*Healthcare Transformation with Informatics and Artificial Intelligence*
*J. Mantas et al. (Eds.)*

*doi:10.3233/SHTI230576*

# Optical Disc Segmentation from Retinal Fundus Images Using Deep Learning

Mohammad Tariqul ISLAM[a], Ferdaus AHMED[b], Mowafa HOUSEH[c] and Tanvir ALAM[c,1]
[a] *Southern Connecticut State University, Connecticut, USA*
[b] *Infosys Ltd, Texas, USA*
[c] *College of Science and Engineering, Hamad bin Khalifa University, Doha, Qatar*

**Abstract.** The optical disc in the human retina can reveal important information about a person's health and well-being. We propose a deep learning-based approach to automatically identify the region in human retinal images that corresponds to the optical disc. We formulated the task as an image segmentation problem that leverages multiple public-domain datasets of human retinal fundus images. Using an attention-based residual U-Net, we showed that the optical disc in a human retina image can be detected with more than 99% pixel-level accuracy and around 95% in Matthew's Correlation Coefficient. A comparison with variants of UNet with different encoder CNN architectures ascertains the superiority of the proposed approach across multiple metrics.

**Keywords.** Retina, Segmentation, CNN, Qatar Biobank

## 1. Introduction

The diagnosis of various diseases such as diabetes [1], cardiovascular disease [2], glaucoma [3], often relies on retinal fundus images, particularly the size, shape, and pathological features of the Optic disc (OD). Ophthalmologists use OD segmentation, which is an integral part of software [4,5], to aid in the diagnosis of these diseases. While deep learning methods have shown promising results in accurately segmenting the OD, there is still room for improvement and different research groups have used different datasets [3-5]. This article proposes combining the three most widely used datasets for OD segmentation to formulate the problem of optical disc detection in retinal images as an image segmentation task. A deep neural network can be trained on this composite dataset, which includes retinal fundus images and corresponding optical disc segmentation masks, to identify the region occupied by the optic disc in unseen retinal images that is occupied by the optic disc.

[1] Corresponding Author: Tanvir Alam. E-mail : talam@hbku.edu.qa.

## 2. Methodology

### *2.1. Dataset*

There exist multiple datasets that are used in the literature for OD segmentation [3]. Among them IDRiD, Drishti-GS1, and RIM-ONE are the most widely used dataset for OD segmentation tasks. In this work, we build a composite dataset by combining these three retinal image datasets: IDRiD, Drishti-GS1, and RIM-ONE.

### *2.2. Proposed Approach*

Given a dataset of retinal image and corresponding optic disc binary masks, we trained a deep neural network, UNet with attention mechanism to predict the masks. The UNet is a U-shaped architecture that downscales the input images to a bottleneck layer in the encoder part of the network before upscaling to the original image size to predict pixel-level class probabilities in the decoder section. Our proposed variant of UNet uses GoogleNet as the encoder along with an attention mechanism to allow it to focus on relevant regions of the retinal image and intermediately generated feature maps that it finds most useful for the segmentation task. We used a pretrained GoogleNet in the UNet encoder that had been trained on the ImageNet-1k dataset. To compare the performance of our proposed network architecture to the same for widely-used alternatives, we experimented with five other CNNs as the encoder for the UNet architecture. Specifically, we fine-tuned a UNet with the following networks as the encoder: AlexNet, VGG, GoogleNet, ResNet, SqueezeNet, and EfficientNet. The backbone networks had been pretrained on the ImageNet-1k dataset that allowed them a head start with the basic features to be learnt for image analysis. The pretrained networks were fine-tuned with a cosine annealing learning rate scheduler with a maximum learning rate of 1e-2 for 40 epochs. We applied weight decay to prevent the networks from overfitting with a regularization constant value of 4e-1. In addition, early stopping with a minimum improvement threshold of 0.01 and a patience of 10 iterations was also used for the same purpose that monitored the mcc metric. The activation function of our choice was ReLU in the hidden layers and per-pixel sigmoid for the output layer. We used Adam optimization and weighted cross entropy loss function with class weights computed from the training set as the inverse of class representation frequency to prevent the network from collapsing to a single class-predicting model. We used nested K-fold cross validation at the image-level to statistically stabilize the outcomes from our experiments.

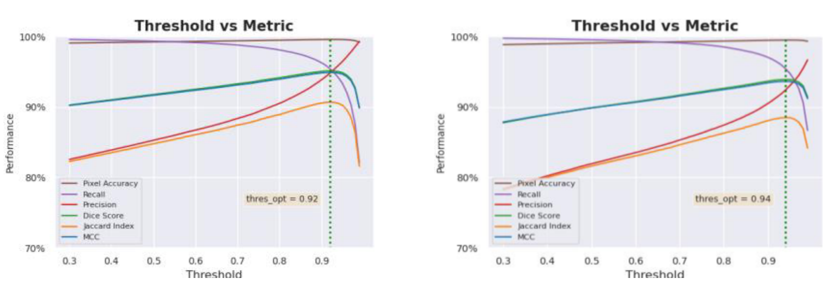


**Figure 1.** Selecting the best threshold for classification (performed on the validation set). We show computations from two folds as samples.

We used accuracy, precision, recall, and Matthew's correlation coefficient as the evaluation metrics. From segmentation-specific categories, we used Dice score and Jaccard's index. To allow the network to learn the true pattern of optical discs, we provided it with class weights that were computed as the inverse of class frequencies in the training set. We show how the values of the metrics changed as the threshold was varied from 0.3 to 0.99 in two of the five folds (Figure 1).

## 3. Results

Table 1 highlights the performance of all the models we tested. Based on the experiments, UNet with GoogleNet encoder performed the best among all the models. Inspecting the MCC column reveals that the Googlenet-based UNet performs the best. Moreover, it outperforms all other candidate models even in other metrics, too.

**Table 1.** Performance of the candidate models on the test set.

| Network | Pixel Accuracy | Recall | Precision | F1-Score | Dice Score | Jaccard Index | MCC |
|---|---|---|---|---|---|---|---|
| Alexnet | 98.68 ± 0.12 | 87.63 ± 3.65 | 84.18 ± 2.26 | 85.78 ± 0.77 | 85.78 ± 0.77 | 75.1 ± 1.18 | 85.16 ± 0.87 |
| VGG16 | 98.63 ± 0.12 | 88.22 ± 3.4 | 82.87 ± 3.84 | 85.32 ± 1.39 | 85.32 ± 1.39 | 74.43 ± 2.12 | 84.73 ± 1.35 |
| ResNet34 | 98.46 ± 0.03 | 90.65 ± 1.06 | 91.65 ± 1.16 | 92.14 ± 0.12 | 92.14 ± 0.12 | 86.92 ± 0.21 | 91.86 ± 0.11 |
| SqueezeNet 1_0 | 99.0 ± 0.08 | 90.25 ± 3.05 | 88.07 ± 1.42 | 89.09 ± 0.97 | 89.09 ± 0.97 | 80.34 ± 1.56 | 88.6 ± 1.0 |
| DenseNet121 | 98.78 ± 0.08 | 89.3 ± 1.74 | 84.78 ± 1.87 | 86.95 ± 0.56 | 86.95 ± 0.56 | 76.92 ± 0.87 | 86.36 ± 0.56 |
| **GoogleNet** | **99.47 ± 0.05** | **94.65 ± 1.31** | **93.8 ± 0.93** | **94.22 ± 0.45** | **94.22 ± 0.45** | **89.07 ± 0.8** | **93.95 ± 0.47** |

## 4. Discussion

Figure 2 shows that our proposed network architecture was able to segment the region corresponding to the optical disc from the retina images taken under a variety of lighting conditions. For example, presence of microaneurysm (row 3) causes the classifier to falsely output high probability for these regions to be a part of the optical disc, as the sporadic darker spots show near the lower right edge on the probability map (column c and d). We can also see that in significantly darker images where the edge of the optical disc region is barely discernible from the rest of the retina, our classifier is prone to predicting more false negatives (the purple region surrounding the optical disc in column e). Similarly, for washed-out images of the retina where the border of the optical disk blends with the background class (e.g., row 2), false positives are more prominent.

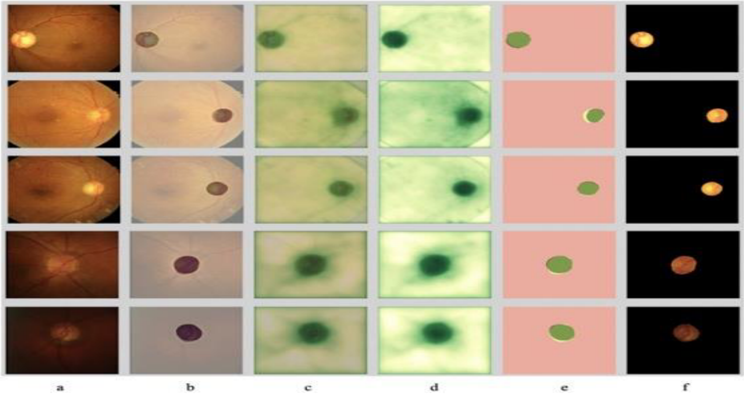


**Figure 2**. From left: (a) Ground truth retinal images, (b) predicted segmentation mask overlaid on original images, (c) prediction probabilities overlaid on original images, (d) predicted class probabilities, (e) ground truth and predicted mask, and (f) extracted optical disc using the predicted mask.

## 5. Conclusion

We propose an image segmentation-based deep learning approach to identify the optic disc in human retinal fundus images. We showed that our proposed approach is better than other candidate models by using multiple evaluation metrics for both general classification tasks and image segmentation. Our future plan includes integrating this method as a diagnostic tool for diseases such as diabetes and DR, using retinal image data gathered from Qatar Biobank (QBB).

## References

[1] Islam MT, Al-Absi HRH, Ruagh EA, Alam T. DiaNet: A Deep Learning Based Architecture to Diagnose Diabetes Using Retinal Images Only. IEEE Access 2021. (accessed April 13, 2023).
[2] Al-Absi HRH, Islam MT, Refaee MA, Chowdhury MEH, Alam T. Cardiovascular Disease Diagnosis from DXA Scan and Retinal Images Using Deep Learning. Sensors 2022;22. https://doi.org/10.3390/s22124310.
[3] Xiong H, Liu S, Sharan RV, Coiera E, Berkovsky S. Weak label based Bayesian U-Net for optic disc segmentation in fundus images. Artif Intell Med 2022;126:102261.
[4] Hasan MK, Alam MA, Elahi MTE, Roy S, Martí R. DRNet: Segmentation and localization of optic disc and Fovea from diabetic retinopathy image. Artif Intell Med 2021;111:102001.
[5] Niemeijer M, Abràmoff MD, van Ginneken B. Fast detection of the optic disc and fovea in color fundus photographs. Med Image Anal 2009;13:859–70.