Big Data and Artificial Intelligence (AI) to Detect Glaucoma

Lei Yang*, Yao Xie *, Md Rafiqul Islam *, Guandong Xu *
*Data Science Institute (DSI), University of Technology Sydney, NSW, Australia
{lei.yang, yao.xie}@student.uts.edu.au, {MdRafiqul.Islam, Guandong.Xu}@uts.edu.au

***Abstract*—Glaucoma is an eye condition that is one of the most prevalent causes of blindness due to damage to the optic nerve. According to the existing studies, it is the second most common reason for vision loss in the world and estimated 80 million people affected with glaucoma by 2020. With the advance development of artificial intelligence (AI), it is necessary to develop a glaucoma detection system for diagnosis. Therefore, the primary objective of this research is to develop a system for detecting glaucoma using the retinal fundus images, which can be useful to determine if the patient was affected by glaucoma or not. Although several methods applied to detect glaucoma in the past decades, it is essential to apply an advance AI technique with a glaucoma detection tool. Thus, in this paper, we divided our task into three-fold: 1) segmentation, 2) classification, and 3) deployment. The U-Net architecture is implemented for segmentation. Pretrained GC-Net model is proposed for classification. Finally, based on the segmentation and classification, we develop a glaucoma detection system for diagnosis. This study uses ACRIMA datasets for the purpose of training and testing. The result of this model is evaluated using various parameters such as accuracy, sensitivity , specificity , f1-score, and auc score. The output is compared to other current deep learning models used for classification, such as ResNet, CNN, Inception V3, and TB-Net. The proposed model achieved 96% accuracy in training and 93% accuracy in testing. Overall, the performance of the proposed model is better in all the analysis. The source code and demo of the proposed framework are publicly available at https://github.com/rafiqulcse/.**

***Index Terms*—Glaucoma Detection, Deep Learning, Segmentation, Classification, Deployment.**

I. INTRODUCTION

Glaucoma is one of the most prevalent causes of blindness, with an estimated 80 million individuals affected by 2020 [3, 6]. It's a long-term eye illness that causes vision loss by steadily damaging the optic nerve. Glaucoma is the silent thief of sight since symptoms do not appear until the disease is advanced. Although glaucoma cannot be cured, therapy can help to delay its progression [9]. However, remarkably, studies show that around 50-90% of glaucoma patients remain undiagnosed around the world. In Australia, it is estimated that approximately 150,000 people with glaucoma are undiagnosed and thus missing the best time for treatment to prevent vision loss [7]. The necessity for early glaucoma identification based on compelling images is critical [2].

Some significant barriers to glaucoma diagnosis have been identified. First of all, glaucoma is primarily asymptomatic in the early stages. The effect of it is so gradual that patients may not notice a change in vision until vision loss arises at an advanced stage. Second, the traditional diagnostic methods for glaucoma, such as measurement of IOP and examination of the optic disc on retinal fundus images, are expensive, labor-intensive, and time-consuming. In the meantime, they must be conducted by specialists with a high level of expertise. However, in developing countries and rural areas, there are insufficient eye specialists and inadequate infrastructures. Last but not least, the lack of cost-effective screening techniques for glaucoma is also a significant challenge in the general population.

Concerning the above challenges in glaucoma diagnosis, it is highly needed to develop an automatic glaucoma diagnosis system that can accurately detect glaucoma without the need for excessive equipment and experienced specialists [10, 11]. It should be less time-consuming and more affordable for the general public. To deal with the challenges, deep convolutional neural networks (DCNNs) would be the ideal solution [5]. Given their strong abilities to extract complex and deep features from the images and the excellent classification accuracies they demonstrated in previous similar studies. Moreover, using DCNNs can also eliminate potential human errors in screening. Therefore, this study mainly focuses on developing an automated diagnosis system based on deep learning. In addition, it would be deployed on a graphical user interface (GUI) to enhance the system's usability further.

The major contribution of this study is the deployment of the system. With the help of the graphical user interface, the complex system can be used by non-technical personnel, and the system's usability is much improved. Meanwhile, this study focuses on improving the performance of the deep learning models by utilizing advanced techniques, including contrast limited adaptive histogram equalization (CLAHE) for improving contrast and U-Net for automatic optic disc (OD) segmentation. This study aims to provide a guideline for future researchers on maximizing their predictive models and ensuring the best possible accuracy for a particular dataset. To achieve this, the following objectives are identified:

- We propose GCNet based deep feature extraction scheme for automatic detection of glaucoma using fundus images.
- We explore the performances of the deep feature set with a GCNet classifier and several deep learning methods.
- We deploy a system for detecting glaucoma using the retinal fundus images, which can be useful to determine if the patient was affected by glaucoma or not.

The remainder of the paper is organized as follows: The

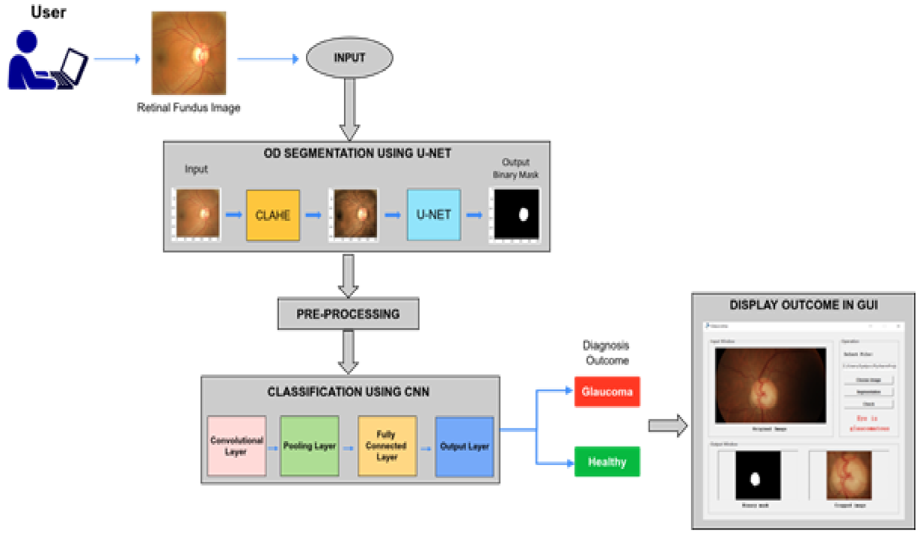

Fig. 1: The architecture of automatic glaucoma detection.

proposed methodology is described in Section 2. This section also describes the data that is used in this study. Section 3 provides the experimental setup and the results with their corresponding discussions in Section 4. Finally, the conclusion is provided in Section 5.

II. Methodology

This study proposes a CNN-based deep feature extraction scheme, 'GCNet', for automatically detecting glaucoma from fundus image data. Figure 1 graphically presents a general architecture of the proposed strategy. The proposed scheme involves four stages such as data collection, segmentation, classification, and deployment. In the data processing stage, the images are resized to 256*256 pixels and grayscale to keep the images consistent and allow the data to be inputted into the segmentation session and classification model. Once the images have been preprocessed, the images are passed into the U-Net segmentation model that has been trained using the ACRIMA dataset. The segmented images are then given into our classification model to classify the sample as either glaucoma or not. Finally, a system is designed to display to result automatically. A detailed description of these steps is provided below.

A. Data Collection

This study collected 2 publicly available retinal fundus image datasets for training and evaluating our models so far. Table 1 below summarizes the information of each dataset:

ACRIMA: The ACRIMA dataset contains 705 labelled retinal fundus images collected from Spain. It is composed of 396 glaucomatous images and 309 normal images. All images in the ACRIMA dataset have been annotated by two glaucoma experts with 8 years of experience. In the version of this dataset that we have obtained, all images were already cropped around the optic disc and renamed.

TABLE I: Summary of Datasets

Dataset	Normal	Glaucoma	Total
ACRIMA	309	396	705
RIM-ONE v3	85	74	159

RIM-ONE v3: RIM-ONE v3 was published in 2015 and contains 85 images of healthy eyes and 74 images of glaucoma patients. Two sets of ground-truths for optic disc and optic cup are provided by experts. This dataset will be used for training the U-Net.

B. Segmentation

Figure 2 shows the UNet architecture of segmentation. The structure of UNet consists of two parts, the encoder section shown in red and the decoder section shown in blue. The encoder section is responsible for understanding the feature in the images and works by having convolution layers paired with max pooling and downsampling of the images to extract the required feature. The next step is to find where in the images these features are located, which is responsible for the decoder section.

The region of interest (ROI) would be the optic disc region in each retinal fundus image. Therefore, a large proportion of the retinal images would become redundant for DCNN supervision. To further improve the performance of classification models, it is necessary to remove the redundant areas on the fundus images as they could potentially become the noise that lowers the performance of the models. The U-Net is the ideal solution to segment the optic disc region automatically. The U-Net is a convolutional neural network that was originally proposed for biomedical image segmentation.

C. Classification

The base model of GC-Net is Xception (Extreme Inception). Xception is the advanced version of Inception architecture. There are 81 layers in-depth in Xception. Two important points

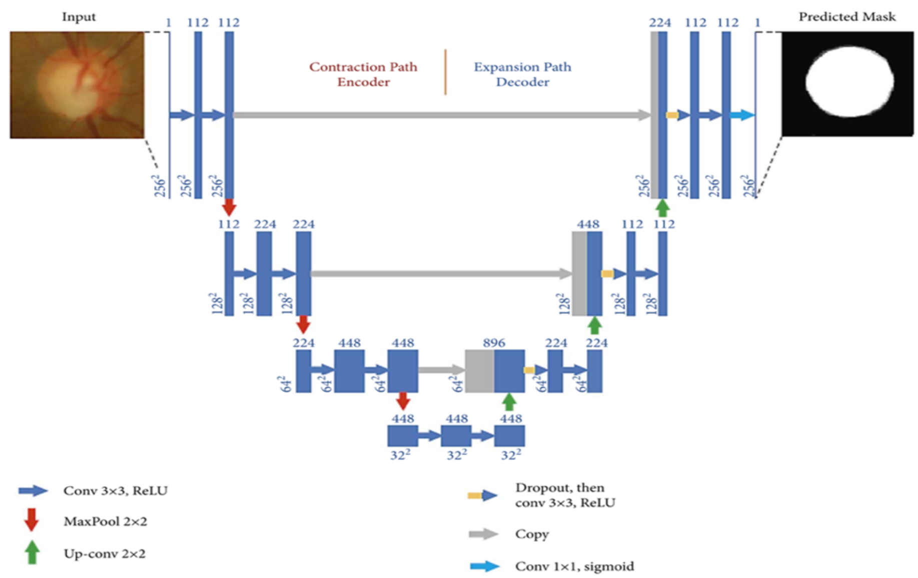

Fig. 2: The U-Net architecture for segmentation

can be found in Xception. The first important point is the pointwise convolution accompanied by a depthwise convolution. More specifically, the 1x1 convolution is put before any nxn spatial convolutions. Therefore, no intermediate ReLU nonlinearity can be found in the model, and the accuracy is higher. The second key point is the residual connections. These connections can skip useless layers. The model with skip connections can have much better accuracy [1, 4, 8].

The size of all input images is (224, 224, 3). The model contains 76 layers. They are twenty-one convolutional layers, ten max-pooling layers, twenty-one batch normalisation layers, eleven activation layers, nine add layers, one dropout layer, and two dense layers.

In the GC-Net, each block includes two convolutional layers, two batch normalization layers, and an activation layer. Besides, max-pooling layers are added in each block to reduce the number of parameters. In addition, each block contains added layers except for block one and the final block, which can minimize data loss. Moreover, each block has a dropout layer to avoid the overfitting problem. At last, the last block contains a global max-pooling layer and a dense layer to give a classification result.

The logistic regression approach is used to classify the glaucomatous and normal images by using the sigmoid activation function. The loss function is binary_crossentropy since the class label was converted into binary numbers, 0 and 1. The optimizer is Adadelta, with a learning rate of 0.001. Moreover, the number of epochs is chosen to be 35 due to the long training time, and the batch size is 16.

D. Deployment

As shown in Figure 4, we design this deployment system as GUI style. GUI has input window, output window and button operation area. The three buttons in the button operation area are the three steps to achieve the entire model effect.

III. Experimental Results

Overfitting and underfitting are two significant issues in machine learning. These problems can be found in comparing the training and test accuracy. In Table II, it is evident that Inception and ResNet have the overfitting problem. The difference between the training accuracy and the test accuracy is quite significant, over 5%. Besides, the underfitting problem can be found in the TB-Net. Both the training and the test accuracy are pretty low.

Different evaluation metrics can reflect the strengths and weaknesses of all models. GC-Net has the highest accuracy, and therefore, the overall performance of GC-Net on the ACRIMA dataset is quite good. The test accuracy of ResNet is the worst among all models. Thus, we can conclude that ResNet has terrible performance on the ACRIMA dataset. However, the ACRIMA dataset is unbalanced, and test accuracy might not reflect the model performance correctly. More glaucoma cases can be found in the dataset. F1 score can consider the distribution of the test data and therefore is introduced here. GC-Net still has the highest F1 score, and ResNet has the lowest F1 score. Sensitivity and specificity are two significant measures in medical cases. Sensitivity pays attention to positive cases, and specificity focuses on negative ones. These measures can help therapists determine which special cases to use.

GC-Net has the highest specificity of 0.95. This model can classify positive cases successfully. However, ResNet's specificity is the lowest among all models. This model has considerable difficulties in predicting positive cases. Inception V3 has the highest sensitivity, near 1. The nearly perfect performance in normal cases can be found in this model. But TB-Net has the lowest sensitivity. TB-Net is weak in predicting normal cases. The ROC curve is the balance between specificity and sensitivity. GC-Net is the closest to the

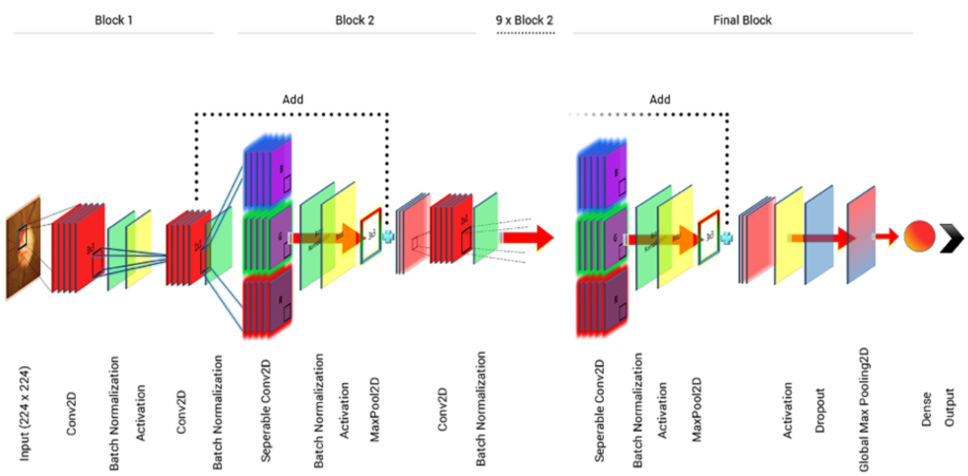

Fig. 3: Pretrained GCNet model for glaucoma detection.

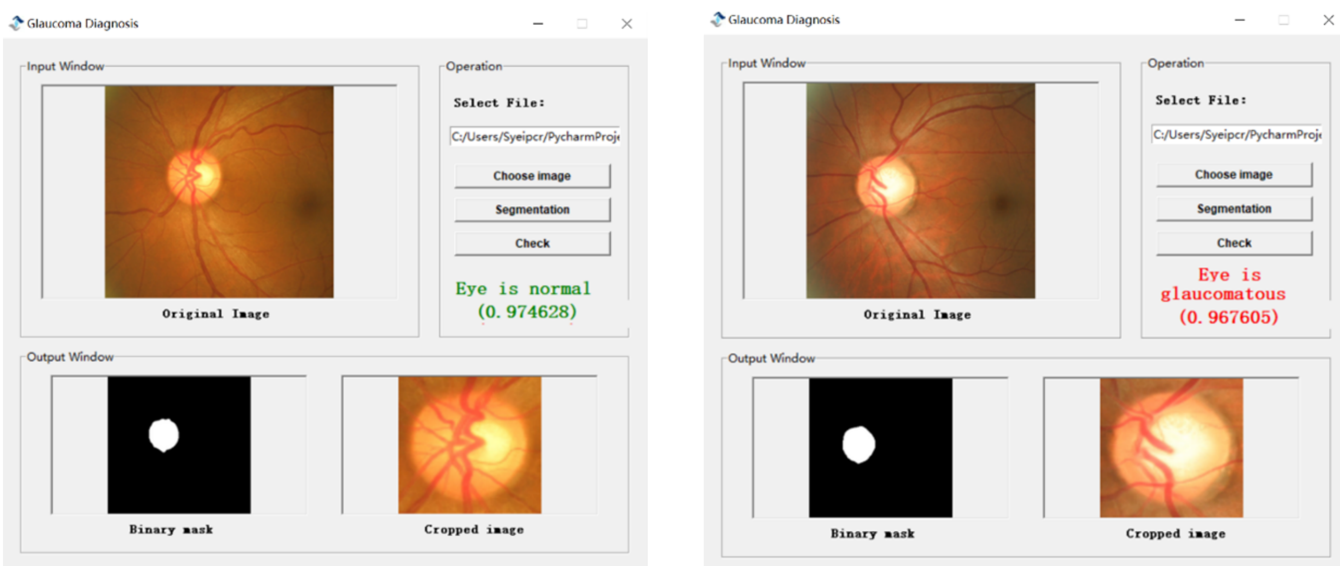

Fig. 4: The deployement system of glaucoma detection.

TABLE II: The performances of the proposed methods

Model	ResNet	CNN	Inception V3	TB-Net	GC-Net
Training Accuracy	0.84	0.94	0.93	0.81	0.96
Test Accuracy	0.78	0.91	0.85	0.80	0.93
F1 Score	0.75	0.90	0.84	0.78	0.93
Specificity	0.68	0.90	0.80	0.82	0.95
Sensitivity	0.95	0.95	0.98	0.78	0.92
ROC AUC Score	0.82	0.92	0.89	0.72	0.93

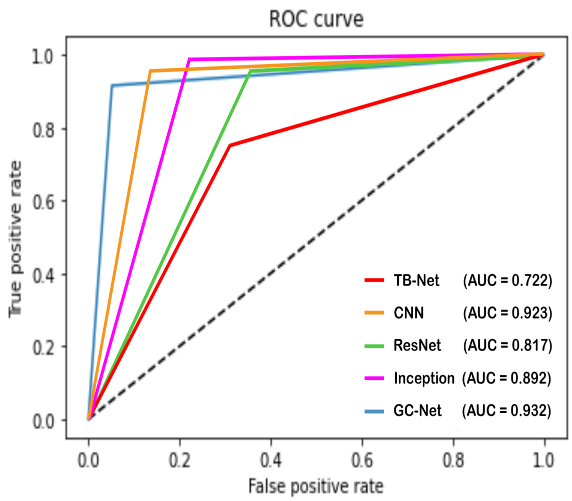

Fig. 5: The ROC curve for different classification models

top-left corner among all models. GC-Net also has the best performance for the ACRIMA dataset. This model has the highest accuracy, and F1 score. We can conclude that the model has good overall performance. Besides, this project focuses more on the glaucoma cases and GC-Net has the highest specificity. GC-Net works better in this respect and matches the project's objective. Even if the sensitivity for GC-Net is not the highest among all models, the sensitivity of GC-Net is near 0.95. It still works well in normal cases.

IV. Discussion

We have included the training and testing accuracy as shown in Table II. The gap between testing accuracy and training accuracy is becoming smaller. These signs indicate that the GC-Net is not overfitting. This is desirable and again proves that GC-Net performs well.

According to research conducted worldwide for people aged between 40 to 80, 1 out of 200 individuals has glaucoma, which increases to 1 out of 8 by the age of 80. According to research, glaucoma is among the world's most known causes of irreversible vision loss after cataracts. It is responsible each year for 12 percent of all cases of blindness. The deep learning algorithm-based U-Net is implemented for optic cup segmentation, while DenseNet-201 is utilized during extraction together with a deep convolution neural network (DCNN). The disease ranks second cause of blindness worldwide, and it must be diagnosed early to reduce loss of vision and general damage while ensuring the appropriate care. The degeneration can lead to two main health issues (first, structural changes that occur in the optical nerve head and nerve fiber layer, and second, concurrent functional failures of the field of vision). For the segmentation, a model called U-NET is implemented. According to the research, the proposed deep learning models for segmenting and classifying glaucoma have been detected from the ACRIMA database using retinal fundus images. During segmentation, the U-NET is utilized while DenseNet-201 is utilized to bring out the image. Lastly, the DCNN is used in classifying the images to detect any signs of glaucoma. During segmentation, the algorithm for deep learning U-Net architecture is used for optic cup segmentation, where afterward the retinal fundus image is handed out as input. The optic cup will be based on the output of the segmentation. Before the image is segmented, the algorithm has to process it. Afterward, the region of interest (ROI) is taken from the retinal fundus image through the use of ground truth from OD to access the closest area of OC. The analysis of the performance of the proposed DCNN in connection with U-Net and model DenseNet-201 is assessed in this section during dataset.

To further assess the effectiveness of the GCNet based model, the ROC curves are drawn for different glaucoma detection models, where the input data was the deep feature set shown in as shown in Figure 5. The corresponding performance measurements in every condition are shown in Table II. We can see that GC-Net has the highest ROC AUC score of 0.932, which is very close to 1, and its curve is most relative to the top left corner. This indicates the GC-Net has the best ability to distinguish between the two classes. Meanwhile, it also has a sensitivity of 0.91, the highest specificity of 0.95, the highest F1 score of 0.93, and the highest accuracy of 0.93, which is good.

Despite achieving accurate classification results during training, testing and inference, these results could be improved. This could be done by creating a custom dataset. The dataset's imbalance has impacted the model's performance and is an area that could benefit from a balanced dataset. Additionally, the proposed custom dataset should also contain retinal images that are proportionally illuminated and have low contrast. Lastly, the addition of further factors in predicting whether a retinal image has glaucoma would be an area of interest. These factors could include age and gender of the individual, medical conditions such as diabetes, and ethnic backgrounds such as race and family history.

V. Conclusion

A glaucoma detection model based on deep learning was suggested in this study. To ascertain if the patient has glaucoma or not, the main goal of this model was to identify glaucoma using retinal fundus images. The ACRIMA dataset was utilized in this suggested deep learning model to assess glaucoma pictures. 30% of the data was utilized for testing, while 70% of the data was used for training. This model utilized the U-Net segmentation model for segmentation and combined DCNN with a pretrained GCNet for feature extraction. The images were classified using the GCNet method to look for signs of glaucoma.

The imbalance data problem will be resolved in the future by enhancing the classifier and lowering the threshold. This model can be used in several medical image segmentation and classification procedures, including identifying breast cancer, brain tumors, diabetic retinopathy, and other diseases.

References

[1] Abbas, Q., 2017. Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. Int J Adv Comput Sci Appl 8, 41–5.

[2] Ahn, J.M., Kim, S., Ahn, K.S., Cho, S.H., Lee, K.B., Kim, U.S., 2018. A deep learning model for the detection of both advanced and early glaucoma using fundus photography. PloS one 13, e0207982.

[3] Almazroa, A., Alodhayb, S., Osman, E., Ramadan, E., Hummadi, M., Dlaim, M., Alkatee, M., Raahemifar, K., Lakshminarayanan, V., 2018. Retinal fundus images for glaucoma analysis: the riga dataset, in: Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications, International Society for Optics and Photonics. p. 105790B.

[4] García, G., Colomer, A., Naranjo, V., 2021. Glaucoma detection from raw sd-oct volumes: A novel approach focused on spatial dependencies. Computer Methods and Programs in Biomedicine 200, 105855.

[5] Gupta, K., Goldbaum, M., Yousefi, S., 2021. Glaucoma precognition based on confocal scanning laser ophthalmoscopy images of the optic disc using convolutional neural network, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2259–2267.

[6] Li, J., Ling, Y., He, J., Jin, P., Zhu, J., Zou, H., Xu, X., Shao, S., Gan, Y., Su, Y., 2021. A gcn-assisted deep learning method for peripapillary retinal layer segmentation in oct images, in: Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XXV, International Society for Optics and Photonics. p. 1163010.

[7] Liao, W., Zou, B., Zhao, R., Chen, Y., He, Z., Zhou, M., 2019. Clinical interpretable deep learning model for glaucoma diagnosis. IEEE journal of biomedical and health informatics 24, 1405–1412.

[8] byul Seo, S., Cho, H.k., 2020. Deep learning classification of early normal-tension glaucoma and glaucoma suspects using bruch's membrane opening-minimum rim width and rnfl. Scientific Reports 10, 1–11.

[9] Xu, J., Chutatape, O., Sung, E., Zheng, C., Kuan, P.C.T., 2007. Optic disk feature extraction via modified deformable model technique for glaucoma analysis. Pattern recognition 40, 2063–2076.

[10] Yadav, K.S., Rajpurohit, R., Sharma, S., 2019. Glaucoma: Current treatment and impact of advanced drug delivery systems. Life sciences 221, 362–376.

[11] Yu, S., Xiao, D., Frost, S., Kanagasingam, Y., 2019. Robust optic disc and cup segmentation with deep learning for glaucoma detection. Computerized Medical Imaging and Graphics 74, 61–71.