Glaucoma Detection Using Computer Vision and Vision Transformers

Aravind Karrothu[1,*], Anilkumar Chunduru[2]

[1,2]Assistant Professor, Dept of Information Technology, GMR Institute of Technology, Rajam, Vizianagaram, Andhra Pradesh-532127

*aravind.k@gmrit.edu.in

Abstract

Vision is one of the gifts from nature to humans. It stores visual messages and signals to the brain. Glaucoma is a high-pressure genetic disorder that relates to the human eye. It is a chronic disease often called "silent -thief of sight" as it has no symptoms. Glaucoma is a major problem for the human eye, often caused by abnormally high pressure in the eye. It leads to damage to the optical nerve and causes blindness. This disease is considered an irreversible disease that results in vision loss. If it is not detected at an early stage, it may cause permanent blindness. There is no proper cure for glaucoma if it was identified at the last stage. Early detection of this disease is very important for providing correct treatment and minimizing vision loss. The common method used for diagnosing glaucoma is through the examination of an eye by an ophthalmologist. Manual glaucoma detection is costly and time-consuming. So, this works presents a very useful computer scientific method for glaucoma detection. In the proposed model Computer Vision (CV) and Vision Transformers (VIT) were used to identify whether the person is affected by glaucoma or not.

Keywords: Glaucoma, Fundus Images, Computer Vision, Vision Transformers, Cup to Disc Ratio.

1. Introduction

Eyes are sensory part of the human body. They can capture the scenes around them and store in a place called brain. The storing is done through the optic nerves. Literally there are millions of them in a bundle form inside anatomy of the eye. Basically, optic nerve carries visual messages and sensory signals to the brain. Such a great thing is done by optic nerve. There are some diseases which affects the optic nerve damage. Glaucoma is disease which is caused by diabetes. This glaucoma can affect a person for vision loss. When a person with diabetes put high effort on his eyes, he/she will get effected immediately to glaucoma disease. The treatment for any disease to diabetic patient is much difficult than a compared to normal patient. Hence, glaucoma is an irreversible disease.

Glaucoma is a disease which is characterized as a loss of retinal cells. Not only with high pressures, but it can also cause through genetically. It is the disease in second place which leads the causing of vision loss. In 2020, the survey said that 57.5 million people were affected to glaucoma. The number may rise at 111.8 million in 2040 [1]. There are some conventional ways to detect the glaucoma in the eyes. For example, tonometry is an instrument which measure Intraocular Pressure (IOP). With the help of this instrument, the IOP is measured [2]. But if it touches the cornea then the cornea may damage. The accuracy of this depends upon thickness of cornea. So, this work presents Fundus imaging in glaucoma detection through CV.

Liu Li, Mai Xu, Hanruo Liu et al., made an Attention-Based Convolutional Neural Networks (AB-CNN) with heat map generation. The heatmaps tracks the required region in fundus image. The authors cropped the fundus images using OpenCV and stored in as their database. The AB-CNN helps to create required ground truth labels for glaucoma detection through cup-to-disc ratio [3].

This works provides an enhancement to the state-of-art model. CV algorithms are the best for preprocessing the dataset. The main theme of the existing work is focusing and heat map generation. Image Enhancement (IE) is a technique which can used to focus on a particular region. With the help IE algorithm, we can focus the region that we are needed. VIT techniques are also the best in focusing the required regions. This work also consists of VIT algorithm to detect the glaucoma in the human eye.

2. Literature Survey

The author’s discussed an Attention-based CNN algorithm for Glaucoma Detection. The algorithm mainly focusses on heatmap generation and cup to disc ratio. The proposed model shows 96.25% accuracy with the LAG dataset and 82.25% with the RIM-ONE dataset. An attention prediction subnet is utilized in AG-CNN to build the focused and cropped maps of fundus pictures then given for glaucoma detection and to measure using statistically [3]. Wheyming Tina Song et al., proposed a robust measurement to detect the glaucoma. They suggested a fundus image-based glaucoma detection methodology that is robust in nature, and automated. The dataset for the proposed approach is 1450 fundus images. The proposed models outperform the state-of-art model. The accuracy of the model is 97% with the 1450 color fundus image dataset [4].

The researchers discussed about CNN in glaucoma detection. They defined glaucoma detection as a classification issue of two classes and learned the representation using CNN. They also investigated the impacts of image size on discriminative capacity and used heatmap analysis to figure out which areas of the image are involved in discrimination. The results of their work shown that input image size is not matter but resolution of the image is mattered. Among the models the ResNet152 shows high accuracy of 95% [5]. The author discussed about automatic glaucoma detection using transfer learning. The pre-trained that they used are ResNet50 and GoogleNet. In terms of Accuracy, Sensitivity and Specificity. For Early Glaucoma Detection some statistical measures were taken. The two performs a great job in measuring the cup-to-disc ratios [6].

The author's Deepak Ranjan Nayak et al., discussed about evolutionary convolutional network for automated detection of glaucoma from fundus images. They used optimized approach at feature extraction point. To evaluate the efficiency of the models they tested with various datasets. The statistical measures are well yielded while using Support Vector Machine (SVM). It also outperforms the state-of-art model [7]. The author's discussed about segmentation of glaucoma. They have used ACRIMA database. The ACRIMA database consists of 705 fundus images, which are used for detecting glaucoma disease in initial stage. In this the photo segmentation method follows the three important steps namely preprocessing, segmentation, feature extraction and it's evaluated the performance in terms of qualitative and quantitative. The performance of the photo segmentation (PS) technique is good in terms of higher disease detection accuracy and lesser disease detection time, compare with the other models [8].

Juan Carrillo, Lola Bautista et al., discussed that Glaucoma is a disease and which is one of the major causes for the permanent blindness. There is no cure to this disease. So, they used the Fundus imaging. It is one of the methods used mostly for detecting glaucoma. They proposed a model like disc segmentation and cup segmentation which detects the glaucoma automatically. The area of cup by area of disc is measured, if the value is greater or equal to 0.6 it is said they are affected by glaucoma [9]. The author's discussed about glaucoma was the optical state which causes blindness if it is not detected at early stage. To overcome this problem, they have some scientific computer aided detection. Transfer learning is state-of-art to work on existing to find the solution for new work. The result of their work is a classification as no glaucoma, early glaucoma, and advanced glaucoma. The model shows a better performance of 80% as compared to the previous works [10].

Shuang Yu, Di Xiao et al., discussed about the detecting the glaucoma at early stage is importing it helps to take treatment at right time and avoid from vision loss. It mainly concentrates on the optic disc and cup of optical fundus images. They have used a modified encoder-decoder model which is the combination of both ResNet-34 as encoding and U-net as decoding. The proposed method in this paper gives robust performance which it avoids the over-fitting and [11]. The author's discussed about Glaucoma is a case where blindness occurs if not treated in soon. It should be detected at early stage. There are a few strategies where that can help to overcome glaucoma. CDR found in retinal fundus images examined so that it detects the presence or absence of glaucoma. They have taken the threshold value at 0.6 which gives the best performance of their proposed model and achieves greater accuracy than state-of-art model. The dataset used is RIMONE v3 for the proposed model and its accuracy is 87.4% [12].

The researchers discussed about Attention CNN. Their work represents on focusing and heatmap generation. There is a great need for fundus images for glaucoma diagnosis. A great deal of attention to how powerful CNN-based glaucoma diagnostics. They have classified the entire dataset as positive glaucoma and negative glaucoma. The proposed model shows 96.25% accuracy with LAG dataset and 82.25% with RIM-ONE dataset [13]. The author's Arkaja Saxena et al., discussed about the so-called disease known as glaucoma which damages the vision system in human body. This condition is an incurable disease that leads to the deterioration of vision. Deep Learning (DL) provides a best classification using some statistical measurements. Differences between glaucoma and non-glaucoma patterns can be determined through DL models. They differentiated the conventional process of detecting the glaucoma and identified some state-of-art DL models. Data sets used for testing by SCES and ORIGA. The performance of model is 88.4% [14].

The researchers stated that the DL models plays a vital role in in Medical Image Segmentation (MIS). They combined two traditional methods and produced a new model named as BoostNet. They divided the entire model into four phases. Finally, they have combined ChampNet and CLAHE models and achieved 95.88% of accuracy [15]. Sonali et al., proposed a technique which filters the noisy area and low contrast areas and creates sharpened images that highlight the object area in an image. The proposed model is known as CLAHE. The evaluation metrics for this model was used are PSNR, SSIM which is same for the Image Super-Resolution (ISR). The model exceeded PSNR improvement by 7.85% and SSIM improvement by 1.19% of the existing model [16]. Viacheslav Voronin et al proposed 3-d block-rooting scheme with application to medical image enhancement. The datasets they have used for this work are NYU, fastMRI, Chest x-ray. They have used Pearson Correlation Coefficient (PCC) and SD for accuracy and performance.

Finally, they show that achieves good performance and can prevent noise increasing during sharpening of image details [17].

3. Methods

3.1. Proposed Architecture

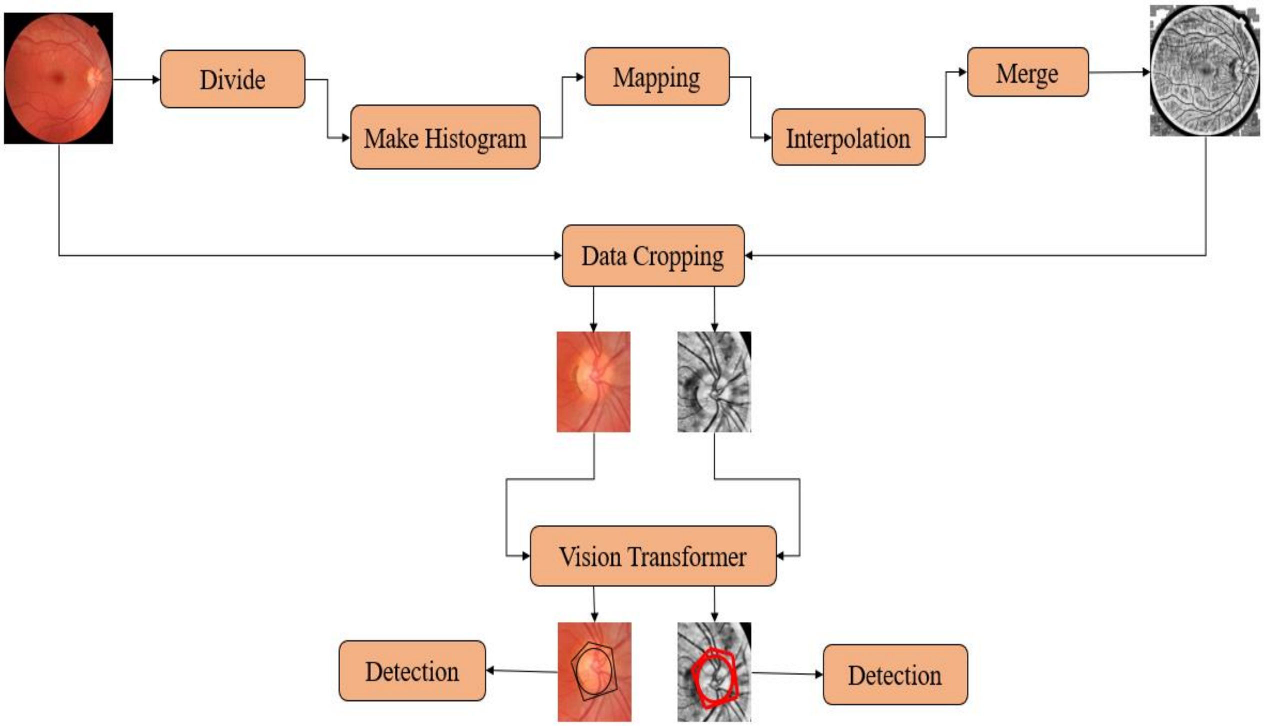

Figure 3.1.1. Glaucoma detection through CLAHE and VIT

After analyzing the literature of this work, we have come to know that the problem statement of our work is a binary classification with some statistical measurements. This statistical measurement helps to identify the cup-to-disc ratios through heatmaps on required area which is to be focused. The cup-to-disc ratio provides the knowledge like detecting the glaucoma in the images to the state-of-art model AB-CNN. The entire work is divided into three modules. Each module has its own purpose in this work. In this work, we are not using any statistical measurements to detect the glaucoma. The first module is all about dataset preparation. To classify our problem statement without statistical measurements, the focus should be on the required region to detect the glaucoma. In the IE technique, CLAHE is one of the models which is used to increase the contrast of the images using adaptive method and histograms. The last module is the VIT to classify the given input images as “GLAUCOMA” or “NOT GLAUCOMA”.

3.2. Dataset Preparation

3.2.1. Data Classification

In this part, we collected the data, and we can perform cropping and split on collected data into various datasets. And this data is in the form of images. These images are labeled as two types of glaucoma or not glaucoma. since we are using a binary image classifier. The glaucomatous image represents the person who suffered from glaucoma disease and the not glaucoma represents the person who is not suffering from glaucoma disease.

3.2.2. Data Gathering

In the state-of-artwork, they have used two types of datasets for fundus image classification. They are LAG dataset and RIM-ONE dataset. The state-of-art models achieved 96% of accuracy with LAG dataset and 85% with RIM-ONE dataset. We finally stick with the RIM-ONE dataset, and it contains full eye fundus images.

3.2.3. Data Cropping

To focus on the required region, we must crop the full fundus images into cropped fundus images using OpenCV. Generally, these fundus images are of two types. They are right eye fundus image and left eye fundus image. We can identify glaucoma at the cup to disk part of the fundus image.

For the right eye fundus image, the cup to disk part is present on the left side of the image so we need to crop the length in such a way that the cup to disk part is clearly cropped. For the left eye fundus image, this cup to disk part is present on the right side of the image so we need to crop accordingly. For both types of images, the width will be the same because the cup to disk part is present at the center of the width side.

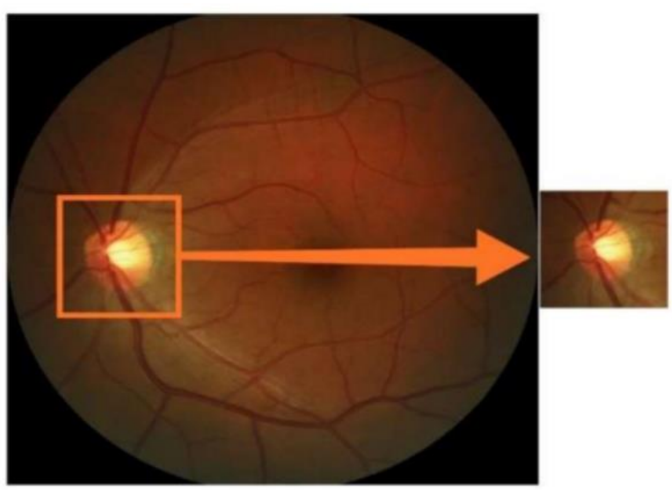

Figure 3.2.1. Sample fundus image cropping

3.3. Preprocessing Module

The classification can be done through RGB images and B&W images. But B &W images are very useful in classification problems. Although to convert a RGB image into B &W is easy. But here the problem statement is about focusing. To focus on the required area, CLAHE is used. This CLAHE model differentiate the brighter features and darker features in the input image. The CLAHE model is unique state-of-art model which is best for B&W images. The CLAHE model is simple 5 stage run. In the first stage, the entire image is divided into regions. The next stage is to make histograms for those divide regions. The next stage is for mapping derived histogram. Later, the comparisons can be done to initial histograms and derived histograms. At last, all regions will be merged and generate new B &W image with more focusing on required area from RGB fundus images.

3.4. Vision Transformers

Vision transformers are transformers that are well-known in the field of deep learning. Prior to the development of vision transformers, we used computer vision to perform complex tasks. With the introduction of vision converters, we have found one very powerful model of computer vision functions. The transformer in machine learning is an in-depth learning model that uses a focused approach, measuring separately each part of the input data. First, mind converts were used in Natural Language Processing (NLP), and more recently they have been used in computer vision. The transformation model uses multi-header attention in Computer Vision without the need for image-specific bias. It has a high degree of accuracy in large databases with reduced training time. The performance of the vision converters depends on the configuration decisions, network depth, and database parameters.

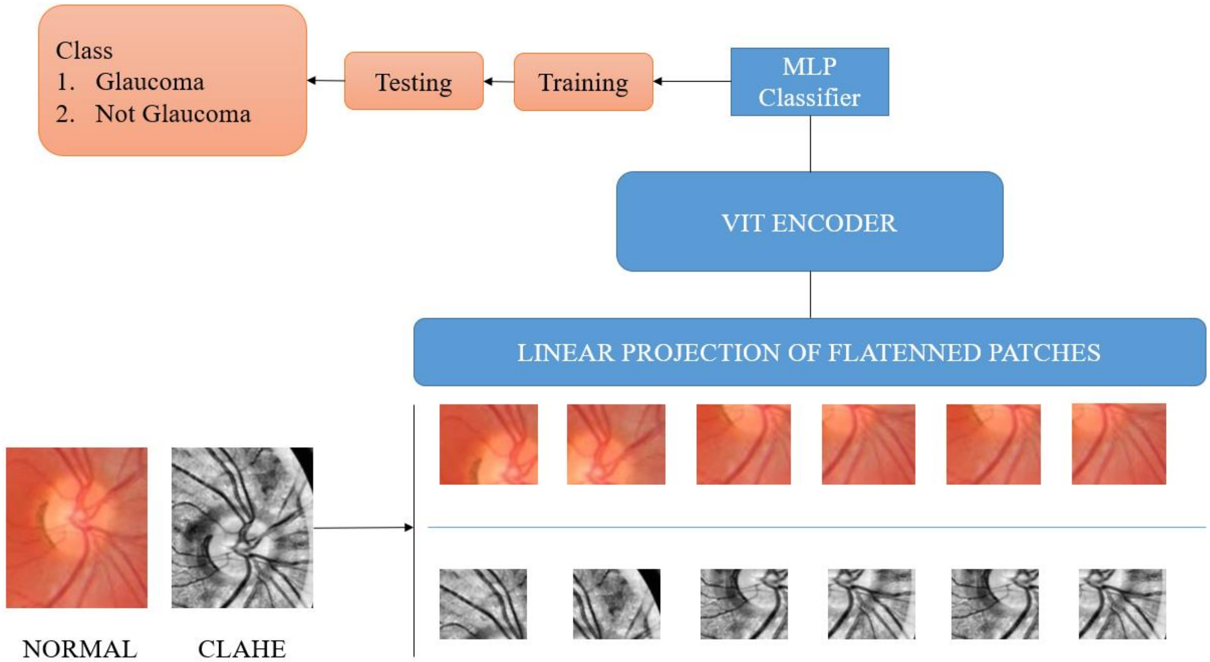

Figure 3.4.1. Architecture of Vision Transformer for glaucoma detection

4. Results and Discussions

4.1. Results in images

4.1.1. CLAHE Results

The fundus images are the cropped RGB images. The RGB images contains 3 main channel colors. The CLAHE model divides the R-RED, G-GREEN, and B-BLUE channels into three subdivisions and increase the contrast of those regions adaptively. These CLAHE images highlights the fluid structure of the glaucoma than compared the RGB image. Figure 4.1.1.1 is the sample glaucoma classed image with highlighting fluid in CLAHE

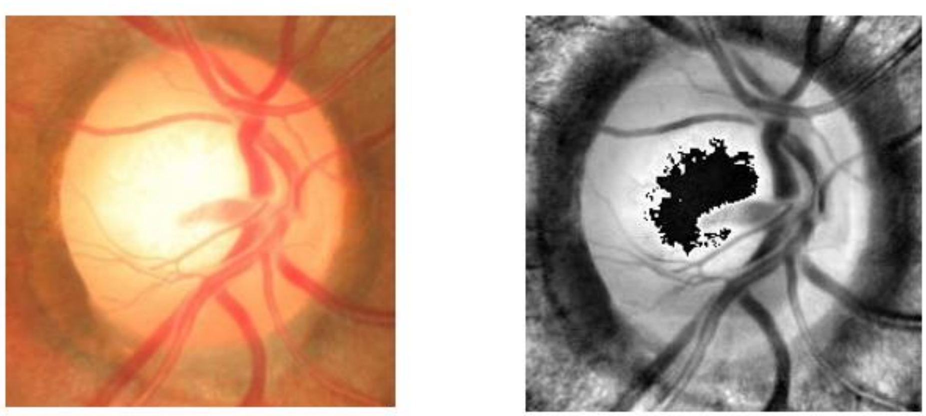

Figure 4.1.1.1. Sample glaucoma classed image with highlighting fluid in CLAHE

4.1.2. Classification of GLAUCOMA and NOT GLAUCOMA through proposed model

The classification confidence is quite given us good results. 4.1.2.1 and 4.1.2.2 are the sample results of both normal test images CLAHE images.

Figure 4.1.2.1. Sample classification of normal test images

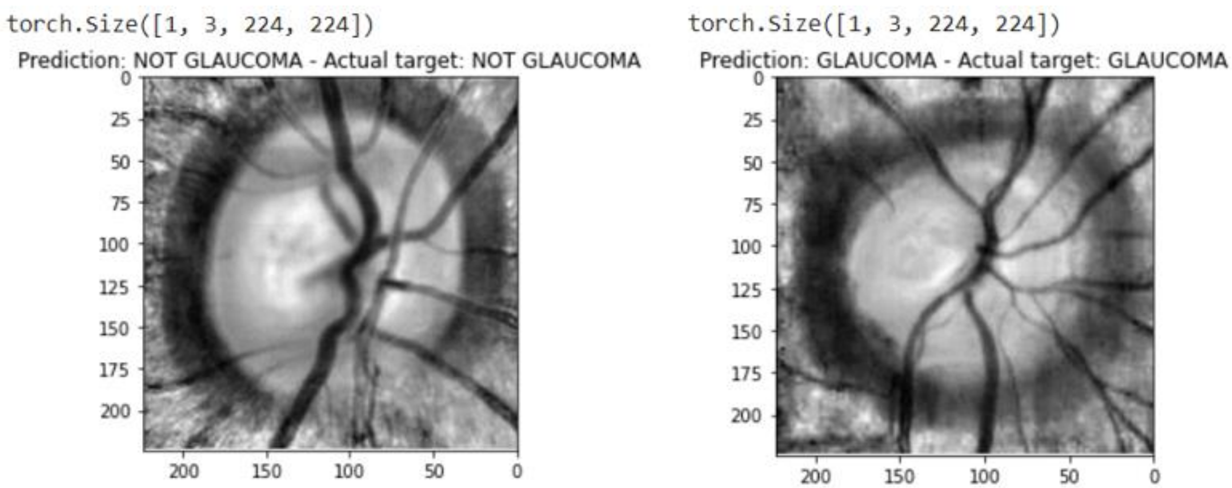

Figure 4.1.2.2. Sample classification of CLAHE test images

4.2. Results comparison

Table 4.2.1. Results of RIM-ONE dataset with state-of-art model and proposed model

S. No.	Model	Accuracy
1	AB-CNN	85.2%
2	VIT + Normal	91.4%
3	**VIT + CLAHE**	**95.7%**

From the above table, the accuracy with CLAHE images is classified as the best of all the models.

4.3. Graphical results

From graphical representation, we can clearly conclude that CLAHE images are more efficient than normal images that shown in the figure 4.3.1. As well as, the training loss of CLAHE images also less than the normal images, that could be showing in the figure 4.3.2.

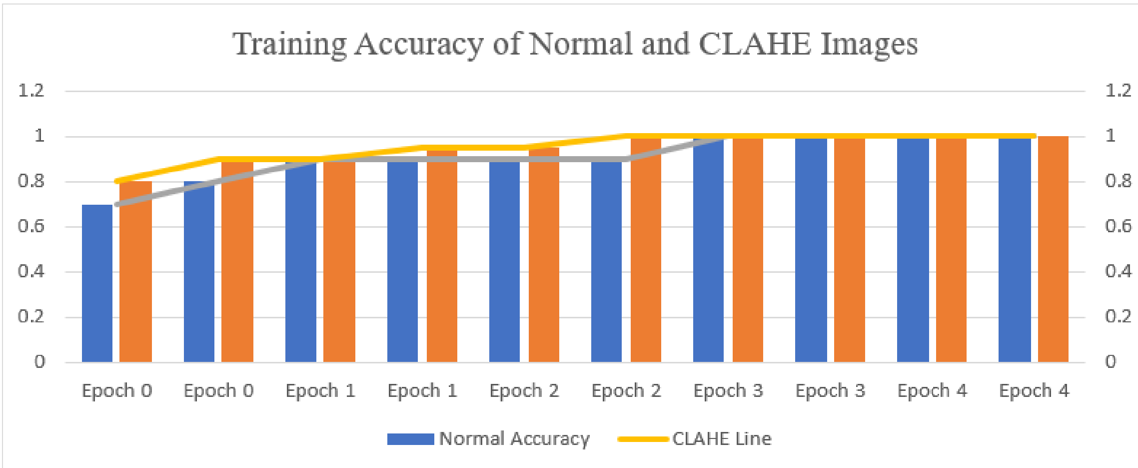

Figure 4.3.1. Graph representation training accuracy of normal and CLAHE images

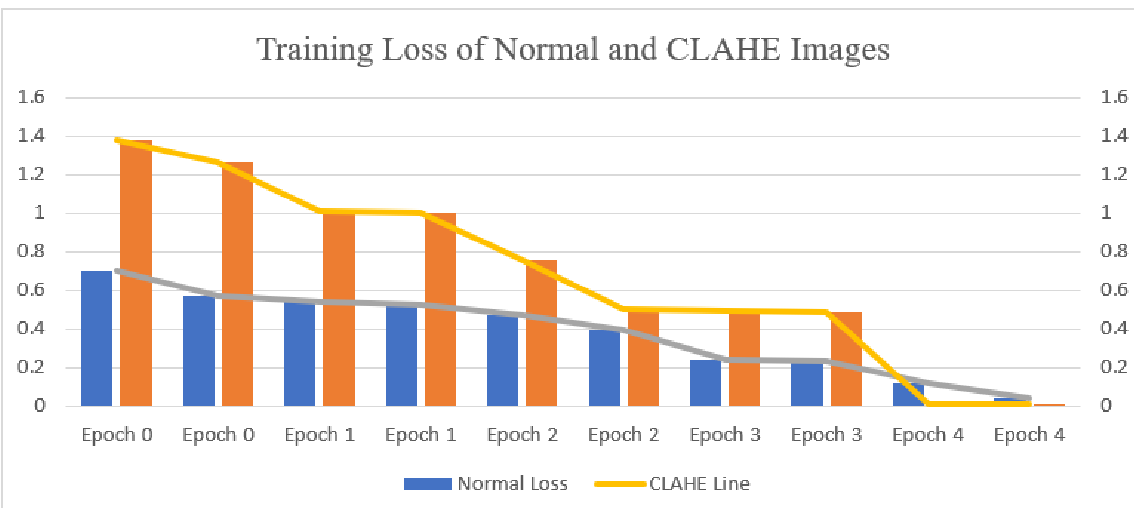

Figure 4.3.2. Graph representation training loss of normal and CLAHE images

5. Conclusions

The main motto of our work is the diagnosing of glaucoma identification with in short span of time (in seconds). For that first we need to gather all the images of a retinal eye scan. By using this model, the diagnosing of glaucoma detection will be able to detect glaucoma without the assistance of a professional ophthalmologist. So, we feel that the proposed method with CLAHE images has been given best accuracy when compared with other models as well as this model is more efficient than manual diagnostics since machines make fewer errors than humans. Converting an input image into B &W, classifies more efficiently than RGB input images.

6. Future Work

The state-of-art IE technique has achieved good result. Further, we want to enhance our dataset with other CV models such as Super-Resolution (SR). The SR images generates a high-quality image which can be used better for classification. We further develop this model into MATLAB environment which is the best for medical projects developments.

List of abbreviations:

CV	Computer Vision
VIT	Vision Transformers
IOP	Intraocular Pressure
AB-CNN	Attention-Based Convolutional Neural Networks
IE	Image Enhancement
SVM	Support Vector Machine
MIS	Medical Image Segmentation
DL	Deep Learning
CLAHE	Contrast-limited adaptive histogram equalization
ISR	Image Super-Resolution
PCC	Pearson Correlation Coefficient
NLP	Natural Language Processing
SR	Super-Resolution

Availability of data and material

The dataset generated and/or analyzed during the current study are available in the [RIM-ONE dataset and LAG dataset] repository,

[https://www.kaggle.com/datasets/lucascunhadecarvalho/rimoner2;
https://github.com/smilell/AG-CNN]

Competing interests

The authors declare that they have no competing interests

Funding

No Funding

Authors' contributions

Both authors have involved for the complete work implementation.

Acknowledgements

Not applicable

References

1. Allison, K., Patel, D., & Alabi, O. (2020). Epidemiology of glaucoma: the past, present, and predictions for the future. Cureus, 12(11).
2. Zeppieri, M., & Gurnani, B. (2022). Applanation Tonometry. In StatPearls. StatPearls Publishing.
3. Li, L., Xu, M., Liu, H., Li, Y., Wang, X., Jiang, L., ... & Wang, N. (2019). A large-scale database and a CNN model for attention-based glaucoma detection. IEEE transactions on medical imaging, 39(2), 413-424.
4. Song, W. T., Lai, C., & Su, Y. Z. (2021). A Statistical Robust Glaucoma Detection Framework Combining Retinex, CNN, and DOE Using Fundus Images. IEEE Access, 9, 103772-103783.
5. Phan, S., Satoh, S. I., Yoda, Y., Kashiwagi, K., & Oshika, T. (2019). Evaluation of deep convolutional neural networks for glaucoma detection. Japanese journal of ophthalmology, 63(3), 276-283.
6. Serener, A., & Serte , S. (2019, October). Transfer learning for early and advanced glaucoma detection with convolutional neural networks. In 2019 Medical technologies congress (TIPTEKNO) (pp. 1-4). IEEE.
7. Nayak, D. R., Das, D., Majhi, B., Bhandary, S. V., & Acharya, U. R. (2021). ECNet: An evolutionary convolutional network for automated glaucoma detection using fundus images. Biomedical Signal Processing and Control, 67, 102559.
8. Raja, P. S., Sumithra, R. P., & Thanusha, G. (2021, March). Automatic Glaucoma Diagnosis Based on Photo Segmentation with Fundus Images. In 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp. 102-105). IEEE.
9. Carrillo, J., Bautista, L., Villamizar, J., Rueda, J., & Sanchez, M. (2019, April). Glaucoma detection using fundus images of the eye. In 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA) (pp. 1-4). IEEE.
10. Serte, S., & Serener, A. (2019, October). A generalized deep learning model for glaucoma detection. In 2019 3rd International symposium on multidisciplinary studies and innovative technologies (ISMSIT) (pp. 1-5). IEEE.
11. Yu, S., Xiao, D., Frost, S., & Kanagasingam, Y. (2019). Robust optic disc and cup segmentation with deep learning for glaucoma detection. Computerized Medical Imaging and Graphics, 74, 61-71.
12. Joshua, A. O., Mabuza-Hocquet, G., & Nelwamondo, F. V. (2020, January). Assessment of the cup-to-disc ratio method for glaucoma detection. In 2020 International SAUPEC/RobMech/PRASA Conference (pp. 1-5). IEEE.
13. Li, L., Xu, M., Wang, X., Jiang, L., & Liu, H. (2019). Attention- glaucoma detection: a large-scale database and CNN model. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10571-10580).

14. Saxena, A., Vyas, A., Parashar, L., & Singh, U. (2020, July). A glaucoma detection using a convolutional neural network. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 815-820). IEEE.
15. Mall, P. K., & Singh, P. K. (2022). BoostNet: a method to enhance the performance of deep learning model on musculoskeletal radiographs X-ray images. International Journal of System Assurance Engineering and Management, 13(1), 658-672.
16. Sahu, S., Singh, A. K., Ghrera, S. P., & Elhoseny, M. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. Optics & Laser Technology, 110, 87-98.
17. Voronin, V., Zelensky, A., & Agaian, S. (2020). 3-D block-rooting scheme with application to medical image enhancement. IEEE Access, 9, 3880-3893.