ARTIFICIAL INTELLIGENCE BASED GLAUCOMA DETECTION USING MACHINE LEARNING APPROACH

Priyanka .C,

Assistant Professor,
Department of Computer Science and Engineering,
KIT-Kalaignarkarunanidhi Institute of Technology,
(Autonomous)
Coimbatore, India.

Pooja. S,

Master of Engineering,
Department of Computer Science and Engineering,
KIT-Kalaignarkarunanidhi Institute of Technology,
(Autonomous)
Coimbatore, India.

***Abstract*— One of the ophthalmological conditions that commonly results in visual loss in today's culture is glaucoma. There is currently no test that has adequate sensitivity and specificity to diagnose glaucoma on its own. However, previous research have evaluated whether anatomical characteristics of the optic nerve can be predictive of glaucomatous damage. This study offers a public dataset containing health information and fundus photos of the same patient's both eyes. Additionally supplied are segmentations of the cup and optic disc as well as patient labelling based on the analysis of clinical data. A neural network was tested on the dataset to distinguish between healthy people and people with glaucoma. In particular, the ResNet-50 has been utilised as the foundation to categorise patients using data from each eye separately as well as utilising the combined data from each patient's two eyes. In order to advance research on the early diagnosis of glaucoma based on the combined examination of both eyes of the same patient, the results give the baseline measurements.**

Keywords*— deep learning, leaf disease detection, plant leaf disease detection, ladies finger, ladies finger leaf disease detection, CNN, convolutional neural network, neural network.

I INTRODUCTION

Glaucoma is an eye condition that, if not treated early enough, can cause lifelong blindness. It ranks as the second most common reason for eye blindness. An example of a contemporary imaging tool used to see within the eye is the fundus camera. The Topcon image net approach, optical coherence tomography, and the retinal nerve fibre layer analyser are a few of the techniques used to identify glaucoma. Optic cup to disc ratio is used to identify glaucoma, although this is due to the high cost and paucity of study in this area. The manner in which the optic cup appears is crucial in determining the extent of the glaucomatous damage. Glaucoma progresses, causing the cup to grow and eventually take up the majority of the disc surface. The optic cup-to-disc ratio evaluates the difference between the diameter of the optic disc's optic cup and the diameter of the entire optic disc. It takes time to manually inspect the optic disc and optic cup. So, an algorithm for automatically identifying glaucoma is devised. Enlargement of the optic cup alone, however, does not indicate glaucoma since genetic factors can induce optic cupping to develop without glaucoma.

Deep learning is the concept of employing algorithms to analyse large amounts of data to identify patterns and/or predict future outcomes. There are several algorithms available, each with a unique set of advantages and disadvantages as well as complexity levels. These methods are simple to use and available in a number of computer languages with various coding requirements (including R and Python). They can substitute more broad instructions for the detailed coding instructions required for your application.

Automated identification of glaucomatous damage and progression on OCT imaging has recently undergone new advancements and enhancements because to the development of unique deep learning (DL) algorithms. In order to enhance the diagnosis of glaucomatous damage on fundus photography, DL algorithms have also been trained using OCT data. This increases the potential value of colour photographs, which can be more easily obtained in a wider range of clinical and screening contexts. This paper summarises ten years of advancements in deep learning models trained using OCT structural data for glaucoma diagnosis and suggests future paths for applying these findings to the study of ageing and the basic sciences.

Without manual involvement, retinal fundus image processing has become a crucial analytical strategy for the early diagnosis of eye-related disorders as glaucoma and diabetic retinopathy. Predicting the bounding box coordinates of the optic disc (OD), which serves as a Region of Interest, is crucial for the analysis and diagnosis of glaucoma and other diseases from retinal images. On complete pictures, a Convolutional Neural Network (CNN) was trained to predict bounding boxes and the corresponding probabilities and confidence levels.

II RELATED WORKS

In 2013, Annu, N., Judith Justin et al., and others [1] proposed a new approach for detection based on textural energy characteristics from DWT. According to the study, five filters are used to create wavelet characteristics. With this technology, detection accuracy may be increased to 95%.

Abhishek Pal et al. (2018) presented the G-EYENET auto encoding system in [2], which is composed of two model system frames and the ROI, or area of interest composed of OD, which is acquired from fundus pictures. It was discovered using a binary map-based, modified U-net CNN.

Xiangyu Chen et al. suggested a six-layered CNN for glaucoma diagnosis in 2015 [3]. Overfitting was a significant issue that was resolved through response-normalization and layer pooling. It was known that the system used data augmentation and dropout tactics to enhance performance.

[4] Alan Carlos de Moura Lima, et al. (2018) compare a number of CNNs to determine which is the best. Each CNN architecture retrieved a significant amount of characteristics since it formed five datasets for each image.

U. Raghavendra et al. exploited the energy spectrum as a source for glaucoma diagnosis in [5] in 2015. Optic disc localization was initially carried out using a search window-based technique. Following that, the Radon Transformation (RT) and Modified Census Transformation were carried out (MCT). The classifier utilised was SVM. This technique was reputed to have a 97% accuracy rate.

S. Maheshwari et al. devised a technique for diagnosing glaucoma in [6] in 2017. EWT was initially used to split down images into different frequency bands. Then, correntropy characteristics were discovered. The value of the feature selection algorithm's t value was then used to rank features. In order to distinguish between the images with and without glaucoma, the picture was classified using a least squares support vector machine classifier. This method improves accuracy by 98.33 percent.

In [7], writers used texton and local configuration pattern-based characteristics to develop a glaucoma diagnosis. They started by doing adaptive histogram equalisation, then convolution operations on pictures using different filter banks, which produced textons. Additionally, Local Configuration Pattern (LCP), a recognisable pattern detected in the picture, was developed; system accuracy was 95.8%.

III METHODOLOGY

3.1. PROPOSED METHODOLOGY

The 256x256 pixel picture input is accepted by the convolution layer. CNN divides the image into 32 filters, each measuring 3 by 3. Next, a pooling layer of 2x2 filters is employed. The filtered value that was selected for the max-pooling procedure has the highest value. The second layer similarly has 32 filters, each measuring 3 by 3. The output is once more fed into a pooling layer with filters that are now 2x2 in size. The ReLU, or corrected Layer Unit, corrects the negative pixel values. The completely linked layer is categorised into the two groups after that. The CNN is then taught to divide the photos into these two groups.

3.2. IMAGE ACQUISITION

Images for the input dataset are initially gathered from a 256x256 dataset that serves as the input for data augmentation.

3.3. IMAGE PREPROCESSING

The primary issue that results in hazy, unresolved photos is fixed during preprocessing. The steps in this procedure include colour space conversion, picture restoration, and image enhancement. Each pixel in a grayscale picture represents one sample of value. As seen in Figure 1, they are made up of various degrees of grey, with black having the palest hue and white having the darkest. Grayscale is sometimes misunderstood to be the same as black and white photographs, which in the context of computer imaging only have two colours, black and white.

Fig. 1. Grayscale fundus image

3.4. DATA AUGUMENTATION

To produce fresh data with multiple orientations, data augmentation is utilised. It is crucial to the equilibrium of two glaucoma classes. For those classes, a horizontal flip operation is used to balance things out.
The first class is Class 0 (No Glaucoma).
The second class is Class1 (Glaucoma).

3.5 CNN ALGORITHM MODEL

We created various CNN architectures using the Inception-v3, VGG19, and ResNet50 models as the foundation. We evaluate the performance of the two datasets independently and for various design models.

CNN designs typically include three major layers. The convolution layer, which extracts the visual characteristics from input pictures, is the fundamental layer of a CNN design. In order to obtain a variety of image feature findings, a convolution layer often uses various convolution kernels as filters. The component that conducts dimensionality reduction for extracted features and data compression to avoid overfitting and improve model fault tolerance is the pooling layer, which is also known as the down-sampling layer. The completely linked dense layer ultimately gives the output of the object classification function after we have utilised max-pooling.

This layer classifies pictures by combining the feature data from each neuron in the top layer. To obtain the best performance, all layers of the CNN architecture—aside from the final weighted layer—were kept in a not-trainable mode.

3.6 DATA COLLECTION

We begin by compiling the finest dataset of images of the eye fundus. On Kaggle and dataverse, we started by hunting for images linked to melanoma. We chose the dataset with the highest number of likes or recommendations out of several datasets with a big number of high-quality pictures and models when looking for glaucoma diseases. The two publicly accessible datasets RIM-ONE [8] and ACRIMA [9] were used to evaluate the proposed technique. Eye fundus pictures are available in three forms in the RIM-ONE collection. There are 169 and 455 monosporic fundus pictures in versions 1 and 2, respectively. The 318 stereo pictures in Version 3 are divided into four types: normal, moderate, early, and deep. selecting, combining, and classifying the top images from the two databases. The collected images and datasets were used to frame and save the data in csv files.

3.7 DATA PREPROCESSING

Data preprocessing is a stage in the data mining and analysis process that converts raw data into a format that computers and machine learning can understand and interpret [10].

To the extent that machine learning models trained on bad data might potentially be damaging to the analysis you're attempting to accomplish — giving you "junk" findings — good, preprocessed data is even more vital than the most potent algorithms.

The act of updating missing information and deleting inaccurate or unnecessary material from a data collection is known as data cleaning.

The most crucial preprocessing stage is dating cleaning since it guarantees that your data is prepared for your downstream requirements.

Data transformation will start the process of transforming the data into the appropriate format(s) you'll need for analysis and other downstream operations. Data cleaning has already started the process of modifying our data. In order to prepare the dataset for inclusion in this study's CNN algorithm training, it was analysed, cleaned up, and converted.

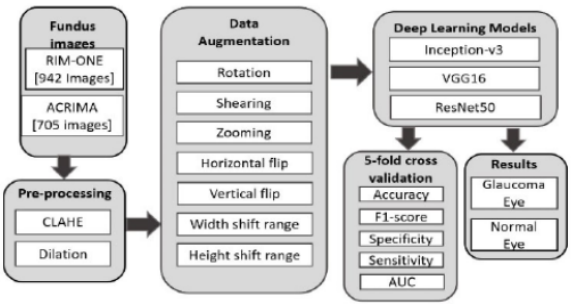

Figure 2: Data flow diagram

IV MODULES

4.1. CNN MODEL

1) The Inception-v3 architecture. The widely used Inception-v3 CNN architecture was created by Google in 2015 and has 48 deep layers. A popular picture recognition model called Inception-v3 is trained on the ImageNet dataset for 1000 classes and has been proven to achieve an accuracy of more than 78.1 percent [11]. The basic CNN architecture was built using the pre-trained

Inception-v3 model from the Keras API. Instead of using the model's top dense layers, we have introduced extra layers in our method. After the Inception-v3 model, the global average pooling layer, the dense layer (512 units), and the Softmax layer were added to minimise the parameters. For the normal and glaucoma categories, it used two neurons in the Softmax layer, respectively. Add a dropout layer with 0.7 rates to reduce model overfitting. Additionally, Inception-v3 contains an additional classifier to help with the vanishing gradient problem. In order to train the model, they employed Adam optimizer and a high learning rate. Figure 3 depicts the improved Inception-v3 architecture.

2) VGG19 Architecture: VGG19 has 19 layers total, with an additional layer of convolution in each of the last three blocks. We improved the pre-trained VGG19 model in this work. Here, the last three newly added layers—dropout layer with 0.5 rates, dense layer (256 units) with ReLU activation function, and finally Softmax layer with two outputs—follow the global average pooling layer. The modified VGG-19 diagram that demonstrates its design is shown in Figure 4.

3) Residual Networks (ResNet) of 50 layers have demonstrated strong performance on picture identification and localization tests. The key issue ResNet50 is seeking to solve is the vanishing gradient issue that arises when using a large number of layers. It has a few skip connections as a result, allowing it to retain information from earlier than two levels. We have adjusted the pre-trained ResNet50 as follows for this project. The dense and softmax layers, the final two layers, are changed first.

The fully connected layer is then replaced by a second, 256-unit fully connected dense layer in the same pre-trained networks. The two classifications of glaucoma and normal are shown in the output. Modified ResNet50 architecture is shown in Figure 5.

The ideal training settings have been found through hyper-parameter tweaking. For internal layers, the relu activation function is utilised, and a batch size of 8 is used for each model's training. Using the ADAM optimizer, the Inception-v3 model is trained at a learning rate of 0.0001. In order to train the VGG19 and ResNet50 models, stochastic gradient descent (SGD) optimizer updates with 0.9 momentum are used. We used binary cross-entropy loss (LossBC E) for each model as the challenge requires binary classification (glaucoma and healthy). For N fundus photos, where p(yi) is the projected probability for the particular image and yi is the actual value actually seen, observing a value of 0 or 1 (yi 0,1). LossBC E determines the deviation, which may either be 0 or 1, between the expected probabilities and the actual class output. In this work, we trained both datasets across 150 epochs.

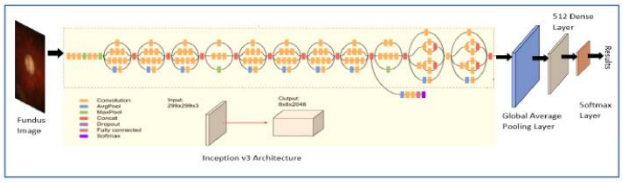

Fig. 3. Inception v3 Architechture

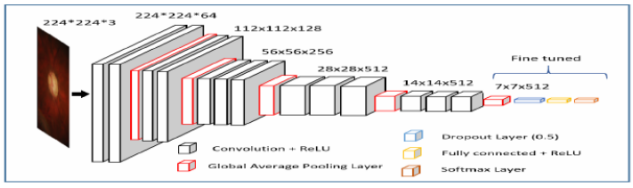

Fig. 4. VGG-19 Architechture

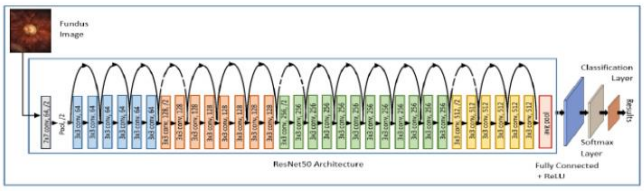

Fig. 5. Resnet50 Architechture

4.2 DATASET

We have categorised all three classes as glaucoma since deep, early, and moderate classes are all associated with the disease. Thus, there are 399 glaucomatous and 543 healthy participants in the RIM-ONE dataset's 942 fundus pictures. There are 309 healthy participants and 396 glaucomatous subjects in the ACRIMA dataset. According to empirical research, splitting the photos into training, testing, and validation images at a ratio of 70:15:15 during the DL process produces the best results. To solve the class imbalance and enhance the number of pictures in the training set for improved performance, data pre-processing techniques are used.

Figure 6(left) displays a sample of a healthy picture, while Figure 6(right) displays a sample of a glaucoma image (right). The optic cup has a regular form and size in Fig. 6(left), and it is bigger and has a different size neuroretinal rim in Fig. 6(right) than in the first picture.

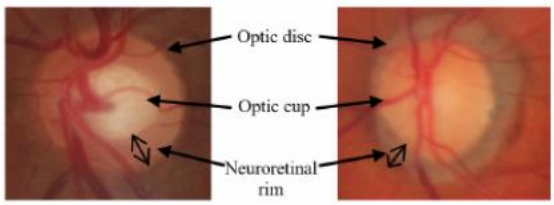

Fig. 6. Healthy vs unhealthy fundus

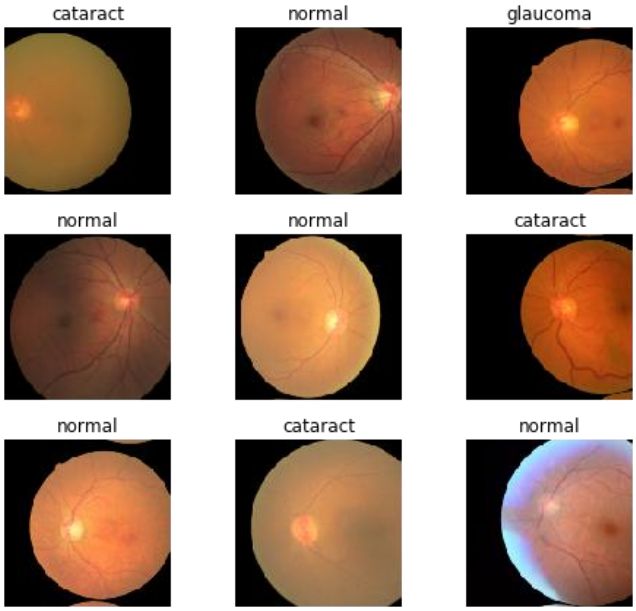

Fig. 7. All classes in the dataset

The collection also includes information on eyes with cataracts. Only the photos of the healthy and glaucoma-affected eye fundi were used to train the model in this investigation.

V RESULT AND DISCUSSION

5.1. ACCURACY AND OBSERVATIONS

Figures 8, 9, and 10 demonstrate, respectively, the findings for the Inception-v3, VGG19, and ResNet50 models' training accuracy, validation accuracy, training loss, and validation loss. We have displayed the findings for each model for the two datasets individually. It can be seen that when the number of epochs rises from 100 to 200, the model's performance improves and it achieves a high degree of accuracy while avoiding overfitting. But generally speaking, having too many or too few epochs might result in overfitting or underfitting, respectively. When an arbitrary high number of epochs is reached, early stopping is used to stop the performance improvements on a holdout validation dataset.

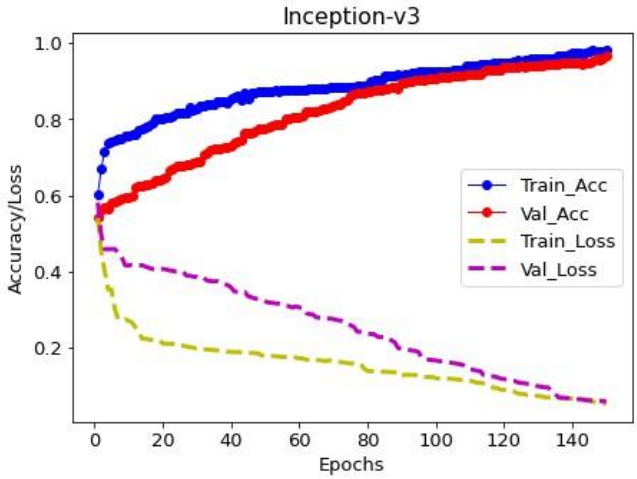

Fig. 8. Accuracy of Inception v3

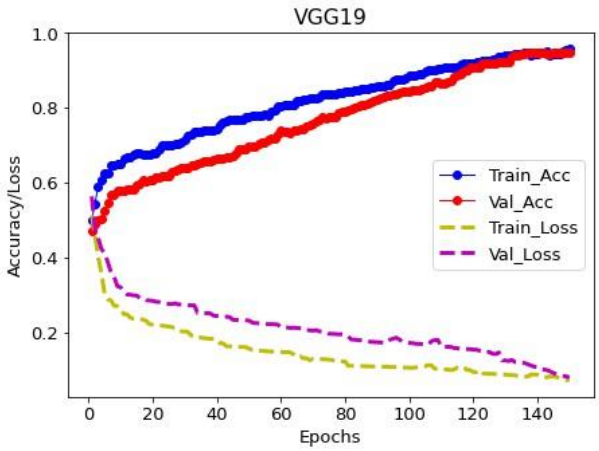

Fig. 9. Accuracy of VGG-19

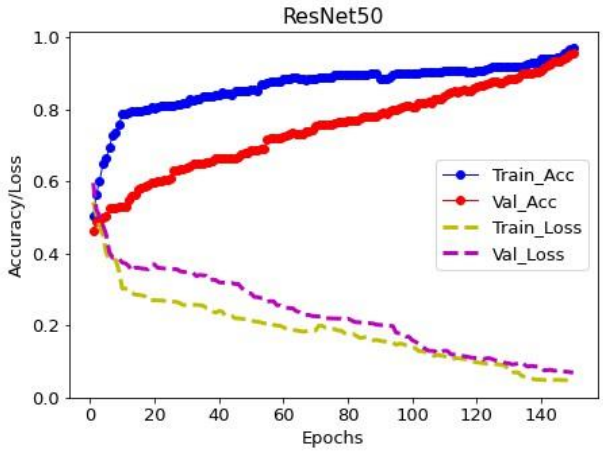

Fig. 10. Accuracy of ResNet-50

Additionally, figures 11, 12, and 13 show, respectively, the confusion matrices of CNN architectures based on the RIM-ONE and ACRIMA datasets. This displays the proportion of accurate and inaccurate predictions made for each class. It is clear that every model has produced more accurate forecasts.

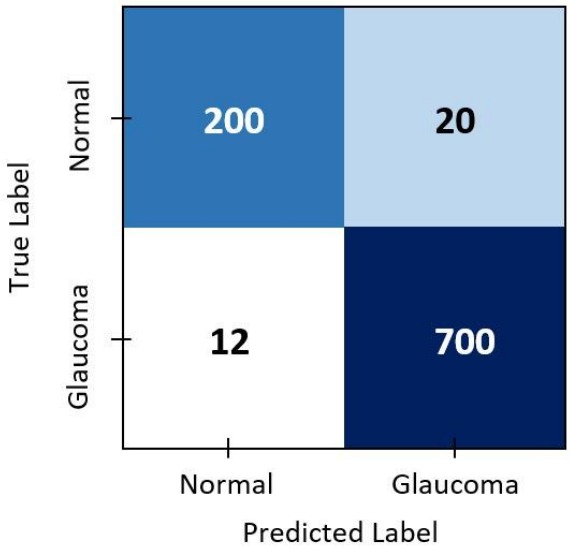

Fig. 11. Confusion Matrix of Inception-v3

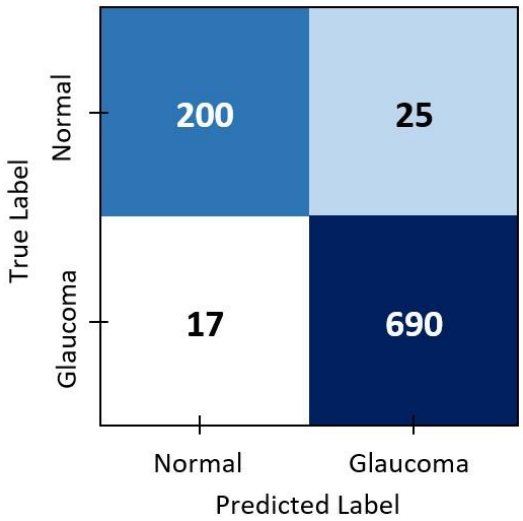

Fig. 12. Confusion Matrix of VGG-19

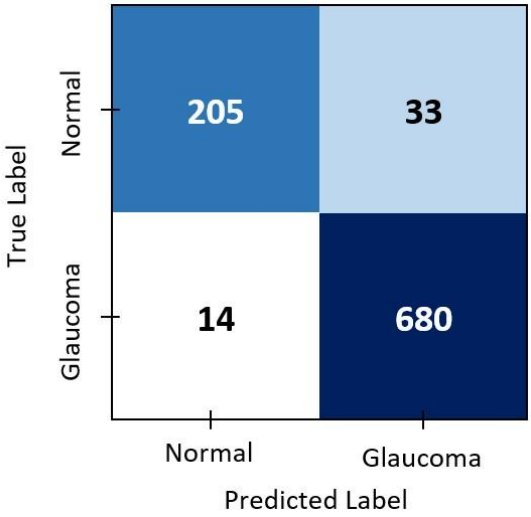

Fig. 13. Confusion Matrix of ResNet-50

Inception-v3 on the RIM-ONE dataset achieves 96.56 percent accuracy and the least amount of test loss when compared to the performance of each CNN architecture. Additionally, Inception-v3 achieves 98.52 percent accuracy with minimal testing loss on the ACRIMA dataset.

VI CONCLUSION AND FUTURE WORK

Millions of people have the devastating eye condition glaucoma each year. The classification of fundus pictures using deep learning was described in this study. To categorise glaucoma and normal pictures, we have suggested three refined architectures based on Inception—v3, VGG19, and ResNet50. During the pre-processing of the data, dilatation and augmentation were used. Inception-v3 outperforms the other convolutional neural network designs in predicting glaucoma with an accuracy of 98.52 percent, according to 5-fold cross-validation-based findings. To get the best results, this study may be expanded to test out more deep learning methods. A glaucoma diagnosis assistance system may also be deployed in real-world situations using deep learning architectures, which can be built to implement in low computing power environments with restricted resources.

In addition, I'll employ the SVM (Support Vector Machine) technique in phase 2 to compare with our present model to increase accuracy, and I'll also create a graphical user interface (GUI) to make it simpler for the general public to use.

REFERENCES

[1] Annu, N., and Judith Justin. "Automated classification of glaucoma images by wavelet energy features." International Journal of Engineering and Technology 5, no. 2 (2013), pp: 1716-1721.

[2] Abhishek Pal, Manav Rajiv Moorthy , A. Shahina , “G-EYENET: a convolutional autoencoding classifier framework for the detection of glaucoma from retinal fundus images”, 25th IEEE International Conference on Image Processing (ICIP), 2018.

[3] Xiangyu Chen, Yanwu Xu, Damon Wing Kee Wong, Tien Yin Wong, Jiang Liu, “Glaucoma Detection based on Deep Convolutional Neural Network”, 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015.

[4] Alan Carlos de Moura Lima, Lucas Bezerra Maia, Roberto Matheus Pinheiro Pereira, Geraldo Braz J´unior, Joao Dallyson Sousa de Almeida, Anselmo Cardoso de Paiva, "Glaucoma Diagnosis over Eye Fundus Image through Deep Features", 25th International Conference on Systems, Signals and Image Processing (IWSSIP), 2018.

[5] U. Raghavendra, Sulatha V. Bhandary, Anjan Gudigar, U. Rajendra Acharya, "Novel expert system for glaucoma identification using non-parametric spatial envelope energy spectrum with fundus images", Bio cybernetics and Biomedical Engineering, Volume 38, Issue 1, 2018, pp:170-180.

[6] S. Maheshwari, R. B. Pachori, and U. R. Acharya, "Automated Diagnosis of Glaucoma Using Empirical Wavelet Transform and Correntropy Features Extracted From Fundus Images," IEEE J. Biomed. Health Inform. 21(3), (2017) pp: 803–813.

[7] U Rajendra Acharya, Shreya Bat, Joel EW Koh, Sulatha V Bhandary, Hojjat Adeli, A novel algorithm, to detect glaucoma risk using texton and local configuration pattern features extracted from fundus images, Computers in Biology and Medicine, vol. 88 (2017),pp: 72-83.

[8] FUMERO BATISTA, Francisco José et al. RIM-ONE DL: A Unified Retinal Image Database for Assessing Glaucoma Using Deep Learning. Image Analysis & Stereology, v. 39, n. 3, p. 161-167, nov. 2020. ISSN 1854-5165. Available at: https://www.ias-iss.org/ojs/IAS/article/view/2346. doi: https://doi.org/10.5566/ias.2346.

[9] Diaz-Pinto, Andres & Morales, Sandra & Naranjo, Valery & Köhler, Thomas & Mossi, José & Navea, Amparo. (2019). CNNs for automatic glaucoma assessment using fundus images: An extensive validation. BioMedical Engineering OnLine. 18. 10.1186/s12938-019-0649-y.

[10] Siddharth Misra, Hao Li, Jiabo He, Chapter 5 - Robust geomechanical characterization by analyzing the performance of shallow-learning regression methods using unsupervised clustering methods, Editor(s): Siddharth Misra, Hao Li, Jiabo He, Machine Learning for Subsurface Characterization, Gulf Professional Publishing, 2020, Pages 129-155, ISBN 9780128177365, https://doi.org/10.1016/B978-0-12-817736-5.00005-3.
(https://www.sciencedirect.com/science/article/pii/B9780128177365000053)

[11] Szegedy C., Vanhoucke V., Loffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; June 2016; Las Vegas, NV, USA.