International Journal of Engineering

An Automatic Optic Disk Segmentation Approach from Retina of Neonates via Attention Based Deep Network

A. Abaei Kashan[a], A. Maghsoudi[b], N. Shoeibi[c], M. Heidarzadeh[d], K. Mirnia*[e]

[a] Department of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran
[b] Department of Mechanical Engineering, Tehran University of Medical Science, Tehran, Iran
[c] Department of Ophthalmology, Mashhad University of Medical Sciences, Mashhad, Iran
[d] Department of Pediatrics, Tabriz University of Medical Science, Tabriz, Iran
[e] Department of Pediatrics, Tehran University of Medical Science, Tehran, Iran

P A P E R   I N F O

A B S T R A C T

Every year, many newborns lose their sight to retinopathy of prematurity (ROP) worldwide. Despite its high prevalence and adverse consequences, periodic examinations can effectively prevent it. The use of an intelligent system enables physicians to avoid medical mistakes while examining newborns. The optic disk (OD) is an integral part of the retina for grading the severity and progression of ROP. Due to the uneven brightness and lack of a defined OD border, the use of retinal images of infants is very challenging for OD diagnosis. This paper provides an innovative model of OD segmentation based on attention gate. Initially, the images were collected and preprocessed and inputted into a novel deep convolutional neural network consisting of attention in skip connections. The architecture is comprised of a two-stage convolutional network. Different outputs are obtained from two individual branches of the original image and image features in the first stage. The outputs were concatenated to transfer into the post-processing stage to identify the area related to the OD. The final results based on the Dice coefficient (Dice) and the Intersection-Over-Union (IoU) were 94.22% and 86.1%, respectively.

1. INTRODUCTION

Retinopathy of prematurity (ROP) is the major leading cause of neonatal blindness worldwide [1], affecting approximately 10% of neonates each year [2]. ROP is caused by abnormal growth of retinal arteries in neonates [1], leading to various visual impairments varying from minor vision defects to retinal detachment and blindness [3]. Since the disease is preventable, early diagnosis makes it treatable and retardates its progression [4]. Periodic examinations are crucial because neonates are not able to express their vision problems. Treatment in the early stages of ROP decreases disease consequences and treatment expenses. Due to the development of neonatal intensive care units, the probability of preterm infants' survival increases [5]; hence, a significant

amount of neonates are at risk of ROP. Therefore, it has heightened the need for intelligent systems in their inspection. Over the past year, due to the coronavirus pandemic, there has been a dramatic increase in artificial intelligence implementation to reduce hospital visitation. An intelligent examination of ROP is based on detecting the anatomic structures in retinal images. The optic disk (OD) is one of the specific structures, and identification of it plays a notable role in determining ROP's zone and stage as a criterion for severity [6]. Accordingly, in this study, we have introduced a method for detecting and segmenting OD in neonatal images. Furthermore, OD detection in adults is used to diagnose glaucoma [7] and papilloma [8]. An optic disk is a bright, circular area (Figure 1) where blood vessels converge. This definition

*Corresponding Author Institutional Email: kmirnia@sina.tums.ac.ir
(K. Mirnia)

is the base of OD detection algorithms, which are discussed subsequently.

Walter et al. [9] segmented OD as a bright area of the retina using a watershed transform. Welfer et al. [10] used a similar procedure, in that they applied adaptive morphology instead of a watershed transform. In a study, Tjandrasa et al. [11] proposed another morphological approach, where the OD area is cropped following image enhancement, and the vessels are removed using thresholding in the cropped area. Then, the OD region is recognized by the Hough transform with a predefined radius. Finally, the Selective Binary-and-Gaussian Filtering Regularized Level Set (SBGFRLS) method is used for segmentation. This method was evaluated on Drive Database and achieved 75.56% accuracy. The mentioned algorithm's performance depends highly on the quality of vessel removal, the assumed OD radius for OD, and the SBGFRLS parameters [11]. Abdullah et al. [12] used a morphological procedure to remove the vessels in a separate investigation. Next, since the algorithm's inputs are acquired from different databases, they should resize the images to obtain a circle with a certain radius through the Hough transform as the OD center. Then, it is used as the initial point for expanding the OD region via search. Although the algorithm is robust to noise and illumination artifacts, some errors occur in pathological structure [12].

Contrary to retinal images, there are scant studies on neonates' retinas. According to Thongnuch and Uyyanonvara's [13] neonatal retina analysis, the contrast histogram equalization could be improved by an average filter in the red channel. Then, closing morphology and the Canny are used to remove vessels and edge detection, respectively. OD edges were segmented through a sequence of morphology operations: filling, erosion, connected component selection, and noise reduction.

Pathan et al. [14] proposed a two-step algorithm to determine OD. First, the authors segmented the vessels by applying the green channel because of its high contrast. Then they extracted the vessels by using a linear directional filter (eight main directions), dilation,
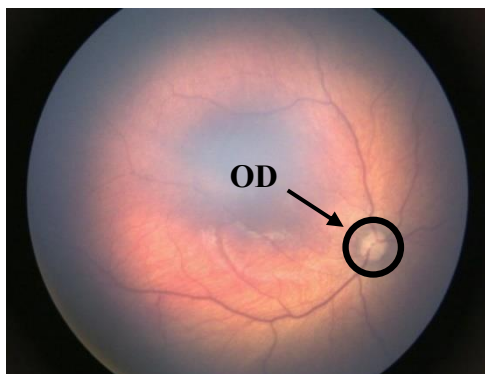
erosion, and thresholding. To improve OD detection, the extracted vessels were removed. Unlike previous stages, the red channel was subtracted from its average. The process was repeated three times to find the existing circles, using the Hough Transform. Eventually, the light intensity threshold between OD and non-OD was obtained from annotated images and used as a threshold in a decision tree [14]. Mookiah et al. [15] found that proper preprocessing is through applying the Limited Adaptive Histogram Equalization to enhance contrast. Later, they separated the vessels through Gabor's two-dimensional Matched filter and closing morphology. Then, the authors determined the location of OD and segmented it by a fuzzy histogram [15]. The Hough transform is an extensively used technique in retinal imaging to detect microaneurysms [16] and OD. Thongnuch and Uyyanonvara [17] indicated that the Canny edge detection method, followed by the Hough transform, has an accuracy of 81.7% in neonatal images. The scientists found that using a deformable contour model with gradient vector flow as an external force enhanced segmentation accuracy by 85.34% [13]. Similarly, Zahoor et al. [18] segmented the OD by utilizing the following steps: morphology to remove vessels, Hough transform and polar transformation to find OD location, and adaptive thresholding to find edges.

There are several studies in the literature on OD segmentation via Thresholding. This method is beneficial in combination with others. For example, Ghadiri et al. [19] initially found the OD kernel, then segmented the OD region using adaptive thresholding, light intensity, and vessel width. Likewise, Septiaria et al. [20] determined OD through thresholding after the removal of vessels.

To solve this issue, many researchers have proposed various segmentation methods, including Superpixel [21, 22], principal component analysis (PCA) [23], and Dolph-Chebyshev matched filter [24].

The past decade was accompanied by the rapid development of deep learning in various fields, such as machine vision [25], object tracking [26], and segmentation [27]. OD segmentation is one of its applications. Sevastopolsky [28] used the U-Net network to segment OD and Optic Cup (OC). Unlike the current U-Net type, the number of filters remained constant after Max Pooling and Up Sampling. As a result, the network training speed increased. Sevastopolsky [28] initially segmented the OD by using the mentioned network and cropped it. The OC was segmented based on this patch [28]. In the same vein, Kim et al. [29] used a pre-trained convolutional neural network (CNN) to classify OD and NOD patches of images. A U-Net model was used to segment OD [29].

Yu et al. [30] initially segmented OD by pre-trained ResNet-34 encoder in the U-Net network and then

Figure 1. Optic disk

separated the optic cup region. To improve the results, they performed morphology [30]. This view is supported by Juneja et al. [31], who utilized red channels in the first CNN and RGB channels to segment OC. In another similar study, Bhatkalkal et al. [32] used Deeplabv3 + and U-Net with Novel Attention Gate inputs in their architecture. Lastly, Conditional Random Fields (CRF) were used for post-processing [32]. A broader perspective has been adopted by Jiang et al. [33] who extracted vessels through a CNN. Then, they cropped 800 × 800 area around the brightest point, as a region of interest (ROI), to reduce the input size of images. OD and OC were segmented in the desired ROI using Faster R-CNN [33]. As a prediction model, Zheng et al. [34] proposed a multi-scale convolutional neural network to provide level set initial contour and evolution parameters. According to the evolution criteria, the existing shape will be modified subsequently. The model was enhanced with previous data and controlled by active contour loss, resulting in more precise contour and edge features in the level set evolution [34]. In a recent literature review on this subject, joint segmentation of OD and OC was defined by Tabassum et al. [35] as a semantic pixel-wise labeling issue.

The retina images of neonates are mostly low contrast. Due to the lack of clarity in the eye media and the difficulties of capturing a moving preterm child, they are often noisy and blurry. Furthermore, the OD's diameter, shape, and color vary significantly across images, making automated identification of OD a difficult process. Due to the high rates of recording artifacts (blurring, reflections), inter-image color fluctuations, and intra-image changes in illumination, the techniques that typically succeed on mature retina images fail on infant fundus images. Therefore, few researchers have addressed neonatal OD segmentation. Nisha et al. [36] proposed assigning an OD score to each pixel of neonates' retina, showing the probable OD localities. The OD area is then segmented employing a region growth approach. The seed growth method begins with the OD center as the first seed. The intensity of the seed pixel is compared to the intensity of adjacent pixels. The zone is then expanded by repeatedly adding nearby pixels, which are comparable to the seed pixel until the stopping threshold is achieved [37]. Since the majority of previous studies did not consider neonate images, a new technique that improves OD segmentation is suggested in this study. Local properties of OD are considered to be the model input to increase the precision of segmentation, contributing to existing knowledge of OD segmentation by providing (1) a novel model architecture with pre-trained VGG-16 model as encoder layer, (2) local image features as input, and (3) significant emphasis on edges property.

The remaining parts of the paper are as follows:
- Collecting data
- Pre-processing
- Data augmentation
- Extracting the OD edge properties
- Designing a CNN architecture
- Illustrating results and discussion

2. METHODOLOGY

Various OD segmentation methods have been developed and proposed; however, most recent studies have used the CNN models for OD because of their capability in extracting features. Therefore, the present study carried out this goal by implementing a novel CNN architecture stated in the following sections.

2. 1. Database   In reviewing the literature, no public databases were found on neonate retinal images. Hence, our colleagues spend much time and effort gathering RetCam imaging data from Khatam-al- Anbya Hospital between 2019 to 2020. Out of the 420 images considered, ROP was recognized on 132 images. The remaining images indicated healthy eyes. The resolution of images was 640×480 and has a 130° field of view (FOV). In contrast to the adult retinal images, the infants' OD's dimensions, direction, and location in each image varied because of imaging conditions.

Data are divided into three sets:
- Training (tuning model parameters)
- Validating (evaluating model performance during the learning process)
- Testing (evaluation of the final model)

Each image is accompanied by an OD ground truth, which an ophthalmologist manually annotated twice with an interval and then averaged.

2. 2. Pre-processing   A preprocessing step was taken to improve the image quality. It is used to reduce noise or brightness variations during image acquisition without missing essential information. In this paper, preprocessing step included zero-padding and cropping operations to acquire a 512×512 resized resolution. Zero-padding is a method often used to ensure that the size of the image input is a power of two. In contrast to the reported cases [38, 39], which reduced the image size to be less costly to run, reducing the image size of neonatal retinal is not possible for two reasons:

i. Neonatal retinal images have a wider FOV, so retinal structures are displayed in a small size. Hence, they may be removed during the resizing process

ii. The neonatal retina is not matured. It contains fewer veins and anatomic structures than adult retina images; therefore, removing or changing some of them during resizing influences processing procedures significantly.

The output of the main processing procedures was affected by these brightness differences. Contrast improvements can be made via image normalization, sometimes defined as Histogram equalization, to ensure that the image's pixel value contains the whole intensity range. Therefore, we normalized the images to improve training time. The technique, used to conduct linear normalizing on a grayscale image, was formulated as:

$$I_N = \frac{I_{raw} - I_{raw,min}}{I_{raw,max} - I_{raw,min}} \tag{1}$$

where I_{raw} is the input image; $I_{raw,max}$ and $I_{raw,min}$ are the maximum and minimum pixel values, respectively; I_N is the normalized image.

2. 3. Data Augmentation   Overfitting is the major practical issue that jeopardizes the learning procedure. Many researchers have proposed various methods to solve this problem. In the present study, data augmentation was used to solve this issue. Augmented data are generated based on problem constraints. The following is the list of procedures applied to augment retinal images: brightness variations, horizontal flips, vertical flips, blur, Gaussian noise, motion blur, and random brightness contrast.

2. 4. Feature Extraction   Despite the vital role of contour borders in the segmentation, few studies have evaluated the impact of the extracted feature on CNN performance [40, 41]. This paper investigated the strengths of two descriptors, namely local binary pattern (LBP) and Histogram of Oriented Gradient (HOG), to extract local characteristics. The considered descriptors are widely used in computer vision applications [42] to identify edges more accurately. LBP, also known as a texture descriptor, reflects the relation among pixels intensity within a given area. For a point (x, y) surrounded by n pixels in radius r, the value of LBP can be computed by the following equations:

$$LBP(n,r) = \sum_1^n g(x) \times 2^{n-1} \tag{2}$$

$$g(x) = \begin{cases} 1 & I_c < I_i \\ 0 & I_c \geq I_i \end{cases} \; i \in \{1,2,\dots,n\} \tag{3}$$

In the above relations, Ic and Ii demonstrate the center point (x, y) and its adjacents light intensity, respectively.

HOG descriptor calculates occurrences of gradient orientation in local divisions of an image to recognize specific objects. As shown in Figure 2, the image is initially partitioned into equal areas, known as blocks. Each block is a set of cells made up of pixels that Gradient intensities are calculated within them and used to organize the histograms. Eventually, the histogram of the cells in each block is located consecutively to obtain an image description.

To concentrate on edge features during training, only boundaries' LBP and HOG values are inputted into the model. For this purpose, mathematical morphological operations are applied. Mathematical morphological operation is an image processing technique that deals with an image's form and shape. Morphological methods are commonly utilized in preprocessing or post-processing steps to filter, narrow, or clean up images. The two fundamental morphological operators are dilation and erosion, where dilation detects the highest value in a region and erosion finds the lowest element in a neighborhood. Various techniques, including morphological gradients, are developed from these functions. It would be utilized for edge detection and segmentation. The morphological gradient is obtained by subtracting erosion from dilation with adjusted kernel size parameters. In the below equations, f presents the input image, k relates to the desired kernel in morphological gradient G (f, k), $\otimes$ demonstrates elementwise multiplication, and $\oplus$ and $\ominus$ illustrate erosion and dilation, respectively. Figure 3 exhibits a preprocessed image, its extracted features, and ground truth.

$$G(f,k) = f \oplus k - f \ominus k \tag{4}$$

$$LBP_{OD} = LBP_{image} \otimes G(f,k)_{OD\ ground\ truth} \tag{5}$$

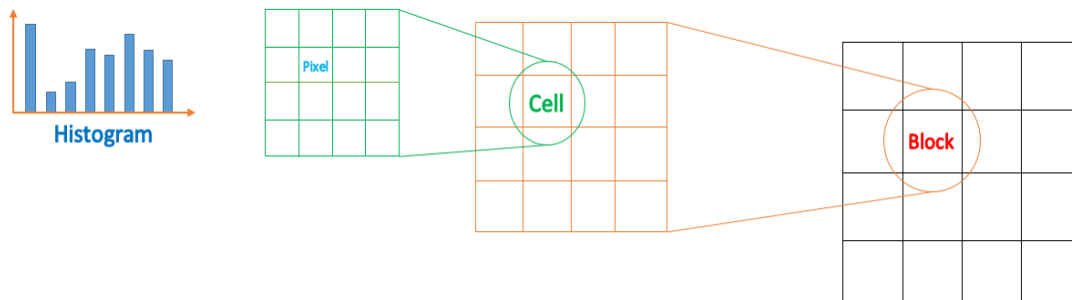$$HOG_{OD} = HOG_{image} \otimes G(f,k)_{OD\ ground\ truth} \tag{6}$$

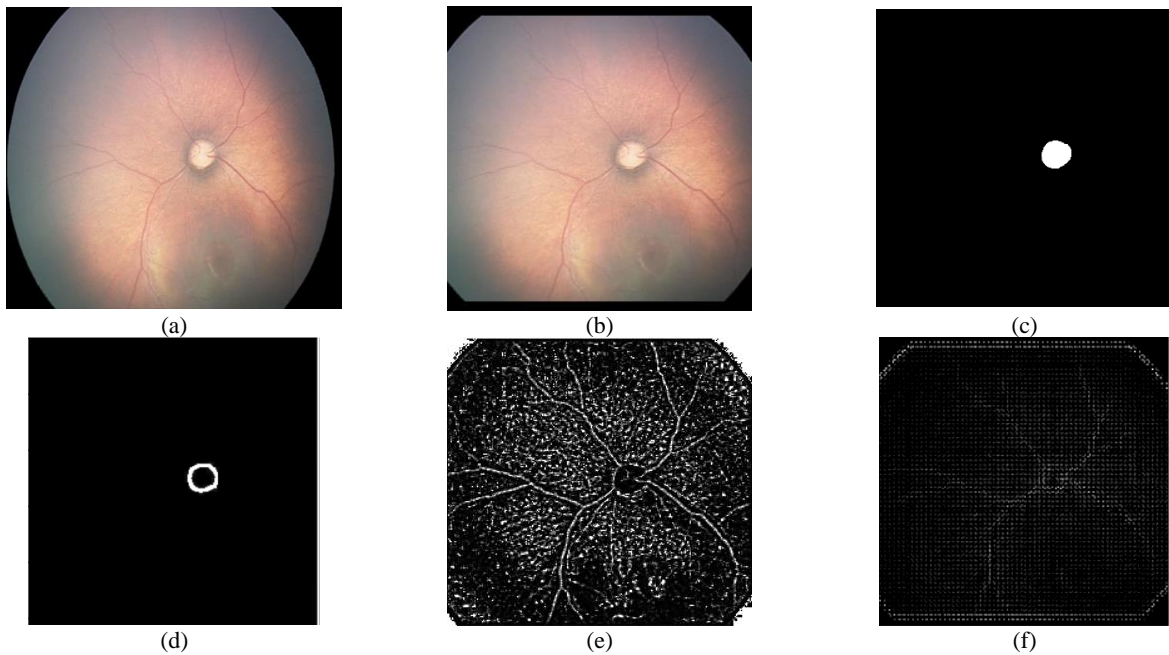**Figure 2.** Subdividing procedure in HOG calculation

Figure 3. (a) Input image (b) Preprocessed image (c) Ground truth (d) Morphological gradient of OD ground truth (e) Local binary pattern (f) Histogram of oriented gradient

2. 5. Model Architecture   In this step, the model's architecture is designed, which greately affects the final outcome of the model. In recent years, there has been increasing literature on biomedical segmentation by using the U-Net architecture; a U-shape convolutional network consists of a contracting path (encoder) expanding path (decoder), presented in Figure 4. The contracting steps include a successive bunch of two convolutional layers, followed by a downsampling operation. Each expanding stage concatenates the up-sampled feature maps from the previous stage with corresponding feature maps from the contracting layers. Due to the capability of the U-Net architecture to train on

small samples in segmentation tasks, various derivations of it have been proposed in recent studies.

As illustrated in Figure 5, the concept of our model architecture is based on the U-Net. To import extracted features, a mini U-Net is parallelized to the main U-Net. In addition, pre-trained VGG-16 weights are substituted in the primary encoder, so only the decoder section is trained. Main model inputs are preprocessed RGB channels that acquire an initial segmentation output. The mini U-Net outcomes are another raw output obtained from the green channel besides LBP and HOG feature maps. In the recommended model, before concatenating the skip connection and upsampling, an attention gate
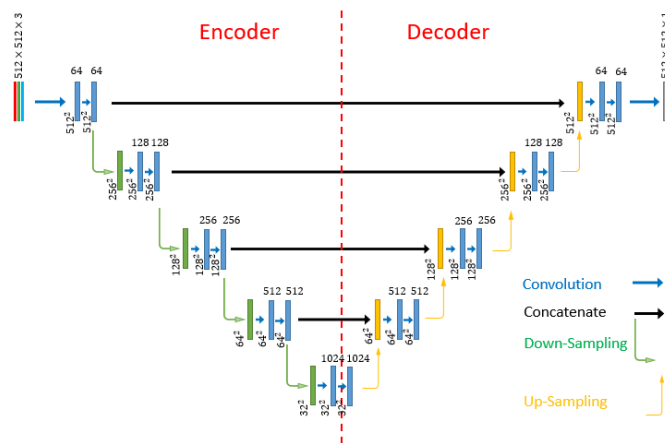

Figure 4. U-Net architecture

(Figure 6) was implemented to suppress irrelevant regions' activations and reduce the excessive characteristics brought across. Ultimately, these results are inputted into the post-processing model to amend each other. Another advantage of the attention gate is that it does not require crop ROI to be inputted into the model, and the entire image is used as the input.

3. RESULTS

Since our segmentation problem has unbalanced foreground and background pixels, a linear combination of dice coefficient and binary cross-entropy loss function was preferred. Let define them by L_{dice} and L_{bce} as below:

$$L_{dice} = \frac{1}{N}\left(1 - \frac{2\sum_{i=1}^{h}\sum_{j=1}^{w}w_{ij}^2 y_{ij}\hat{y}_{ij}+\varepsilon}{\sum_{i=1}^{h}\sum_{j=1}^{w}w_{ij}^2 y_{ij}+\sum_{i=1}^{h}\sum_{j=1}^{w}w_{ij}^2\hat{y}_{ij}+\varepsilon}\right) \tag{7}$$

$$L_{bce} = \frac{1}{N}\left(\sum_{i=1}^{h}\sum_{j=1}^{w}\hat{y}_{ij}\log(y_{ij}) + (1-\hat{y}_{ij})\log(1-y_{ij})\right) \tag{8}$$

$$L_{total} = L_{dice} + L_{dice} \tag{9}$$

where N is the number of training images, w_{ij}^2 is the gained weight from ground truth, ε is minimum amount prevent from zero division, and $\hat{y}_{ij}$ and y_{ij} are the target and predicted value, respectively.

Nadam was selected, as an optimizer, to train the proposed model. The momentum in Nadam is the opting motivation since the momentum prevents from getting stuck in local minimums.

IoU and dice coefficient are selected criteria to measure segmentation accuracy. The equations that describe them are stated as follows:

$$Dice = \frac{2\times TP}{TP+TN+FP+FN} \tag{10}$$

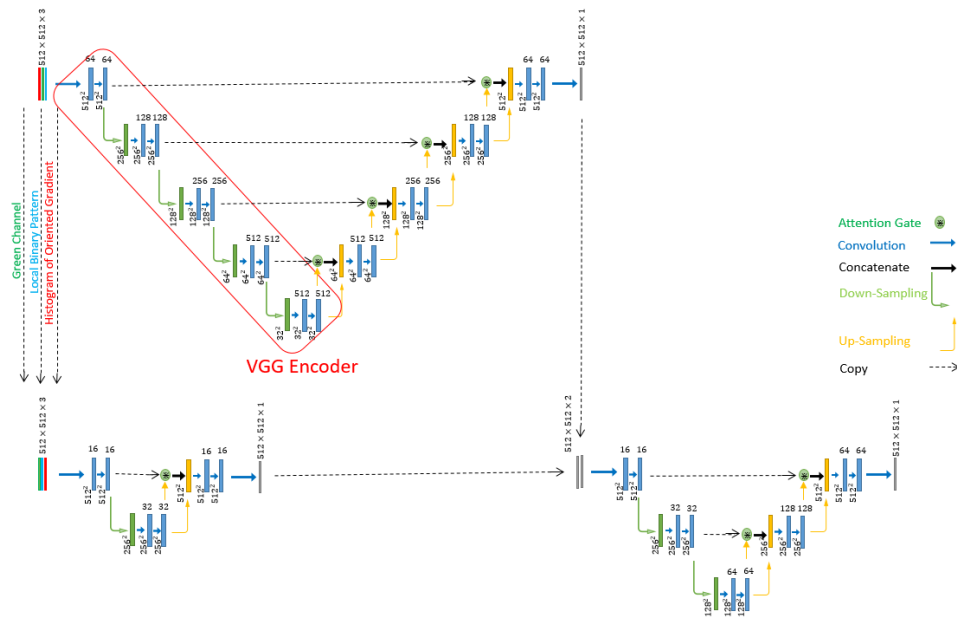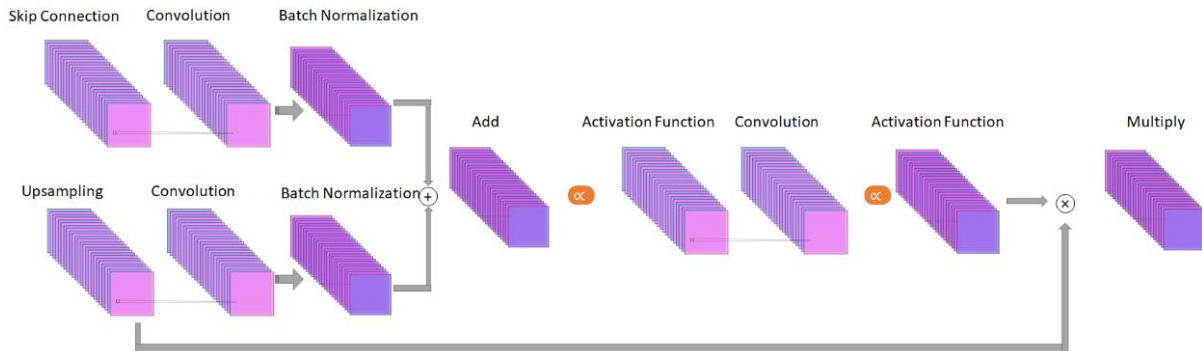**Figure 5.** Proposed model

Figure 6. Attention block

$$IoU = \frac{TP}{TN+FP+FN} \qquad (11)$$

Parameters used in Equations (9) and (10) are presented in Table 1.

The previous sections have revealed that the suggested model consists of modified U-Net models and a VGG-16 encoder. Hence, the authors have also considered the original U-Net, pre-trained U-Net encoder, and proposed model without attention. These comparisons can highlight model advantages. It appears from Figure 7 that the U-Net model and pre-trained U-Net performance individuality are not very accurate for OD segmentation on the database; by contrast, the recommended model without attention gate has performed admissible. Moreover, it is apparent from Figure 7 that accentuating on edge features and model architecture could improve the results significantly.

A more detailed comparison is presented in Figure 8. As shown in this figure, our method functions the same as the absence of attention gate on typical images, and they beat the original U-Net and pre-trained U-Net. As illustrated in Figure 9, the suggested method performs significantly superior to the architecture without attention to challenging images. In these images, OD boundaries have low contrast and cause problems due to the illumination and retina structure. The recommended technique could dominate this issue by emphasizing on edge features throughout the training time. The most striking result from the comparison was the lack of contour recognition on the image by the original U-Net.

4. DISCUSSION

Due to the growing rate of premature infants' birth and their inability to express vision problems, researchers' consideration has been drawn to intelligent diagnosis. Hence, the objective of the paper is to contribute to ROP computer-aided systems through segmenting OD. In this study, a novel adaptation of U-Net attention architecture was designed. In addition to the main image channels (RGB), the model inputs contain extracted local features (LBP and HOG) focused on OD boundaries. Pre-trained weights are employed in the primary U-Net encoder to lessen training time cost and enhance the operation. The findings obtained in this research are consistent with previous studies [40, 41], which have evaluated the effect of local features as model inputs. Figure 8 indicates extreme improvements within modifications initiated from the original U-Net toward the proposed model.
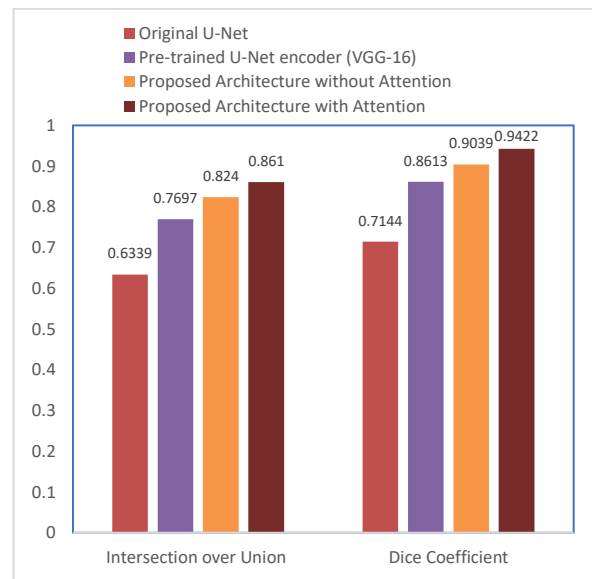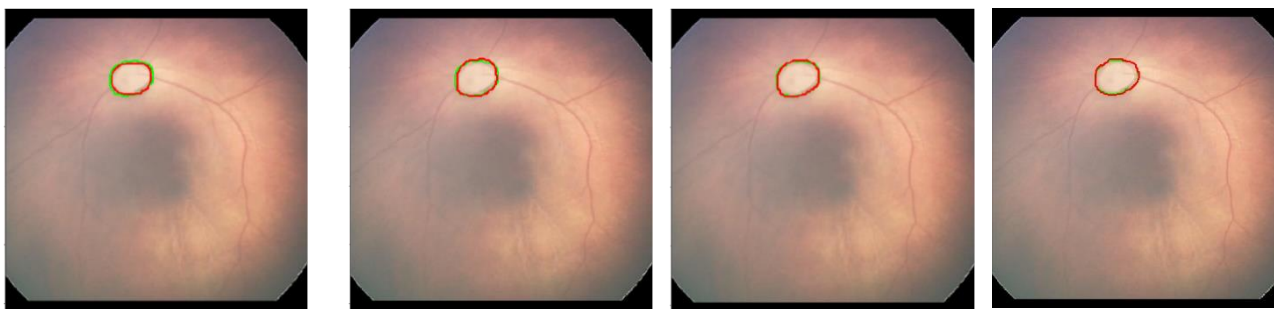
TABLE 1. Parameters definition

Expected Output \ Model output	Non Optic Disk	Optic Disk
Optic Disk	False Negative (FN)	True positive (TP)
Non Optic Disk	True Negative (TN)	False Positive (FP)

Figure 7. Comparison of segmentation accuracy in different models

| Original U-Net | Pre-trained U-Net encoder (VGG-16) | Proposed architecture without attention gate | Proposed Model |

Figure 8. Segmentation results (Ground truth and segmented object are plotted in green and red respectively)

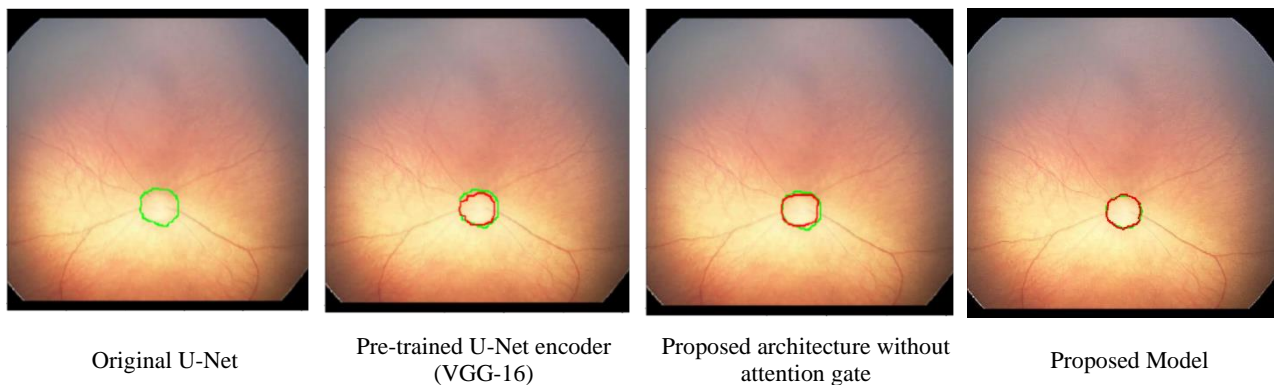| Original U-Net | Pre-trained U-Net encoder (VGG-16) | Proposed architecture without attention gate | Proposed Model |

Figure 9. Various segmentation methods comparison on challenging image (Ground truth and segmented object are plotted in green and red respectively)

As a consequence of the shortage of infants' retina imaging equipment, there is no publicly available database. To the best of the authors' knowledge, this study is the first research on dealing with neonatal OD segmentation; hence, the paper's main study limitation was the lack of similar studies to compare the results and make a proper assessment. It enables readers to review the advantages and disadvantages of different techniques.

Since the peripapillary choroid pigmentation is usually different in adults than infants, OD borders detection is more complex in neonates [43]. Moreover, the OD color is mostly pink or reddish because some premature infants have primary vitreous vessels covering the OD's surface. Furthermore, the OD boundaries of infants are blurred compared to adults and relatively irregular. In neonatal retinal imaging, numerous captured images are slightly blurred due to infants' lack of cooperation (crying and intense shaking) during imaging. Besides, the eyes' blood vessels of infants are forming and growing more rapidly than adults; consequently, OD characteristics are in a broader range that adds to recognition challenges. Consequently, fundus data acquire from adults are not proper diagnostic criteria for neonatal fundus screening.

To provide a valid database, semi-blurred images are also collected, which reduces accuracy. Considerably more work will be required to be done to automatically detect anatomical structures of the retina. These attempts can push forward the study to determine the ROP zone and stage. Moreover, it could be regulated and modified regarding adult retina and utilized in glaucoma diagnosis. Further, it is recommended to perform more general studies on other local features.

5. CONCLUSION

The progression of neonatal intensive care units has led to pay heed to ROP as a serious, worldwide public health concern. This study set out to determine the infants' OD boundary by CNN. The novel architecture results show a reliable response to OD segmentation. The illumination robustness of the method is achieved from importing the extracted features and executing the attention gate. On the other hand, the significant limitation of this study is the deficiency of similar investigation, the current outcomes append to a growing body of literature on segmenting, and ROP automated examination. It is recommended to conduct more studies on adding further properties to the models and determining ROP severity.

6. ETHICAL ISSUES

The ethical code for this research is IR.TUMS.CHMC.REC.1398.038.

7. ACKNOWLEDGEMENT

8. REFERENCES

1. Chen, J. and Smith, L.E., "Retinopathy of prematurity", *Angiogenesis*, Vol. 10, No. 2, (2007), 133-140, https://doi.org/10.1016/S0140-6736(13)60178-6

2. Lumley, J., "1 the epidemiology of preterm birth", *Bailliere's Clinical Obstetrics Gynaecology*, Vol. 7, No. 3, (1993), 477-498, https://doi.org/10.1016/S0950-3552(05)80445-6

3. Hartnett, M.E., "Pathophysiology and mechanisms of severe retinopathy of prematurity", *Ophthalmology*, Vol. 122, No. 1, (2015), 200-210, https://doi.org/10.1016/j.ophtha.2014.07.050

4. Tasman, W.S., "Revised indications for the treatment of retinopathy of prematurity: Results of the early treatment for retinopathy of prematurity randomized trial", *Arch Ophthalmol*, Vol. 121, (2003), 1684-1695, https://doi.org/10.1001/archopht.121.12.1684

5.   Stensvold, H.J., Klingenberg, C., Stoen, R., Moster, D., Braekke, K., Guthe, H.J., Astrup, H., Rettedal, S., Gronn, M. and Ronnestad, A.E., "Neonatal morbidity and 1-year survival of extremely preterm infants", *Pediatrics*, Vol. 139, No. 3, (2017), https://doi.org/10.1542/peds.2016-1821

6.   Aslam, T., Fleck, B., Patton, N., Trucco, M. and Azegrouz, H., "Digital image analysis of plus disease in retinopathy of prematurity", *Acta Ophthalmologica*, Vol. 87, No. 4, (2009), 368-377, https://doi.org/10.1111/j.1755-3768.2008.01448.x

7.   Naqvi, S.S., Fatima, N., Khan, T.M., Rehman, Z.U. and Khan, M.A., "Automatic optic disk detection and segmentation by variational active contour estimation in retinal fundus images", *Signal, Image Video Processing*, Vol. 13, No. 6, (2019), 1191-1198, https://doi.org/10.1007/s11760-019-01463-y

8.   Fatima, K.N., Akram, M.U. and Bazaz, S.A., "Papilledema detection in fundus images using hybrid feature set", in 2015 5th International Conference on IT Convergence and Security (ICITCS), IEEE. (2015), 1-4.

9.   Walter, T., Klein, J.-C., Massin, P. and Erginay, A., "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina", *IEEE Transactions on Medical Imaging*, Vol. 21, No. 10, (2002), 1236-1243, https://doi.org/10.1109/TMI.2002.806290

10. Welfer, D., Scharcanski, J., Kitamura, C.M., Dal Pizzol, M.M., Ludwig, L.W. and Marinho, D.R., "Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach", *Computers in Biology Medicine*, Vol. 40, No. 2, (2010), 124-137, https://doi.org/10.1016/j.compbiomed.2009.11.009

11. Tjandrasa, H., Wijayanti, A. and Suciati, N., "Optic nerve head segmentation using hough transform and active contours", *Telkomnika*, Vol. 10, No. 3, (2012), 531, https://doi.org/10.12928/telkomnika.v10i3.833

12. Abdullah, M., Fraz, M.M. and Barman, S.A., "Localization and segmentation of optic disc in retinal images using circular hough transform and grow-cut algorithm", *PeerJ*, Vol. 4, (2016), e2003, https://doi.org/10.7717/peerj.2003

13. Thongnuch, V. and Uyyanonvara, B., "Automatic optic disk detection from low contrast retinal images of rop infant using gvf snake", *Suranaree J Sci Technol*, Vol. 14, No. 3, (2007), 223-234, doi.

14. Pathan, S., Kumar, P., Pai, R. and Bhandary, S.V., "Automated detection of optic disc contours in fundus images using decision tree classifier", *Biocybernetics Biomedical Engineering*, Vol. 40, No. 1, (2020), 52-64, https://doi.org/10.1016/j.bbe.2019.11.003

15. Mookiah, M.R.K., Acharya, U.R., Chua, C.K., Min, L.C., Ng, E., Mushrif, M.M. and Laude, A., "Automated detection of optic disk in retinal fundus images using intuitionistic fuzzy histon segmentation", *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, Vol. 227, No. 1, (2013), 37-49, https://doi.org/10.1177/0954411912458740

16. ASADI, A.S., Hassanpour, H., Shahiri, M. and Ghaderi, R., "Detection of microaneurysms in retinal angiography images using the circular hough transform", (2012).

17. Thongnuch, V. and Uyyanonvara, B., "Automatic detection of optic disc from fundus images of rop infant using 2d circular hough transform", *Sirindhorn International Institute of Technology, Thammasat University, Thailand*, Vol., (2006).

18. Zahoor, M.N. and Fraz, M.M., "Fast optic disc segmentation in retina using polar transform", *IEEE Access*, Vol. 5, (2017), 12293-12300, https://doi.org/10.1109/ACCESS.2017.2723320

19. Ghadiri, F., Bergevin, R. and Shafiee, M., "An adaptive thresholding approach for automatic optic disk segmentation", arXiv preprint arXiv:.05104, (2017).

20. Septiarini, A., Harjoko, A., Pulungan, R. and Ekantini, R., "Optic disc and cup segmentation by automatic thresholding with morphological operation for glaucoma evaluation", *Signal, Image Video Processing*, Vol. 11, No. 5, (2017), 945-952, doi. https://doi.org/10.1007/s11760-016-1043-x

21. Rehman, Z.U., Naqvi, S.S., Khan, T.M., Arsalan, M., Khan, M.A. and Khalil, M., "Multi-parametric optic disc segmentation using superpixel based feature classification", *Expert Systems with Applications*, Vol. 120, (2019), 461-473, https://doi.org/10.1016/j.eswa.2018.12.008

22. Dutta, M.K., Mourya, A.K., Singh, A., Parthasarathi, M., Burget, R. and Riha, K., "Glaucoma detection by segmenting the super pixels from fundus colour retinal images", in 2014 international conference on medical imaging, m-health and emerging communication systems (MedCom), IEEE. (2014), 86-90.

23. Morales, S., Naranjo, V., Angulo, J. and Alcañiz, M., "Automatic detection of optic disc based on pca and mathematical morphology", *IEEE Transactions on Medical Imaging*, Vol. 32, No. 4, (2013), 786-796, doi. https://doi.org/10.1109/tmi.2013.2238244

24. Dharmawan, D.A., Ng, B.P. and Rahardja, S., "A new optic disc segmentation method using a modified dolph-chebyshev matched filter", *Biomedical Signal Processing Control*, Vol. 59, (2020), 101932, https://doi.org/10.1016/j.bspc.2020.101932

25. Sezavar, A., Farsi, H. and Mohamadzadeh, S., "A modified grasshopper optimization algorithm combined with cnn for content based image retrieval", *International Journal of Engineering*, *Transactions A: Basics*, Vol. 32, No. 7, (2019), 924-930, https://doi.org/10.5829/ije.2019.32.07a.04

26. Feizi, A., "Convolutional gating network for object tracking", *International Journal of Engineering*, *Transactions A: Basics*, Vol. 32, No. 7, (2019), 931-939, https://doi.org/10.5829/ije.2019.32.07a.05

27. Bagheria, F., Tarokh, M. and Ziaratbanb, M., "Semantic segmentation of lesions from dermoscopic images using yolo-deeplab networks", *International Journal of Engineering*, *Transactions B: Applications*, Vol. 34, No. 2, (2021), 458-469, doi. https://doi.org/10.5829/ije.2021.34.02b.18

28. Sevastopolsky, A., "Optic disc and cup segmentation methods for glaucoma detection with modification of u-net convolutional neural network", *Pattern Recognition Image Analysis*, Vol. 27, No. 3, (2017), 618-624, https://doi.org/10.1134/s1054661817030269

29. Kim, J., Tran, L., Chew, E.Y., Antani, S. and Thoma, G.R., "Optic disc segmentation in fundus images using deep learning", in Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications, International Society for Optics and Photonics. Vol. 10954, (2019), 109540H.

30. Yu, S., Xiao, D., Frost, S. and Kanagasingam, Y., "Robust optic disc and cup segmentation with deep learning for glaucoma detection", *Computerized Medical Imaging Graphics*, Vol. 74, (2019), 61-71, doi. https://doi.org/10.1016/j.compmedimag.2019.02.005

31. Juneja, M., Singh, S., Agarwal, N., Bali, S., Gupta, S., Thakur, N. and Jindal, P., "Automated detection of glaucoma using deep learning convolution network (g-net)", *Multimedia Tools Applications*, Vol. 79, (2020), https://doi.org/10.1007/s11042-019-7460-4

32. Bhatkalkar, B.J., Reddy, D.R., Prabhu, S. and Bhandary, S.V., "Improving the performance of convolutional neural network for the segmentation of optic disc in fundus images using attention gates and conditional random fields", *IEEE Access*, Vol. 8,

(2020), 29299-29310, https://doi.org/10.1109/access.2020.2972318

33. Jiang, Y., Xia, H., Xu, Y., Cheng, J., Fu, H., Duan, L., Meng, Z. and Liu, J., "Optic disc and cup segmentation with blood vessel removal from fundus images for glaucoma detection", in 2018 40th annual international conference of the ieee engineering in medicine and biology society (EMBC), IEEE. (2018), 862-865.

34. Zheng, Y., Zhang, X., Xu, X., Tian, Z. and Du, S., "Deep level set method for optic disc and cup segmentation on fundus images", *Biomedical Optics Express*, Vol. 12, No. 11, (2021), 6969-6983, https://doi.org/10.1364/boe.439713

35. Tabassum, M., Khan, T.M., Arsalan, M., Naqvi, S.S., Ahmed, M., Madni, H.A. and Mirza, J., "Cded-net: Joint segmentation of optic disc and optic cup for glaucoma screening", *IEEE Access*, Vol. 8, (2020), 102733-102747, https://doi.org/10.1109/access.2020.2998635

36. Nisha, K., Sreelekha, G., Sathidevi, P., Mohanachandran, P. and Vinekar, A., "A robust algorithm for automated detection of optic disc in retina of premature infants", in TENCON 2019-2019 IEEE Region 10 Conference (TENCON), IEEE. (2019), 684-689.

37. Alais, R., Dokládal, P., Erginay, A., Figliuzzi, B. and Decencière, E., "Fast macula detection and application to retinal image quality assessment", *Biomedical Signal Processing Control*, Vol. 55, (2020), 101567, https://doi.org/10.1016/j.bspc.2019.101567

38. Wang, L., Liu, H., Lu, Y., Chen, H., Zhang, J. and Pu, J., "A coarse-to-fine deep learning framework for optic disc

segmentation in fundus images", *Biomedical Signal Processing Control*, Vol. 51, (2019), 82-89, https://doi.org/10.1016/j.bspc.2019.01.022

39. Ke, P., Cai, M., Wang, H. and Chen, J., "A novel face recognition algorithm based on the combination of lbp and cnn", in 2018 14th IEEE International Conference on Signal Processing (ICSP), IEEE. (2018), 539-543.

40. Touahri, R., AzizI, N., Hammami, N.E., Aldwairi, M. and Benaida, F., "Automated breast tumor diagnosis using local binary patterns (lbp) based on deep learning classification", in 2019 International Conference on Computer and Information Sciences (ICCIS), IEEE. (2019), 1-5.

41. Badaghei, R., Hassanpour, H. and Askari, T., "Detection of bikers without helmet using image texture and shape analysis", *International Journal of Engineering*, *Transactions C: Aspects*, Vol. 34, No. 3, (2021), 650-655, https://doi.org/10.5829/ije.2021.34.03c.09

42. Hellström, A., Hård, A., Chen, Y., Niklasson, A. and Albertsson-Wikland, K., "Ocular fundus morphology in preterm children. Influence of gestational age, birth size, perinatal morbidity, and postnatal growth", *Investigative Ophthalmology Visual Science*, Vol. 38, No. 6, (1997), 1184-1192.

43. Feng, X., Nan, Y., Pan, J., Zou, R. and Shen, L., "Comparative study on optic disc features of premature infants and term infants", (2020), https://doi.org/10.21203/rs.3.rs-30626/v1

Persian Abstract

چکیده

بیماری رتینوپاتی نارسی مهمترین علت نابینایی در نوزادان بوده که هر سال تعداد بسیار زیادی از نوزادان به آن مبتلا می‌شوند. با توجه به گسترش زیاد بیماری و عوارض نامطلوب آن تشخیص صحیح و به موقع رتینوپاتی نارسی بسیار ارزشمند است. تشخیص به موقع بیماری نیازمند افزایش معاینات نوزادان بوده که موجب افزایش نیاز به پزشک متخصص شده که اجرای آن دشوار است. به همین علت لزوم وجود سیستم‌های هوشمند برای کمک به پزشکان بیش از پیش احساس می‌گردد. در طراحی سیستم‌های هوشمند تشخیص رتینوپاتی نارسی یکی از مهمترین قسمت های شبکیه برای تعیین میزان شدت و پیشروی این بیماری سر عصب بینایی که در این پژوهش به آن پرداخته شده است. در این مقاله یک مدل نوآورانه یادگیری عمیق بر مبنای مکانیزم توجه برای شناسایی محل و ناحیه سر عصب بینایی معرفی شده است. در روش ارایه شده، ابتدا تصاویر را پیش پردازش کرده و سپس تعداد آنها افزایش داده می‌شود. پس از آن برای تعیین محل دقیق ناحیه مربوط به سر عصب بینایی با بهبود شبکه U-Net به یک معماری بهینه دست می‌یابیم. معماری به کار رفته یک شبکه کانولوشنی دو مرحله‌ای است. در مرحله اول با استفاده از تصویر و ویژگی‌های تصویر به دو خروجی متفاوت دست کرده که ورودی مرحله دوم شبکه می‌باشد. مرحله دوم به عنوان پس پردازش، با بررسی خروجی‌ها و استفاده از مکانیزم توجه یک نتیجه نهایی فراهم نموده که براساس معیار ضریب تشابه سورنسون و معیار ژاکارد به ترتیب دارای دقت ۹۴/۲۲ درصد و ۸۶/۱ درصد میباشد.