developed due to the increased aqueous humor and blockage of the drainage system between the iris and cornea. As a result, the optic nerve head, which sends visual stimulus from our eyes to the brain, is damaged, causing visual field loss and ultimately blindness. Glaucoma is considered as the sneak thief of vision because it is difficult to diagnose early, and its regular screening is highly recommended to distinguish the neurological disorder. The detection of glaucoma is costly and time-consuming and not only there always remains a good possibility of human error but also this detection method is dependent upon the availability of the resources (experienced ophthalmologists and expensive instruments). In this work, an automatic glaucoma classification technique has been developed by utilizing multiple deep learning approaches. First, a new private dataset of 634 color fundus images has been collected and annotated by two eye specialists, a pediatric ophthalmologist and a glaucoma and refractive surgeon, from Bangladesh Eye Hospital, Bangladesh. Next, various deep learning models (EfficientNet, MobileNet, DenseNet, and GoogLeNet) have been used to detect glaucoma from fundus images. The model with EfficientNet-b3 architecture achieved the best results with test accuracy, F1-score, and ROC AUC of 0.9652, 0.9512, and 0.9574, respectively, for the cropped optic cup and disc fundus photographs. We also constructed a new dataset by segmenting the blood vessels from retinal fundus images employing a U-net model trained on High-Resolution Fundus Image Database. The MobileNet v3 model trained on this dataset achieved a satisfactory test accuracy of 0.8348 and an F1-score of 0.7957. This impressive result suggests that blood vessel segmentation of fundus images can be utilized as an alternative to detect glaucoma automatically.

INDEX TERMS EfficientNet, fundus image, glaucoma, image classification, image segmentation, MobileNet, U-net.

I. INTRODUCTION

Glaucoma is one of the most well-known causes of irreversible blindness across the globe. Glaucoma is an optic neuropathy in which damage to the retinal ganglion cells causes permanent vision loss [1]. Structural changes of the retina, especially in the optic nerve head (ONH) region, cause this eye disease [2]. Open-angle glaucoma (OAG) is perhaps the most frequently observed glaucoma type. It is

initiated by gradual congestion of the drainage system (angle between iris and cornea), leading to enlargement of the optic cup region and higher ocular pressure [3]. Angle-closure glaucoma (ACG) is another class of glaucoma generated by clogged drainage canals and an abrupt, rapid build-up of ocular pressure inside the eyes [4]. World Health Organization (WHO) announced that glaucoma is the second main reason for affecting vision and blindness globally. It might affect anyone at any age, but in aged people, it is more common. For people above the age of 60, glaucoma is one of the leading sources of blindness. Glaucoma affects more

The associate editor coordinating the review of this manuscript and approving it for publication was Khin Wee Lai¹.

than three million Americans, 2.7 million of whom are aged 40 and older [5]. There are no warning signs for certain cases of glaucoma. The effect is so progressive that people cannot detect a loss in vision until the disease is at an advanced stage, and hence it is considered the sneak thief of sight. Around 80 million people worldwide have glaucoma in 2020, and this figure is projected to grow to over 111 million by 2040 [5].

Glaucoma has no cure, but an early diagnosis can prevent significant loss of vision [6]. There are several methods to detect and identify glaucoma because of the various eye features of each individual [7]. The conventional techniques for detecting glaucoma are developed on the validation of six main factors: Tonometry, Ophthalmoscopy, Visual Field Testing, Gonioscopy, Nerve Fiber Analysis, and Pachymetry, which have been described briefly in the following paragraphs.

Tonometry: Tonometry, also known as intraocular pressure (IOP), is a standard test for measuring the pressure inside the eye [2]. Typical ocular pressure ranges between 12 and 22 mmHg. Glaucoma is more likely in people who have higher than average eye pressure. But having a higher pressure than average does not always indicate glaucoma. Conversely, glaucoma may develop in a person with lower eye pressure rather than in higher pressures.

Ophthalmoscopy: It aids in detecting glaucoma damage by examining the form and color of the optic nerve [8]. More tests are needed if the intraocular pressure (IOP) does not lie within the standard range or if the ocular nerve, which sends visual stimulus from our eyes to the brain, appears abnormal.

Visual Field Testing: Visual field testing is also called perimetry. It generates a diagram of the entire range of vision [9]. This inspection will assist in deciding whether the eyesight has been impaired by glaucoma or not.

Gonioscopy: Gonioscopy is a procedure that involves softly touching the layer of the eye with a unique reflective instrument to check the angle where the cornea touches the iris [10]. The doctor will decide what form of glaucoma is occurring and how serious the glaucoma is based on whether this angle is open or closed.

Pachymetry: Pachymetry is a technique for determining the cornea's thickness [11]. When tonometry is done, the thickness of the cornea appears to estimate the swelling of the cornea.

Nerve Fiber Analysis: Nerve fiber analysis is a recent glaucoma screening technique that assesses the width of the nerve fiber surface [12]. Narrower areas could indicate glaucoma-related damage.

Glaucoma diagnostic methods focused on medical image processing are currently gaining traction over more traditional studies [13], [14]. Several characteristics of the ocular retinal structure must be noted in these instances, including the optic nerve head (ONH), cup, retinal nerve fiber surface, peripapillary atrophy, and so on. Figure 1 shows a fundus image of a normal or healthy eye with its various

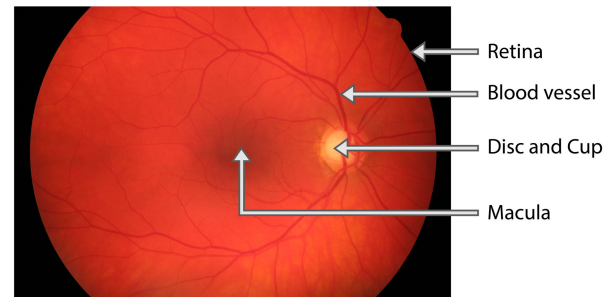


FIGURE 1. Labelled fundus image of a healthy eye.

elements, e.g., retina, blood vessel, optic cup, disc, and macula.

This paper implements automatic glaucoma detection techniques and brings forth the following contributions:

- A novel dataset of retinal fundus images referred to as BEH, labeled glaucomatous and/or healthy (nonglaucomatous) fundus pictures has been introduced. The dataset contains 634 color fundus images (463 normal/nonglaucoma and 171 glaucomatous), which have been obtained from Bangladesh Eye Hospital, Dhaka, and carefully annotated by two ophthalmologists. We have used four different deep learning algorithms, EfficientNet, MobileNet, DenseNet, and GoogLeNet, to diagnose glaucoma from the cropped optic cup and disc fundus images.
- Another significant contribution of this work is to develop a new dataset by segmenting the blood vessels from retinal fundus images applying the U-net model. The experiment results indicate that it is possible to identify glaucoma from blood vessel segmented fundus images.
- In this paper, we also measured the training timing for these two separate models and found the training time to be significantly lower for blood vessel segmented fundus images and the accuracy of the model to be slightly compromised.

To the best of our knowledge, this is the first time a comprehensive study of automatic glaucoma detection has been performed in this work by applying multiple deep learning-based techniques on both optic cup and disc segmented and blood vessel separated retinal fundus images.

II. RELATED WORKS

Conventional diagnosis of glaucoma by manual inspection of fundus images needs extensive qualified and skilled persons, and it depends hugely on their subjective judgments [15]. Consequently, this diagnosis process is costly and time-consuming and depends on human error and resource availability [16]. That's why medical images and artificial intelligence-based automated investigation of glaucoma have been extensively studied in recent years [17]. Some contemporary and significant automatic and qualitative glaucoma screening works from retinal image analysis have been described briefly in the following paragraphs.

Automated retinal image processing for glaucoma detection has been studied extensively in recent times, with mixed results. The techniques differ from conventional machine learning to advanced and sophisticated deep learning techniques. Glaucoma has been detected using both public and hybrid databases. The majority of the researchers tried to diagnose glaucoma by combining various public retinal image datasets. For instance, in order to detect glaucoma from the feature extraction of the optic cup and disc, J. Civit-Masot *et al.* used a combination of the two open-source fundus image datasets, RIM-ONE V3 and DRISHTI [18]. First, the authors performed data preprocessing, offline and online augmentations by trimming, resizing, adjusting brightness, zooming, and rotating the images. Next, they implemented the convolutional neural network U-net architecture to segment the optic cup and disc from the fundus images and extracted its features. Finally, the convolutional neural network, MobileNet V2, has been used to classify the fundus images into healthy and glaucoma ones. The authors obtained 0.84-0.93 Dice coefficient scores for the cup and disc segmentation system. This work reported an area under the curve (AUC) of 0.93 for the classification scheme, which involves trivial computational resources compared to other existing classification methods. Manually segmenting the optic cup and disc to estimate the cup-to-disc ratio (CDR) is a complex and time-consuming process. Thereupon, M. Alghamdi and M. Abdel-Mottaleb [19] developed an automated glaucoma diagnostic framework based on three different learning methods of convolutional neural network (CNN) models and evaluated these models' results with ophthalmologists. Based on both labeled and unlabeled input, they utilized transfer convolutional neural network model (TCNN), semi-supervised convolutional neural network technique (SSCNN) with self-learning, and semi-supervised convolutional neural network model with autoencoder (SSCNN-DAE). Their models implemented on open-source datasets, RIM-ONE and RIGA, demonstrated convincing results and confirmed the effectiveness of deep learning models on glaucoma detection. According to the authors, the TCNN, SSCNN, and SSCNN-DAE achieved 91.5 percent, 92.4 percent, and 93.8 percent overall accuracy, respectively. In the paper [20], J. Gómez-Valverde *et al.* experimented and developed a transfer learning-based CNN model for automatic glaucoma classification. They have used color fundus images from DRISHTI-GS and RIM-ONE datasets. They also collected some additional images from different campaigns in Barcelona, Spain, and combined all three datasets. Then they preprocessed the images and fined tuned five different CNN models with transfer learning technique. Among these, the VGG-19 model showed the best result with an AUC of 94% with a sensitivity and specificity score of 87.01% and 89.01%, respectively. In [21], A. Diaz-Pinto *et al.* utilized five distinct ImageNet-trained architectures, e.g., VGG16, VGG19, ResNet50, InceptionV3, and Xception, for detecting glaucoma, where any feature extraction or estimations of geometric optic nerve

head (ONH) structures like CDR are not required. The authors used a combination of five open-source datasets containing 1,707 fundus images from which the ACRIMA dataset was introduced and made accessible to the public through the authors. ACRIMA database consists of 396 and 309 glaucoma and healthy (nonglaucoma) images, respectively, and achieved an AUC of 0.7678 with a 70.21% accuracy in the test set. Finally, the authors obtained varying AUC scores of 0.8354, 0.8041, 0.8575, and 0.7739 for the remaining open-source datasets utilized in this work, i.e., HRF, DRISHTI-GS1, RIM-ONE, sjchoi86-HRF, respectively. The authors of the paper [22] proposed an automated two-fold glaucoma detection scheme to ease the burden of eye doctors. The implemented deep learning model, DeepLabv3+ distinguished and obtained the optic disc (OD) from the entire image in the first step. Next, three approaches of deep CNNs have been used in the second phase to recognize normal and glaucoma cases in the fragmented OD region. Finally, the authors tested their algorithms on five public datasets containing 2,787 fundus images and suggested that a mixture of MobileNet and DeepLabv3+ models is the best option for optic disc segmentation. They gained an accuracy of 97.37%, 90.00%, 86.84%, 99.53%, and 95.59% for RIM-ONE, ORIGA, DRISHTI-GS1, ACRIMA, and REFUGE datasets, respectively.

Only individual public or private fundus image datasets have been utilized in some literature to detect glaucoma automatically. In [23], N. A. Mohamed and his team developed a superpixel classification-based glaucoma detection system and tested it on the RIM-ONE database. To train their classifier, they first preprocessed the images to improve the contrast and removed noise. Next, they applied different filters and went through some illumination corrections. Then, on these images, they applied the SLIC algorithm and segmented the images to differentiate the optical disc and optical cup. Finally, they used SVM and tried with linear and RBF kernels with 5-fold cross-validation for superpixel classification. Lastly, their proposed system has achieved 98.6% accuracy with 97.6% sensitivity and 92.3% specificity. Salam *et al.* [24] employ a local private dataset of 100 retinal fundus images. They propose an algorithm that provides a computer-aided framework for automatic glaucoma detection that helps ophthalmologists diagnose glaucoma patients with high precision. The algorithm combines structural (cup-to-disc ratio) and non-structural (texture and intensity) features that enhance accuracy. Using CDR calculation, the proposed model takes preprocessed images of the fundus and extracts the optic cup and optic disc. The classifier is trained and tested by extracting non-structural intensity and textural features. Then CDR and features were incorporated to identify the image as glaucoma. Finally, the proposed system is able to get an average sensitivity of 100% and specificity of 87%.

Some authors have utilized a combination of public and private datasets for automatic glaucoma screening. The authors proposed a unique automatic glaucoma testing method in [25], where clinically computed and image-based

attributes have been extracted and compared to detect glaucoma consecutively. As the number of glaucomatous fundus images is almost three times higher than the normal class, the Adaptive Synthetic (ADASYN) algorithm has been utilized to reduce the imbalance in the ORIGA dataset. An improved U-Net neural network titled CP-FD-UNet++ is used to extract the optic cup and disc segmentation features. The algorithm implemented on the open-source ORIGA dataset gave a better performance than other existing methods with accuracy and AUC of 0.843 and 0.901, respectively. H. Fu *et al.* implemented a new deep learning technique called novel disc-aware ensemble network (DENet) for automatic glaucoma imaging in [26]. The DENet architecture includes four deep channels corresponding to the different levels and components of the color fundus photographs. This work used three distinct datasets (ORIGA, SCES and SINDI) for glaucoma classification. The ORIGA dataset was used for training the network and the authors used the private SCES and SINDI datasets for testing. They performed seven different glaucoma inspection approaches on SCES and SINDI datasets, i.e., Airpuff IOP, Wavelet, Gabor, GRI, Superpixel, DeepCDR, and DENet. The proposed glaucoma screening method, DENet achieved the highest accuracy on both datasets with 0.91 and 0.81 AUC scores on the SCES and SINDI datasets, respectively. Recently R. Zhao *et al.* have designed two methods to estimate the CDR values as the parameter works as a vital indicator in glaucoma screening from fundus images. The authors have demonstrated the efficiency of their proposed system on the private Direct-CSU (collected from a public hospital of Hunan, China) and public ORIGA datasets to train their model. The first method that the authors have used is a segmentation-based technique. In this technique, they first separated the VCD and the VDD and then took the geometry measurement to estimate the ratio. The second approach is a semi-supervised approach where they have used MFPPNet CNN architecture and combined a random forest model by applying a regression approach to separate the VCD and the VDD. Both the models showed excellent results, and their proposed architecture achieved an AUC score of 0.90 on the Direct-CSU and 0.80 on the ORIGA datasets.

A small number of works investigate the automatic identification of glaucoma by employing unconventional datasets and detection techniques. In a recent work [27], the authors established a completely new large-scale glaucoma (LAG) dataset consisting of 11,760 images of the fundus corresponding to 4,878 samples of glaucoma and 6,882 nonglaucoma/healthy classes utilizing human attention levels. The paper is mentioned as unconventional because it utilizes an attention-based glaucoma database rather than using fundus images. This paper proposed a unique attention-based convolutional neural network named AG-CNN to automatically recognize glaucoma and localize pathological areas on the fundus images. The experiment established accuracy of 96.2% and AUC of 0.983 over the validation set.

III. MATERIALS AND METHODS

A. DATASET

A significant contribution of this work is to present unique retinal images from a specialized hospital of all the ocular diseases located in Dhaka, Bangladesh, Bangladesh Eye Hospital (BEH). The color fundus images were taken by a Topcon Retinal Camera TRC-50DX, which is considered the gold standard for retinal imaging [28]. The retinal images were obtained over a timespan of two years, i.e., from 2019 to 2020, and from various Bangladeshi patients aged between 35 to 80. The diagnosis of glaucoma in the optic nerve head (ONH) was performed by two ocular disease professionals, a pediatric ophthalmologist and a glaucoma and refractive surgeon. Finally, the obtained dataset contains 463 normal (nonglaucoma) and 171 glaucoma color fundus images. Figure 2 illustrates some examples of retinal fundus images from the collected BEH dataset.

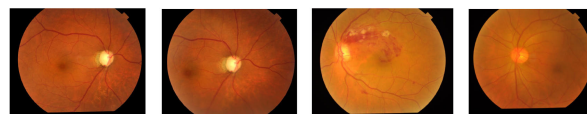


FIGURE 2. Samples of BEH dataset fundus images.

Two different models were prepared for this work by incorporating color fundus samples from Bangladesh Eye Hospital (BEH) dataset and the open-source ACRIMA dataset, which consists of 705 ocular fundus photographs [21]. The information of these constructed combined datasets used in this work has been described in the subsequent paragraphs.

1) DATASET-1

Dataset-1 comprises cropped fundus images from the BEH dataset and ACRIMA dataset. The fundus images are cropped to contain the cup and disc portion of the entire color fundus photographs. We used an online tool, “BIRME” to crop the fundus images in large batches. This online tool facilitates us to crop and resize bulk fundus images efficiently. Images were cropped 300×300 pixels in terms of height and width. The cup and disc are placed in the middle, and the CDR ratio is set to 1:1. Dataset-1 consists of 210 glaucoma samples (69 samples from the BEH dataset and 141 samples from the ACRIMA dataset) and 369 normal samples (319 and 50 samples from the BEH and ACRIMA dataset, respectively). Figure 3 illustrates samples of Dataset-1 fundus images.

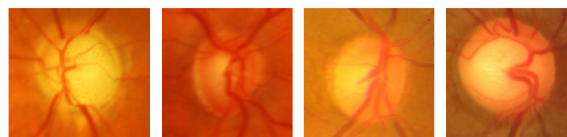


FIGURE 3. Samples of Dataset-1 fundus images.

2) DATASET-2

Dataset-2 is an updated version of Dataset-1, constituting cropped and blood vessel segmented fundus images from the

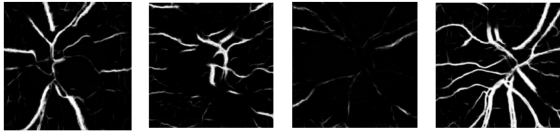


FIGURE 4. Samples of Dataset-2 blood vessel segmented fundus images.

BEH dataset and ACRIMA dataset. The fundus images are cropped to contain the cup and disc portion of the fundus image. Dataset-2 comprises 210 glaucoma samples (69 samples from the BEH dataset and 141 samples from ACRIMA dataset) and 369 normal samples (319 samples from BEH and 50 samples from ACRIMA dataset). The samples in Dataset-2 were blood vessel segmented fundus images using a U-net model that was trained on a dataset of High-Resolution Fundus (HRF) Image Database [29], which has been demonstrated in Figure 4.

B. DATA PREPROCESSING

In this research, the datasets were preprocessed with different conventional augmentation techniques. Dataset-1 was preprocessed with various augmentation techniques to increase the dataset size and gain balance and variability in the dataset. Brightness and contrast augmentation is essential as there was a deviation in the saturation, contrast, and color temperature among the fundus images. The dataset contains fundus images from both the left and right eye, which differentiates the orientation of the blood vessel. Random horizontal flip was used to reduce the influence of left/right distinction. Additionally, random rotation was applied to augment the fundus images, so it is less susceptible to the orientation of the cup, disc, and blood vessels. The augmentation techniques used in this research with the values of its various parameters are illustrated in Table 1.

TABLE 1. Augmentation techniques and its parameters.

Augmentation technique	Parameter
Random brightness adjustment	-
Random contrast adjustment	-
Random horizontal flip	probability: 0.50
Random vertical flip	probability: 0.50
Random rotation	range: [-45°, +45°]
Normalization	mean:[0.25, 0.25, 0.25], st. dev.:[0.90, 0.90, 0.90]



FIGURE 5. Samples of augmented fundus images from Datasets 1 and 2.

Dataset-1 contained 424, 40, and 115 fundus image samples in the training, validation, and test set, respectively. Dataset-1 was augmented by a factor of 15 for the glaucoma

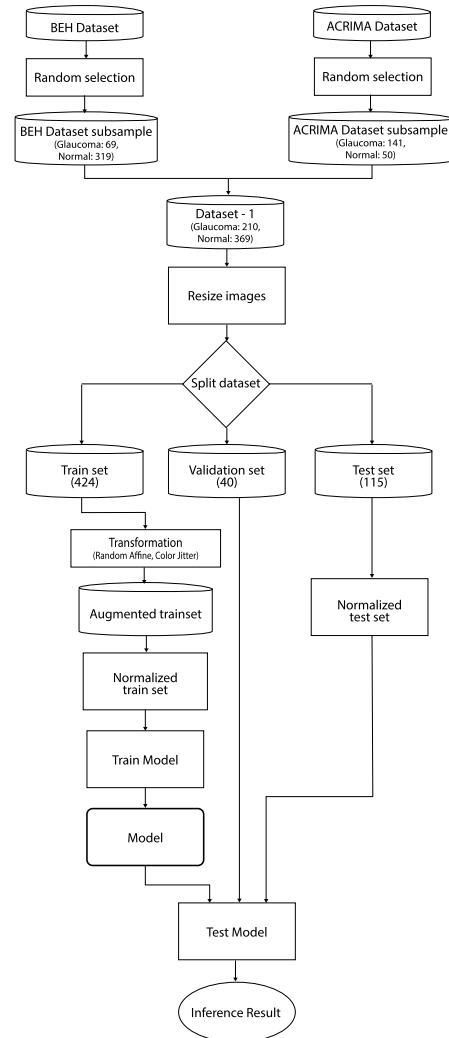


FIGURE 6. Architecture of model for Dataset-1.

class and by a factor of 11 for the normal class. Finally, the augmented Dataset-1 contains 5236 fundus image samples (glaucoma class: 2200 and normal class: 3036) in the train set. The models were trained on 424 samples per epoch. The entire working sequences for the training, validation, and test fundus images for Dataset-1 has been illustrated in Figure 6.

Next, Dataset-2 was preprocessed with the same augmentation techniques and parameters as Dataset-1. Dataset-2 contains 424, 40, and 115 fundus image samples in the training, validation, and test set, respectively. The images were augmented by a factor of 15 for the glaucoma class and 11 for the normal class. Lastly, the augmented Dataset-2 contains 5236 fundus image samples (glaucoma class: 2200 and normal class: 3036) in the training set. The models were trained on 424 samples per epoch. The working flowchart for Dataset-2 has been portrayed in Figure 7.

It is worth mentioning that, only the glaucoma samples were augmented by varying different factors to balance glaucoma and normal classes for various datasets used in this work. Samples of augmented fundus images from Datasets - 1 and 2 have been shown in Figure 5.

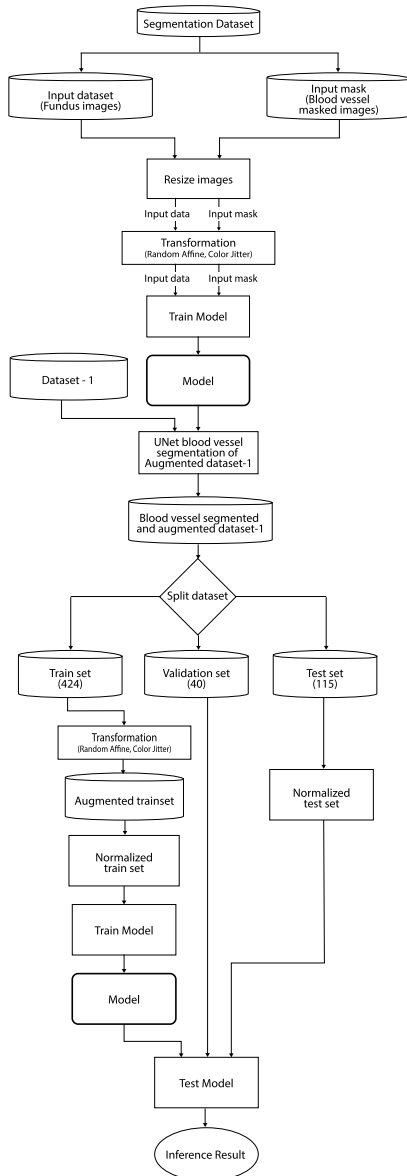


FIGURE 7. Architecture of model for Dataset-2.

Finally, Table 2 illustrates the number of normal/healthy and glaucoma images of Datasets - 1 and 2 before and after the applied augmentation techniques.

C. SEGMENTATION SYSTEM

Segmentation of retinal blood vessels is considered an effective technique for diagnosing ocular diseases, especially diabetic retinopathy detection [30], [31]. In this paper, segmentation of blood vessels was attempted by two convolutional network architectures, semantic pixel-wise segmentation (SegNet) and U-net. SegNet consists of subsequent encoding and decoding networks, which is widely used for semantic segmentation [32]. Unfortunately, the SegNet network overfitted the segmentation process and did not achieve satisfactory results. The U-net technique obtained convincing results with better classification accuracy. That’s why the

TABLE 2. Number of normal and glaucoma images of Datasets 1 and 2 before and after augmentation.

Dataset	Train		Validation		Test	
	Glaucoma	Normal	Glaucoma	Normal	Glaucoma	Normal
Dataset-1	148	276	20	20	42	73
Dataset-1 Augmented	2220	3036	20	20	42	73
Dataset-2	148	276	20	20	42	73
Dataset-2 Augmented	2220	3036	20	20	42	73
HRF Dataset	7	7	2	2	6	6

results for the U-net blood vessel segmentation method have been reported in this paper.

1) U-NET

U-net [33] is a powerful CNN technique that is widely used for semantic segmentation. It consists of a sequence of convolutional layers followed by nonlinear activation function (ReLU) and max pooling layers, and finally, the architecture follows an encoder-decoder type model. As we move deeper into the model architecture, it contracts the spatial features by down-sampling. It tries to map all the features to achieve a single output vector for detected segments. It also has an additional expansion path that performs up-convolutions and concatenation that uses the corresponding cropped feature map with high-resolution features from the contraction path to generate the final output segmentation map [34]. It has 23 convolutional layers, and the final layer uses a 1 by 1 convolution to map each 32 component feature vectors to the desired number of classes. The U-net architecture used to create the blood vessel segmented dataset in this work contains 32 filters, 4 layers and a spatial dropout rate of 0.3. For training the U-net model, Adam optimizer and ReLU activation function are utilized, and binary cross-entropy is employed as the loss function, which is expressed as:

$$L_{bce} = \sum_i y_i \log o_i + (1 - y_i) \log (1 - o_i) \quad (1)$$

where $i = 2$ for binary classification problems, y_i denotes model output’s i^{th} predicted scalar value, and the corresponding true label is depicted as y_i .

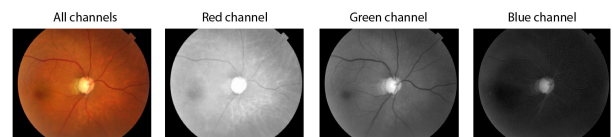


FIGURE 8. Fundus image separated into red, green, and blue color channels.

As stated earlier, the U-net model was trained on the HRF dataset. The HRF dataset contains 45 full-size fundus images, divided into fundus images of healthy/nonglaucoma, glaucoma affected, and diabetic retinopathy patients. The dataset also includes corresponding blood vessel segmented masks

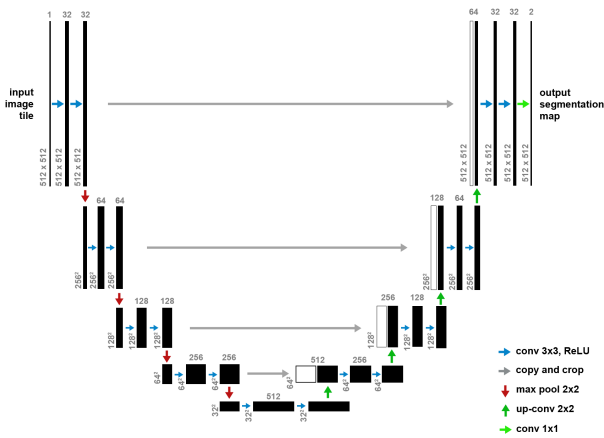


FIGURE 9. U-net architecture used for segmenting blood vessels from fundus images.

and eye masks for all of the fundus images. The 45 full-size fundus images and their corresponding blood vessel segmented images were selected for training the U-net model. When the fundus images were separated into individual red, green, and blue channels (as depicted in Figure 8), the blood vessels were found to be most prominent on the green channel. Therefore, the green channel was selected to be the input for the U-net model. Finally, the input fundus images and their corresponding masks were resized to 512×512 in dimension. The U-net architecture used for segmenting blood vessels from fundus images has been illustrated in Figure 9.

TABLE 3. Hyperparameters used to augment input and segmented images for training the U-net model.

Augmentation technique	Parameter
Random rotation	range: $[-5^\circ, +5^\circ]$
Random width shift	range: 3.5
Random height shift	range: 3.5
Random shear	range: 40
Random zoom	range: [0.4, 0.5]
Random horizontal flip	probability: 0.5

Table 3 demonstrates the hyper-parameters utilized to augment input and segmented images for training the U-net model. In this research, the U-net model was trained for a total of 22 epochs with 100 steps per epoch.

TABLE 4. Hyperparameters used to train the U-net model.

Parameter	Value
No. of filters	32
No. of layers	4
Dropout	0.3
Dropout change per layer	0.0
Batch normalization	True

Different hyperparameters used to train the U-net model have been exhibited in Table 4.

Figures 10 and 11 demonstrate the samples of original and cropped input fundus images and the corresponding

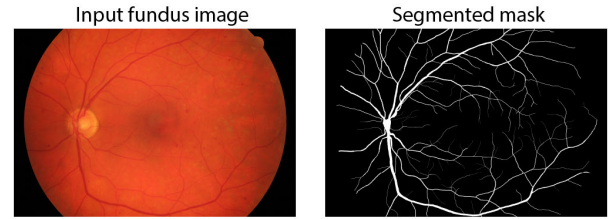


FIGURE 10. Sample of input fundus image and the corresponding segmented fundus image from the HRF dataset.

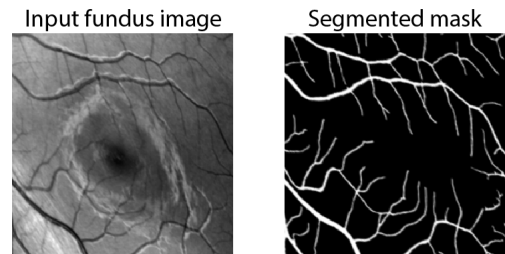


FIGURE 11. Sample of cropped input fundus image from HRF dataset and the corresponding segmented fundus image.

segmented fundus images from the HRF dataset, respectively.

$$IoU = \frac{\text{Area of intersection}}{\text{Area of union}} \quad (2)$$

Intersection over Union (IoU) is a widely used evaluation metric used for measuring the accuracy of annotation in image segmentation, object detection, and object localization, which is expressed in (2). IoU provides a definitive metric of how effectively the U-net model is able to segment the blood vessels from the fundus photographs.

The images of the dataset were randomly cropped to make it similar to Datasets - 1 and 2. Next, the dataset was also augmented to improve the model performance. Finally, after training for 22 epochs, the IoU, validation IoU, IoU threshold, and validation IoU threshold were 0.3397, 0.3472, 0.4198, and 0.5139, respectively.

Samples of input fundus image (green channel), segmented fundus image obtained from the U-net model, and overlay of the segmentation on the input fundus image have been depicted in Figure 12.

D. CLASSIFICATION SYSTEM

1) MobileNet

MobileNet is a CNN architecture with a reduced number of parameters but high classification accuracy optimized for mobile computer vision models [35]. A typical network has a computational price of 300 million multiply-adds and utilizes 3.4 million parameters [14]. In this work, we use MobileNet v2 because it outperforms MobileNet v1 with significant model size and computational cost. The architecture is built with a combination of 3×3 downsampling convolutional layers to extract features and uses the average pooling layer before the fully classified layer. It combines depth-wise convolution and pointwise convolution to employ depth-wise

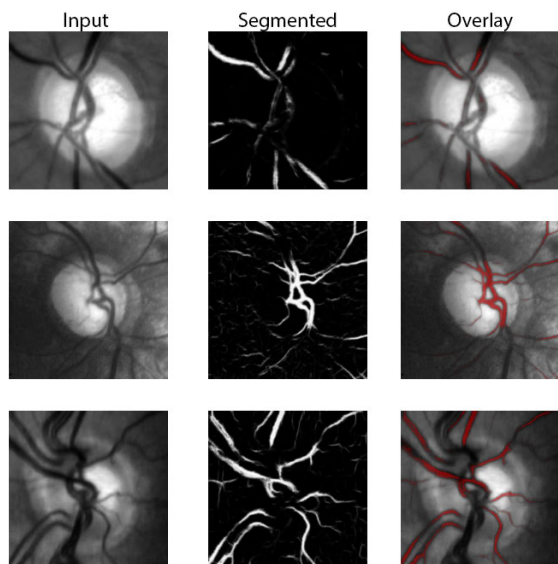


FIGURE 12. Samples of input fundus image (green channel), segmented fundus image, and overlay of the segmentation on the input fundus image.

separable convolutions architecture to reduce the number of parameters extensively. MobileNet algorithm basically tries to separate the filter's depth and spatial dimension. For this reason, it is an excellent choice for training models with a restricted resource for an on-device or embedded application.

2) EfficientNet

EfficientNet [36] is the CNN architecture that significantly outperformed other popular CNN models when trained on ImageNet Dataset. It has a complicated architecture with a stack of convolutional layers and inverted bottleneck MBConv layers [37]. The MBConv layers also apply the scaling method to uniformly scale the network width, depth, and resolution based on a set of scaling coefficients. This unique operation allows the model to use more computational resources by controlling the parameters alpha, beta, and gamma to increase depth, width, and image size, respectively. The MBConv layer is a layer having an inverted residual bottleneck, which is designed with the combination of normal convolutional layers with pooling and depthwise separable convolution layer. The EfficientNet has a compound coefficient ϕ , used to scale the network in all dimensions uniformly. If the image size is larger than the model will need more layers to detect fine-grained patterns. Thus, increasing the ϕ values allows us to scale all three dimensions while maintaining a balance among all the network dimensions. The model has a total of eight variations from $b_0 - b_7$. The base model b_0 is almost similar to the MNasNet architecture, and b_0 is scaled up using the compound scaling to achieve b_1 to b_7 . For training the glaucoma classification data, pre-trained EfficientNet b3 was used in this research. EfficientNet b3 is composed of 7 blocks, with 40, 24, 32, 48, 96, 136, 232, 384, 1536 channels for stage 1 to 9, respectively. EfficientNet b0, b3 and b7 consist of 5.3 million, 12 million and 66 million

TABLE 5. Parameters used to train the EfficientNet b3 model.

Parameter	Value
Width coefficient	1.2
Depth coefficient	1.4
Resolution	300
Dropout rate	0.3

parameters, respectively. Finally, EfficientNet b3 was chosen as training time is lower than EfficientNet b7.

Table 5 shows the parameters for the EfficientNet b3 model.

3) DenseNet

DenseNet is a complex convolutional neural network that employs a dense structure to link all layers effectively to each other, resulting in dense interconnections between layers [38]. Each layer takes extra inputs from all prior levels and passes on its own feature maps to all following levels to maintain the feed-forward nature. DenseNet provides several compelling advantages, including eliminating the vanishing-gradient problem, improved feature propagation, feature repetition, and a significant reduction in the number of parameters. Also, DenseNet produces considerable gains over the state-of-the-art deep learning models while using minimal computation to reach excellent performance.

4) GoogLeNet

In 2014 Google published its individual network that proved to perform better than VGG in the ImageNet database. GoogLeNet is built on inception models, which creates a situation where the total number of parameters is minimal [39]. Instead of having a single convolution in an inception module, it includes the composition of convolution of different filter sizes and average pooling. Finally, the output of this model is the concatenation of all these compositions. GoogLeNet has 22 layers in its architecture.

IV. RESULTS AND DISCUSSION

This section of the paper discusses the results of the proposed glaucoma detection algorithm. In this work, two distinct datasets (Datasets 1 and 2) have been created from the cropped optic cup and disc and blood vessel segmented color fundus images of the private BEH and public ACRIMA databases. U-net, the robust CNN architecture, has been utilized to construct the blood vessel segmented dataset. Finally, a wide range of CNN approaches, e.g., MobileNet, EfficientNet, DenseNet, and GoogLeNet, have been used to classify the images into two classes.

The following hyper-parameters were used for training all the models on Datasets - 1 and 2, which have been demonstrated in Table 6. For training the models on Datasets - 1 and 2, Adam optimizer with the parameters mentioned above in Table 6 was used, and cross-entropy loss function was used as the loss function.

TABLE 6. Hyper-parameters used to train all the models on Datasets 1 and 2.

Parameter	Value
Batch-size	13
Learning rate	0.001
Learning rate decay	0.96
Batch normalization	True

In Dataset - 1 and Dataset - 2, the full-size fundus images were cropped to only the cup and disc portion. Initial attempts to train different neural network architectures on full-size fundus images were unsatisfactory. Doing so would result in the model over-fitting to the training dataset and poor accuracy in the test dataset. According to ophthalmologists and other research papers, it is possible to detect whether glaucoma is present from the cup and disc segmentation of the fundus image [18], [22]. Training different neural network architectures on the cropped fundus images proved to be satisfactory in diagnosing glaucoma.

Finally, to validate the efficiency of the proposed system of automatic glaucoma detection, various performance metrics are evaluated, e.g., confusion matrix, validation accuracy graph, precision, recall, F1-score, TPR vs. FPR graph, train/validation/test accuracy, ROC AUC, etc. The mentioned metrics are extensively used to evaluate and assess computer vision and deep learning-based classification problems [40]. Specificity or true negative rate indicates the fraction of negatives that are accurately inferred, which is expressed as:

$$specificity = \frac{TN}{TN + FP} \tag{3}$$

Precision or positive predictive value indicates the proportion of positive and negative inferences, which can be calculated as:

$$precision = \frac{TP}{TP + FP} \tag{4}$$

Recall or sensitivity or true positive rate indicates the proportion of positives that are correctly inferred.

$$recall = \frac{TP}{TP + FN} \tag{5}$$

For any classification problem, accuracy is defined as the ratio of the correct predictions to the total number of cases, which is articulated as:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

The balance between the precision and recall is measured by the F1-score, which is expressed as:

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \tag{7}$$

The true positive rate versus the false positive rate is plotted in the ROC curve for various threshold values. The AUC coefficient measures the total area of the ROC curve.

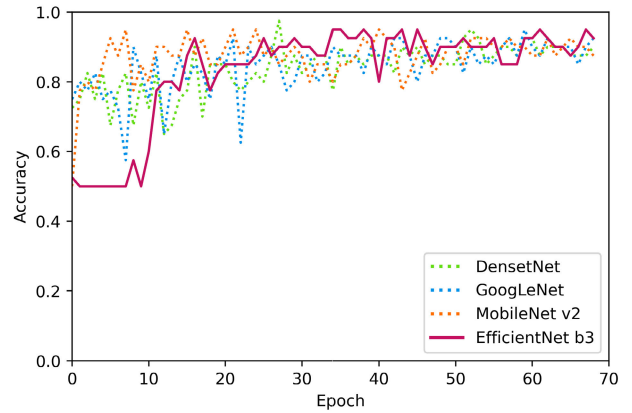


FIGURE 13. Validation accuracy graph of different models trained on Dataset-1.

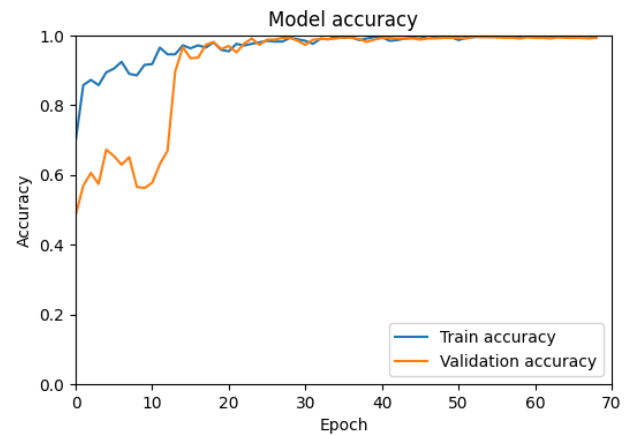


FIGURE 14. Train-validation accuracy graph of EfficientNet b3 model trained on Dataset-1.

A. DATASET-1 MODEL

The input images were resized, normalized, and split into the training, validation, and test set. Dataset-1 was used to train with DenseNet, GoogleLeNet, MobileNet v2, and EfficientNet b3 networks. The neural network models were trained for 70 epochs with the parameters and configurations mentioned earlier. The results of validation accuracies versus epochs for different models trained on Dataset-1 can be seen in Figure 13. EfficientNet b3 delivered the best test result with a training accuracy of 0.9929, validation accuracy of 0.9500, and test accuracy of 0.9652. The model had a ROC AUC score of 0.9574 and a specificity of 0.9863. The glaucoma class had a precision of 0.9750 and recall of 0.9286. The nonglaucoma class had precision and recall of 0.9600 and 0.9863, respectively. The training and validation accuracies of the EfficientNet b3 model trained on Dataset-1 have been depicted in Figure 14.

Figure 15 displays the confusion matrix of the EfficientNet b3 classifier trained on Dataset-1. It can be observed from the figure that, from a total of 42 glaucoma images, 39 and 3 cases are classified correctly and incorrectly, respectively. Conversely, the normal (nonglaucoma) cases classification performs better, i.e., only one image is classified incorrectly out of 72 images.

		Predicted Label	
		Glaucoma	Normal
Actual Label	Normal Glaucoma	39	3
	Normal	1	72

FIGURE 15. Confusion matrix of EfficientNet b3 model trained on Dataset-1.

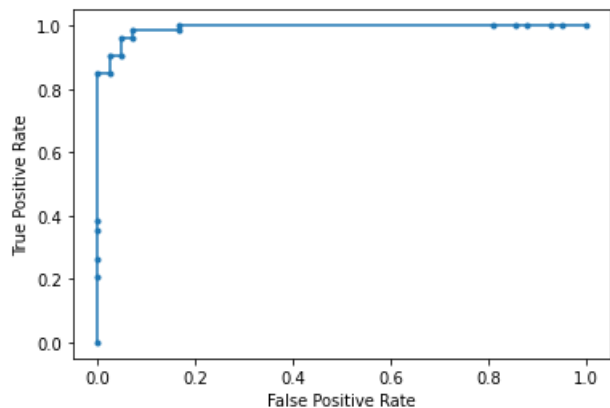


FIGURE 16. TPR vs. FPR graph of EfficientNet b3 model trained on Dataset-1.

Next, the TPR vs. FPR graph (ROC curve) of the EfficientNet b3 model trained on Dataset-1 has been plotted in Figure 16.

Finally, Table 7 illustrates various performance metrics of different CNN methods for Dataset-1, e.g., training, validation, and test accuracy, ROC AUC, precision, recall, and F1-score of normal and glaucoma classes.

B. DATASET-2 MODEL

Dataset-2 is similar to Dataset-1 in terms of size of training, validation, and test set. It is worth mentioning that, Dataset-2 is a blood vessel segmented variant of Dataset-1. Dataset-2 contains cropped and blood vessel segmented fundus image samples.

The HRF dataset was used to train a U-net model to segment the blood vessels. The U-net model takes as input one channel of full or cropped fundus images of dimension (512, 512, 1), and the model generates a blood vessel segmented image of dimension (512, 512, 1). Samples from Dataset-1 were segmented with the U-net model to create Dataset-2.

Blood vessel segmentation of full fundus image and cropped fundus image was satisfactory. The U-net model was able to segment the blood vessels from the fundus images. However, attempts to perform cup-disc segmentation were unsatisfactory; the U-net model was unable to precisely segment the cup and disc section of the fundus images.

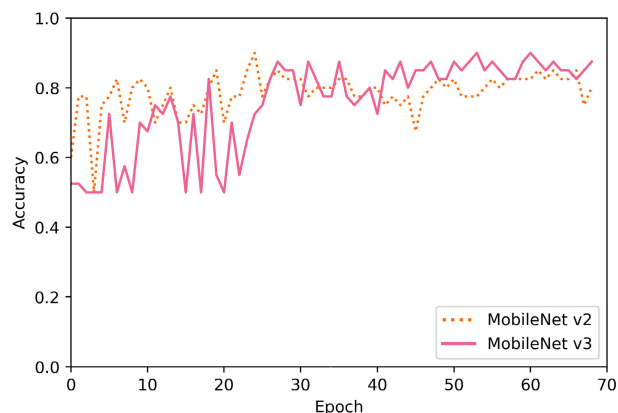


FIGURE 17. Validation accuracy graph of different models trained on Dataset-2.

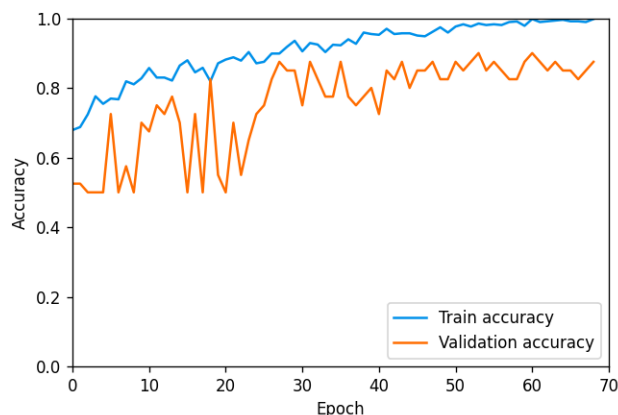


FIGURE 18. Train-validation accuracy graph of MobileNet v3 model trained on Dataset-2.

Performance results achieved by MobileNet v2, and MobileNet v3 models for Dataset-2 and HRF dataset have been demonstrated in Table 8.

The test accuracy vs. epoch graph on Dataset - 2 with MobileNet v2 and MobileNet v3 models, can be observed in Figure 17. MobileNet v3 delivered the best performance, i.e., training, validation, and test accuracies of 0.9976, 0.9000, and 0.8348, respectively. The model had a ROC AUC score of 0.8446 and a specificity of 0.8082. The glaucoma class had a precision of 0.7255 and a recall of 0.8810. The healthy class delivers precision and recall of 0.9219 and 0.8082, respectively. Figure 18 illustrates the training and validation accuracy graph of the MobileNet v3 model trained on Dataset-2.

It should be noted that Dataset-2 was also trained on DenseNet, GoogLeNet, SqueezeNet, InceptionNet, small AlexNet, and VGG architectures. The test accuracies of these models were not as satisfactory as MobileNet v3 (Large), and that’s why these results are not reported in the paper. Except for GoogLeNet, the accuracy of the model was approaching that of MobileNet v3 (Large).

The confusion matrix of the MobileNet v3 classifier trained on Dataset-2 has been represented in Figure 19. It can be observed that, glaucoma and normal/healthy images have been correctly classified by a percentage of approximately

TABLE 7. Illustration of various performance metrics for Dataset-1.

Dataset	Architecture	DenseNet	GoogLeNet	MobileNet v2	EfficientNet b3
Dataset-1	Train accuracy	0.9552	0.9929	0.9882	0.9929
	Validation accuracy	0.9500	0.9500	0.9500	0.9500
	Test accuracy	0.8696	0.9652	0.9304	0.9652
	ROC AUC	0.8417	0.9524	0.9098	0.9574
	Precision (glaucoma)	0.8857	1.0000	0.9722	0.9750
	Recall (glaucoma)	0.7381	0.9048	0.8333	0.9286
	F1-score (glaucoma)	0.8052	0.9500	0.8974	0.9512
	Precision (normal)	0.8625	0.9481	0.9114	0.9600
	Recall (normal)	0.9452	1.0000	0.9863	0.9863
	F1-score (normal)	0.9020	0.9733	0.9474	0.9730

TABLE 8. Performance results of MobileNet v2 and MobileNet v3 models for Dataset-2 and HRF Dataset.

Dataset	Architecture	MobileNet v2	MobileNet v3
Dataset-2	Train accuracy	0.9717	0.9976
	Validation accuracy	0.8500	0.9000
	Test accuracy	0.7652	0.8348
	ROC AUC	0.7038	0.8446
	Precision (glaucoma)	0.8000	0.7255
	Recall (glaucoma)	0.4762	0.8810
	F1-score (glaucoma)	0.5970	0.7957
	Precision (normal)	0.7556	0.9219
	Recall (normal)	0.9315	0.8082
	F1-score (normal)	0.8344	0.8613
HRF Dataset	Train accuracy	1.0000	1.0000
	Validation accuracy	1.0000	0.7500
	Test accuracy	0.8333	0.8333
	ROC AUC	0.8333	0.8333
	Precision (glaucoma)	0.7500	0.7500
	Recall (glaucoma)	1.0000	1.0000
	F1-score (glaucoma)	0.8571	0.8571
	Precision (normal)	1.0000	1.0000
	Recall (normal)	0.6667	0.6667
	F1-score (normal)	0.8000	0.8000

		Predicted Label	
		Glaucoma	Normal
Actual Label	Normal Glaucoma	37	5
	Normal	14	59

FIGURE 19. MobileNet v3 model's confusion matrix trained on Dataset-2.

88% and 81%, respectively. Next, the ROC curve of the MobileNet v3 model trained on Dataset-2 has been exhibited in Figure 20.

The HRF dataset was also used to train a model with MobileNet v2 and MobileNet v3 architectures. The models were trained and tested on the small samples in the HRF dataset. The train, validation, and test samples contained 14, 4, and 12 images, respectively, which is shown in Table 2. The training samples were largely augmented to compensate for the limited number of samples.

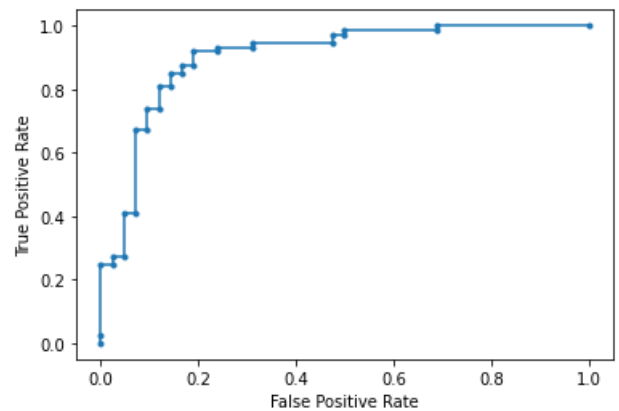


FIGURE 20. ROC curve of MobileNet v3 model trained on Dataset-2.

Test accuracies of 0.8333 and 0.8333 were achieved on MobileNet v2 and MobileNet v3, respectively, which is notable as the HRF dataset contains full-size fundus images. The MobileNet algorithms delivered similar results with a training accuracy of 1.0000 and validation accuracy of 1.0000 and 0.7500. The models both had ROC AUC score of 0.8333. For both models, the glaucoma class achieved precision and recall of 0.7500 and 1.0000, respectively. The normal class had a precision of 1.0000 and a recall of 0.6667. Finally, these convincing results inferred that it is possible to detect glaucoma from full-size fundus images by blood vessel segmentation.

Table 9 summarizes the performance metrics (AUC, accuracy, specificity and sensitivity) of the proposed system and other state-of-the-art glaucoma detection works on different open-source datasets. The implemented system of this work, EfficientNet b3, outperformed most of the other methods, especially those that utilized a combination of multiple datasets.

C. TRAINING TIME RESULTS

The training of the models of this work has been executed on the Tesla T4 GPU of Google Colab with 15109 MiB memory. We trained for 70 epochs for both datasets. We obtained the

TABLE 9. Performance comparison of the proposed system with other existing works.

Author	Dataset	Network	AUC	Accuracy	Specificity	Sensitivity
[18]	Combining RIM-ONE and DRISHTI	MobileNet v2	0.93	0.86	0.82	0.89
[21]	HRF, DRISHTI-GS1, RIM-ONE, ACRIMA and sjchoi86-HRF	Xception	0.9605	0.8977	0.8580	0.9346
[26]	SCES Dataset	DENet	0.9183	0.8429	0.8380	0.8478
	SINDI Dataset	DENet	0.8173	0.7495	0.7115	0.7876
		DeepCDR	0.7929	0.7585	0.7647	0.7522
[19]	RIM-ONE and RIGA	SSCNN-DAE		0.938	0.989	0.905
[22]	RIM-ONE	Combination of DeepLabv3+ and MobileNet	0.100	0.9737		
	ORIGA					
	DRISTI-GS1					
	ACRIMA					
	REFUGE					
[25]	ORIGA	UNet++	0.901	0.843		0.894
[27]	LAG	AG-CNN	0.983	0.962	0.967	0.954
	RIM-ONE					
[41]	Direct-CSU	MFPPNet CNN	0.88			
	ORIGA					
[23]	RIM-ONE	SLIC Algorithm	0.91	0.986	0.923	0.976
[20]	DRISHTI-GS, RIM-ONE and other created from a glaucoma screening	VGG19 TL	0.94	88.05	0.8901	0.8701
This work	ACRIMA	EfficientNet b3	0.9891	0.9905	0.9783	1.000

TABLE 10. Comparison of elapsed training time for Datasets 1 and 2.

Dataset	Model	GPU	Time		
			Best validation accuracy epoch	Time elapsed	Overall time elapsed (Epoch 70)
Dataset-1	MobileNet v2	Tesla T4 GPU	53	474.217 s	620.807 s
Dataset-2	MobileNet v2	Memory 15109 MiB	60	317.225 s	371.154 s

best validation accuracy at epochs 53 and 60 for Dataset-1 and Dataset-2, respectively. According to Tabel 10, overall elapsed training time at epoch 70 for Dataset-1 and Dataset-2 are approximately 621 and 371 seconds, respectively. Dataset-2 was prepared by segmenting fundus images with a model of U-net architecture. The training time of the U-net model for 10 epochs on the HRF dataset is 1985 seconds. Finally, all the 5411 fundus images were segmented with the U-net model in approximately 3000 seconds. It is worth noting that the U-net model's training and segmentation of the fundus images were carried out on the same computer. We can examine from Table 10, Dataset-2 requires much lesser training time than Dataset-1 by a factor of approximately 0.60. From here, we can conclude that even though Dataset-1 has better accuracy and follows a much more conventional approach compared to Dataset-2, still if we want to deploy our proposed system in an embedded system, Dataset-2 can be an excellent option.

V. CONCLUSION

Glaucoma is an irreversible neurodegenerative illness that damages the optic nerve and is responsible for vision loss

and blindness. The conventional manual detection of glaucoma by eye specialists is costly and time-consuming and depends on human error, experienced ophthalmologists, and expensive instruments. This paper attempts to implement an automatic glaucoma diagnosis system based on deep learning approaches. This study uses a private dataset comprising 463 normal (nonglaucoma) and 171 glaucoma color fundus images, which have been collected from Bangladesh Eye Hospital (BEH), Dhaka. A set of cropped fundus images have been constructed from these private and public ACRIMA datasets to contain the optic cup and disc portion. Significantly, cropped and blood vessel segmented fundus images are created using a U-net model trained on High-Resolution Fundus (HRF) Image Database. Finally, multiple CNN approaches, e.g., MobileNet, EfficientNet, DenseNet, and GoogLeNet, have been used as the glaucoma classifier networks. The EfficientNet b3 model offers the best performance for the cropped fundus images with accuracy and ROC AUC of 0.9652 and 0.9512, respectively. Alternatively, the MobileNet v3 network exhibits the highest accuracy and ROC AUC for the blood vessel segmented fundus images. The impressive results obtained from the proposed system are believed to help ophthalmologists examine and detect glaucoma more quickly and economically. In the future, the proposed system can be turned into a more robust model by training with more data comprising public and private eye fundus images and incorporating synthetic images. The proposed system and eye fundus photographs can be utilized to diagnose more complex

ocular diseases, such as diabetic retinopathy, age-related macular degeneration (AMD), amblyopia. The glaucoma color fundus images from Bangladesh Eye Hospital (BEH) and the programing code implementations can be found at: <https://github.com/mirtanvirislam/Deep-Learning-Based-Glaucoma-Detection-with-Cropped-Optic-Cup-and-Disc-and-Blood-Vessel-Segmentation>.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ishtiaque Anwar (Glaucoma Surgeon, Bangladesh Eye Hospital, Dhaka, Bangladesh) and Dr. Shahedara Begum (Pediatric Ophthalmologist and Strabismus Surgeon, Bangladesh Eye Hospital) for their contributions to classify and labeling of the ocular color fundus images used in this research.

REFERENCES

- [1] P. J. Foster, R. Buhrmann, H. A. Quigley, and G. J. Johnson, "The definition and classification of glaucoma in prevalence surveys," *Brit. J. Ophthalmol.*, vol. 86, no. 2, pp. 238–242, Feb. 2002.
- [2] R. Weinreb, T. Aung, and F. Medeiros, "The pathophysiology and treatment of glaucoma: A review," *JAMA, J. Amer. Med. Assoc.*, vol. 311, pp. 1901–1911, May 2014.
- [3] N. Zhang, J. Wang, Y. Li, and B. Jiang, "Prevalence of primary open angle glaucoma in the last 20 years: A meta-analysis and systematic review," *Sci. Rep.*, vol. 11, no. 1, pp. 1–12, Dec. 2021.
- [4] X. Sun, Y. Dai, Y. Chen, D.-Y. Yu, S. J. Cringle, J. Chen, X. Kong, X. Wang, and C. Jiang, "Primary angle closure glaucoma: What we know and what we don't know," *Prog. Retinal Eye Res.*, vol. 57, pp. 26–45, Mar. 2017.
- [5] K. Allison, D. Patel, and O. Alabi, "Epidemiology of glaucoma: The past, present, and predictions for the future," *Cureus*, vol. 12, pp. 1–19, Nov. 2020.
- [6] L. Storgaard, T. L. Tran, J. C. Freiberg, A. S. Hauser, and M. Kolko, "Glaucoma clinical research: Trends in treatment strategies and drug development," *Frontiers Med.*, vol. 8, p. 1492, Sep. 2021.
- [7] A. S. Barbone, M. Meftah, K. Markiewicz, and K. Dellimore, "Beyond wearables and implantables: A scoping review of insertable medical devices," *Biomed. Phys. Eng. Exp.*, vol. 5, no. 6, Oct. 2019, Art. no. 062002.
- [8] L. Cohen and L. Pasquale, "Clinical characteristics and current treatment of glaucoma," *Cold Spring Harbor Perspect. Med.*, vol. 4, pp. 1–16, Jun. 2014.
- [9] B. Cvenkel and M. Kolko, "Current medical therapy and future trends in the management of glaucoma treatment," *J. Ophthalmol.*, vol. 2020, pp. 1–14, Jul. 2020.
- [10] A. Heijl, "Glaucoma treatment: By the highest level of evidence," *Lancet*, vol. 133, pp. 1–3, Apr. 2014.
- [11] N. Grabow, V. Senz, and K.-P. Schmitz, "Medical device innovations for cardiovascular, ophthalmologic and otolaryngologic applications," *Current Directions Biomed. Eng.*, vol. 5, no. 1, pp. 117–118, Sep. 2019.
- [12] E. Ansari, J. Pavicic-Astalos, F. Ayan, A. King, M. Kinsella, E. Ng, and A. Nita, "Treatment of open-angle glaucoma and ocular hypertension with preservative-free tafluprost/timolol fixed-dose combination therapy: UK and Ireland results from the visionary study," *Adv. Therapy*, vol. 38, pp. 1436–1451, Apr. 2021.
- [13] A. Thompson, A. Jammal, and F. Medeiros, "A review of deep learning for screening, diagnosis, and detection of glaucoma progression," *Transl. Vis. Sci. Technol.*, vol. 9, pp. 42–60, Jan. 2020.
- [14] D. Mirzania, A. C. Thompson, and K. W. Muir, "Applications of deep learning in detection of glaucoma: A systematic review," *Eur. J. Ophthalmol.*, vol. 31, no. 4, pp. 1618–1642, Jul. 2021.
- [15] N. H. Butt, M. H. Ayub, and M. H. Ali, "Challenges in the management of glaucoma in developing countries," *Taiwan J. Ophthalmol.*, vol. 6, no. 3, pp. 119–122, Sep. 2016.
- [16] S. Kumar, A. Giubilato, W. Morgan, L. Jitskaia, C. Barry, M. Bulsara, I. J. Constable, and K. Yogesani, "Glaucoma screening: Analysis of conventional and telemedicine-friendly devices," *Clin. Experim. Ophthalmol.*, vol. 35, no. 3, pp. 237–243, Apr. 2007.
- [17] A. S. Mursch-Edlmayr, W. S. Ng, A. Diniz-Filho, D. C. Sousa, L. Arnold, M. B. Schlenker, K. Due nas-Angeles, P. A. Keane, J. G. Crowston, and H. Jayaram, "Artificial intelligence algorithms to diagnose glaucoma and detect glaucoma progression: Translation to clinical practice," *Transl. Vis. Sci. Technol.*, vol. 9, pp. 1–21, Jan. 2020.
- [18] J. Civit-Masot, M. J. Dominguez-Morales, S. Vicente-Diaz, and A. Civit, "Dual machine-learning system to aid glaucoma diagnosis using disc and cup feature extraction," *IEEE Access*, vol. 8, pp. 127519–127529, 2020.
- [19] M. Alghamdi and M. Abdel-Mottaleb, "A comparative study of deep learning models for diagnosing glaucoma from fundus images," *IEEE Access*, vol. 9, pp. 23894–23906, 2021.
- [20] J. J. Gómez-Valverde, A. Antón, G. Fatti, B. Liefers, A. Herranz, A. Santos, C. I. Sánchez, and M. J. Ledesma-Carbayo, "Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning," *Biomed. Opt. Exp.*, vol. 10, no. 2, pp. 892–913, 2019.
- [21] A. Diaz-Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: An extensive validation," *Biomed. Eng. OnLine*, vol. 18, no. 1, pp. 1–19, Dec. 2019.
- [22] S. Sreng, N. Maneerat, K. Hamamoto, and K. Y. Win, "Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images," *Appl. Sci.*, vol. 10, no. 14, p. 4916, Jul. 2020.
- [23] N. A. Mohamed, M. A. Zulkifley, W. M. D. W. Zaki, and A. Hussain, "An automated glaucoma screening system using cup-to-disc ratio via simple linear iterative clustering superpixel approach," *Biomed. Signal Process. Control*, vol. 53, pp. 1–10, Aug. 2019.
- [24] A. A. Salam, T. Khalil, M. U. Akram, A. Jameel, and I. Basit, "Automated detection of glaucoma using structural and non structural features," *SpringerPlus*, vol. 5, no. 1, pp. 1–21, Dec. 2016.
- [25] F. Guo, W. Li, J. Tang, B. Zou, and Z. Fan, "Automated glaucoma screening method based on image segmentation and feature extraction," *Med. Biol. Eng. Comput.*, vol. 58, no. 10, pp. 2567–2586, Oct. 2020.
- [26] H. Fu, J. Cheng, Y. Xu, C. Zhang, D. Wong, J. Liu, and X. Cao, "Disc-aware ensemble network for glaucoma screening from fundus image," *IEEE Trans. Med. Imag.*, vol. 37, no. 11, pp. 2493–2501, Nov. 2018.
- [27] L. Li, M. Xu, H. Liu, Y. Li, X. Wang, L. Jiang, Z. Wang, X. Fan, and N. Wang, "A large-scale database and a CNN model for attention-based glaucoma detection," *IEEE Trans. Med. Imag.*, vol. 39, no. 2, pp. 413–424, Jul. 2020.
- [28] M. Kimura, S. Kusuhara, M. Tagami, and M. Nakamura, "Impaired retinal circulation during axitinib treatment for metastatic renal cell carcinoma," *Case Rep. Ophthalmology*, vol. 10, no. 1, pp. 5–10, Jan. 2019.
- [29] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," *Int. J. Biomed. Imag.*, vol. 2013, pp. 1–11, Dec. 2013.
- [30] M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. Rudnicka, C. Owen, and S. Barman, "Blood vessel segmentation methodologies in retinal images—A survey," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 407–433, 2012.
- [31] T. A. Soomro, A. J. Afifi, L. Zheng, S. Soomro, J. Gao, O. Hellwich, and M. Paul, "Deep learning models for retinal blood vessels segmentation: A review," *IEEE Access*, vol. 7, pp. 71696–71717, 2019.
- [32] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Jan. 2017.
- [33] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351, 2015, pp. 234–241.
- [34] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, vol. 37, Jul. 2015, pp. 448–456.
- [35] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.
- [36] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 2019, *arXiv:1905.11946*.
- [37] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, "Classification of remote sensing images using EfficientNet-B3 CNN model with attention," *IEEE Access*, vol. 9, pp. 14078–14094, 2021.
- [38] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

- [39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [40] G. Shao, L. Tang, and H. Zhang, "Introducing image classification efficacies," *IEEE Access*, vol. 9, pp. 134809–134816, 2021.
- [41] R. Zhao, X. Chen, X. Liu, Z. Chen, F. Guo, and S. Li, "Direct cup-to-disc ratio estimation for glaucoma screening via semi-supervised learning," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 4, pp. 1104–1113, Apr. 2020.



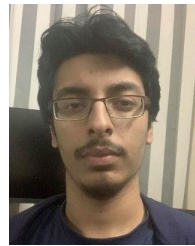
MIR TANVIR ISLAM received the B.Sc. degree in computer science and engineering from the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh. He is currently working as a Machine Learning Engineer with NybSys Ltd., Dhaka. His current research interests include computer vision, data science, machine learning, and computer networks security.



SHAFIN T. MASHFU received the Bachelor of Science degree in computer science and engineering from North South University, Dhaka, Bangladesh. He is currently a Data Analyst working at Sheba Platform Ltd. His research interests include big data analytics, data warehousing, ETL processes, and machine learning.



ABRAR FAISAL received the B.Sc. degree in computer science and engineering from the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh. His current research interests include computer vision, cyber security, machine learning, and web networks.



SADMAN CHOWDHURY SIAM received the bachelor's degree in computer science and engineering from North South University, Dhaka, Bangladesh. He is currently working on computer vision and face detection-related problems and recently published his work at a conference. His current research interests include deep learning, cybersecurity, and computer vision.



INTISAR TAHMID NAHEEN received the B.Sc. degree in electrical and electronics engineering from the Islamic University of Technology, Gazipur, Bangladesh, in 2014, and the M.Sc. degree in information and communication engineering from Technische Universität Darmstadt (TU Darmstadt), Germany, in 2018. Currently, he is working as a Lecturer with the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh. His current research interests include artificial intelligence (AI), reinforcement learning (RL), computer vision, 3D object detection, augmented reality, and application of machine learning in the power sector.



RIASAT KHAN (Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from the Islamic University of Technology, Bangladesh, in 2010, and the M.Sc. and Ph.D. degrees in electrical engineering from New Mexico State University, Las Cruces, USA, in 2018. Currently, he is working as an Assistant Professor with the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh. His research interests include biomedical engineering, cardiac electrophysiology, computational bioelectromagnetics, model order reduction, and power electronics.

...