Deep Learning Approach to Enhance Accuracy for Early Detection of Glaucoma

Pranita Niraj Palsapure
Electronics and Communication
Acharya Institute of Technology
(Visvesvaraya Technological University)
Bengaluru, India

Anu H A
Electronics and Communication
Acharya Institute of Technology
(Visvesvaraya Technological University)
Bengaluru,India

Ashmitha G
Electronics and Communication
Acharya Institute of Technology
(Visvesvaraya Technological University)
Bengaluru, India

Abhishek Reddy B H
Electronics and Communication
Acharaya Institute of Technology
(Visvesvaraya Technological University)
Bengalru, India

Mainak Jana
Electronics and Communication
Acharaya Institute of Technology
(Visvesvaraya Technological University)
Bengaluru, India

***Abstract*—Diabetes is a medical disorder when the blood sugar (glucose) level cannot be controlled by the body. This can occur if the body can't properly use the insulin it produces or if the body doesn't produce enough insulin. Diabetes can lead to major health issues and increase your chance of developing a number of eye illnesses if it is not properly managed. The advancement of machine learning algorithms has made early detection of various eye illnesses using an automated method significantly more advantageous than manual detection. The ocular illness that lead to visual loss is Glaucoma which do not have any symptoms. Early detection can help to reduce disease-related vision loss. This study proposes a segmentation using UNet model (which is a U-shaped encoder-decoder network architecture, which consist of four encoder blocks and four decoder blocks that are connected via a bridge) on fundus images followed with data augmentation. The CNN (Convolution Neural Network) model is then trained using pre-processed fundus image. The proposed model was using IEEE dataset named REFUGE (Retinal Fundus Glaucoma Challenge). In an evaluation after 100 epochs, the accuracy is 98%. The proposed model outperforms existing deep learning model for early detection of glaucoma.**

Keywords*— *Glaucoma, Diabetic retinopathy, Convolutional neural network

I. Introduction

Glaucoma is one of the most common eye diseases that can lead to irreversible blindness if left undiagnosed and untreated. Glaucoma is a group of disease that damage the optic nerve and can lead to vision loss. The prevention of blindness depends heavily on the early identification and treatment of these diseases. The standard method for determining the presence of glaucoma requires a thorough eye examination that includes a visual acuity test, a dilated eye exam, and an intraocular pressure measurement. These tests are time-consuming and demand trained workers, making it challenging to execute them in locations with limited resources. With the purpose of detecting glaucoma, We have number of imaging techniques such as fundus photography, optical coherence, OCT, and fluorescein angiography, provide detailed images of the eye that can be analyzed by medical professionals.

Furthermore, the development of machine learning algorithms has enabled the automation of the analysis of these images, allowing for faster and more accurate detection of glaucoma. These algorithms can identify signs of this diseases, such as microaneurysms, hemorrhages and exudates and optic nerve head cupping and thinning of the retinal nerve fiber layer. The development of machine learning algorithms for the detection ofglaucoma is a promising area ofresearch, with the potential to improve early detection and prevent blindness in millions of people worldwide. This paper will review the current state-of-the-art techniques for the detection of glaucoma, with a focus on machine learning-based approaches.

II. Literature Review

Dr V Geetha and et al. [1] proposed CNN model that uses ARGALI method and a resolution of 256*256 fundus images. There are four convolutional layers and two fully connected layers. On the ORIGA and SCES datasets, this model was evaluated and trained, yielding AUC (Area Under ROC curve) values of .822 and .882 respectively. When compared to the state of art approach, which produced values of.809 and.859 for the aforementioned datasets, respectively, the results obtained were better.

Arkaja Saxena and et al. [2] proposed a model which has six layers of max pooling and fully connected layers later detection bySoftMax is done. Here region of interest is taken for the images, this makes the execution faster. By the region of interest the optic nerve is detected and marked and bright fringe is removed from images. These steps are called ARGALI approach.

S G Athira Lakshmi and et al. [3] proposed method used ANN. Hardware implementation is done by raspberry pi. Feature extraction is done using hybrid algorithm. The ISNT rule and Cup disc ratio is taken into consideration for the analysis. Image processing is done in two stages.

R.Geetha Ramani and et al. [4] performed Fundus picture preprocessing, measurement calculations, feature relevance analysis, and the creation of classification rules—which together make up the Knowledge Base—are all conducted. To get rid of the noise in the image, average filtering was used. The resulting image was divided into its green channel (G).

For the image's Green channel (G), histogram equalisation (H) was applied. The Hand G pictures were used to derive measurement information. In each image, 32 measures were taken. They included histogram-based measurements, statistical measurements, and measurements based on the Grey-Level Co-occurrence Matrix (GLCM).

Raveenthini. M and et al. [5] the proposed approach involves pre processing fundus images using median filter. By using the Contrast Limited Adaptive Histogram Equalization (CLAHE) approach, fundus imaged contrast is improved. After that, nonlinear features like FD, HOS, and other entropies are retrieved. Following the application of the relevant image transforms, such as the Radon Transform (RT) and 2D-Variational Mode Decomposition (2D-VMD), on previously processed images, these features are retrieved.

L.-P. Cen et al. [6] using 249,620 fundus photoslabelled with 275,543 labels from diverse sources, we constructed a deep learning platform (DLP) capable of detecting several common referable fundus diseases and ailments (39 classes). Our DLP acquired an area under the receiver operating characteristic curve (AUC) of 0.9984, sensitivity of 0.978, specificity of 0.996, and frequency-weighted average F1 score of 0.923.

Sang Phan and et al. [7] proposed a method which included 465 photos of poor quality to create an accurate approach, coupled with 363 photographs of good quality. They utilised heatmap visualization while training and testing on various already-existing CNN models. Prior to training, the photos were set to 256 x 256 and 512 x 512 resolutions.

Romany F Mansour and et al. [8] proposed for the identification of cup and disc region detection and boundary marking, a multilayer neural network which includes the convolution layer, ReLU layer, a pooling layer, SoftMax later, and pixel classification layer has been proposed. Later to determine the cup-disc ratio a perceptron-based convolution multilayer network for the detection of glaucoma.

Vishnu Raja and et al. [9] proposed a SBEFCM model for classification glaucoma using ensemble learning. here they have detected the cup and disc boundaries using spatially weighted ellipse fitting model followed by blood vessel segmentation to the input fundus images.

A Shantini and et al. [10] proposed a CNN model which included many layers in it. The different layers include the ReLU and dense layers and SoftMax classifiers used in end for classification. The dataset used here is the ACRIMA, they have used total of 705 images for training and testing. In this 396 were infected and the rest was normal. The algorithm used was VGG 16 and they have got an accuracy of 94.4%.

III. PROPOSED METHODOLOGY

Input Images from the dataset have different pixel values, so by preprocessing those images are scaled down to 1000x1000 pixel required to our model. After preprocessing we follow the procedures as shown in below block diagram of our model. According to that we done data augmentation Process. After that loading the training, testing and validation dataset can be done. We use UNet model for image segmentation ang AUC based classification. The detailed block diagram of our model is shown in Fig. 1.

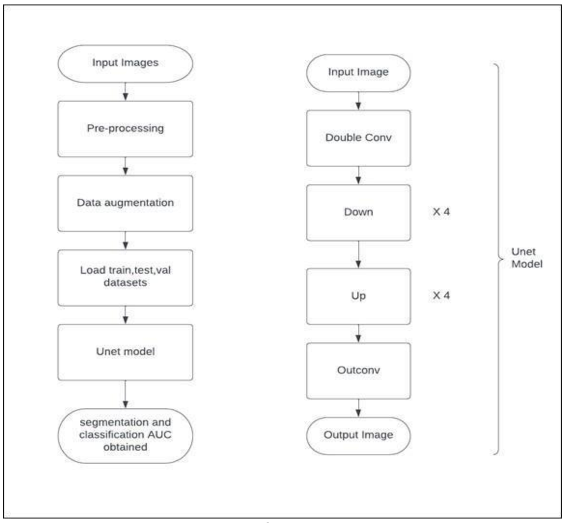

Fig. 1. Block Diagram of Proposed Methodology

A. *Dataset Information and Preprocessing*

Retinal fundus images are a crucial tool for assessing the condition of the retina and are frequently employed in the diagnosis of ocular conditions such glaucoma, diabetic retinopathy, and age-related macular degeneration. In our approach we are using IEEE dataset named REFUGE. It has 1200 annotated retinal fundus images in that we take 400 images for testing, training and validation respectively. Preprocessing the images of dataset includes scalling down the images into pixel size of 1000x1000 to form a standardized dataset and increases the saturation by a factor of 2. The preprocess_img and preprocess_seg functions are used to preprocess images and segmentation maps, respectively. The preprocess_img function takes an image as input and performs two transformations on is it one is transforms.functional.adjust_saturation: This function adjusts the saturation of the image by a given factor. Saturation refers to the intensity of the colors in the image. Afactor of increases the saturation by a factor of 2, making thecolors more vivid. second one is transforms.functional.crop: This function crops the image by a given amount from each edge. In this case, the image is cropped upto 500 pixels fromthe top and buttom, 1000 pixels from the left and right edges.The resulting image has a size of 1000 x 1000 pixels. The preprocess_seg function is similar, but it only performs the cropping transformation on the segmentation map seg. For displaying image we are loading an image using matplotlib.image, displaying it using matplotlib.pyplot, and then applying some preprocessing steps to it using PIL(python image library) and transforms.functional before displaying the preprocessed image. The original image before preprocessing and cropped image after preprocessing is shown below as Fig.2 and Fig.3 respectively.

Fig. 2. Original image Before Preprocessing

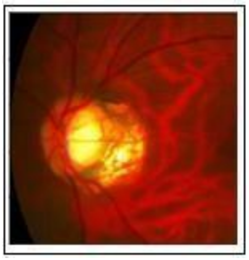

Fig.3. Preprocessed image After Preprocessing

B. Image Augmentation

Image augmentation is the process of artificially expanding the dataset. This is helpful when we have the dataset with very few data samples. Usually there are two image augmentation technique are popular because of 90% accuracyin that we use keras image augmentation technique. The generate_images function is used to augment the training data for an image segmentation task. It does this by generating rotated versions of the images and corresponding masks, and appending these augmented images and masks to the training dataset object. The purpose of augmenting the training data is to help the model generalize better and reduce overfitting, by exposing it to different variations of the same image. This can lead to improved performance on the validation and test sets. By taking predicted segmentation map as a input we get refined segmentation map where only the biggest connected component is retained. The input prediction is expected to be a PyTorch tensor. we need to converts the tensor to a numpy array using the .numpy() method. Then, for each image in the batch, it uses the label function from the scipy.nd image module to identify the connected components in the segmentation map. Thefunction then computes the bincount of the labels and retainsonly the largest connected component. Finally, the largest connected component is converted back to a PyTorch tensor and stacked with the other tensors to form the output. The output is a PyTorch tensor with the same shape as the input, where each pixel value is either 0 or 1, depending on whether it belongs to the largest connected component or not.

C. Procedure of Segmenting Image

In this paper, the picture segmentation process is carried out using the UNet Model. The prominent deep learning architecture known as the U-Net, which is used for image segmentation tasks, is implemented in PyTorch as the UNet class. A structure of encoders and decoders makes up the UNet , where the encoder downsamples the input image using max pooling, and the decoder upsamples the encoded image using bilinear upsampling and concatenates it with the corresponding feature maps from the encoder.
This allows the network to retain spatial resolution and fine details in the output segmentation mask.[2]

The UNet class has several components:

DoubleConv: This is a block oftwo convolutional layers with batch normalization and ReLU activation, used as the basic building block of the U-Net. Down: This block downsamples the input image by applying max pooling, followed by a DoubleConv block.
Up: This block upsamples the input image using bilinear upsampling, concatenates it with the corresponding feature maps from the encoder, and applies a DoubleConv block. OutConv: This is a single convolutional layer used to generate the final output segmentation mask.
In the forward method of the UNet class, the input image is passed through the encoder, which consists of an inc DoubleConv block followed by three Down blocks. The feature maps from the encoder are then passed through the decoder, which consists of three Up blocks followed by an OutConv block. Finally, the output of the last convolutional layer is passed through the output_layer submodule to obtain the segmentation map. The output is passed through a sigmoid activation function to obtain values between 0 and 1. The refine_seg function takes a segmentation map and returns only the biggest connected component of the map. The parameters for training a neural network for an image segmentation task are learning rate-1e-4, batch size-8 used for training, and the number of worker threads-8 for loading the data, and the total number of training epochs we considered is 100.

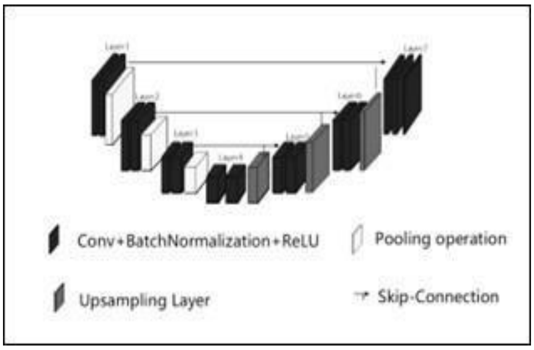

Fig. 4. Architecture of Unet Model

D. Several Evaluation Metrics For Image Analysis Task.

1. Computing the Dice similarity coefficient, a common metric for measuring the overlap between two binary segmentation masks. We consider two tensors input and target, and returns the mean Dice coefficient over a batch of images.
2. Computing the Dice coefficient for a single pair of input and target binary segmentation masks. We takes two tensors input and target, and returns the Dice coefficient.
3. Computing the vertical diameter of the "fattest" area of a binary segmentation mask. We takes a numpy array as binary_segmentation, and returns the maximum diameter.
4. Computing the vCDR (vertical cup-to-disc ratio) from a pair of binary segmentation masks representing the OD (optic disc) and OC (optic cup). We take two numpy arrays od and oc, and get the vCDR.
5. Computing the prediction error, predicted vCDR, and ground truth vCDR for the vCDR metric. We takes in four numpy arrays predicted OD and OC, obtained OD and OC, representing the predicted and ground truth binary segmentation masks for the optic disc and optic cup, and get a tuple containing the vCDR error, predicted vCDR, and ground truth vCDR.
6. Computing the area under the ROC curve (AUC) classification score for a binary classification task. We consider two numpy arrays classified predicted vCDR and classified ground truth vCDR representing the predicted and ground truth class labels, and get the AUC score.
7. Computing the fovea localization error metric, which measures the mean root squared error between the predicted and ground truth fovea locations.

E. Convolution Neural Network

The CNN model will be trained using the input fundus images from the datasets. We trained a model on a segmentation task and using the predicted segmentation to compute vCDR, which is then used to classify images into glaucoma or non-glaucoma cases. The training and validation sets are used to train and validate the model, respectively. Then we iterates over the training and validation sets and computes the loss and the segmentation and classification metrics for each batch of images, and accumulates them over the entire set. We also trains a logistic regression model on the predicted vCDRs and the corresponding class labels of the training set, and computes the AUC score of the logistic regression model on the predicted vCDRs of the validation set. At the end of each epoch, the model parameters will updated and prints the training and validation loss and metrics, and saves the model if the validation AUC is better than the previous best validation AUC. The We also calculated the segmentation Dice coefficient for the optic disc and cup, as well as the vCDR error, which measures the difference between the predicted and ground-truth vCDRs. The refine_seg function is used to clean up the predicted segmentation maps, and the compute_vCDR_error function computes the vCDR and the vCDR error from the predicted and ground-truth segmentation maps. Convolution layers and max pooling layers will be combined in the early layers[1]. Convolutional layers perform a convolution operation on the input and transmit the outcome to the subsequent layer. To maintain the spatial size of the input and output volumes, we employ the convolution layer, pooling layer, hidden layer, feature pooling layer, and activation functions. The convolution layer uses zero padding. In our paper we use the softmax layer as classifier and ADAM(adaptive moment estimation) as optimizer. The segmentation loss is defined to be BCE (Binary Cross Entropy). The end layers will be fully connected layers with dropout to lessen the weight of neurons in portions of the images that are not essential and image augmentation to increase the retinal dataset by producing various variations of the dataset. This improves the CNN model's accuracy. Additionally, we intend to use the RF-CNN-F technique, which combines CNN features with random forest and is more efficient than CNN alone. The output at completely connected layers isregarded as the diagnosis or classification of disorders.

F. Discussion

In this section we discuss about the performance of the proposed solution using various metrics such as training and validation loss values for every 20 epoch as shown in table 1, and corresponding graph of losses percent verses epoch is shown in Figure 5, training and validation dice score for both optic disc and optic curve segmentation for every20 epoch as shown in table 2 and table 3 respectively, and their graph of dice score percent versus epoch shown in Figure 6 and Figure 7, training and validation values of vCDR error for every 20 epoch asshown in table 4 and its graph of vCDRerror percent versus epoch shown in Figure 8.

Table 1: Training and Testing Losses

Epoch	Losses	
	Training %	Validation%
20	0.0088	0.0131
40	0.0015	0.0100
60	0.0006	0.0105
80	0.0004	0.0115
100	0.0003	0.0166

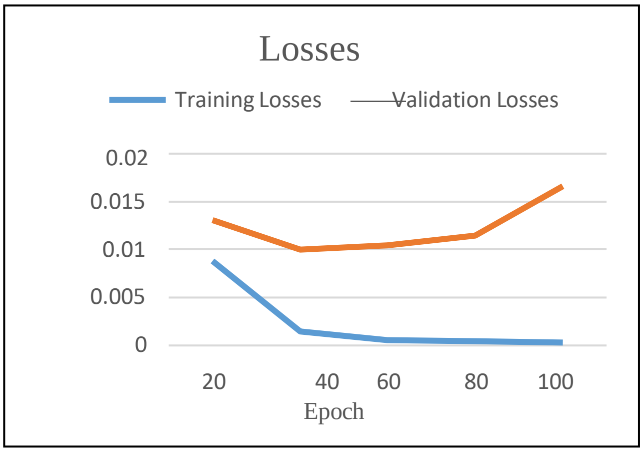

Fig. 5. Losses Analysis

Table 2: Training and Testing Dice Score of OD

Epoch	OD Segmentation Dice Score	
	Training %	Validation%
20	0.9787	0.9065
40	0.9904	0.8699
60	0.9939	0.8924
80	0.9952	0.8684
100	0.9967	0.8280

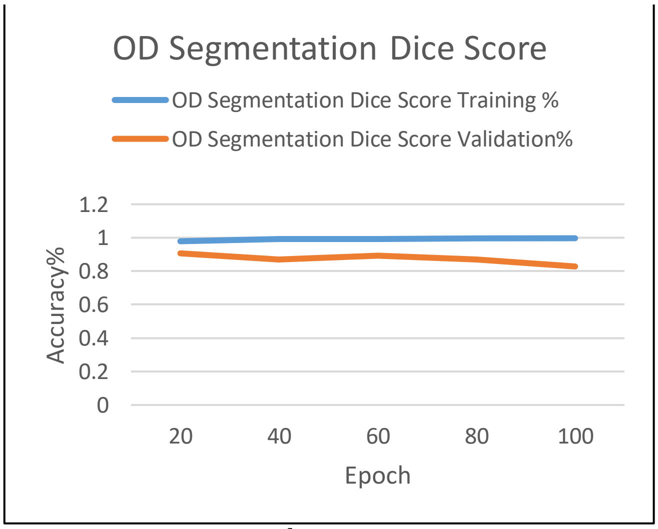

Fig. 6. OD Dice score Analysis

Table 3: Training and Testing Dice Score of OD

Epoch	OC Segmentation Dice Score	
	Training %	Validation%
20	0.9381	0.8270
40	0.9708	0.8176
60	0.9804	0.7850
80	0.9854	0.8024
100	0.9889	0.7764

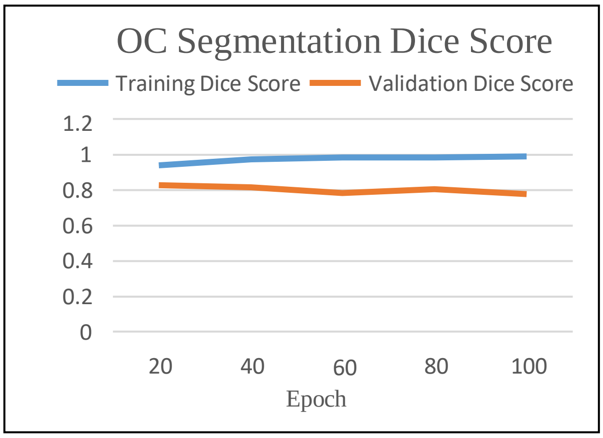

Fig.7. OC Dice score Analysis

Table 4: Training and Testing vCDR

Epoch	vCDR error	
	Training %	Validation%
20	0.0274	0.4651
40	0.0139	0.4402
60	0.0089	0.2395
80	0.0061	0.4462
100	0.0053	1.1671

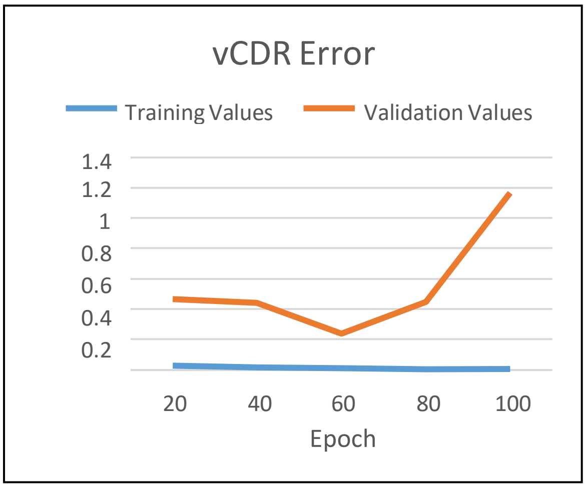

Fig. 8. vCDR Analysis

IV. Result

In the paper [1] they used the ACRIMA dataset with ReLu as activation function, softmax as classifier and ADAM optimizer but they achieve only 86% accuracy. In paper[2] they used SCES and ORIGA as dataset and VGG16 architecture as CNN model, ReLu as activation function and sofmax layer as classifier and achieved only 94.4% accuracy. In paper [3] they used histogram equilization and random transfer technique for preprocessing and artificial neural network as classifier and By taking sensitivity, specificity and accuracy into consideration achieve 90.3% accuracy. The dataset that we utilised in our model was also used in publication [4], but the preprocessing method was more complicated because it involved GLCM (Gray Level CoOccurrence Matrix) calculations, histogram-based measurements, and data mining for classification. In paper [5], they used SVM and RBF kernel functions to reach a maximum accuracy of 85% for the categorization of three .different conditions: normal, DR, and glaucoma. In study [6], they use Kaggle, a freely accessible database with 1000 photos divided into several subgroups of normal and other retinal fundus images, to identify 39 fundus disorders. The proposed model in this paper is improved by considering every metric results for 20 epochs .we got best AUC result for 1, 2, 4, 5, 18, 20 epoch and their training and validation values are shown in table 5 and corresponding graph of accuracy percent versus epoch is shown in Figure 9. The constructed machine learning is tested using test photos, andthe estimated total accuracy is 98%. Further research focuseson applying the suggested framework to diagnose more eye illnesses. Added features are included throughout for improved performance.

Table 5: Training and Testing of Best AUC

Epoch	Best AUC	
	Training %	Validation%
1	0.7865	0.4513
2	0.8924	0.6566
4	0.9484	0.6828
5	0.9526	0.8315
18	0.9507	0.8782
20	0.9466	0.9436

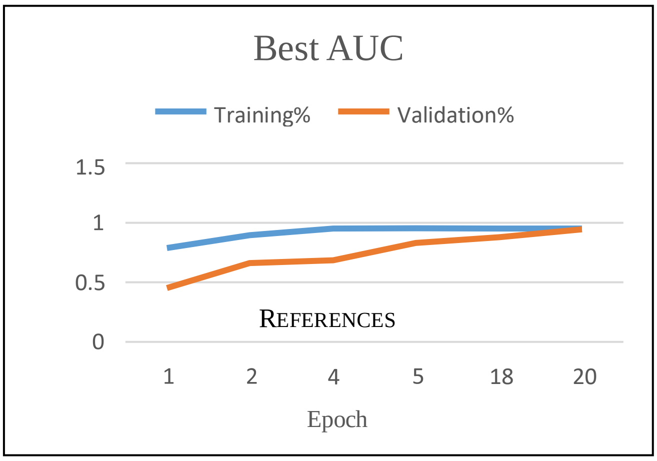

Fig. 9. Best AUC Analysis

V. Conclusion

In this paper We have defined the three datasets for the training, validation, and testing phases. Make sure that the data has been preprocessed, and We have also apply data augmentation techniques to the training set, such as rotation, flipping, and resizing. Also, We split the training set into smaller batches to fit the model into the GPU(Graphics processing unit) memory. We use the PyTorch DataLoader to load and preprocess the data efficiently, ADAM for optimization and ReLu for maintain non-linearity and sigmoid activation function to obtain Dice score between 0 and 1.

In the future, this model may be enhanced to detect additionaleye-related disorders such as Diabetic Retinopathy by employing a larger collection of images for training and testing.

References

[1] S. Yedukrishnan, V. Geetha, and V. J. Jayalakshmi, “Glaucoma detection in fundus image using CNN,” 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA). IEEE, Oct. 08, 2021

[2] A. Saxena, A. Vyas, L. Parashar, and U. Singh, “A Glaucoma Detection using Convolutional Neural Network,” 2020 International Conference

[3] S. G. A. Lekshmi and G. M. Rajathi, “Detection of Glaucoma and Diabetic Retinopathy using Image Processing Technique by Raspberry pi,” Indian Journal of Science and Technology, vol. 12, no. 29. Indian Society for Education and Environment, pp. 1–6, Aug. 01, 2019

[4] R. Geetha Ramani, L. Balasubramanian, and S. G. Jacob, “Automatic prediction of Diabetic Retinopathy and Glaucoma through retinal image analysis and data mining techniques,” 2012 International Conference on Machine Vision and Image Processing (MVIP). IEEE, Dec. 2012.

[5] Raveenthini. M and L. R, “Combined Diagnosis of Diabetic Retinopathy and Glaucoma Using Non-Linear Features,” 2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP). IEEE, May 24, 2021.

[6] L.-P. Cen et al., “Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks,” Nature Communications, vol. 12, no. 1. Springer Science and Business Media LLC, Aug. 10, 2021

[7] S. Phan, S. Satoh, Y. Yoda, K. Kashiwagi, and T. Oshika, “Evaluation of deep convolutional neural networks for glaucoma detection,” Japanese Journal of Ophthalmology, vol. 63, no. 3. Springer Science and Business Media LLC, pp. 276–283, Feb. 24, 2019

[8] R. F. Mansour and A. Al-Marghilnai, “Glaucoma detection using novel perceptron based convolutional multi-layer neural network classification,” Multidimensional Systems and Signal Processing, vol. 32, no. 4. Springer Science and Business Media LLC, pp. 1217–1235, May 18, 2021.

[9] D. Stalin David et al., “Enhanced Detection of Glaucoma on Ensemble Convolutional Neural Network for Clinical Informatics,” Computers, Materials & Continua, vol. 70, no. 2. Computers, Materials and Continua (Tech Science Press), pp. 2563–2579, 2022.

[10] A. Shanthini, G. Manogaran, G. Vadivu, K. Kottilingam, P. Nithyakani, and C. Fancy, “Threshold segmentation based multi-layer analysis for detecting diabetic retinopathy using convolution neural network,” Journal ofAmbient Intelligence and Humanized Computing. Springer Science and Business Media LLC, Mar. 13, 2021.

[11] S. Keel, J. Wu, P. Y. Lee, J. Scheetz, and M. He, “Visualizing Deep Learning Models for the Detection of Referable Diabetic Retinopathy and Glaucoma,” JAMA Ophthalmology, vol. 137, no. 3. American Medical Association (AMA), p. 288, Mar. 01, 2019.

[12] M. R. Mathews and S. M. Anzar, “Residual Networks and Deep Densely Connected Networks for the Classification of retinal OCT Images,” 2022 International Conference on Connected Systems & Intelligence (CSI). IEEE, Aug. 31, 2022.

[13] G. An et al., “Glaucoma Diagnosis with Machine Learning Based on Optical Coherence Tomography and Color Fundus Images,” Journal of Healthcare Engineering, vol. 2019. Hindawi Limited, pp. 1–9, Feb. 18, 2019

[14] R. S. Latha, G. R. Sreekanth, R. C. Suganthe, B. Bizu, K. Suvalakshmi, and K. Venkatachalam, “Diagnosis of Diabetic Retinopathy and Glaucoma from Retinal Images Using Deep Convolution Neural Network,” 2022 International Conference on Computer Communication and Informatics (ICCCI). IEEE, Jan. 25, 2022.

[15] A. Li, J. Cheng, D. W. K. Wong, and J. Liu, “Integrating holistic and local deep features for glaucoma classification,” 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, Aug. 2016.

[16] R. S. Latha, G. R. Sreekanth, R. C. Suganthe, B. Bizu, K. Suvalakshmi, and K. Venkatachalam, “Diagnosis of Diabetic Retinopathyand Glaucoma from Retinal Images Using Deep Convolution Neural Network,” 2022 International Conference on Computer Communication and Informatics (ICCCI). IEEE, Jan. 25, 2022.

[17] F. Alzami et al., "Adaptive Hybrid Feature Selection-Based Classifier Ensemble for Epileptic Seizure Classification," in IEEE Access, vol. 6, pp. 29132-29145, 2018.

[18] L. Qiao, Y. Zhu and H. Zhou, "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for NonProliferative Diabetic Retinopathy Based on Deep Learning Algorithms," in IEEE Access, vol. 8, pp. 104292-104302, 2020.