CDED-Net: Joint Segmentation of Optic Disc and Optic Cup for Glaucoma Screening

MUNAZZA TABASSUM1, TARIQ M. KHAN2, (Member, IEEE), MUHAMMAD ARSALAN3, (Graduate Student Member, IEEE), SYED SAUD NAQVI2, MANSOOR AHMED1, HUSSAIN AHMED MADNI1, AND JAWAD MIRZA2, (Senior Member, IEEE)

1Department of Computer Science, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan
2Department of Electrical and Computer Engineering, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan
3Division of Electronics and Electrical Engineering, Dongguk University, Seoul 04620, South Korea

Corresponding authors: Munazza Tabassum (munazzatsd1993@gmail.com) and Tariq M. Khan (tariq045@gmail.com)

ABSTRACT Glaucoma is an eye disease that can cause loss of vision by damaging the optic nerve. It is the world’s second leading cause of blindness after cataracts. Early diagnosis of glaucoma is a key to prevent permanent blindness as it has no noticeable symptoms in its early stages. Color fundus photography is used for examining the optic disc (OD) which is an important step in the diagnoses of glaucoma. This is done by estimating the cup-to-disc ratio (CDR). In this paper, we proposed a Cup Disc Encoder Decoder Network (CDED-Net) for the joint segmentation of optic disc (OD) and optic cup (OC). We have eradicated the pre-processing and post-processing steps to reduce the computational cost of the overall system. Segmentation of (OD) and OC is modeled as a semantic pixel-wise labeling problem. The model was trained on the DRISHTI-GS, RIM-ONE and REFUGE datasets. Experiments show that our CDED-Net system achieves state-of-the-art OD and OC segmentation results on these datasets.

INDEX TERMS Glaucoma diagnosis, OD and OC segmentation, deep convolutional neural network, semantic segmentation.

I. INTRODUCTION

Glaucoma is a disease caused by the deterioration of nerve fibers that gradually damages the optic nerve, which is a part of an eye that carries impulses from the retina to the brain, interpreted as images. It is damaged by an increased intraocular pressure (IOP) which is fluid pressure in an inner portion of an eye. If the eye cannot drain excess fluid, pressure begins to build up, which damages the optic nerve fibers and thickens the layers of the retinal nerve fibers (RNFL). This causes the optic cup (OC) to become larger than optic disc (OD) usually known as cupping. It may also lead to the thinning and distortion of the retinal pigment epithelium in the area around the optic nerve, called peripapillary atrophy (PPA), usually associated with high myopia [1]. Glaucoma is the second largest cause of irreversible blindness in the world after cataracts [2]. Hence, early diagnose of glaucoma is necessary to avoid permanent blindness.

Glaucoma is mainly classified into two types, open-angle, and closed-angle glaucoma. Open-angle glaucoma is the most common type of glaucoma where the cornea-iris drainage angle remains open. Patients with this type of glaucoma have no early symptoms and are unaware of their conditions until significant vision is lost. Closed-angle glaucoma occurs when the drainage angle gets blocked by the parts of iris. Excessive fluid cannot be drained through the angle that leads to increased pressure within an eye. Patients with this type of glaucoma have noticeable symptoms that include redness of the eye, sudden ocular pain, elevated intraocular pressure, and sudden decreased vision [3].

Clinical glaucoma treatment includes various ophthalmologist examinations. Tonometry measures the increased pressure within an eye. Normal-pressure ranges between 12-22 mmHg. Glaucoma cases emerge when pressure exceeds 20mmHg. Ophthalmoscopy is used for examining the appearance of the optic nerve for diagnosing glaucoma. If there is any unusual change in optic nerve shape and color, patients will be referred to go through another test.
known as Perimetry. Perimetry test enables doctors to determine whether the vision of patients have been affected by glaucoma. Pachymetry test is used for examining corneal thickness as it can influence the pressure readings with an eye. Gonioscopy is another test performed by doctors to determine whether the angle for fluid drainage is open or narrow. Patients with narrow-angle have high risks of glaucoma. Physical analysis of an eye using these techniques require trained and professional experts. These techniques are expensive, time-consuming and can lead to inter-observer variability [4]. However, a computer-aided diagnosis system is a reliable and efficient method to facilitate glaucoma detection. For instance, Fundus photography can be used for structural analysis of obtained retinal images. It makes the diagnosis process cost-effective and robust. Moreover, using computer-aided tools can help to reduce the subjective analysis of retinal images by medical experts. Studies show different methods for early glaucoma diagnosis. Among which examining the optic nerve head is the most effective method for glaucoma diagnosis. ONH is divided into the OD and OC regions [5], [6].

An important parameter for detecting glaucoma is cup-to-disc ratio or cupping. This is the ratio between the vertical diameter of the OC to the vertical diameter of the OD [1]. A healthy eye has a normal cup size with several optical nerve fibers. But due to an increased IOP, nerve fibers start to degenerate and expand the area of the cup, causing an increase in CDR. CDR value for a normal eye is less than 0.6, whereas glaucoma eye has a CDR value greater than 0.6. Cup-to-disc ratio alone for detecting glaucoma is not sufficient, especially in cases where patients can genetically have a large OC or are diagnosed with myopia [7]. Another important factor suggested to detect glaucoma is observing the structure of neuroretinal rim which is a circular part between disc and cup comprising of nerve fibers. For screening of glaucoma, the ISNT rule is used, which states that in the normal eye, the width of the rim is thickest in the inferior region followed by superior, nasal, and temporal areas of the rim. But for glaucoma, this rule is violated as an increase in OC size decreases the NRR region [8].

Accurate segmentation of OD and OC is a very important task and may lead to several challenges. Fundus cameras may result in retinal images with poor contrast, low resolution, and varying illumination that have made the segmentation process more tedious. Noise is another factor that can be introduced in fundus images in the form of reflections and other bright objects, such as exudates, which appear as high-intensity bright objects similar to OD and OC. This can lead to the false detection of concerned candidates. Exact detection of OC boundary is also a difficult task due to blood vessels present in the layer beneath retina [9]. In some cases, noise can also be introduced in segmented images in the form of extra information that needs to be eliminated [10]. To cope with all these problems, existing studies have added pre-processing and post-processing stages along with segmentation. In addition to this, some authors have also considered localization as a pre-requisite in which region of interest (ROI) around the OD is extracted from images before the segmentation process [11]–[14]. This is not a generalized approach as it requires a bounding box of OD and OC and cannot be applied to unseen full eye fundus images. Performing these additional steps takes a lot of time which makes the processing slow and increases the overall cost of the system. Recent state-of-the-art methods ensure competitive performance on the task of joint OC and OD segmentation [15]–[18]. However, this performance is achieved through parameter heavy generative adversarial networks [15], [16] and ensemble networks [18]. The method of Fu et al. [17] employs a relatively less parameter heavy U-Net [19] based architecture. However, their method also relies on the accurate localization of the OD, which is achieved through a heuristic method. Therefore, the motivation of this research work is the segmentation of OD and OC region from the retinal fundus images using the deep learning-based model and to eliminate the additional processing blocks that increase the overall computational cost of the system.

In this work, a novel Cup Disc Encoder Decoder Network (CDED-Net) architecture is proposed for joint OC and OD segmentation. Owing to the efficient design of the encoder-decoder network, the proposed method is independent of the localization step and any pre- or post-processing steps. The encoder network is intelligently designed to promote feature propagation and reuse inspired by DenseNets [20]. This enables the proposed method to ensure boundary preservation resulting in robust segmentation. The decoder network inspired by the technique of Badrinarayanan et al. [21] allows the reuse of information from the encoder, thus alleviating the need to learn to upsample. This results in reducing the parameters of the network without comprising accuracy. The number of encoder-decoder layers is kept low to ensure semantic information preservation and parameter reduction. The result is an architecture with an order of magnitude fewer parameters as compared with the U-Net and DeepLab based state-of-the-art networks. The contributions of this work are three-fold:

1) We propose a novel encoder-decoder architecture that is computationally less expensive both in the training and test phases as compared with state-of-the-art approaches. This is achieved through the shallower network structure and reuse of information at the decoder stage. Furthermore, careful design of the encoder-decoder architecture eradicates the need for any pre- or post-processing steps.

2) We present an intelligent encoder design to ensure boundary preservation through the iterative restoration of semantic information at each encoder stage.

3) Our experimental results demonstrate noticeably better performance in terms of both OC and OD segmentation in comparison to recent state-of-the-art methods, hence enabling more robust Glaucoma evaluation.
The paper is organized as follows. Section 2 explains the relevant literature. In Section 3 and 4 methodology and network architecture for OD and OC segmentation are explained. Section 5 discusses the experimental results and detailed comparative analysis with state-of-the-art techniques. This is followed by a detailed discussion of the proposed approach in Section 6. Finally, the last section concludes the study in section 7.

II. RELATED WORK

For glaucoma diagnosis, segmenting OD and OC is a necessary step to determine the accurate cup-to-disc ratio. Existing literature has proposed several OD and OC segmentation approaches broadly classified as image processing and deep learning-based methods. Image processing techniques include thresholding and active contour-based methods, edge and region-based methods, pixels and superpixels approaches. Thresholding is the simplest technique for segmentation that is used to obtain a binary image. To get the segmented image, a value is selected as a threshold to allocate colors below and above thresholds. In [22], adaptive thresholding method is used for segmentation. To make segmentation independent of quality, the Otsu thresholding technique is used over the red channel image for the OD. While for the OC, green channel of ROI image is used and features are extracted and added which gives an intensity value that is considered as a threshold for segmentation.

In [23], an algorithm has been proposed to obtain ROI in both red and green channels of an image. The Gaussian window is used for plotting histogram against red channel image to obtain threshold for OD segmentation. A similar process is done with a green channel image for the OC. ISNT rule based on the width of the neuroretinal rim has been used for glaucoma screening in [7]. In [24], authors have proposed an approach using mean, median and Otsu thresholding for segmentation. But the color variation between patients reduces the robustness of such systems. In [9], the intensity-based thresholding technique was proposed for disc while geometrical based features have been used to remove noisy pixels.

For OC boundary, vessel tracking bending based approach has been used for tracking and observing sudden changes in blood vessels. This work achieves a high f-score value of 0.9485 indicating a good classification of fundus images. However, techniques based on thresholding are not efficient with low contrast images [25]. Some papers have used edge and region-based methods for segmentation as regional information gives robustness against intensity and contrast variations. To overcome the problems with large datasets, [26] has presented a regional information-based method using intensity normalization for obtaining cup boundary. [27] has proposed a method for eliminating peripapillary atrophy using different operations to obtain accurate disc boundary. This approach makes the segmentation more accurate as non-disc structures are excluded. Some papers have used active contours and deformable based algorithms for defining disc and cup boundaries by minimization of the energy function. An enhanced segmentation method was proposed for OD and OC in [28] by calculating CDR and CAR. To overcome classification errors, an adaptive deformable model has been proposed for capturing shape variation and irregularity in [29]. In [30], an active contour-based technique has been proposed for segmenting disc. Nonetheless, the active contour-based system can be stuck in a local minimum due to noise and pathologies present within an image, and performance is extremely dependent on contour model initialization [25]. Some authors have employed superpixel based techniques for segmentation. Another approach based on superpixel classification has been used by [31] for obtaining segmented disc and cup. But this proposed method has overestimated and underestimated the cup identification for small and large sizes. Preprocessing is an important module for providing good quality retinal images that can assist in the segmentation process. In the proposed work [8], fundus images were preprocessed to eliminate noise and illumination issues. For OD and OC segmentation, Simple Linear Iterative Clustering (SLIC) algorithm based on a superpixel approach has been used to obtain superpixels from input images, which were then classified into an OD, OC and background regions using a classifier. Observing the structure of the retinal nerve fiber layer (RNFL) is also considered as an indication of glaucoma [32]. A framework for detecting glaucoma using statistical features was suggested by [33] and K-nearest neighbor was used for image classification.

Machine learning-based techniques have shown good segmentation results. In [34] author has proposed a method to efficiently and accurately locate OD in images containing noise and other abrasions. An automated regression-based method has been proposed in [35] to obtain the correct boundary of the OC and OD. Machine learning approaches are highly dependent on the type of hand crafted features that have been manually extracted representing a specific dataset. This task of manually extracting features is tedious and time-consuming.

In recent years, deep learning-based approaches have been widely used that can automatically learn complex features through training. Authors in [36] have presented a reformed version of original U-Net CNN for OD and OC segmentation, where the input image is passed through the network’s contracting and expansive path with upsampling layers to increase image dimensions. This method results in good quality segmented OD and OC with the lowest prediction error. In [37], authors have designed an ensemble learning architecture inspired by a convolution neural network for segmentation of OD and OC. The entropy sampling approach was used for designing a learning architecture for convolutional filters by identifying the most informative points. This proposed work is useful when the small dataset is available. Authors in [37] have analyzed the capability of using off-the-shelf CNN architectures namely Overfeat and VGG-S as feature extractors. Preprocessing techniques have been
applied to fundus images that include contrast enhancement and vessel inpainting to analyze the performance of these networks. In [10], authors have proposed a fully convolutional neural network (FCN) that consists of a VGG-16 encoder and decoder with upsampling layers and skip connections for performing segmentation. A weighted loss with the mask was used to prioritize pixels and a filtering module was applied to crop disc and cup boundaries. A deep-learning based approach using M-Net architecture has been used in [17]. Original fundus images are transferred into the polar coordinate system using polar transformation, which is then fed into M-net which is a multi-label deep network to produce probability maps for regions containing disc and cup. This network was evaluated on the ORIGA dataset. A fully automated 18 layer CNN for feature extraction has been used in [4]. A method is proposed in [14] using DenseNet with a fully convolutional network FC-DenseNet on fundus images for pixel-wise classifying input images. This method has some limitations as the OD center is not automatically localized and has relatively long training time. For discriminating early and advanced stage glaucoma, authors in [38] have constructed a complex and deep learning CNN architecture with multiple layers. Some authors have proposed transfer learning techniques for glaucoma detection. Authors in [39] have evaluated five different ImageNet trained neural network architectures as an alternative for glaucoma detection algorithms. High accuracy, sensitivity, and specificity were achieved by analyzing these models on publicly available datasets. In [11], OD and OC have been segmented simultaneously using a modified deep fully convolutional network (FCN) with pre-processing. In [40], segmentation was performed using an enhanced U-net architecture. For OC segmentation, images were cropped and scaled-down and then fed into a modified U-net convolutional network with more convolutional layers and fewer parameters. The same process except cropping was then applied for OD segmentation. The method was evaluated on DRISHTI and RIM-ONE dataset. A new segmentation method has been proposed in [13] which combines a pre-trained ResNet-34 model as an encoder with traditional U-Net as a decoder. This network was trained on the RIGA dataset. A novel RNN architecture (RACE-net) was proposed for the segmentation of biomedical images which model the boundaries of objects as an evolving level set curve [41]. Some authors have proposed domain adaptation methods for generalizing deep networks trained in the source domain to perform efficiently in different target domains of varying appearances. A similar method is proposed in [15], using a patch-based Output Space Adversarial Learning architecture (pOSAL) for the segmentation of disc and cup. In another paper [16], a multi-label DCNN GL-net has been proposed for segmentation. In this model, skip connections were used in the Generator to facilitate the mixture of low and high-level features and were tested on the Drishti dataset.

### III. MATERIALS AND METHODOLOGY

#### A. IMAGE DATASETS

In this study, we have used three publicly available glaucoma diagnostic databases that contain retinal fundus images from the DRISHTI-GS, RIM-ONE and REFUGE database for pixel-wise semantic segmentation of both the OD and OC.

#### B. DRISHTI-GS

The DRISHTI-GS dataset consists of 101 retinal fundus images that are captured with dilated pupils, optical disc centered, and 30 degree FOV. The Aravind Eye Clinic, Madurai, India collects these images and annotate them. These images are 2896 × 1944 pixels in size, and are stored in uncompressed PNG format. DRISHTI-GS1 offers average OD and OC borders based on 4 professional manual labeling [42].

#### C. RIM-ONE

RIM-ONE is an open database of retinal fundus images and is consisting of 159 retinal images. It contains 85 healthy eye images and 74 Glaucomatous eye images. All images were collected from 3 Spanish hospitals and were annotated by two ophthalmologists [43].

#### D. REFUGE

The Retinal Fundus Glaucoma Challenge (REFUGE) dataset consists of 400 training retinal fundus images stored in JPEG format, with 8 bits per color channel. These images are obtained from patients sitting upright by ophthalmologists using a Zeiss Visucam 500 fundus camera with a resolution of 2124 × 2056 pixels. The images are focused on the posterior end, with both the macula and the OD visible, to allow the ONH to be assessed. It contains 360 normal eyes and 40 glaucomatous eyes images along with their ground truths [44].

### IV. NETWORK ARCHITECTURE

#### A. DESIGNING AND LEARNING

SegNet is a fully convolutional neural network architecture developed for pixel-wise semantic segmentation. The motivation behind this network was the need for architecture for understanding road and indoor scenes [21]. Basic SegNet architecture is composed of an encoder network and a corresponding decoder network followed by a pixel-wise classification layer. The encoder network’s structure is topologically the same as the 13 convolution layers in the VGG16 network [45]. The purpose of the decoder network is to map the encoder feature maps of low resolution to full input resolution for pixel-wise classification. As compared to most of the fully convolutional networks (FCN) architectures like FCN-AlexNet, FCN-VGG16, FCN-GoogLeNet [46]. SegNet has been proved to be efficient in terms of memory, computational time and accuracy. SegNet is considering fewer trainable parameters because of the elimination of fully
connected layers which are responsible for a huge number of parameters.

SegNet-Basic architecture [21] is based on VGG-16 architecture (13 convolutional layers) without any skip connections which guarantees the feature-reuse policy to reduce the vanishing gradient problem [47]. As OD and OC are with few pixels in the image, so during the continuous convolutional process the features of OC can vanish completely. Therefore, we proposed a Cup Disc Encoder Decoder Network (CDED-Net) for the joint segmentation of OD and OC. To preserve possible information features in our proposed CDED-Net architecture, we employed dense connections between convolutional layers of the encoder, while in the up-sampling decoder section we followed the original SegNet structure as shown in Figure 1. Unlike SegNet-Basic architecture, our proposed network is a balanced network which consists of four dense blocks in encoder with two convolutional layers in each block. The proposed network improves the results by feature empowerment by concatenation and reducing information loss by over convolutional process. The bottleneck layers used inside encoder control the number of channels after feature concatenation to reduce memory consumption. Table 1 shows the valuable differences from SegNet-Basic architecture. The network is capable of joint OD and OC segmentation through encoder-decoder framework without performing any pre- or post-processing and ROI extraction steps, which results in low computational cost. After obtaining the segmented OD and OC, CDR can be estimated for Glaucoma diagnosis.

**Proposed Encoder** in the network handles the down-sampling of the image to provide the classes with tiny representation. In the encoder, there is a total of 8 convolutional layers with 4 blocks (Encoder Convolutional-Block-1 to Encoder Convolution-Block-4 as in Table 2), each having two convolutional layers that are densely connected by the help of depth-wise concatenation. Since encoder contains convolutional layers with dense connectivity, it is called as Light-dense-encoder. Each encoder performs convolution with filter banks and produces a set of feature maps. Dense encoder concatenates the output feature maps of the layer with the feature maps of incoming layer and stores the max-pooling indices (per windows) for the feature maps they produce. The diminished features in each block are compensated by the depth-wise concatenation. A batch normalization layer is then used for normalizing these feature maps followed by an element-wise rectified-linear non-linearity (ReLu) max \(0, x\) activation function. The problem of depth increment is very important to handle in dense networks. To reduce
the number of input feature-maps, a $1 \times 1$ convolution can be introduced as a bottleneck layer before each $3 \times 3$ convolution. These layers are used as dimension reduction modules to improve computational efficiency. Therefore, bottleneck layers are used in our network to handle the number of channels which allows for increasing the depth and width of our networks without significant performance degradation [20], [48].

**Decoder network** performs the upsampling of the feature maps produced by the last encoder. In the decoder, there are also 4 convolutional blocks (Decoder Convolutional-Block-1 to Decoder Convolution-Block-4 as in Table 3). Upsampling is performed by using the max-pooling indices archived by their corresponding encoders. The feature maps are sparse in nature whose density is increased through convolution using a trainable filter bank. Finally, the dense feature maps are batch normalized. The same process is repeated with each decoder feeding the next decoder. Feature maps produced are consistent in channel size and number to their equivalent encoder inputs. However, there is an exception with the last decoder corresponding to the first encoder. The input to the first encoder is a 3 channel RGB image, whereas the output produced by the last decoder has a multi-channel feature map. Finally, a trainable soft-max classifier is used for classifying each pixel independently and output a K channel image of probabilities. Here K represents the number of output classes. In our case, K is equal to 3, i.e., OD, OC, and the background. The prediction of the resulting segmentation relates to the class having the highest probability at every pixel.

### B. FEATURE EMPOWERMENT BY CONCATENATION AND REDUCTION IN INFORMATION LOSS

It has been advocated in the past that information flow between the layers of the convolutional neural network improves its performance and training process [20]. In encoder-decoder architectures, due to the repetitive convolution process, there exists a semantic gap between the features of the encoder and the decoder. To eradicate this issue, the U-Net architecture [19] copy and concatenate features from the encoder to the decoder through concatenation paths. This methodology of passing information between the encoder and decoder stages has been questioned before in the context of medical images [49], [50], as the first encoder stage outputs low-level feature information. This flow of information from first encoder and the last decoder stage of the U-Net architecture via a concatenation path is shown in Figure 2(a).

---

#### TABLE 2. CDED-Net encoder with feature map size of each encoder convolutional block (EConv and Pooling indicate encoder convolution layer and pooling layer respectively). Convolutional layers with ** indicate batch normalization (BN) and ReLU.

| Block                  | Name/Size               | No. of filters | Output feature map size | No of parameters |
|------------------------|-------------------------|----------------|-------------------------|------------------|
| Encoder convolution block-1 | EConv - $1.1^**/3 \times 3$ | 64             | $500 \times 560 \times 64$ | 1792             |
|                        | EConv - $1.2^*/3 \times 3$ | 64             | -                       | 36928            |
| Pooling-1              | Pool - $1/2 \times 2$   | -              | $250 \times 280 \times 64$ | -                |
| Encoder convolution block-2 | EConv - $2.1^*/3 \times 3$ | 128            | $250 \times 280 \times 128$ | 73856            |
|                        | EConv - $2.2^*/3 \times 3$ | 128            | -                       | 147584           |
| Pooling-2              | Pool - $3/2 \times 2$   | -              | $125 \times 140 \times 128$ | -                |
| Encoder convolution block-3 | EConv - $3.1^*/3 \times 3$ | 256            | $125 \times 140 \times 256$ | 295168           |
|                        | EConv - $3.2^*/3 \times 3$ | 256            | -                       | 590080           |
| Pooling-3              | Pool - $3/2 \times 2$   | -              | $62 \times 70 \times 256$  | -                |
| Encoder convolution block-4 | EConv - $4.1^*/3 \times 3$ | 512            | $62 \times 70 \times 512$  | 1180160          |
|                        | EConv - $4.2^*/3 \times 3$ | 512            | -                       | 2359808          |
| Pooling-4              | Pool - $4/2 \times 2$   | -              | $31 \times 35 \times 512$  | -                |

#### TABLE 3. CDED-Net decoder with feature map size of each decoder convolutional block (DConv and Unpool indicate decoder convolution layer and un-pooling layer respectively). Convolutional layers with ** indicate BN and ReLU.

| Block                  | Name/Size               | No. of filters | Output feature map size | No of parameters |
|------------------------|-------------------------|----------------|-------------------------|------------------|
| Unpooling-4            | Unpool-4                | -              | $62 \times 70 \times 512$  | -                |
| Decoder convolutional block-4 | DConv - $4.2^*/3 \times 3$ | 512            | -                       | 2359808          |
|                        | DConv - $4.1^*/3 \times 512$ | 256            | $62 \times 70 \times 256$  | 1179904          |
| Unpooling-3            | Unpool-3                | -              | $125 \times 140 \times 256$ | -                |
| Decoder convolutional block-3 | DConv - $3.2^*/3 \times 3$ | 256            | -                       | 590080           |
|                        | DConv - $3.1^*/3 \times 256$ | 128            | $125 \times 140 \times 128$ | 295040           |
| Unpooling-2            | Unpool-2                | -              | $250 \times 280 \times 128$ | -                |
| Decoder convolutional block-2 | DConv - $2.2^*/3 \times 3$ | 128            | -                       | 147584           |
|                        | DConv - $2.1^*/3 \times 3$ | 64             | $250 \times 280 \times 64$ | 73792            |
| Unpooling-1            | Unpool-1                | -              | $500 \times 560 \times 64$ | -                |
| Decoder convolutional block-1 | DConv - $1.2^*/3 \times 3$ | 64             | -                       | 36928            |
|                        | DConv - $1.1^*/3 \times 3$ | 3              | $500 \times 560 \times 3$ | 1731             |
We tackle the semantic gap problem through the feature re-use policy that can be observed from Figure 2 (b). The information from the earlier layers of the encoder stage is passed to the later layers through dense connections, where the feature information is concatenated to fill the semantic gap. We argue that the advantages of the dense connections in the proposed encoder architecture are two-fold: 1) it helps our encoder architecture to transfer the useful semantic information to the decoder stage, thus closing the semantic gap, 2) information flow between the layers is enhanced, which in turn improves the accuracy and training efficiency of the network. To maintain the computational feasibility, bottleneck layers are introduced after concatenation to reduce the number of feature maps.

Other noticeable differences between the U-Net and proposed CDED-Net are presented in Table 4. While the number of convolutional layers is almost similar in the encoder and decoder, the U-Net architecture does not perform batch normalization both at encoder and the decoder stages.

In conclusion, the proposed CDED-Net includes a concatenation layer to retain the semantic information from previous layers and bottleneck layers to reduce the number of input feature maps. In contrast, the U-Net architecture does not include concatenation and bottleneck layers.

C. MODEL TRAINING

Model is trained on three publicly available datasets i.e. DRISHTI-GS, RIM-ONE and REFUGE that contain their ground truth. Initially, we randomly selected 70% images for training, and the remaining 30% images for testing. For training purposes, images were first resized to the dimensions of $500 \times 560$ for Drishti. And for Rim-one, images were resized to the dimensions of $570 \times 429$. For the REFUGE dataset, we selected 300 fundus images as the training set with ground truth labels (including 30 glaucoma cases) and the rest as the test set. Since datasets contain less number of images, therefore, data augmentation was required to improve model accuracy and performance. For this purpose, an augmentation scheme was performed on training data. In augmentation, we used angle rotation from 0-360 with the difference of 5 degrees along with brightness, and then we increased and decreased the brightness to cater to the contrast issue.

The model was trained on the computer with Intel(R) Xeon(R) W-2133 CPU 3.60GHz processor, 32GB RAM, and Nvidia 2080TI GPU. Stochastic Gradient Descent with Momentum (SGDM) with L2 regularization was used with a weight decay of 0.005 for model training. An initial learning rate of 1e-3 was employed for the segmentation network. The training batch size was set at 4 and takes 25 epochs. It takes around 6h to train the model with a test time of 0.01 sec on the Drishti dataset. On Rim-one, the model was trained in 5.5h with a test time of 1 sec.

V. RESULTS

A. EVALUATION PARAMETERS

Various assessment metrics such as Dice coefficient (F-measurement), Jaccard (overlapping), accuracy,
sensitivity, and specificity were used to evaluate the performance of the proposed method for segmenting OD and OC relative to ground truth. These parameters are defined as follows:

\[
\text{Dice}(DC) = \frac{(2 \times TP)}{(2 \times TP) + FP + FN} \tag{1}
\]

\[
\text{Jaccard}(Jc) = \frac{TP}{TP + FP + FN} \tag{2}
\]

\[
SN = \frac{TP}{TP + FN} \tag{3}
\]

\[
SP = \frac{TN}{TN + FP} \tag{4}
\]

\[
\text{Acc} = \frac{(TP + TN)}{(TP + FN + TN + FP)} \tag{5}
\]

where TP, FP, TN, and FN refer to true positive, false positive, true negative, and false negative, respectively. We have also calculated balance accuracy (A) and overlapping error (E) between ground truth and segmented region. The parameter A can be expressed as

\[
A = \frac{1}{2} (Sen + Sp), \tag{6}
\]

where Sen and Sp are the sensitivity and specificity. The overlapping error can be written as

\[
E = 1 - \frac{\text{Area}(S \cup G)}{\text{Area}(S \cap G)}, \tag{7}
\]

where S and G represent the segmented region and ground truth respectively.
After the compilation of results, obtained by the proposed algorithm, they are compared with state-of-the-art techniques. Previously several techniques have been implemented on different datasets namely SCES, ORIGA and some other datasets designed by the researchers, for the detection of glaucoma disease which is not publicly available. In this research, the proposed method is tested on three publicly available datasets, i.e., DRISHTI-GS, RIM-ONE and REFUGE. The comparison is made with other state-of-the-art techniques that have performed additional processing and localization (ROI extraction) steps before the segmentation process or have achieved comparable performance through extensive networks.

Figure 3 and Figure 4 represent OD and OC segmentation results on the DRISHTI-GS and RIM-ONE datasets. In [36], Sevastopolsky et al. proposed a method based on the U-Net convolutional network to segment out the OD and OC. In the case of OD, the achieved dice and Jaccard are 94% and 89%. Whereas for OC, it is 82% and 69% on Rim-one. For OC, attained dice and Jaccard are 85% and 75% and for OD values are 95% and 89% respectively on the Drishti dataset. Zilly et al. proposed an ensemble learning-based CNN model and to reduce the computational complexity, entropy sampling is
performed and a framework is designed using boosting for convolutional filters. This algorithm achieved dice and Jaccard of 94.2% and 89% for OD and 82.41% and 80.2% for OC on Rim-one. In the case of the Drishti dataset, the acquired dice and Jaccard for OD are 97.3% and 91.4% and for OC; dice and Jaccard are 87.1% and 85% respectively [12]. Cheng et al. presented a superpixel based classification model [31] for the segmentation purpose and the achieved dice and Jaccard are 89.2% and 82.93%, and in case of OC, 74.4% and 73.2% respectively which is much lesser than our proposed method. Aquino et al. performed the segmentation of OD using a template-based method [55]. In this method, morphological and edge detection based techniques were used; followed by CHT that is used to detect the boundary of OD. The attained dice and Jaccard are 90.1% and 84.2% respectively. [56] has used the Ant Colony optimization method for OC segmentation. This method achieved a lower Jaccard value of 75.7% for cup segmentation. Zhou et al. presented a locally statistical active contour model [32] for joint segmentation of OD and OC [51]. They achieved 95.5% and 84.7% dice values for disc and cup on Drishti and 85.3% and 78.5% values on rim-one. RNN architecture (RACE-net) has achieved 97% and 87% values for dice on disc and cup for Drishti [41]. Wang et al. have proposed a patch-based adversarial network and have achieved 85.8% and 96.5% dice for OD and OC respectively for Drishti and 78.7% and 86.5% cup and disc dice values on rim-one dataset [15]. Their dice value for the cup is much lower than achieved by the proposed architecture. Xu et al. have designed a U-shaped convolutional neural network with multi-scale input and multi-kernel modules (MSMKU) for disc and cup segmentation but their results were lower than the proposed model. Some results taken from the paper [54] are deep learning methods for OC and OD segmentation that includes U-Net [36], ensemble CNN [12], Generative Adversarial Network (GAN) [52], encoder-decoder based CE-Net [53] and modified U-net M-net [17]. Their results are shown in Tables 5 and 6. Our proposed method has achieved the highest dice value for the cup i.e. 92.4% and 86.22% and comparable dice value for disc i.e. 95.97% and 95.82% on both Drishti and Rim-one datasets. Our method has achieved state-of-the values for Jaccard on disc and cup when compared with other methods. Table 5 and 6 gives a comparison of these techniques with the proposed model on both datasets. We have also evaluated our model in terms of sensitivity and specificity. The proposed model achieves state-of-the-art performance on both datasets for disc and cup. It achieves sensitivity and specificity

### Table 5. Results of the proposed method for OD and OC segmentation on Drishti-GS dataset (%) compared with the existing methods.

| Author            | Method                             | OC            |                  |                  | OD            |                  |                  |
|-------------------|------------------------------------|---------------|-----------------|-----------------|---------------|-----------------|-----------------|
|                    |                                    | DC(P) | JC(O) | Sensitivity | Specificity | DC(P) | JC(O) | Sensitivity | Specificity |
| M. Tabassum et al. | Modified U-net CNN               | 85.21 | 75.15 | 84.76        | 98.81        | 90.43 | 83.5 | 91.56       | 99.69         |
| Zilly et al. [12] | Ensemble CNN                     | 87.1  | 85   | -            | -            | 97.3  | 91.4 | -            | -             |
| Sedai et al. [35] | Regression based method           | 85    | -    | -            | 95            | -    | -    | -            | -             |
| Zhou et al. [52]  | Locally statistical active contour model | 84.7  | -    | -            | 95.5          | -    | -    | -            | -             |
| Son et al. [53]   | Generative Adversarial Network (GAN) | 86.43 | 77.48 | 85.39        | 99.07        | 95.27 | 91.85 | 97.47       | 99.77         |
| Chakravarty et al. [41] | Recurrent Neural Network (RNN)-based RACE-net | 87    | -    | -            | 97            | -    | -    | -            | -             |
| Fu et al. [17]    | M-net                             | 86.18 | 77.3 | 88.22        | 98.62        | 96.78 | 93.86 | 97.11       | 99.91         |
| Wang et al. [15]  | Patch-based Output Space Adversarial Learning | 85.8 | -    | -            | 96.5          | -    | -    | -            | -             |
| Gu et al. [54]    | Encoder-Decoder CE-Net           | 88.18 | 80.06 | 88.19        | 99.09        | 96.42 | 93.23 | 97.59       | 99.9          |
| Xu et al. [55]    | U-shaped convolutional neural network | 89.2  | 82.30 | -            | 97.8          | 94.9  | -    | -            | -             |
| Proposed          | CDED-Net                          | 92.4  | 86.32 | 95.67        | 99.81        | 95.97 | 91.83 | 97.54       | 99.73         |

### Table 6. Results of the proposed method for OD and OC segmentation on Rim-one dataset (%) compared with the existing methods.

| Author            | Method                             | OC            |                  |                  | OD            |                  |                  |
|-------------------|------------------------------------|---------------|-----------------|-----------------|---------------|-----------------|-----------------|
|                    |                                    | DC(P) | JC(O) | Sensitivity | Specificity | DC(P) | JC(O) | Sensitivity | Specificity |
| Cheng et al. [31] | Superpixel classification          | 74.4  | 73.2  | -            | -            | 89.2  | 82.93 | -            | -             |
| Aquino et al. [56] | Template-based approach            | -    | -    | -            | -            | 90.1  | 84.2 | -            | -             |
| Array et al. [57] | Ant colony optimization            | -    | -    | 75.7          | 99.76        | 95    | 89    | 95.02       | 99.73         |
| M. Tabassum et al. | Modified U-net CNN                | 82    | 69   | 75.45        | 99.76        | 95    | 89    | 95.02       | 99.73         |
| Zilly et al. [12] | Ensemble CNN                      | 82.4  | 80.2  | -            | 94.2          | 89    | -    | -            | -             |
| Zhou et al. [52]  | Locally statistical Active contour model | 78.5 | -    | -            | 85.3          | -    | -    | -            | -             |
| Son et al. [53]   | Generative Adversarial Network (GAN) | 82.5  | 71.65 | 81.42        | 99.65        | 95.32 | 91.22 | 94.57       | 99.87         |
| Fu et al. [17]    | M-net                             | 83.48 | 73   | 81.46        | 99.67        | 95.26 | 91.14 | 94.81       | 99.86         |
| Gu et al. [54]    | Encoder-Decoder CE-Net            | 84.35 | 74.24 | 83.52        | 99.77        | 95.37 | 91.15 | 95.02       | 99.86         |
| Wang et al. [15]  | Patch-based Output Space Adversarial Learning | 78.7 | -    | -            | 86.5          | -    | -    | -            | -             |
| Xu et al. [55]    | U-shaped convolutional neural network | 85.64 | 75.86 | 85.15        | 99.71        | 95.61 | 91.72 | 95.21       | 99.87         |
| Proposed          | CDED-Net                          | 86.22 | 75.32 | 95.17        | 99.81        | 95.82 | 91.01 | 97.34       | 99.73         |
of 95.67%, 99.81% respectively for the OC, and 97.54% and 99.73% for the OD segmentation on Drishti. On Rim-one, it achieves sensitivity and specificity of 95.17% and 99.81% for cup and 97.34% and 99.73% for OD segmentation which are higher than other state-of-the-art methods.

Another useful insight from Table 5 and 6 is that the proposed method is preferable for the OC segmentation, while also offering comparable performance for the OD segmentation.

For the OD and OC, we compare the performance of the proposed method with a recent work [57] in terms of overlapping error (E) and balance accuracy (A) in Table 7 on the DRISHTI-GS and RIM-ONE datasets. It is important to mention here that these performance measures were not analyzed by the methods presented in Table 5 and Table 6. In [57], the authors proposed a fully convolutional network (FCN) for the depth estimation using a pre-training scheme called pseudo-depth reconstruction and a dilated residual inception (DRI) block for multi-scale extraction of features. The guided network was used for OD and OC segmentation using a depth map as a guide and evaluated E and A values for different experiments as shown in Table 7. It can be observed from the results that the proposed CDED-Net outperforms all the variants reported in [57] in terms of E and A measures for the OC segmentation task. Also, the proposed method performs better than all the compared methods on the task of OD segmentation in terms of A, with comparable performance in terms of E.

Table 8 has been reproduced from [58] that report the intersection over union (IoU) and mean intersection over union (MIoU) results of different methods on the REFUGE dataset. We trained our model as suggested by Liu et al. where they have selected 300 images as the training set and the remaining 100 images as the testing set to have a direct comparison of our proposed approach with the existing state-of-the-art methods on REFUGE dataset. As shown in Table 8, our model has achieved state-of-the-art performance with 0.8837, 0.8111 and 0.8705 scores. Such encouraging results demonstrate the efficacy of our model for the segmentation of both the OD and OC.

To determine the presence of glaucoma, CDR is used for the evaluation of fundus images. CDR is a clinical parameter that is usually measured manually by the ophthalmologist. AUC curve is used for assessing the efficiency of the proposed algorithm in detecting glaucoma presence and is based on the sensitivity and specificity of ground truth and experimental results. Figure 6 represents the AUC curves for both Drishti and Rim-one datasets. We obtain an average (AUC) value of 0.963 on Drishti based on our segmentation and have compared with other techniques as shown in Table 9. A high AUC value indicates that the percentage of correctly classified images of the overall system is high. Reference [59] proposed a multistage deep learning-based model for segmenting OD and OC on which morphological features were applied and disease was identified. They achieved AUC of 0.82 on Drishti, less than the AUC attained in our model. Reference [37] generated feature vector for two different CNN, i.e. Overfeat and VGG-S along with preprocessing methods to detect Glaucoma. They observed that the performance of OverFeat features was better than VGG-S features and obtained an AUC of 0.7212 with vessel inpainting when tested on the DRISHTI dataset. While AUC, without vessel inpainting was 0.76 which is again lesser than the AUC value achieved in our method. Reference [60] developed a method that fuses image-based features and those derived from the segmented OD and OC based on a semi-supervised approach. They obtained an AUC of 0.78 when tested their method, whereas we have achieved an average AUC of 0.963 on the DRISHTI dataset. We have also computed AUC separately for OD and OC i.e 0.957 and 0.969 on Drishti and 0.909 and 0.987 on Rim-one respectively. The proposed model achieves state-of-the-art accuracy.
for OD and OC segmentation on both datasets as shown in Table 10.

The ROC curve of the proposed method for OD and OC segmentation for the REFUGE dataset is presented in Figure 7. It can be observed that the average AUC performance of the proposed method is significantly high for OC segmentation, while the average AUC for OD segmentation also exhibits competitive performance of the proposed method. It is important note that the qualitative and quantitative performance of the proposed method is consistent on all three datasets, thus exhibiting it ability to generalize on unseen data.

TABLE 10. Results of the proposed method for OD and OC segmentation in terms Accuracy and AUC on Drishti and Rim-one dataset.

|         | Drishti | Rim-one |
|---------|---------|---------|
| Accuracy| OD      | OD      |
|         | 99.71   | 99.66   |
| AUC     | 0.957   | 0.969   |

VI. DISCUSSION

In this research, we have developed a Cup Disc Encoder Decoder Network (CDED-Net) based architecture for pixel-wise classification of input images. To improve the efficiency of our network, the number of convolutional layers has been reduced in both encoder and decoder. For glaucoma diagnosis, accurate segmentation of OD and OC is a necessary step. For this, we have introduced dense connections between layers in the encoder. Feature maps of the first convolution layer are passed through dense connections and are concatenated with the feature maps of the next layer. This helps to store vital edge information that would otherwise be lost in the repeated convolution process. The dense connectivity between layers of encoder helps the network to perform correct segmentation of both OD and OC and leads to faster convergence. To assess our model’s efficiency and robustness, we evaluated the model on DRISHTI-GS and RIM-ONE datasets which are publicly available. Our model performed well on both datasets when evaluated with dice and Jaccard score for OD and OC segmentation. It is worse than [12], [41] and [54] on dice value for Drishti. However, it is worthwhile to point out that our model has achieved higher dice value for OC segmentation which is a more challenging task due to the presence of blood vessels. It is observed that the underlying architecture is the traditional or modified U-Net model for most previous deep neural networks for segmentation [17], [36], [54]. Although U-Net based architectures perform well on segmenting OD and OC, this requires a computationally heavy network with more trainable parameters and system memory.
Some authors have used general adversarial networks [15], [52] to reduce the performance degradation caused by domain shift. These extensive networks for segmentation tasks increase the number of required parameters. Moreover, during the training phase, a large number of unlabeled target images are needed which may not be readily available. Whereas in this paper, a new memory-efficient architecture is proposed to robustly and jointly segment OD and OC without any pre- or post-processing and ROI extraction blocks with fewer parameters required. The model achieves state-of-the-art performance for OC and OD segmentation in terms of dice, Jaccard, sensitivity, and specificity value for both DRISHTI-GS and RIM-ONE dataset. We have also evaluated our model on REFUGE dataset and have compared intersection over union (IoU) and mean intersection over union (MiIoU) results with the existing methods. The model has achieved state-of-the-art performance with the highest scores which demonstrates the efficacy of our proposed model for the segmentation of both the OD and OC.

**ACKNOWLEDGMENT**

The authors would like to thank the teams of DRISHTI-GS, RIM-ONE and REFUGE for maintaining and keeping these databases active and making them easily accessible for the researchers to carry out retinal image analysis research. Furthermore, they would also like to thank the experts for taking the time to establish the ground truth images.

**REFERENCES**

[1] Y. Hagiwara, J. E. W. Koh, J. H. Tan, S. V. Bhandary, A. Laude, E. J. Ciaccio, L. Tong, and U. R. Acharya. “Computer-aided diagnosis of glaucoma using fundus images: A review,” Comput. Methods Programs Biomed., vol. 165, pp. 1–12, Oct. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260718305492.

[2] H. A. Quigley, “The number of people with glaucoma worldwide in 2010 and 2020,” Brit. J. Ophthalmol., vol. 90, no. 3, pp. 262–267, Mar. 2006. [Online]. Available: http://bio.biij.nim.ch/content/90/3/262.abstract

[3] M. I. Rizzo, A. Greco, A. De Virgilio, A. Gallo, L. Tavarniti, M. Fusconi, M. Conte, G. Pagliuca, R. Turchetta, and M. de Vincentis. “Glaucoma: Recent advances in the involvement of autoimmunity,” Immunologic Res., vol. 65, no. 1, pp. 207–217, Feb. 2017. doi: 10.1007/s12026-016-8837-3.

[4] U. Raghavendra, H. Fujita, S. V. Bhandary, A. Gudigar, J. H. Tan, and U. R. Acharya. “Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images,” Inf. Sci., vol. 441, pp. 41–49, May 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002002551830074X.

[5] A. Almazroa, R. Burman, K. Rahaemifar, and V. Lakshminarayanan,”Optic disc and optic cup segmentation methodologies for glaucoma image detection: A survey,” J. Ophthalmol., vol. 2015, pp. 1–28, Nov. 2015.

[6] M. Naveed, A. Ramzan, and M. U. Akram, “Clinical and technical perspective of glaucoma detection using OCT and fundus images: A review,” in Proc. 1st Int. Conf. Next Gen. Comput. Appl. (NextCom), Jul. 2017, pp. 157–162.

[7] W. Ruengkittipinyo, W. Kongprawechnon, T. Kondo, P. Bunnum, and H. Kaneko, “Glaucoma screening using rim width based on ISNT rule,” in Proc. 6th Int. Conf. Inf. Commun. Technol. Embedded Syst. (IC-ICTE), Mar. 2015, pp. 1–5.

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

[9] S. M. A. Issac, and M. K. Datta, “An automated and robust image processing algorithm for glaucoma diagnosis from fundus images using novel blood vessel tracking and bend point detection,” Int. J. Med. Informat., vol. 110, pp. 52–70, Feb. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S138650561730429X

[10] V. G. Edupuganti, A. Chawla, and A. Kale, “Automatic optic disc and cup segmentation of fundus images using deep learning,” in Proc. 25th IEEE Int. Conf. Image Process. (ICIP), Oct. 2018, pp. 2227–2231.

[11] P. Qin, L. Wang, and H. Lv, “Optic disc and cup segmentation based on deep learning,” in Proc. IEEE 3rd Int. Conf. Netw., Electron. Autom. Control Conf. (ITNEC), Mar. 2019, pp. 1835–1840.

[12] J. Zilly, J. M. Buhmann, and D. Mahapatra,”Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation,” Computerized Med. Imag. Graph., vol. 55, pp. 28–41, Jan. 2017.

[13] S. Yu, D. Xiao, S. Frost, and Y. Kanagasingam, “Robust optic disc and cup segmentation with deep learning for glaucoma detection,” Computerized Med. Imag. Graph., vol. 74, pp. 61–71, Jun. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S089561118305573

[14] B. Al-Bander, B. Williams, W. Al-Nuaimy, M. Al-Taee, H. Pratt, and Y. Zheng, “Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis,” Symmetry, vol. 10, no. 4, p. 87, 2018.

[15] S. Wang, L. Yu, X. Yang, C.-W. Fu, and P.-A. Heng, “Patch-based output space adversarial learning for joint optic disc and cup segmentation,” IEEE Trans. Med. Imag., vol. 38, no. 11, pp. 2485–2495, Nov. 2019.

[16] Y. Jiang, N. Tan, and T. Peng, “Optic disc and cup segmentation based on deep convolutional generative adversarial networks,” IEEE Access, vol. 7, pp. 64483–64493, 2019.

[17] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao,”Joint optic disc and cup segmentation based on multi-label deep network and polar transformation,” IEEE Trans. Med. Imag., vol. 37, no. 10, pp. 1597–1605, Jun. 2018.

[18] H. Fu, J. Cheng, Y. Xu, C. Zhang, D. W. K. 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.

[19] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Medical Image Computing and Computer-Assisted Intervention—MICCAI (Lecture Notes in Computer Science), vol. 9351, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds. Cham, Switzerland: Springer, 2015.

[20] 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. 4700–4708.

[21] 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, Dec. 2017.

[22] A. Agarwal, S. Gullia, S. Chaudhary, M. K. Dutta, C. M. Travieso, and J. B. Alonso-Hernandez, “A novel approach to detect glaucoma in retinal fundus images using cup-disc and rim-disk ratio,” in Proc. 4th Int. Work Conf. Bioinspired Intell. (WBOI), Jun. 2015, pp. 139–144.
A. Issac, M. Parthasarathi, and M. K. Dutta, “An adaptive threshold based algorithm for optic disc and cup segmentation in fundus images,” in Proc. 2nd Int. Conf. Signal Process. Integr. Netw. (SPIN), Feb. 2015, pp. 143–147.

T. Khalil, M. U. Akram, S. Khalid, and A. Jameel, “Improved automated detection of glaucoma from fundus image using hybrid structural and textural features,” IET Image Process., vol. 11, no. 9, pp. 693–700, Sep. 2017.

N. Thakur and M. Juneja, “Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma,” Biomed. Signal Process. Control, vol. 42, pp. 162–189, Apr. 2018. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S1566006718300211

G. D. Joshi, J. Sivashwamy, K. Karan, and S. R. Krishnasad, “Optic disk and cup boundary detection using regional information,” in Proc. IEEE Int. Symp. Biomed. Image. From Nano Macro, 2010, pp. 948–951.

J. Cheng, J. Liu, D. W. K. Wong, F. Yin, C. Cheung, M. Baskaran, T. Aung, and T. Yin Wong, “Automatic optic disc segmentation with peripapillary atrophy elimination,” in Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc., Aug. 2011, pp. 6224–6227.

G. D. Joshi, J. Sivashwamy, and S. R. Krishnasad, “Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment,” IEEE Trans. Med. Imag., vol. 30, no. 6, pp. 1192–1205, Jun. 2011.

M. S. Haleem, L. Han, J. V. Hemert, B. Li, A. Fleming, L. R. Pasquale, and G. D. Joshi, J. Sivaswamy, and S. R. Krishnadas, “Optic disk and cup boundary segmentation from monocular color retinal images,” IEEE J. Biomed. Health Informat., vol. 42, no. 1, pp. 20, Dec. 2017, doi: 10.1109/JBHI.2017.0859-4.

J. R. H. Kumar, A. K. Pediredla, and C. S. Seelamantula, “Active discs for segmented glaucoma detection,” in Proc. Global Conf. Signal Inf. Process. (GlobalSIP), Dec. 2015, pp. 225–229.

J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, F.-M. Tan, D. Tao, C.-Y. Cheng, T. Aung, and T. Y. Wong, “Superpixel classification based optic disc and optic cup segmentation for glaucoma screening,” IEEE Trans. Med. Imag., vol. 32, no. 6, pp. 1019–1032, Jun. 2013.

A. Septiariini, A. Harjoko, R. Pulungan, and R. Ekantini, “Automatic detection of retinal nerve fiber layer by texture-based analysis for glaucoma evaluation,” Healthcare Informat. Res., vol. 24, no. 4, pp. 335–345, 2018.

A. Septiariini, D. M. Kharina, A. H. Kridalaksana, and H. Handani, “Automatic glaucoma detection method applying a statistical approach to fundus images,” Healthcare Informat. Res., vol. 24, no. 1, pp. 53–60, 2018.

K. Akyol, B. Şen, and C. S. Bayar, “Automatic detection of optic disc in retinal image by using keypoint detection, texture analysis, and visual dictionary techniques,” Comput. Math. Methods Med., vol. 2016, pp. 1–10, May 2016.

S. Sedai, P. K. Roy, D. Mahapatra, and R. Garnavi, “Segmentation of optic disc and optic cup in retinal fundus images using shape regression,” in Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Aug. 2016, pp. 3260–3264.

A. Sevastopolsky, “Optic disc and cup segmentation methods for glaucoma detection with modification of U-net convolutional neural network,” Pattern Recognit. Image Anal., vol. 27, no. 3, pp. 618–624, Jul. 2017, doi: 10.1134/S1054661817030269.

J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, “Active discs for segmented glaucoma identification,” in Proc. 12th Int. Symp. Med. Inf. Process. Anal., May 2016.

J. M. Ahn, S. Kim, K.-S. Ahn, S.-H. Cho, K. B. Lee, and U. S. Kim, “Superpixel based optic disc and optic cup segmentation with deep learning from unbalanced samples,” in Proc. 24th Int. Symp. Comput.-Based Med. Syst. (CBMS), Jun. 2011, pp. 499–512, 2019.

Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, “CE-net: Context encoder network for 2D medical image segmentation,” IEEE Trans. Med. Imag., vol. 38, no. 10, pp. 2281–2292, Oct. 2019.

Y.-L. Lu, S. Lu, H.-X. Li, and R.-R. Li, “Mixed maximum loss design for optic disc and optic cup segmentation with deep learning from imbalanced samples,” Sensors, vol. 19, no. 20, pp. 4401–4410, Oct. 2019.

A. Aquino, M. E. Gérguez-Arias, and D. Marín, “Detecting the optic disc and cup boundary in digital fundus images using morphological, edge detection, and feature extraction techniques,” IEEE Trans. Med. Imag., vol. 29, no. 11, pp. 1860–1869, Nov. 2010.

S. Liu, J. Hong, X. Lu, X. Jia, Z. Lin, Y. Zhou, Y. Liu, and H. Zhang, “Joint optic disc and cup segmentation using semi-supervised conditional GANS,” Comput. Biol. Med., vol. 115, Dec. 2019, Art. no. 103485. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482519305539

S. M. Shankaranarayana, K. Ram, K. Mitra, and M. Sivaprakasam, “Fully convolutional networks for monocular retinal depth estimation and optic disc-cup segmentation,” IEEE J. Biomed. Health Informat., vol. 23, no. 4, pp. 1417–1426, Jul. 2019.

J. Liu, J. Hong, X. Lu, X. Jia, Z. Lin, Y. Zhou, Y. Liu, and H. Zhang, “Joint optic disc and cup segmentation using semi-supervised conditional GANS,” Comput. Biol. Med., vol. 115, Dec. 2019, Art. no. 103485. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482519305543

O. Perdomo, V. Andrearczyk, F. Meriaudeau, H. Müller, and F. A. González, “Glaucoma diagnosis from eye fundus images based on deep morphometric feature estimation,” in Computational Pathology and Ophthalmic Medical Image Analysis (OMIA) (Lecture Notes in Computer Science), vol. 11039, D. Stoyanov et al., Eds. Cham, Switzerland: Springer, 2018.

A. Chakravarty and J. Sivashwamy, “Glaucoma classification with a fusion of segmentation and image-based features,” in Proc. IEEE 13th Int. Symp. Biomed. Imaging (ISBI), Apr. 2016, pp. 689–692.
TARIQ M. KHAN (Member, IEEE) received the B.S. degree in computer engineering from the COMSATS Institute of Information Technology, Islamabad, Pakistan, the M.Sc. degree in computer engineering from the University of Engineering Technology, Taxila Pakistan, and the Ph.D. degree in electronic engineering from Macquarie University, Sydney, Australia, in 2016. He is currently working as an Assistant Professor with the Department of Electrical Engineering, COMSATS University Islamabad, Pakistan. He is interested in both Digital Image Processing (with an emphasis on biometrics) and VLSI. His research interests include most aspects of image enhancement, pattern recognition, medical image analysis, scene understanding, and deep learning methods for image analysis.

MUHAMMAD ARSALAN (Graduate Student Member, IEEE) received the B.S. degree in computer engineering from COMSATS University Islamabad, Pakistan, in 2012, and the M.S. degree in computer science from the NCBA&E, Lahore, Pakistan, in 2016. He is currently pursuing the Ph.D. degree in electronics and electrical engineering with Dongguk University, Seoul, South Korea. His research interests include computer vision and deep learning. He helped to perform the experiments and analysis.

SYED SAUD NAQVI received the B.Sc. degree in computer engineering from the COMSATS Institute of Information Technology, Islamabad, Pakistan, in 2005, the M.Sc. degree in electronic engineering from The University of Sheffield, U.K., in 2007, and the Ph.D. degree from the School of Engineering and Computer Science, Victoria University of Wellington, New Zealand, in 2016. He is currently working as an Assistant Professor with the COMSATS Institute of Information Technology, Islamabad, Pakistan. His research interests include saliency modeling, medical image analysis, scene understanding, and deep learning methods for image analysis.

MANSOOR AHMED is currently working as an Assistant Professor with the Department of Computer Science, COMSATS Institute of Information Technology (CIIT), Islamabad, Pakistan. Before joining CIIT, he was a Postdoctoral Fellow with Indiana University Bloomington, USA. His research interests include Information security and privacy, distributed computing, knowledge-based systems, data provenance, and semantic Web technologies. He was awarded senior researcher fellowship scholarship by Indiana University, USA, for his postdoctoral studies. He was also awarded Higher Education Commission (HEC) scholarship for higher studies (Ph.D.) in Austria.

HUSSAIN AHMED MADNI received the Master of Science degree in computer science from the Department of Computer Science, COMSATS University Islamabad (CUI), Islamabad, Pakistan, in 2019. His research interests include deep neural networks, computer vision, and pattern recognition.

JAWAD MIRZA (Senior Member, IEEE) received the B.Sc. degree in electrical (telecommunications) engineering from the COMSATS Institute of Information Technology, Islamabad, Pakistan, in 2007, the M.Sc. degree in communications engineering from the University of Manchester, Manchester, U.K., in 2009, and the Ph.D. degree from the Victoria University of Wellington, Wellington, New Zealand, in 2016. He was a Postdoctoral Research Associate with the Wolfson School Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough, U.K. He is currently an Assistant Professor with the Department of Electrical Engineering, COMSATS University Islamabad, Islamabad. His research interests include massive multiple-input-multiple-output and full-duplex communications.

* * *