Journal of Information Systems Engineering and Management
2025, 10(31s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

An Enhanced RNN-LSTM Model for Image Classification using Deep Learning Techniques

Dr. V. Balaji Shanmugam[1], Dr. M. Rajesh Babu[2], Dr. V. Kalpana[3], Dr. S. Uma[4], Dr. S. Mohana Sundaram,[5], Devika Anil[6]

[2]Assistant Professor, Department of CSE, PPG Institute of Technology, Coimbatore, India
Email: balajishanmugam.cse@gmail.com
[2]Professor & Associate Dean, School of Computing, Rathinam Technical Campus, Coimbatore, India
Email: drmrajeshbabu@gmail.com
[3]Associate Professor, Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai, India
Email: kalpanavadivelu@gmail.com
[4]Professor, PG In charge, Department of CSE, Hindusthan College of Engineering and Technology, Coimbatore, India
Email: druma.cse@hicet.ac.in
[5]Assistant Professor [SG], Department of IT, Government College of Engineering, Erode, India
Email: smohanirtt@gmail.com
[6]PG Student, Department of CSE, Hindusthan College of Engineering and Technology, Coimbatore, India
Email: devikash419@gmail.com
Corresponding author email id: balajishanmugam.cse@gmail.com

ARTICLE INFO

Received: 18 Dec 2024

Revised: 10 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

Glaucoma is a leading ocular disease which majorly damages the optic nerve head (ONH) of the retina. The main cause for the glaucoma is intraocular pressure of the eye and due to this it may leads to complete or partial vision loss. The regular screening and the early detection of glaucoma is the only one solution to avoid from vision loss. The Computer Aided Diagnosis (CAD) technique helps in the diagnosis of glaucoma in the early stage using retinal fundus images. The proposed enhanced RNN methodology is applied to classify the images from normal and glaucoma images. The segmented region of the optic disc and optic cup obtained from the enhanced CNN technique is considered for RNN-LSTM classification model. This CNN-RNN-LSTM model helps in increasing the performance of the system to classify the images. The DRISHTI-GS database with ground truth images including data augmentation process helps to train and test the model. This helps to increase the performance of the model and finally achieved with a 96% accuracy.

Keywords: optic nerve head (ONH), Computer Aided Diagnosis (CAD), RNN-LSTM, *Model Predictive*

1. INTRODUCTION

Nandhini [1] proposed an ensemble of deep learning models to achieve better prediction performance when compared to that of the individual models. Glaucoma [2][3] is a leading retinal disease that occurs due to the progressive loss in the Retinal Ganglion Cell. This leads to irreversible vision problems and some time it results in complete blindness if it is not treated in the proper period. The World Health Organization (WHO) declares that glaucoma is the second leading eye disease which causes vision loss. When the pressure inside the eye increases, then it leads to the blockage in the fluid present in the internal part of the eye. This pressure to the eye is called Intraocular Pressure (IOP) [4] and this is the main cause of glaucoma. Due to this IOP, there is a damage to the optic nerve on the retina, which helps to transfer the information from retina to the brain. Because of this, there is a deterioration in the optic nerve fiber which leads to the retinal nerve fiber layer becoming more thicken. This is known as 'cupping' and this leads to the cause for the progression of glaucoma. There is a parameter called Cup-to-Disc-Ratio (CDR) [5] which is used to detect the glaucomatous changes in the retinal part of the eye. The CDR is the distance between the Optic Disc (OD) and Optic Cup (OC) [6] boundary of the Optic Nerve Head (ONH) present on the retina. According to clinical results, if the CDR value is within 0.5 then it a healthy eye and if it exceeds that value then it says that it is the glaucomatous eye.

1.1 Clinical Diagnosis of Glaucoma

An ophthalmologist conducts various tests during the diagnosis of glaucoma. But after the age of 40 years, regular eye checkups can prevent vision loss during glaucoma. There are some tests which were regularly conducted during the diagnosis of glaucoma and few of them are described below,

Tonometry is one of the tests mainly conducted to check the intraocular pressure of the eye. The equipment called a tonometer helps to check the pressure inside the eye. During the test conduction, the eye drop will be applied to numb the eye. After the test, if the person is having more than the 20mm Hg value then it is diagnosed as a further test for glaucoma. The ophthalmoscopy helps to check the optic nerve during the diagnosis of glaucoma. The eye drop is applied to the dilation of the pupil to check the color and shape of the optic nerve. If the pressure inside the eye is within the range but the color and shape of the optic nerve are abnormal, then the patient has to go with a test called perimetry to measure the visual field of the eye. Pachymetry is also one of the tests to check the thickness of the cornea because this problem occurs due to the high pressure on the eye.

The above discussed clinical diagnosis of glaucoma is very time consuming and there may be a variation on the inter or intra measurement values. The digital fundus images are used to measure the progress of many retinal diseases like Glaucoma, Diabetic Retinopathy (DR), Age-related Macular Degeneration (AMD), etc. Those images are captured with the help of the camera called ad fundus camera. Using these images, the main parts of the eye such as optic disc, optic cup, macula, fovea, blood vessels can be easily visualized and also the camera is very less expensive and easy to handle. These fundus images used to measure the parameters verified during the diagnosed glaucoma like disc diameter, cup diameter, cup-to-disc ratio, optic nerve head, etc. Finally, digital fundus images are used as a very effective tool for the diagnosis of glaucoma. Hence Computer-Aided Diagnosis (CAD) [7] using fundus images is very much helpful for the diagnosis of glaucoma by applying many techniques and methodologies.

1.2 CAD for Glaucoma

CAD plays a major role in the detection of glaucoma because it takes less amount of time to diagnose glaucoma for a greater number of patients. For feature extraction and segmentation mainly the techniques like Discrete wavelet transform (DWT), Higher Order Spectra [HOS], Threshold based technique, and machine learning techniques like Fully convolutional neural network, Regions with Convolutional Neural Network (RCNN) etc. are used. To classify the images from normal and glaucomatous, there are many classification methodologies applied in the existing system commonly KNN (K-nearest neighbour), SVM (support vector machine), NB (Naïve Bayes), ANN (Artificial Neural Network), CNN (Convolutional Neural Network) etc. Mainly in the glaucoma detection commonly there are two stages like segmentation and classification. Wei Zhou et al [8] proposed a method for image segmentation by using Local Feature Spectrum Analysis (LFSA) and Dictionary Selection for image segmentation and for classification K-nearest Neighbour for image classification using the MESSIDOR database and gives the performance of 99.83% detection rate. Anum A. Salam et al [9] proposed a method for image segmentation by applying Local Binary Pattern (LBP) and Gabor Wavelet Transform and for classifying the images Support Vector Machine is used with DIARETDB1 database, finally achieved a result of Sensitivity-100%, Specificity-87%, Accuracy-100%. U Raghavendra [10] stated a method Artificial Neural Network (ANN) and Support Vector Machine (SVM) for image segmentation and K-Nearest Neighbour for image classification by using the dataset from Kasturba Medical College and achieved a result of Sensitivity-98.00%, Specificity-98.30%, Accuracy-98.13%. Deepthi K Prasad et al [11] proposed a method for segmentation using Morphological Hough Transform Algorithm and classification by applying Naive Bayes Classifier, K-Nearest Neighbour algorithm using the database of High-Resolution Fundus (HRF) database, achieved with a result of Sensitivity-86.41%, Specificity-78%, Accuracy-78.19%. These existing handcrafted techniques are very difficult and also it is very time-consuming methods. Here, these techniques show the restriction to work on the huge database. To solve this problem the deep learning features are applied to enhance the performance of the system and to achieve the accurate results. Ahmad et al. [12] and Khan et al. [13] used CDR and ISNT rule to classify the images and Ahmad et al. used 80 images from DMED database, Messidor dataset and FAU data library and resulted with 97.6% accuracy but Khan et al. [14] applied the same technique with 50 images from DMED database along with same dataset mentioned above resulted with 94% accuracy. This shows that the reducing the number of images to train and test the images effects the accuracy. Xu et al. [15] proposed a method called reconstruction-based method for both classification of images and also for localization of OD and OC. Here, the codebook is generated randomly for the images which are manually labelled and based on this OD will be located. The AUC from this method is 0.823. This shows that there is no better classification methodology than the segmentation methods in the existing system. Li et al. [16] proposed a method with the combination of holistic features and local features for segmentation and CNN for classification of images. Using this method, the maximum AUC is 0.8384 but the same method with manual classification using ORIGA database gives 0.8390 of AUC. Chen et al. [17] proposed a deep convolutional neural network for image classification using 99 train images and test images in ORIGA database. Here, they applied a greater number of complex models to get high representative data. The AUC from this method is 0.838 and the same author in other publication used CNN with six layers and applied the same method mentioned above, achieves 0.831 AUC. The proposed classification method used enhanced RNN approach which helps to get more accurate result compare to existing system and also helps to detect the glaucoma in the early stage.

2. DEEP LEARNING FRAMEWORK

In the recent days for the classification of images deep learning based technologies are applied. Mainly deep learning models helps to extract the image features automatically. These techniques can also be applicable for feature extraction and segmentation process. In this proposed methodology the enhanced Recurrent Neural Network (RNN) is used to detect the glaucoma automatically. For many reasons the RNN is used in this proposed system as a classifier. One is that RNN is a discriminative model. Second is, the extracted features from the images are learned in the supervised manner. This helps us to build a very good representative system to classify the images based on learned features.

2.1 RNN Architecture

The Recurrent Neural Network is a type of Artificial neural network with the collection of many internal nodes. These nodes in the network are represented with the neurons. The input image is given to the developed neuron network model and receive a output from the model by passing intermediate nodes. There are three layers in the RNN, which are input layer, hidden layer and output layer. The data can be manipulated once the data is in hidden layer. Here, the nodes in the network are loop within the node. Before the data goes to the next layer, it undergoes many iterations within that node, this is why it is called as ‘recurrent’. The short-term memory is the main disadvantage of RNN, due to this the long sequence data cannot be processed in the proper manner. There is a chance of losing some information that stored previously in the node’s memory. In this proposed methodology, the fundus images given as an input to the model and these images are stored in the matrix form. The rows and columns data are processed in the sequential format by using this RNN model. To overcome from the problem of RNN, there is one of the unit cells of RNN called LSTM (Long Short-Term Memory), which gives the solution for this problem. Here, the long sequence data can be stored in a longer time. Hence to work on the large image sequences we are applying both the concept of enhanced RNN and LSTM for the detection of glaucoma in the early stage. The Figure 1 shows the architecture of the RNN.

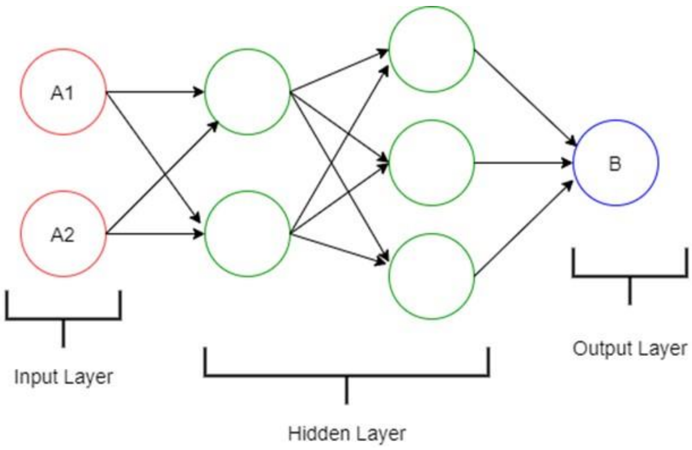


Figure 1: Internal Architecture of RNN

The internal parts of LSTM are,

a. **Input gate:** The input enters into the model through this input gate. Here, there is a default activation function called tanh activation function plays a major role in it. There is a gate which decides, which data will be considered and which is to be rejected to enters into the next gate. But in the proposed methodology, the SoftMax activation function will be used.

b. **Memory cell:** All the incoming data are stored in the memory cell and it's a place where all the action performs. The data decided to process for next step by the gates and keeps trac of all the information flows through the model.
c. **Output gate:** This output gate helps to filter and to regulate the output of the function. The role is to decide what the next hidden state information should carry based on the previous input, current hidden state value and newly applied cell state value.
d. **Forget gate:** This gate used to decide which information should be carried further and which data to be removed. Based on the hidden state value and current input state value is send to the activation function and which gives the resultant as 0 or 1. If the output is 0 the data should not be consider further means the data will be thrown and if the output is 1 the data will be keep it for further processing.

The extracted features and the classified segmented parts of the optic disc and optic cup by applying enhanced deep learning CNN model output is given as an input for the proposed classification model to classify the fundus images to diagnose the glaucoma. The suggested combination of the proposed method and RNN is capable to learn the number of image sequences. This sequential event can model time dependencies which will enhance the accuracy of the prediction level.

3. PROPOSED METHODOLOGY

In the proposed methodology, to detect the glaucoma automatically, the deep learning models are applied to segment and classify the fundus images. The steps which are used diagnose the glaucoma are discussed below,

3.1 Pre-Processing

To enhance the image and to improve the image quality the pre-processing is the initial step for all the image processing concept. As because we are using fundus images to detect the glaucoma disease, the gaussian noise and the salt and pepper noise is commonly identified noise. To remove these types of noises, the gaussian filter is applied. To enhance the image, some of the mathematical morphological operations like erosion, dilation and CLAHE is applied to the image. The complete RGB is considered for image processing rather than separating the image into three channels. To detect the glaucoma, localizing the optic disc part of the optic nerve head plays a major role. Because of more amount of blood vessels on the retinal part, it creates a problem to identify the optic disc of the retina. The dilation morphological operations used in the proposed methodology to solve this problem. It helps to lighten the blood vessels on the optic disc part and add additional pixels to the borders of the internal parts of the image. Hence, the dilation will be applied twice to get make more enhancement in the image. The erosion helps to contrast the input RGB image and also smoothen the image.

Especially, the medical image processing is very challenging task because of its low contrast and poor-quality images. Hence, the enhancement of an image helps the machine to analyse the image properly and also to get more accurate results. The Histogram Equalization (HE) is used to change the contrast of an image by varying the intensity levels of an image. The Eq.1 used to vary the intensity of an image is show below,

$$I_k = T(r_k) = \sum_{j=1}^{k} p_r(r_j) = \sum_{j=1}^{k} \frac{n_j}{n} \quad \text{-------------} \quad (1)$$

Here, I_k is the intensity value that can be changed, which is corresponding to the value r_k is the image which we want to enhance, $p_r(r_j)$=1,2,3,4----L denoted the different intensity level. The noise can also be reduced by applying the HE, but while increasing the intensity levels, there may be chance that it may fall within the range of scale of image. Due to this problem, in the proposed method the Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied. Here the complete image global histogram is calculated. The Whole input image is divided into multiple parts or regions, in which each sub-region is called as tiles. For every sub-region, the separate histogram is calculated. While combining these sub-regions the bilinear interpolation is used to remove the extra added boundaries in the image and this helps to increase still contrast in the image.

- **Shape Detection**

Here the shape of multiple images on the retinal part of the eye is detected by applying Edge Histogram Descriptor (EHD) and by this it helps to detect the shape of optic nerve head part where the optic disc and optic cup is located. Here, the complete retinal fundus image is given as an input. Then the whole image is divided into 4*4 sub-images, in which each sub-image is considered as a region. For every region the edge histogram is generated with 5 bins of horizontal, vertical, 45-degree diagonal, 135-degree diagonal, and non-directional edge types. Other than these directional edges, the image is further divided into small region to get the non-overlapping image blocks called as monotonous block. If the region is having these type of blocks, then the edge type is incremented by one otherwise it will not increase the histogram bins. Finally, after calculation of bins in every block, the total sum of the normalized block is calculated and then the result is quantized. Since, there are 5 bins and 16 sub-regions (4*4) ,16*5=80 bins were generated.

3.2 Cropping

After localizing the optic nerve head part of the retina, only the ONH part will be cropped to reduce the computational time. The complete image is not given to the model to extract the features, only the cropped image features will be extracted and processed. The image with size 2896 × 1944 is reduced to 512*512 by cropping technique. This helps to get the more accurate result and also helps to reduce the computational time. The cropped input image is shown in the Figure 2.

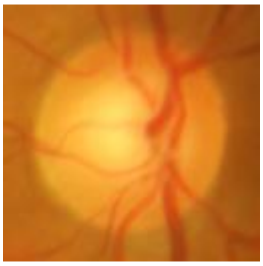


Figure 2: Cropped image with size 512x512

3.3 Feature Extraction

In the proposed method the features of edges and the contour shape of the image is extracted. To extract the edge features of an image Sobel edge detection algorithm is used and to extract the shape feature of an image, watershed algorithm is used. Using these features the boundary of the optic disc and the optic cup regions are located properly. The ROI (Region of Interest) cropped image is given as an input fundus image for the next process.

3.3.1 Sobel Edge detection algorithm

The Sobel edge detection technique is one of the edge detection techniques in which the boundaries of the given image can be easily detected, hence the optic disc and optic cup boundaries can be properly localized. This helps to calculate the cup to disc ratio accurately. Mainly this technique helps to extract the pixels located on the edges of the image. The pixels on the edges of an image contain high frequency information, hence high pass filter is applied to detect the edges of the image. In first order derivative magnitude is used as an operator to extract the edges information. The kernel used is shown in the **Figure 3.**

A1	A2	A3
A4	A5	A6
A7	A8	A9

(a)

-1	-2	-1
0	0	0
1	2	1

(b)

-1	0	1
-2	0	2
-1	0	1

(c)

Figure 3: (a)The image with size 3x3 (b) 0^{o} Sobel Kernel (c) 90^{o} Sobel Kernel

This technique is based on the gradient based detection technique, and the operator is a linear operator. Here, there is a convolution kernel which it extracts the information in pixel by pixel and line by line in two different ways with two different angles. In x-axis, the kernel moves line by line and pixel by pixel in 0^o in y-axis and returns the gradient value G_x, the kernel moves line by line and pixel by pixel in 90^o and returns the gradient value G_y. The pixels on the edges from lighter to brighter with their vector values are applied; hence the vascular edges besides the disc and cup are easily extracted. The Figure 4 gives the result after applying the Sobel Edge Detection Algorithm to the input cropped fundus image. The image high lightens the edges of optic disc and optic cup.

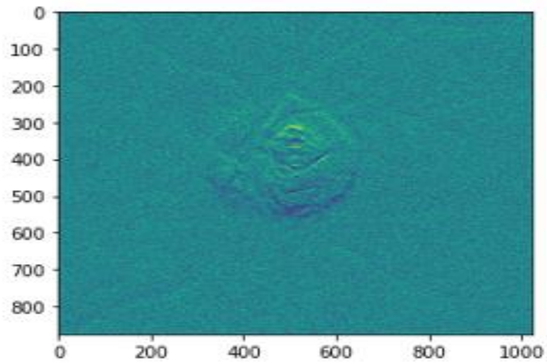


Figure 4: Result obtained from the Sobel Edge detection algorithm

3.3.2 Watershed Algorithm

The watershed algorithm can be applicable for both segmentation and feature extraction process. Here, in the proposed method, to extract the feature of contour shape, the watershed algorithm is used. By applying this, the boundaries and contour shape features of the optic disc and optic cup will be easily extracted. In this technique, the complete cropped image is segmented into a different region based on their contour shapes. Then in every region the pixels are collected and named it as super pixel. Then that whole pixel is considered as surface and in that plane surface it identifies 'catchment basins' and the 'watershed lines' on it. Here, the regions are created by considering light pixels are low and dark pixels are high. The resultant image after applying the Watershed algorithm to the input image is shown in the Figure 5.

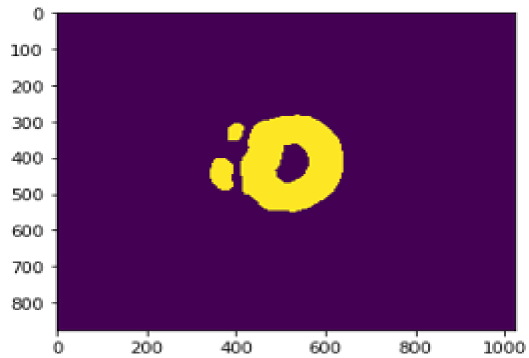


Figure 5: Contour shape detection of the input fundus image by watershed algorithm

3.4 Segmentation using Enhanced CNN

In the proposed methodology, the deep learning CNN model is applied to automatically diagnose the glaucoma with high performance. The CNN consists of many layers such as input layer, max pooling layer, convolution layer, rectified linear input, average pooling layer etc. Based on the input size and the application, the layers used in the model will be varied. There is no condition that we have to use all the layers, the layers may be increased or decreased. In the proposed model, two CNN models were developed for the segmentation of optic disc and optic cup separately. The cropped image with size 128*128 is given as an input for the segmentation model. The framework is developed with 39 layers of CNN model for optic cup segmentation and the same 39 layers for optic cup segmentation is developed. By applying the deep learning process with increased number of layers helps to train more number of parameters in each layer and this helps in getting more accuracy.

The 39 layers in the model is divided into convolutional (19 layers), drop out (11 layers), Merge layer (1), down sampling (4 layers), and op-sampling (4 layers).The filter size used here is (1X1) and (3X3) and depending upon the layers, the number of filters are used is 1,8,16,32,64,128.Firstly,to the input image, the convolution operation is applied by the convolution layer. The window will be generated by the user and which is moved towards the input image to generate a feature map for the next layer. This feature map will be considered as a input for the next layer. The drop out layer randomly select some data during the training and drops it from the network to avoid overfitting problem in the model. The drop out data is called as ignoring unit or a neuron and this drop out layer is applied after the convolution layer. The dropped-out data is not further considered in the training phase. The input image size will be reduced by max-pooling layer. Which collects the feature maps from the previous layer and select the maximum value from every cluster to generate a new feature map with reduced size. This layer helps to consume low memory and fasten the system during training data. The up-sampling is used to recollect the lost information from down-sampling method. In the proposed application, the most efficient information about the optic disc and optic cup will be regained when the up-sampling is comes with the output layer.

The ReLu is an activation function which is mainly applied in the convolutional layer. The thresholding is applied for the input of the convolutional layer i,e., the value which is less than zero will be zero. This helps to select only most required feature for the next level and also helps to avoid redundancy of data but it will not change the size of the input what it received from the previous layer. The final output layer is resulted with the binary mask of the optic disc represented with 0's and 1's. The output image predicted pixel is compared with the actual image corresponding pixel. If there is a match in the pixel value then it represents with black color and if it does not match with the pixel value then it represents with the white pixel. Finally, the predicted optic disc mask from the first optic disc model is given as an input to the next optic cup model to predict the optic cup mask. The same layers what we used for the optic disc CNN model is applied for the optic cup model and predicted the optic cup mask from the input image. From the output of predicted optic disc and optic cup, the cup-to-disc ratio is calculated. This helps to analyze whether the image is glaucoma image or normal image. The sample result for few images of the database obtained from the segmentation model is shown in the Figure 6 and Figure 7.This result is taken to the proposed classification model to classify the images. This result gives the predicted mask of the optic disc and the optic mask. From this we calculate the CDR value.

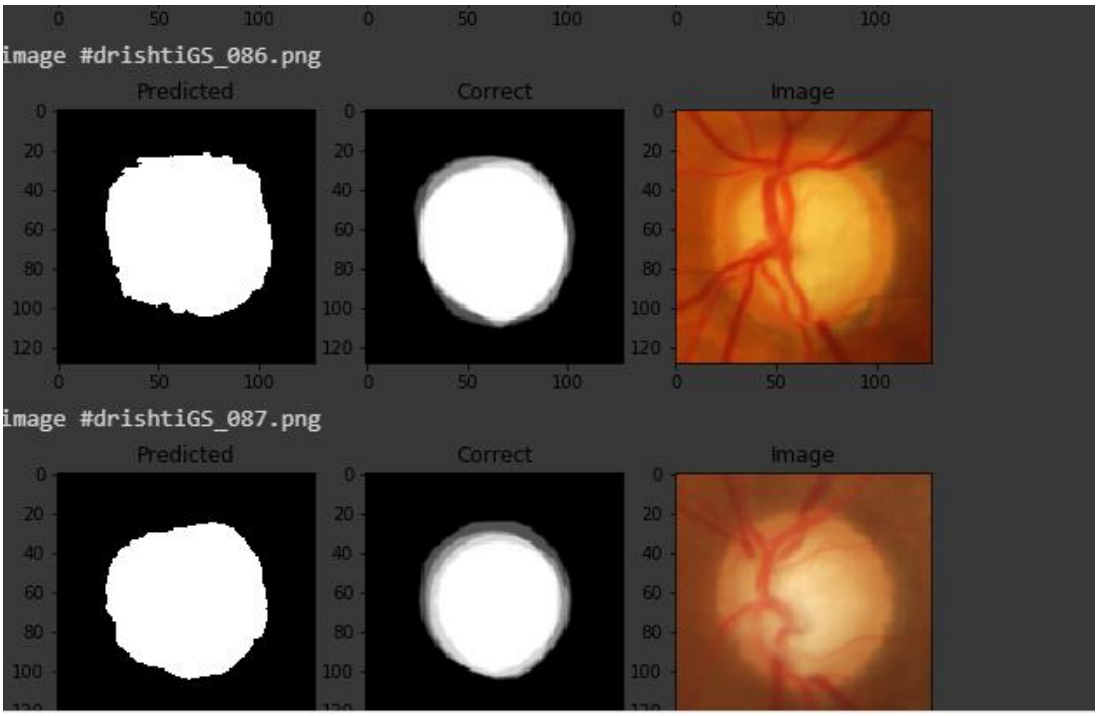


Figure 6: The optic disc mask obtained from the CNN optic disc segmentation model

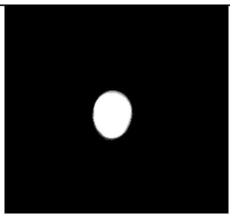
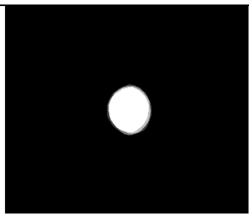
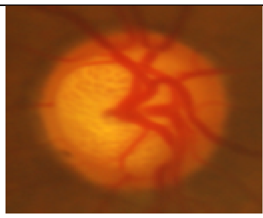
Image	Predicted OC mask	Actual OC mask	Actual Image
Drishti-GS_001			

Figure 7: The optic cup mask obtained from the CNN optic cup segmentation model

3.5 Proposed Classification model

The modified version of the Recurrent Neural Network is used to classify the fundus images by taking the output from the image segmentation process using enhanced CNN methodology. The extracted features and segmented results will be imported into a recurrent neural network for classifying the glaucoma disease. Here, the output depends on the present inputs and the previous step's neuron state. Also, it is capable of learning the number of image sequences. Also, the recurrent networks along with recurrent connections amongst hidden units, which reads an entire sequence and then produces a single output. These sequential events can model time dependencies which will enhance the accuracy of the prediction level. Once done with the feature extraction, selection and segmentation process, three directional recurrent unit's layers is constructed on the outputs of RNN for classification of retinal images. Here the diagnosis results separated into two such as normal and glaucoma. The resulted RNN model will use the learning capabilities toward segmented glaucoma-related tissues like the optic cup, optic disc region and then calculates the optic curve area. The suggested combination of the proposed method and RNN is capable to learn the number of image sequences. This sequential event can model time dependencies which will enhance the accuracy of the prediction level.

The output of Segmented masked region is given to RNN and LSTM, which classifies the images as glaucoma and normal. The modified RNN classifier is used by considering different layer (3 – dense, 3- dropout and one batch normalization) and same combination of layers are considered to LSTM network. Based on the segmented results we can get the details of classification results i.e. glaucoma or not. Hence it is not required to consider CDR ratio. Especially, the cup-to-disc ratio compares the diameter of the "cup" portion of the optic disc with the total diameter of the optic disc. This portion has been covered in deep learning process itself. So, we have neglected it. The layers used in the model are discussed below,

i. **Simple RNN layer:** In this layer, the input image values are arranged in the form of sequences. Here the data is taken in the term of step values. Hence, the input sequence values are defined as various step values. This layer also helps to reshape the data. The input layer with size 128X128 is given as an input to the Simple RNN layer. Here the input sequences are divided into 100 units and extracted 13,300 parameters.

ii. **Dense Layer:** The dense layer helps to represent the data int the matrix format. The values are in the matrix are the trainable parameters and every time it gets updated when it does the backpropagation. It also helps to change the vector dimensions during the training process. Mathematically it applies scaling, rotation and translation functions to the input image vector.

iii. **Dropout Layer:** The dropout layer helps to drop some of the parameters during training period which are not useful by setting the input vector parameter to zero. Once the parameters are set to zero, it will not get updated during backpropagation.

iv. **Batch Normalization Layer:** The Batch Normalization mainly helps in the data distribution of the model and helps to fix the learning rate which is suitable to the model. Here, the data is grouped into batches and it gives a chance to all batches to participate in the training process. This process mainly helps to increase the accuracy of the model.

v. **SoftMax Activation Function:** SoftMax is an activation function used at the last layer in the model. It helps to map the output within the range [0,1]. It converts the vector values into probability distributed value, in such a way that the total sum will be equal to 1. This helps to decide whether the given input image is glaucomatous image or not

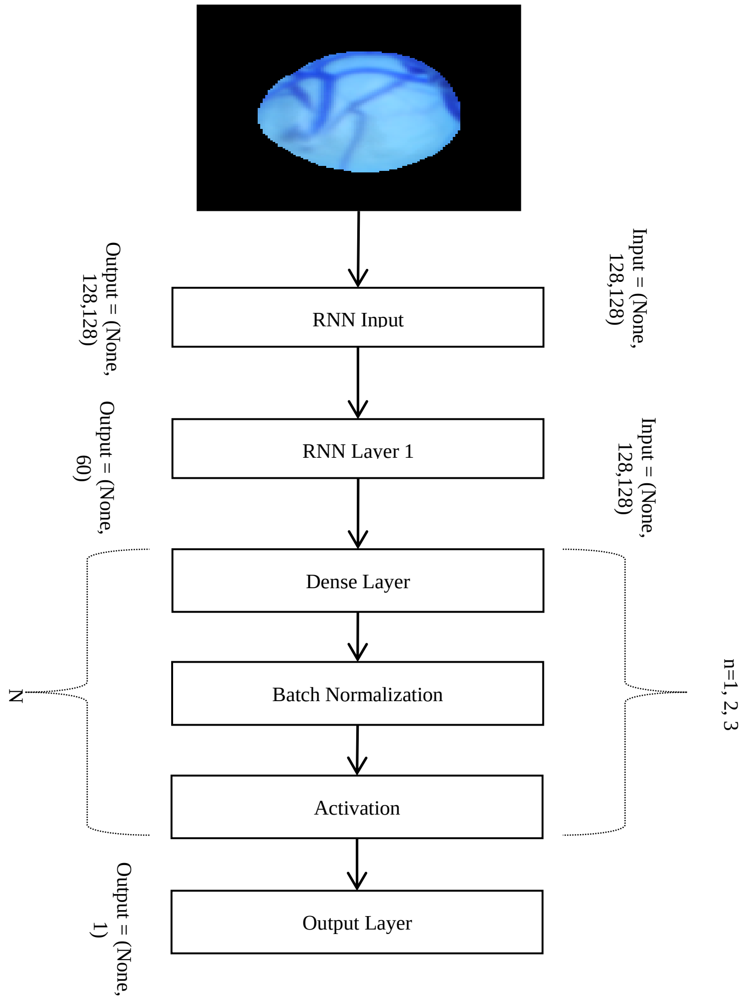


Figure 8: Block Diagram of the Proposed Classification model.

In the shared Figure 8, the first layer represents the input layers of RNN which 128 * 128-pixel neurons (Input as image), second layers is the simple RNN hidden layers which is also consists of 128 * 128 neurons, third layers is simple dense layer which 60 neurons after that we apply batch normalization technique which avoid over fitting. After that every dense layer will have activation function and fourth layer is final output layer which has one neuron based on the that the classification is made whether it is glaucoma or not (threshold will be varied based on the activation function). The layer details with the input batches made in every layer and the number of neurons used in each layer is explained in the Figure 9.

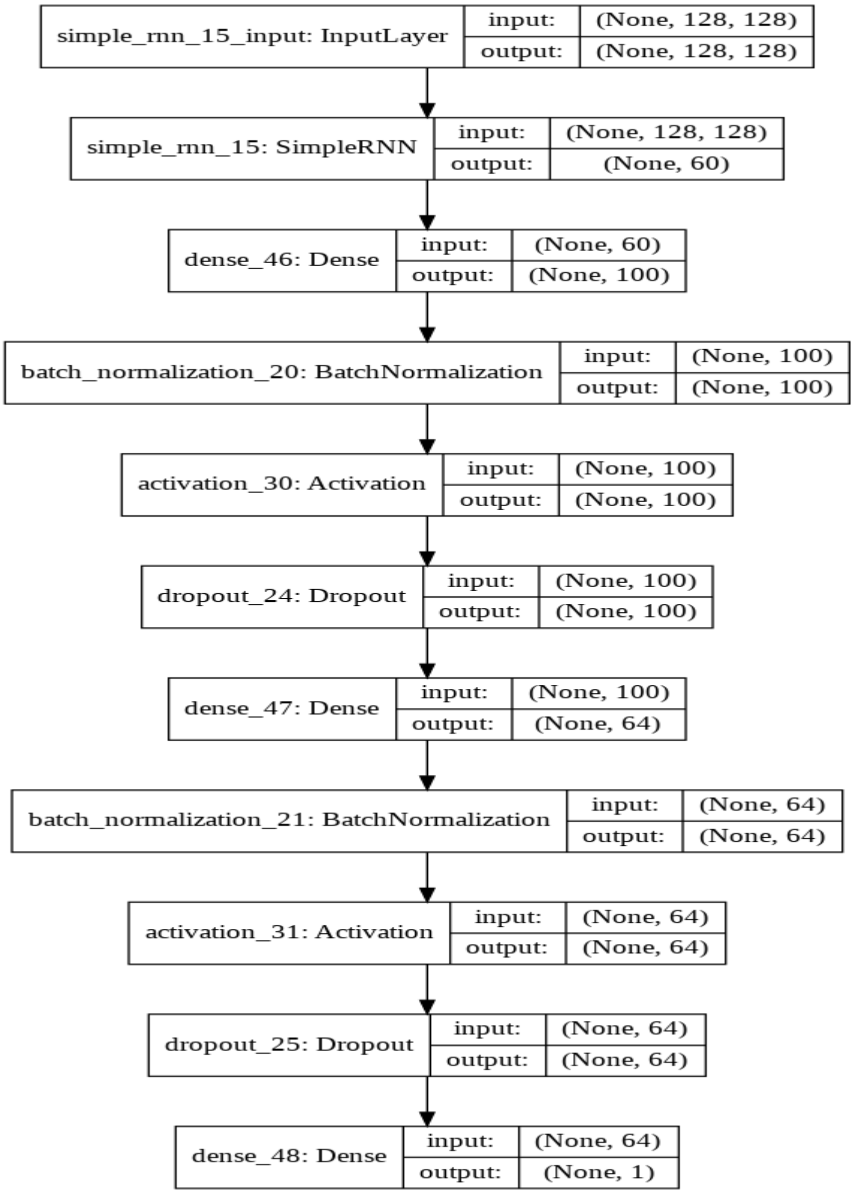


Figure 9: Proposed Architecture of the Classification RNN model

Here, the input size (None,128,128) is given as an input to the input layer of the RNN, then passing parameters to next layers finally the output (None,1) is achieved and this helps to decide whether the image is glaucoma or healthy image. The output shape and the parameters extracted in every layer are listed in the Table 2 and totally 17,205 parameters are extracted.

Table 2: Summery of the model

Layer (Type)	Output Shape	Param #
simple_rnn_4 (Simple RNN)	(None,100)	13300
dense_10(Dense)	(None,32)	3232
batch_normalization_8	(None,32)	128
dropout_10(Dropout)	(None,32)	0

dense_11(Dense)	(None,16)	528
dropout_11(Dropout)	(None,16)	0
dense_12(Dense)	(None,1)	17
dropout_12(Dropout)	(None,1)	0
activation_4(Activation)	(None,1)	0

4. RESULTS AND DISCUSSIONS

The NVIDIA GeForce graphics is used to train the model and is implemented in python using tensor flow and keras libraries. The DRISHTI-GS database with data augmentation with various angles is used to train and test the model. Every input batch made in each layer contains few glaucoma images and which helps to avoid biasing towards the healthy image classes. The proposed CNN-RNN model is validated by working with various learning rate like 0.1,0.01,0.001, 00001.If the learning rate is too small, the system become very slow and if the learning rate is high, the system result come static. Hence, the more accurate result is obtained with the learning rate 0.001.

The classification performance of the model is verified based on some of the performance measures which are discussed below,

Recall or Sensitivity: The correct prediction out of total amount of samples used in the model called as recall.

$$\text{Recall}=TP / (TP+FN)$$

Precision: The correct predictions made out of total amount of correctly predicted class called precision.

$$\text{Precision}=TP / (TP+FP)$$

Specificity: The actual negative prediction out of total amount of samples used in the model.

$$\text{Specificity}=TN/(TN+FP)$$

F1-score: It is used to calculate the performance of the model by doing the harmonic mean with the precision and recall value is called as F1-score.

$$\text{F1-score}=2.\frac{precision*recall}{precision+recall}$$

Accuracy: The test correctness is measured with the help of accuracy.

$$\text{Accuracy} = (TN+TP) / (TN+TP+FN+FP)$$

Where, TN = True Negative FN = False Negative TP = True Positive FP = False Positive

The DRISHTI-GS database consists of 100 images including both healthy and glaucoma images and by doing data augmentation process with multiple angles, the number of samples increased with 100 more images. The proposed model trained and tested with total 200 images including both healthy and normal images. Out of 192 healthy images, 190 images are correctly identified as healthy images and it is called as True Positive(TP),2 are misclassified as glaucoma images and it is called as False Positive(FP), Out of 8 images 5 images are properly classified as glaucoma images and it is called as False Negative (FN),3 of them are wrongly classified as healthy images and it is called as True Negative(TN). These values are represented with the help of the matrix called as Confusion Matrix and which is shown in the Figure. The classification report containing all the performance measures is shown in the Table 3

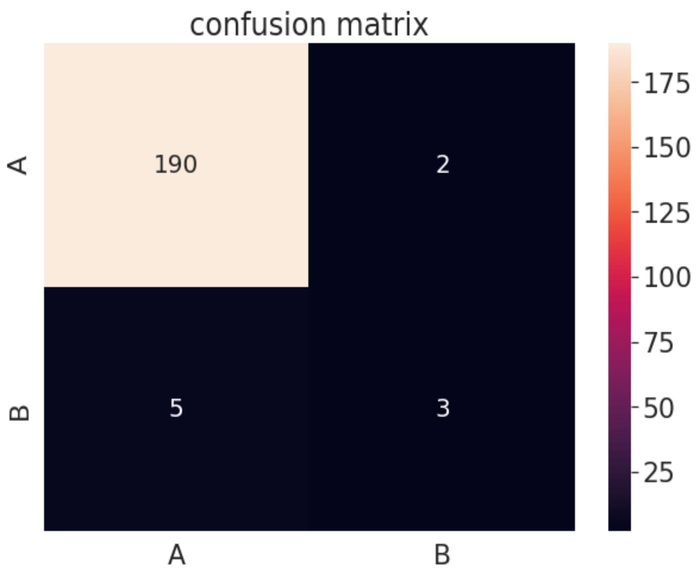


Figure 10: Report of Confusion matrix

Table 3: Classification Report

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	4
1	0.98	0.97	0.974	10
Accuracy	0.96			

The proposed CNN-RNN image classification model used in the diagnosis of glaucoma retinal disease and achieved a good result with an accuracy of 96% and it is compared with the existing classification approaches used in the detection of glaucoma. The comparison result is shown in the Table 4 and which is graphically presented in the Figure 11.

Table 4: Comparison result of proposed method with the existing approaches.

Author	**Approach**	**Accuracy**
Proposed approach	CNN-RNN	0.96
Juneja et al (2019) [1]-	G-net CNN	0.93
Artem [2] -	U-Net CNN	0.85
Zilly et al. (a) [3]	CNN with softmax logistic regression classifier - DRISHTI-GS data set	0.87
Zilly et al. (b) [4]	CNN with softmax logistic regression classifier	0.83

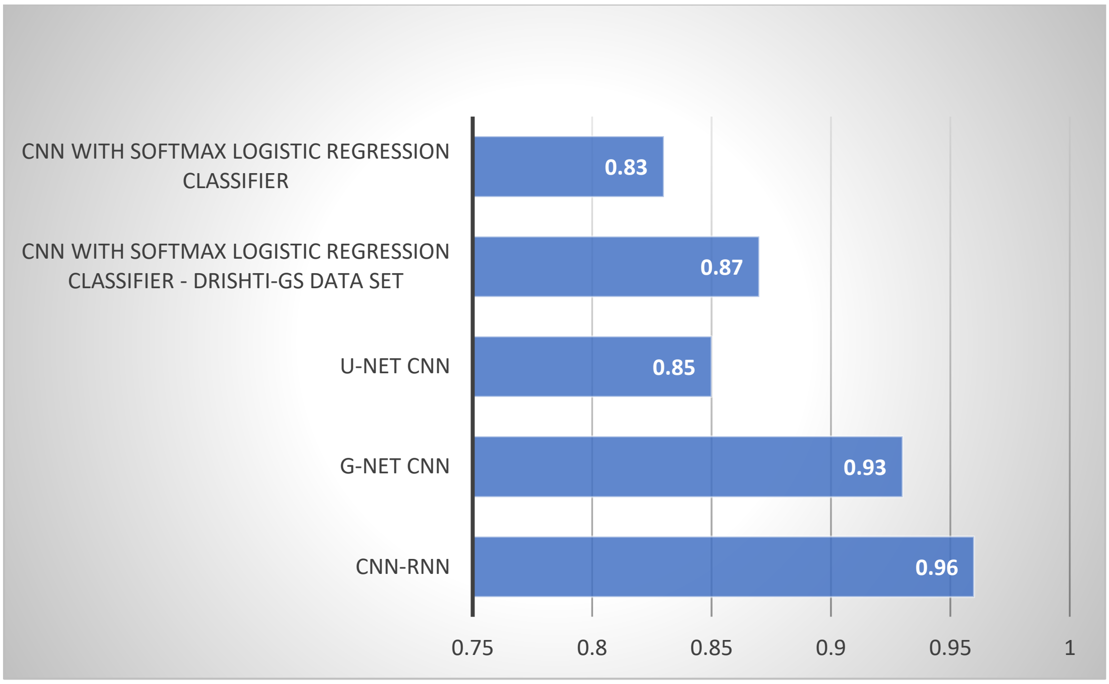


Figure 11: Comparison graph obtained by comparing the existing approach and proposed approach with the Accuracy metrics

5. CONCLUSION

Glaucoma is one of the severe retinal diseases which can leads to blindness. The large amount of people all around the world affected by this disease. In recent technology, the CAD tools help the ophthalmologist in the early diagnosis of glaucoma. Our proposed method gives a better accuracy, sensitivity, specificity of 96%,97%,98% respectively, by using a smaller number of images to train the model. The novel RNN-LSTM model can efficiently diagnose the glaucoma and normal class images. This deep learning technique helps to automatically extract the greater number of features in every layer and which increases the performance of the model in classifying the images. The combination of enhanced CNN and RNN model in segmentation and classification of images helps the ophthalmologist to diagnose the disease in early stage.

REFERENCES

[1] Nandhini, T., Rajesh Babu, M., Rajalakshmi, S., Rajasekar, M., and Sivakumar, T., “Rectal Fundus Image Recognition Using Deep Learning Ensemble CNN Models”, Journal of Information Systems Engineering and Management, Vol.10, No. 26s, 2025.

[2] Anindita Septiarini, Dr, Dyna M. Khairina, MKom, Awang H. Kridalaksana, MKom, Hamdani Hamdani, MCs (2022), “Automatic Glaucoma Detection Method Applying a Statistical Approach to Fundus Images”, Health Informatics Research.

[3] Julian Zilly Joachim M. Buhmann Dwarikanath Mahapatra (2020),” Glaucoma Detection Using Entropy Sampling And Ensemble Learning For Automatic Optic Cup And Disc Segmentation”, Computerized Medical Imaging and Graphics, volume 55.

[4] Niharika Thakur∗, Mamta Juneja (2018), "Survey on segmentation and classification approaches of the optic cup and optic disc for diagnosis of glaucoma, "Biomedical Signal Processing and Control, ELSEVIER.

[5] Shuang Yua,_, Di Xiaoa, Shaun Frosta, Yogesan Kanagasingama,(2019), "Robust Optic Disc and Cup Segmentation with Deep Learning for Glaucoma Detection," Computerized Medical Imaging and Graphics.

[6] Ahmed Almazroa (2015), "Optic Disc and Optic Cup Segmentation Methodologies for Glaucoma Image Detection: A Survey," Journal of Ophthalmology.

[7] U Raghavendra , Hamido Fujita , Sulatha V Bhandary , Anjan Gudigar , Jen Hong Tan , U Rajendra Acharya (2018),” Deep Convolution Neural Network for Accurate Diagnosis of Glaucoma Using Digital Fundus Images”, *Information Sciences,* DOI: 10.1016/j.ins.2018.01.051

[8] Wei Zhou (2018), Automatic Optic Disc Detection in Color Retinal Images by Local Feature Spectrum Analysis, Hindawi Computational and Mathematical Methods in Medicine Volume.

[9] Anum A. Salam (2016), Automated Detection of glaucoma using structural and non-structural features , Springer Plus.

[10] U Raghavendra (2018), "Deep Convolution Neural Network for Accurate Diagnosis of Glaucoma Using Digital Fundus Images", Information Sciences.

[11] Deepthi K Prasad (2018), "Improved Automatic Detection Of Glaucoma Using Cup-To-Disk Ratio And Hybrid Classifiers", ICTACT Journal On Image And Video Processing, Volume: 09, issue: 02.

[12] Ahmad H, Yamin A, Shakeel A, Gillani SO, Ansari U. Detection of glaucoma using retinal fundus images. In: Robotics and emerging allied Technologies in Engineering (iCREATE), 2014 international conference; 2014. p. 321–4. IEEE.

[13] Khan F, Khan SA, Yasin UU, ul Haq I, Qamar U. Detection of glaucoma using retinal fundus images. In: Biomedical engineering international conference (BMEiCON), 2013 6th; 2013. p. 1–5. IEEE.

[14] Khan MW, Sharif M, Yasmin M, Saba T. CDR based glaucoma detection using fundus images: a review. Int J Appl Pattern Recog. 2017;4(3):261–306.

[15] Xu Y, Lin S, Wong DWK, Liu J, Xu D. Efficient reconstruction-based optic cup localization for glaucoma screening. In: International conference on medical image computing and computer-assisted intervention: Springer; 2013. p. 445–52.

[16] Li A, Cheng J, Wong DWK, Liu J. Integrating holistic and local deep features for glaucoma classification. In: Engineering in medicine and biology society (EMBC), 2016 IEEE 38th annual international conference of the; 2016. p. 1328–31. IEEE.

[17] Chen X, Xu Y, Yan S, Wong DWK, Wong TY, Liu J. Automatic feature learning for glaucoma detection based on deep learning. In: International conference on medical image computing and computer-assisted intervention: Springer; 2015. p. 669–77.