International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep learning-Assisted Glaucoma Diagnosis and Model Design

J.Surendiran[1], M. Meena[2]

Submitted: 25/10/2022 **Revised:** 20/12/2022 **Accepted**: 03/01/2023

Abstract:

Purpose:

To evaluate the key factors used for classification and methods to improve the glaucomatous optic neuropathy (GON) in the form of systematic study along with fundus pictures as constraint sample be detected.

Methods:

Retrospectively, 940 fundus photographs from Eye hospital were gathered. The clinical and demographic details were noted together with the constitutional and functioning metrics of the pictures with GON. Convectional neural networks (CNNs) were built using transfer learning based on VGGNet to recognize GON. When CNN classifier had least rating the concluding classification will be construct with the cup-to-disc would be created by a extractor named support vector Machine. Scores were used to prevent missing instances with advanced GON. The TVGH dataset was used to create the CNN classifier, which was subsequently improved by fusing the training pictures from the TVGH and Drishti-GS datasets. CNN's primary characteristics for classification were identified using the class activation map (CAM). Classifier's performance were evaluated with Area under receiver operating characteristic curve (AUC) and the diagnostic accuracy will enable to ccomparing the ensemble model.

Results:

While the comparison to ensemble model's accuracy rate of 92.8 percent, the CNN classifier's accuracy, sensitivity, and specificity were 95.0 percent, 95.7 percent, and 94.2 percent, respectively, in 187 EH test images, respectively, in 187 EH test pictures, and the AUC was 0.992 as opposed to the ensemble model's accuracy rate of 92.8 percent. The accuracy of the CNN, the fine-tuned CNN, and the ensemble model for the Drishti-GS test pictures was 33.3 percent, 80.3 percent [33], and 80.3 percent, respectively. Neither moderately nor severely ill photos were incorrectly classified using the CNN classifier. Class-discriminate areas discovered by CAM co-localized with established GON traits.

Conclusions:

When enormous image datasets are not readily accessible for developing evident learning model towards glaucoma diagnosis, The combo model or a personalized CNN classifier might be as practicable designs.

Keywords: *Glaucoma, GON, CNN, enormous, TVGH, classifier, diagnosis, characteristic*

1. Introduction

Glaucoma is the kind of factor which contributing the permanent blindness. The prevention of glaucoma blindness depends on early diagnosis and care. Yet, since most kinds of glaucoma are asymptomatic, early identification is challenging. Even in wealthy nations the rate for glaucoma is less than 50% under worldwide diagnosis system. Even though population-based screening may detect glaucoma in its asymptomatic stage, it is presently not practical owing to its poor cost-efficiency [6]. However, tele glaucoma and robotic glaucoma detection combined may make it possible to conduct widespread glaucoma screenings for those at high risk [7].

Fundus picture was mostly resting on the eyecup-to-eyedisc ratio (CDR), RNFL or parapapillary atrophy anomalies, in certain circumstances to spotting out the automated glaucoma before the expansion of cavernous learning, Disc haemorrhage or inadequate accuracy were the important diagnostic information were over looked by the CDR-based technique in the state of in corrected disc size. Convolutional neural networks (CNN) and their variants, which have been used in computer vision applications recently, have improved medical image analysis via their increased ability to extract either high-level, refined features or low-level, coarse visual data [14]. The CNN-based algorithm successfully distinguished between the healthy fundus and the glaucomatous fundus with a high diagnostic rate [15–17], but its generalization is still unknown because the majority of studies only confirmed the algorithm's effectiveness in images from the same dataset [15, 16].

An elaborate analytic procedure of glaucoma is highly demanding for analyzing glaucoma with the construction of large Fundus image datasets, which involves both structural and functional evaluation [18]. However, training pictures

1 HKBK CE, Bangalore – 560045, INDIA
ORCID ID: 0000-0001-5502-0920
2 VISTAS, Chennai – 600117, INDIA
ORCID ID: 0000-0001-8315-2728
** Corresponding Author Email: surenjaya1981@gmail.com*

count would impact the diagnostic performance and extending learning principle model. The mannered dataset pictures and generic diagnostic traits like larger CDR, to the model were helps to intensify the relevancy for the CNN classification of Glaucoma diagnosis with attainable learning process.

A few training images were construct the performance of a CNN algorithm were explained in the present study and also it handle with the benefits of using CDR estimation which were considered to classified CNN algorithm's to assessed formation of images from various resources as a detection process of glaucomatous change and as a key feature to extracted as an ensemble model.

2. Materials And Methods

2.1. Database

The present research included 465 eyes without glaucoma and 469 human eyes with (POAG) primary open angle glaucoma, with fundus pictures in JPEG format in both cases. These images were obtained from the ophthalmology department's photograph database at Taipei Veterans General Hospital in the past (TVGH). POAG patients' fundus images were taken during frequent follow-up visits, as well as patients who attended our department for ocular screening or concerns unrelated to optic disc and retina abnormalities. The TVGH Institutional Review Board authorised this research, which adhered to the Helsinki Declaration principles. Because the research was retrospective in nature and employed anonymized clinical data and images for analysis, formal informed permission was not required.

Disc changes, RNFL loss, recurring "visual field (VF) abnormalities, and an ordinary open-point on gonioscopy diagnosed POAG of afflicted eyes. The Intraoscular pressure (IOP) in normal eyes was less than 21 mmHg, and the discs seemed to be devoid of RNFL anomalies. Three fundus cameras were used to capture the images: There are three kinds of digital retinal cameras: the Canon CX-1 hybrid mydriatic/non-mydriatic digital retinal camera, the Canon CR-DGi NM fundus camera, and the Canon CR-2 PLUS AF non-mydriatic digital retinal camera. Each of the highlighted graphics made the disc easily visible. Two glaucoma experts (CL and YK) reviewed the photos to locate the vertical CDR and verify the diagnosis. Prior to further examination, all pictures were conventionally anonymized.

Training set and a test set have separated photograph, each comprising 763 images and 181 images were used to evaluate and as a model training by using stratified sampling. Table 1 contains clinical information on the photos. The TVGH-CNN CNN model was constructed using training pictures from the TVGH dataset. Additionally, pictures from the Drishti-GS Dataset [21], Training and test data sets are considered as an open-access dataset and to diagnostic our model performance it was utilized with 50 and 51 images,

Table 1. Clinical and demographic characteristics of the image dataset from Taipei Veterans General Hospital.

	Primary Open Angle Glaucoma	Normal
Number of eye images	469	464
Age (x)	58.7±13.82	61.2±194
Gender (Female in %)	168 (35.07%)	227 (48.30%)
(CDR)	0.801±0.11	0.41±0.21
Deviation of Mean in (dB)	6.88 ± 6.02	
Thickness of RNFL in Avg (μm)	71.74 ± 10.33	

3 Proposed Detection System

Fig. 1 depicts the process of our detection system. The fundus pictures underwent pre-processing, and an extracted region of interest (ROI) with a 256 256 pixel area centred on the disc. Images with improperly extracted ROI were immediately removed. By the usage of CNN model, the categorization process was carried out. Each picture was identified as normal or glaucoma in the final layer of this CNN model using a probability value that served as the CNN classifier's confidence score. If the confidence score fell below a particular threshold, The superior and inferior neuroretinal rim have an increased vertical CDR and the classification was based on structuralism features, Support vector machine (SVM) were forwarded to booster up with the picture as a classifier and the ensemble model helps to find the people who had increased CDR will avoid missing people with mild to severe conditions.

3.1 **Image Pre-processing and ROI Ex**traction

The brightness and clarity of each picture were adjusted in the first pre-processing of the fundus images. The anatomy of the optic disc and the blood vessels were enhanced using the un-sharp masking by Gaussian blur approach. Using the idea of guided backpropagation, ROI recognition focused at the optic disc was put into practise to eliminate duplicate topographies [22]. Images that had the optic disc in the centre of the picture after Region Of Interest extraction were deemed to be qualifying images. Histogram were distributed on the green and red channel, invalid instances with improper ROI were automatically removed. Images that had the disc centred at the ROI would have a spire in the middle and was used in the study that followed in (Figure 2). Last but not least, through bear on rotation, transferring, sheering, and rising using the Keras package were enhanced the dataset. (https://keras.io/).

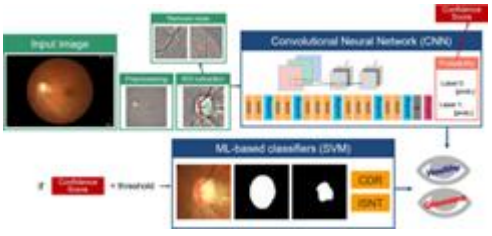

Fig: 1 Overall Framework Before submitting the vessels for ROI extraction, the original color pictures were contrasted. Images with noise the CNN model were eliminate the incorrect ROI selection. ML-based classifier helps to score the classification of Image and also forwarded. , that used the ISNT rules of the inferior and superior neuro-retinal rims were thin and CDR value as important characteristics. CNN stands for "convolutional neural network," while ML stands for "machine learning,"

ROI stands for "region of interest," and SVM stands for "support vector machine."

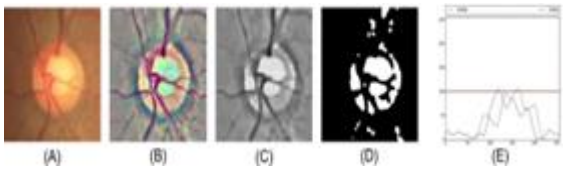

Fig: 2 Noisy conditions with poor ROI selection, (A) colour fundus image; (B) a preprocessed fundus image; (C) further preprocessing on the red and green channels; (D) binary presentation of the image; and (E) a histogram of the projection on x- and y-axes, displaying the best ROI choice with the maxima on both axes in the middle. Region of Interest, or ROI.

3.2 CNN's Detection

Preprocessed and qualifying pictures, ReLU, pooling, FC, and a softmax loss layer were used to train our CNN model. Cross-entropy and the softmax layer were combined to give an estimate of how likely it is that each image can be put into a category at the end of processing. The stochastic gradient descent was also employed to reduce he dropping of function during optimization.

In Oxford University, The Visual Geometry Group Transfer was framed VGGNet and pre-trained CNN model for accomplishing the learning procedure.. Retraining the CNN model created using the VGG Net's learned weights entails deleting top layers FC layers and replacing them with a sequence of convectional layers with a 3 x 3 filter size the FC layers and replacing them (Fig. 3). To avoid over fitting and to promoted the convergence by the data augmentation and drop out layers in Regularization techniques.

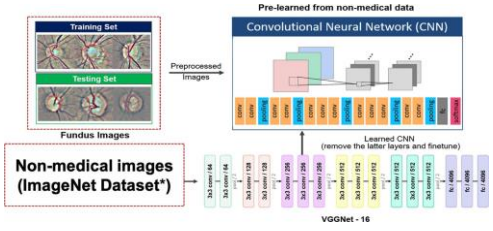

Fig: 3 Shows the VGGNet architecture and transfer learning. The CNN model was started using the learned weight of VGGNet with a 3 x 3 transfer learning approach. Three-size filter Convolutional layers, pooling layers, and the fc layer were all present in the final model. VGGNet, created by the Visual Geometry Group at the University of Oxford; CNN; conv; and fc; completely linked. ImageNet large-scale visual recognition challenge, Russakovsky O, et al. IJCV. 115, 211– 252 (2015). (2015).

3.3 Cross-validation of the CNN model

To assess the effectiveness and prevent the CNN model from being over fit, we used 5-fold cross validation. By using an 80\20 split as a set in each folds of an individual training and validation, the merged photos were also divided. Both set of test photos and the validation sets were used to assess each fold's performance. Unless there was a considerable difference in the performance of each fold and final model were chosen as a best performance to fold as a parameters.

3.4 Visualization of Class activation map (CAM)

"CAM was used to find the properties recognised by the proposed CNN model came under the division of glaucomatous and normal discs" [26]. Using "global average pooling (GAP) on the convolutional feature maps and an FC layer that provide output in right manner" (Fig. 4), [27] the weights of the output layer were projected back on the convolutional feature maps.

3.5 Making use of an SVM classifier with a larger CDR

The SVM classifier was trained using features of the glaucomatous disc, such as an increased CDR and thinning when the neurorational rim were in the form of the superior or inferior. The green and red channels were added to the previously processed images first. The red channel which described the disc margin, while the cup area was calculated by the green channel. The higher ranking and second class junctional points of the disc border were moved along with the vertical meridian were utilized to calculate the diameter of the vertical disc. K-means clustering was used to figure out the area of the cup after the duplicated data were taken out by subtracting the mean and standard deviation.

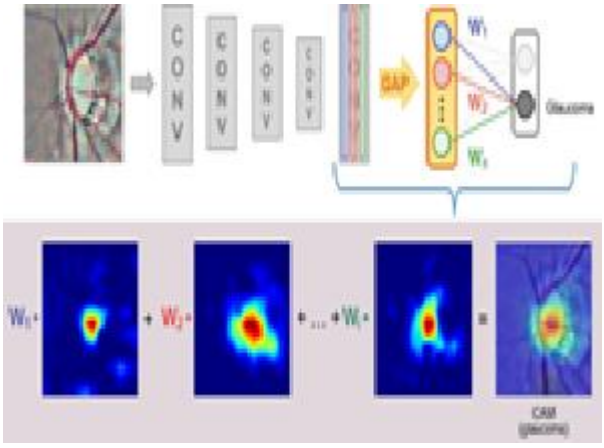

Fig: 4 shows the class activation map that is connected to the top layer of the suggested CNN model. CAM, which was created using GAP and the multiplication of each channel by the specified weight from the fully connected (fc) layer, highlighted the key locations for the final categorization (W1, W2. . .Wn). CNN is for convolutional neural network; CONV stands for convolutional; and GAP stands for global average pooling.

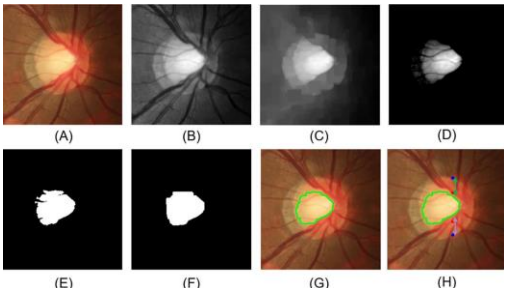

Fig: 5 shows the disc margin, cup area, vertical cup-to-disc ratio, and superior and inferior neuroretinal rim-to-disc ratios. (A) Color fundus image after ROI extraction; (B) green and red channel processing; (C) vasculature removal; (D) mean and standard deviation removal; K-means clustering and morphological noise reduction Cup border (green line), superior and inferior rims, vertical disc diameter (blue points), and cup diameter (red points) (purple arrow). ROI: Region of Interest.

The processed picture was divided into two halves and produced through green channels using a K-value of 2 [28]. The width of superior and the inferior neurorational rim was calculated the division of rim-to-disc and the diameter of the vertical disc were calculated by the CDR with the length of the vertical cup. The three characteristics were supplied to the RBF kernel-based SVM classifier in order for it to categorise them. Grid searches and the Scikit-Learn SVM model were used for prediction [29].

3.6 Statistical analysis

The final count of the test images were divided the categorization of healthy and glaucomatous, sensitivity, specificity, and accuracy were guided and also utilize to evaluate the performance of our model. To learn the module Python sick were used for presenting The ROC curve and area under the RPC curve (AUC), false positive rate were represented by the X-axis and the true Positive rate were represented by the Y-axis.Because the bulk of the measurements from the test images was not normally distributed, the proper and erroneous predictions between the images were compared and the clinical characteristics of these things were offered by the non-parametric tests. The data and the Mann-Whitney U test were categories as for continuous variables and as a data with the Chi-Square test. SPSS was used to conduct the analysis.

4. Result

4.1 Analytic performance of the CNN model

The diagnostic accuracy, sensitivity and specificity were categorize with 95.0%,95.7% and 94.2% in the total set of 187 with the Utilization of TVGH-CNN model with the test photographs from the TVGH data set and it also detecting the Glaucomatous eyes with 0.992 which was described in the table 2 and the figure 6. 51 test images from the Drishti-Gs data set were classified worse job as a model of TVGH-CNN with 33% right forms,it indicate the model was not as good in generalization.so CNN-TVGH made two plans to solve the restrict issue :

- By using training images from the data set of Drishti-GS to optimize the TVGH-CNN as primarily.
- Second, a collective model that used CDR for classification was applied when the CNN classifier's confidence score was low.

The test -pictures in the “Drishti-GS” were under the classification of CNN model would improve when the 50 training pictures of dataset in the Drishti-GS were mutually added with the data set of TVGH training method which helps to improve the fine-tuned CNN model “(T&D-CNN MODEL)”.accuracy, sensitivity and specificity to diagnostic the Drishti-GS under the T&D-CNN model’s were enhanced with 80.3%,73.6%and 100% and finally an AUC of 0.937 .From the Table 2 and the Figure 7 we can analyze the datasets of Drishti-GS into TVGH which integrate and determine the optimum images which were restricted the training images under the T&D-CNN MODEL we can also evaluate the training image preprocessing ,enlargement and ROI extraction and its function.The results revealed that using ROI increased the model's performance. As shown in Table 3, Drishti-GS suggests that preprocessing and augmentation dataset performance were improved with few images.

4.2 The diagnostic performance of the ensemble model

In order to prevent leaving out individuals who were moderately to severely ill, an ensemble model for categorisation was also created. This model used a CDR-based SVM classifier when the CNN classifier's likelihood dropped below 0.85. The CNN classifier's twofold arrangement likelihood is shown by the likelihood value. A variety of threshold values, starting at 0.5 and increasing by 0.05 increments up to 1.0, were evaluated to determine the threshold value by measuring the sensitivity and specificity of the ensemble model.

On test photos from the “**TVGH and Drishti-GS datasets**”, the collaborative model had the maximum sensitivity and specificity at a threshold value equal to 0.849. The SVM classifier was trained using the “TVGH dataset”. and the diagnosis precision for the test pictures from TVGH and Drishti-GS was 76.7 percent and 72.5 percent, respectively. by employing the collaborative model.

Table 2. Diagnostic capability of the CNN model, SVM model, collaborative model and fine-tuned CNN model in identifying glaucomatous disc.

	Dataset of test images							
	TV				Drishti-GS			
Models	Accuracy	Specificity	Sensitivity	AUC (95% CI)	Accuracy	Specificity	Sensitivity	AUC (95% CI)
TVGH-CNN model[a]	95.1%	94.1%	95.6%	0.991 (0.984–1.0)	33.2%	100%	10.0%	0.763 (0.632–0.894)

SVM model [a]	76.8 %	78.2 %	75.6 %	0.807 (0.740–0.872)	72.4 %	76.8 %	71.1 %	0.782 (0.617–0.950)
Collaborative model	92.7 %	95.5 %	90.5 %		80.2 %	92.4 %	76.4 %	
Fine-tuned T&D-CNN model [b]	93.8 %	95.5 %	92.5 %	0.985 (0.972–0.998)	80.2 %	100 %	73.5 %	0.936 (0.874–1.0)

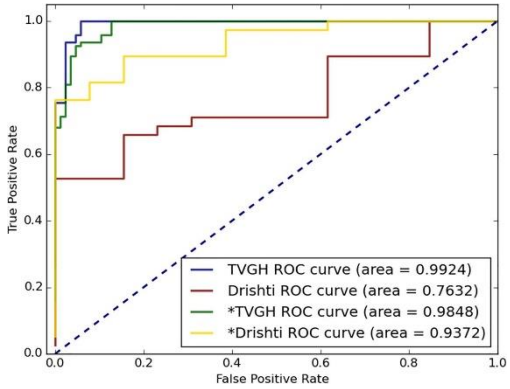

Fig: 6 CNN models' ROC curves

The CNN models' ROC curves for differentiating between a healthy disc and a disc with glaucoma are shown in Figure 6. The output of the TVGH-CNN model, which was built using TVGH-training pictures to calculate TVGH test pictures, was shown by the TVGH-ROC curve. The outcomes displayed by the CNN model created using the Drishti-GS training pictures to anticipate the Drishti-GS test images were supplied by the Drishti-ROC curve. To predict the "TVGH and Drishti test images", It utilized the combined TVGH and Drishti training datasets. Convolutional neural networks are referred to as CNNs, and ROC stands for receiver operating characteristic. Improved diagnosis accuracy of the Drishti GS test pictures but comparable diagnostic precision to the updated T&D - CNN model. However, nothing changed for the better.

4.3 Factors influencing the CNN classifier's performance

Identify the likely factors that may lead to incorrect categorization using our system, we examined the clinical characteristics for the proof graphs with accurate and incorrect predictions. (Table 4).

The glaucoma photographs that were properly diagnosed as having the disease had a higher CDR and a worse visual field defect than the photos that were incorrectly labelled. The deviation of all GON photographs that were wrongly categorised and was less than -6 dB, as was the case with all of the pictures. The CDR values of images of healthy eyes misclassified as having glaucoma were higher than those of properly labelled photo figure 7 shows example CAMs with superimposed images that distinguish between patterns in healthy and glaucomatous eyes. The parapapillary atrophy region and cup in the affected eyes (A, B) and the nerve fibre bundles retinal rim in the healthy eyes were where the hot spots (red colour) were decolour (C, D). The overlay images with the disc as the CAM's focal point allowed us to discern between the CAM patterns in glaucomatous (E) and healthy eyes (F). CAMs are class activation mappings.

Table 3. Impact of image processing on the performance of the proposed CNN model in identifying glaucomatous discs using a mixed training dataset including training images of Drishti-GS and TVGH datasets.

	Dataset of test images					
	TVGH			Drishti-GS		
MTVds included in image processing	Specificity	Sensitivity	AUC (95% CI)	Specificity	Sensitivity	AUC (95% CI)
All [a]	95.4 %	92.6 %	0.984 (0.972–0.998)	100 %	73.6%	0.937 (0.874–1.0)
All except ROI selection	93.4 %	89.4 %	0.978 (0.960–0.996)	100 %	55.2%	0.887 (0.789–0.984)
All except preprocessing	97.7 %	94.6 %	0.993 (0.985–1.0)	84.6 %	84.2%	0.897 (0.794–0.999)
All except augmentation	95.4 %	96.8 %	0.990 (0.997–1.0)	92.3 %	68.4%	0.925 (0.849–1.0)

We performed pre-processing, augmentation, and ROI selection. The augmentation was done using the 55 training images from the "Drishti-GS dataset", with a final image count of 30.000 after rotating, shifting, sheering, and zooming. Taipei Veterans General Hospital is referred to as TVGH. Convolutional nTVl neA convolution red to as CNN. Area under the curve The area referred to as AUC.

Table 4. Clinical characteristics for the test images and relationship with the prediction of the deep learning algorithm.

	POAG			Control		
	CNN prediction			CNN prediction		
	Correc	Incorre	P-	Correct	Incorre	P-

	t N = 87	ct N = 7	value[a]	N = 83	ct N = 4	value[a]
Age (years)	58.85± 15.71	62.00±2 .52	0.389	62.26± 19.55	51.75± 23.26	0.297
Sex (Female, %)	33 (37.9%)	1 (14.28%)	0.210	49 (59.03 %)	4 (100%)	0.101
CDR	0.80±0 .11	0.67±0. 21	0.037	0.41±0. 14	0.57±0. 10	0.018
Visual field						
Mean deviation (dB)	- 6.49±5 .47	- 2.16±2. 27	0.015			
PSD (dB)	6.28±3 .72	2.97±1. 51	0.011			
Average RNFL thickness (μm)	72.79± 11.06	80.71±1 0.92	0.750			

4.4 Visualization Of Predictive Features

CAM was used to reveal the attributes that were used by our CNN model to categorise the depict categorize areas were highlighted in blue, whilst critical regions for classification were marked in red. Different CAMs were found in glaucomatous and healthy eyes. The hot spots were more commonly seen in the parapapillary atrophy and cup areas in the images of the glaucomatous eyes than they were in the retinal nerve bundles or at the neuroretinal boundary in the images of the healthy eyes. The overlapping images revealed the distinction between the two groups (Fig 7).

5. Discussion

According to this study, test pictures obtained by a desktop fundus camera using colour fundus and CNN-based automated glaucoma diagnosis had high accuracy when compared to the calibre of the dataset. How generalizable the classification model was, however, depended on the depending of the training data. A limited number of images from the desired test population were added to the training dataset, considerably enhancing the CNN detection model's accuracy. With this modification, CNN fared equally well or better in terms of diagnostic performance than the ensemble model that included CDR. CAM demonstrated that our CNN model accurately caught features of the optic disc that enable individuals to perceive glaucomatous changes.

For deep learning to successfully categorise reties categorize, the quantity and quality of the training data are essential [19, 20]. Contrary to the diagnosis and grading of diabetic retinopathy, which only uses fundus images, the diagnosis of glaucoma necessitates a range of morphological and functional tests [18]. It is extremely difficult to gather tens of thousands of glaucomatous fundus images with a precise diagnosis and thorough profiles of functional and anatomical factors. This issue was overcome in some studies aiming to build CNN classifier a using a large number of fundus images by separating referable GON but not definite GO definite findings of the fundus photos and agreement between clinicians [16, 30].

This method has a basic issue in that it is difficult to assess how well any one method works as a diagnostic tool for photos with different degrees of disease intensity. According to a research by Cristopher et al. [15], the performance of the deep learning system changed depending on how severe the GON was. Their research confirmed our theory, showing that glaucomatous eyes at an earlier stage were more likely to be mistaken for normal. Deep learning algorithms may therefore perform differently in terms of identifying glaucoma and are more likely to miss early GON depending on the severity of the condition. [15]. Due to the fact because model was developed using a limited number of fundus images, its generalizability was low. The lack of diversity in our database may be the cause of the poor generalizability of our model; however, it is unclear whether using more images from diverse origins would improve the generalizability of a deep learning model because the CNN model's validation on other databases has not been thoroughly investigated [15, 16]. Phene et al. showed that a deep learning approach with 58033 fundus images as its training data had variable diagnostic accuracy in recognising GON on two validation datasets with an AUC of 0.940 and 0.858 [17].

Liu et al. [30] used 241032 images from 68013 people to train a CNN classifier to recognise glaucoma. On images of people of different ethnicities and of differing ethnicities CNN classifier still continued to perform badly an AUC decreasing from 0.996 with internal validation set to 0.923 and 0.823, respectively. Even with the largest picture collection available to date, this is true. The performance of a deep learning model in diagnosing diabetic retinopathy peaked at about 60000 images, according to Gulshan et, al's research [19]. Therefore, in order to increase the generalizability on classifier for glaucoma diagnosis, strategies other than adding more pictures to the training dataset should also be considered.

We recommended the following two methods to broaden the applicability of our model: In order to improve this mixed model's diagnostic performance for photos of varied origins, the CNN model was initially fine-tuned using a small sample of images from the target test dataset. When there is no accessible large-scale glaucoma image data sets or a common deep learning algorithm for glaucoma diagnosis, this strategy, which is not innovative but beneficial, tries to construct clinic-based glaucoma screening algorithms utilizing a small number of pictures from each clinic.

The study by Diaz-Pinto et al. [31] produced similar findings. It was based on diagnostic performance of the varied models depending images which were under the training, the test dataset was used to train the model, with an AUC ranging from 0.9605 to 0.7678, respectively. This was done in order to test the generalizability of their CNN model using five publicly available datasets.

The ensemble model was included in the second choice that included the CNN classifier's confidence score was low based the CDR.Raising GON is the primary intention of the ensemble model was to use to prevent missing occurrences. However, the results demonstrated that both the ensemble

model and the tweaked CNN model functioned diagnostically. Therefore, clear diagnoses of GON is not available when the large image dataset is applicable, at first hundreds of high-quality images were built with CNN model and then it could be fine-tuned using some images from the target test population to increase the model's diagnostic power in the target test population.

The CAM demonstrated that the class-discriminatory areas of the CNN model co-localized with the well-known GON characteristics. While photographs of healthy eyes exhibited hot spots close to the neuroretinal rim and nerve bundles, the majority of photos that were correctly diagnosed as glaucoma images had hot areas of the CAM in the cup region. This finding supported Christopher et al assertion that the superior and inferior portions of the neuroretinal rim are the critical elements for employing a deep learning system to identify between eyes with GON and healthy eyes [15]. These findings helped the CNN model classify glaucoma more accurately and may allow for additional advancements.

There are certain limitations to our research. Since the training dataset was only gathered at one medical institution with a limited number of images, the generalizability of our model may be in question. On test images from various sources, our model performed badly as expected, however this may be corrected by using the revised model. More research is needed to determine if deep learning algorithms for glaucoma detection, particularly for population-based glaucoma screening, are generalizable because the fundus images utilised for glaucoma diagnosis may differ in layout, quality, and visibility of the RNFL. The use of a ROI centred at the ONH to develop a model to reduce the size and complexity of the pictures is the second limitation of our study and may result in information loss and difficult pre-processing for clinical application.

The photos in table 3 were the comparison of sensitivity and specificity in the model which increasing stage of ROI extraction. Other deep learning technique did not examine the performance on algorithms based on the addition to VGGNet , the TVGH-CNN model's performance and generalizability did not significantly differ when utilising things like “AlexNet, GoogleNet, or Xception (data not shown)” based on CNN algorithms, There is just a little difference between VGGNet, Inception, and ResNet50 in distinguishing GON, according to Christopher et alresearch, which validates our findings, with similar AUC values of 0.89, 0.91, and 0.91 [15].

Last but not least, our model might not be relevant to instances with these cases of features with severe or abnormal disc appearances and cataract other than drusen were not included in the dataset. No measurements, counts, or considerations of disc size, disc haemorrhage, or peripapillary atrophy were made in this investigation, therefore it is uncertain how these factors affected the CNN classifier's performance.

6. Conclusions

The small number of fundus images were up skill through the proposed extensive learning model, correctly diagnosed all instances of mild to moderate disease and distinguished GON with high accuracy on photos of comparable quality. By carefully adjusting the model to raise its diagnostic accuracy for images from multiple sources, it could be able to develop glaucoma screening based on clinic with the CNN classifier. Through the human perception the identification of GON were framed while similar properties was caught accurately with the model CAM demonstration. Glaucoma screening is aided with the practical method which were detected by the deep learning -aided GON based on the mentioned findings, but a learning model's generalizability has to be established before it can be used to the intended test people.

Acknowledgements

This research was supported/partially supported by [Name of Foundation, Grant maker, Donor]. We thank our colleagues from [Name of the supporting institution] who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper. We thank [Name Surname, title] for assistance with [particular technique, methodology], and [Name Surname, position, institution name] for comments that greatly improved the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

S. S S, J. Surendiran, N. Yuvaraj, M. Ramkumar, C. N. Ravi and R. G. Vidhya, "Classification of Diabetes using Multilayer Perceptron," 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), 2022, pp. 1-5, doi: 10.1109/ICDCECE53908.2022.9793085.

Ernest PJ, Schouten JS, Beckers HJ, Hendrikse F, Prins MH, Webers CA. An evidence-based review of prognostic factors for glaucomatous visual field progression. Ophthalmology. 2013; 120(3):512–19. https://doi.org/10.1016/j.ophtha.2012.09.005 PMID: 23211636

Chen PP. Blindness in patients with treated open-angle glaucoma. Ophthalmology. 2003; 110(4):726– 733. https://doi.org/10.1016/S0161-6420(02)01974-7 PMID: 12689894

Surendiran, J., Theetchenya, S., Benson Mansingh, P. M., Sekar, G., Dhipa, M., Yuvaraj, N., & Alene, A. (2022). Segmentation of Optic Disc and Cup Using Modified Recurrent Neural Network. BioMed Research International, 2022.

Mitchell P, Smith W, Attebo K, Healey PR. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. Ophthalmology 1996; 103(10):1661–9. https://doi.org/10.1016/s0161-6420(96) 30449-1 PMID: 8874440

Burr JM, Mowatt G, Hernández R, Siddiqui MA, Cook J, Lourenco T, et al. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evalu- ation. Health Technol Assess. 2007; 11(41): iii-iv, ix-x, 1–190. https://doi.org/10.3310/hta11410 PMID: 17927922

Thomas SM, Jeyaraman MM, Hodge WG, Hutnik C, Costella J, Malvankar-Mehta MS. The effective- ness of teleglaucoma versus in-patient examination for glaucoma screening: a systematic review and meta-analysis. PLoS One. 2014; 9(12):e113779. https://doi.org/10.1371/journal.pone.0113779 PMID: 25479593

Hertzog LH, Albrecht KG, LaBree L, Lee PP. Glaucoma care and conformance with preferred practice patterns. Examination of the private, community-based ophthalmologist. Ophthalmology. 1996; 103 (7):1009–13. https://doi.org/10.1016/s0161-6420(96)30573-3 PMID: 8684788

D.Saravanan, J.Surendiran,"A new framework for video data retrieval using hierarchical clustering technique", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-6S3, September 2019.

Badalà F, Nouri-Mahdavi K, Raoof DA, Leeprechanon N, Law SK, Caprioli J. Optic disk and nerve fiberlayer imaging to detect glaucoma. Am J Ophthalmol. 2007; 144(5):724–32. https://doi.org/10.1016/j. ajo.2007.07.010 PMID: 17868631

O'Neill EC, Gurria LU, Pandav SS, Kong YX, Brennan JF, Xie J, et al. Glaucomatous optic neuropathy evaluation project: factors associated with underestimation of glaucoma likelihood. JAMA Ophthalmol. 2014; 132(5):560–6. https://doi.org/10.1001/jamaophthalmol.2014.96 PMID: 24699817

J.Surendiran, Dr.S.V.Saravanan, F.Elizabeth Catherine,"GLAUCOMA DETECTION USING FUZZY C- MEAN (FCM)", IJPT| Sep-2016 | Vol. 8 | Issue No.3 | 16149-16163

Garway-Heath DF, Ruben ST, Viswanathan A, Hitchings RA. Vertical cup/disc ratio in relation to opticdisc size: its value in the assessment of the glaucoma suspect. Br J Ophthalmol. 1998; 82(10):1118– 24. https://doi.org/10.1136/bjo.82.10.1118 PMID: 9924296.

Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural net-works. ArXiv: 1512.07108v6. 2017 [cited 2017 Oct 19]. Available from: https://arxiv.org/abs/1512.07108.

J.Surendiran, Dr.S.V.Saravanan, K.KManivannan,"Detection of Glaucoma Based on Color Moments and SVM Classifier Using K Mean Clustering", IJPT| Sep-2016 | Vol. 8 | Issue No.3 | 16139-16148.

Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glauco- matous optic neuropathy based on color fundus photographs. Ophthalmology. 2018; 125(8):1199– 1206. https://doi.org/10.1016/j.ophtha.2018.01.023 PMID: 29506863

Phene S, Carter Dunn R, Hammel N, Liu Y, Krause J, Kitade N, et al. Deep learning to assess glaucomarisk and associated features in fundus images. ArXiv: 1812.08911v1. 2018 [cited 2018 Dec 21]. Avail- able from: https://arxiv.org/abs/1812.08911.

Jonas JB, Aung T, Bourne RR, Bron AM, Ritch R, Panda-Jonas S. Glaucoma. Lancet. 2017; 390(10108):2183–2193. https://doi.org/10.1016/S0140-6736(17)31469-1 PMID: 28577860

Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and valida-tion of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016; 316(22):2402–2410. https://doi.org/10.1001/jama.2016.17216 PMID: 27898976

Krause J, Gulshan V, Rahimy E, Karth P, Widner K, Corrado GS, et al. Grader variability and the impor-tance of reference standards for evaluating machine learning models for diabetic retinopathy. Ophthal- mology. 2018; 125(8):1264–1272. https://doi.org/10.1016/j.ophtha.2018.01.034 PMID: 29548646

Sivaswamy J, Krishnadas S, Chakravarty A, Joshi GD, Ujjwal, Syed TA. A comprehensive retinal imagedataset for the assessment of glaucoma from the optic nerve head analysis. JSM Biomed Imaging Data Pap. 2015, 2(1):1004.

Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M. Striving for simplicity: the all convolutional net. ArXiv: 1412.6806. 2014 [cited 2014 Dec 21]. Available from: https://arxiv.org/abs/1412.6806.

Simonyan K, Zisserman A. Very Deep Convolutional networks for large-scale image recognition. ArXiv: 1409.1556. 2014 [cited 2014 Sep 4]. Available from: https://arxiv.org/abs/1409.1556.

Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? ArXiv: 1411.1792v1. 2014 [cited 2014 Nov 6]. Available from: https://arxiv.org/abs/1411.1792v1

Orlando JI, Prokofyeva E, Fresno MD, Blaschko MB. Convolutional neural network transfer for auto- mated glaucoma identification. Proceedings of the 12th International Symposium on Medical Informa- tion Processing and Analysis. 2016 Dec 5–7; Tandil Argentina; SPIE; 2016. Vol. 10160; pp. 1–10. https://doi.org/10.1117/12.2255740.

Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localiza- tion. ArXiv: 1512.04150. 2015 [cited 2015 Dec 14]. Available from: https://arxiv.org/abs/1512.04150.

Lin M, Chen Q, Yan S. Network in network. ArXiv: 1312.4400. 2013 [cited 2013 Dec 16]. Available from: https://arxiv.org/abs/1312.4400.

Ayub J, Ahmad J, Muhammad J, Aziz L, Ayub S, Akram U, et al. Glaucoma detection through optic disc and cup segmentation using K-mean clustering. Proceedings of 2016 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube).2016 April 11–12; Quetta Pakistan; IEEE; 2016. https://doi.org/10.1109/ICECUBE.2016.7495212.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. JMLR. 2011; 12:825–2830. http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

Liu H, Li L, Wormstone IM, Qiao C, Zhang C, Liu P, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. JAMA Ophthalmol. 2019; 137(12):1353–

1360. https://doi.org/ 10.1001/jamaophthalmol.2019.3501

Diaz-Pinto A, Morales S, Naranjo V, Kohler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomed Eng Online. 2019; 18(1):29. https://doi.org/10.1186/s12938-019-0649-y PMID: 30894178

Ko, Yu-Chieh, et al. "Deep Learning Assisted Detection of Glaucomatous Optic Neuropathy and Potential Designs for a Generalizable Model." Deep Learning Assisted Detection of Glaucomatous Optic Neuropathy and Potential Designs for a Generalizable Model | PLOS ONE, journals.plos.org, 14 May 2020, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0233079.