Glaucoma classification using deep convolutional neural network architectures

Data Science Project Proposal

March 2023

Table of Contents

[LIST OF FIGURES 3](#_Toc129961309)

[LIST OF ABBREVIATIONS AND ACRONYMS 3](#_Toc129961310)

[LIST OF APPENDICES 3](#_Toc129961311)

[ABSTRACT 4](#_Toc129961312)

[INTRODUCTION 4](#_Toc129961313)

[Project aim 5](#_Toc129961314)

[Stakeholders 5](#_Toc129961315)

[related works 5](#_Toc129961316)

[Shortcomings and gaps in previous work 6](#_Toc129961317)

[Not considering glaucoma variations 6](#_Toc129961318)

[Lack of diversity in data 6](#_Toc129961319)

[Data 6](#_Toc129961320)

[Methods 6](#_Toc129961321)

[Pre-processing 7](#_Toc129961322)

[Data split 7](#_Toc129961323)

[Proposed CNN models with Transfer Learning 7](#_Toc129961324)

[VGGNet19: 7](#_Toc129961325)

[ResNet-50: 7](#_Toc129961326)

[GoogLeNet (InceptionV3): 7](#_Toc129961327)

[Xception (extreme inception): 8](#_Toc129961328)

[Ensemble learning: 8](#_Toc129961329)

[Transfer learning: 8](#_Toc129961330)

[Evaluation and performance 8](#_Toc129961331)

[Work plan 8](#_Toc129961332)

[Risk assessment 9](#_Toc129961333)

[Biased datasets 9](#_Toc129961334)

[Computational resources 9](#_Toc129961335)

[Expected results 9](#_Toc129961336)

[Conclusions 9](#_Toc129961337)

[References 11](#_Toc129961338)

LIST OF TABLES

[Table 1: Public glaucoma datasets 7](#_Toc129965022)

LIST OF FIGURES

[Figure 1 & 2: healthy eye vs glaucomatous eye 4](https://d.docs.live.net/ba5b25d88e4fe846/Documents/DSM060/CW2/CW2_Glaucoma_Research.docx#_Toc130162003)

[Figure 2: U-Net architecture for OD and OC segmentation 7](#_Toc130162004)

[Figure 3: CNN architecture 8](#_Toc130162005)

[Figure 4: VGG-Net 19 Model Architecture 9](#_Toc130162006)

[Figure 5: ResNet50 Model Architecture 9](#_Toc130162007)

[Figure 6: Inception V3 Model Architecture 10](#_Toc130162008)

[Figure 7: Xception Model Architecture 10](#_Toc130162009)

[Figure 8: Transfer Learning Model Architecture 11](#_Toc130162010)

LIST OF ABBREVIATIONS AND ACRONYMS

|  |  |
| --- | --- |
| **Abbreviation** | **Description** |
| AI | Artificial Intelligence |
| AUC | Area Under Curve |
| CDR | Cup-to-Disc-Ratio |
| CLAHE | Contrast Limited Adaptive Histogram Equalization |
| CNN | Convolutional Neural Network |
| CV | Cross Validation |
| DL | Deep Learning |
| FCN | Fully Convolutional Network |
| IOP | Intraocular Pressure |
| LR | Logistic Regression |
| MAPNet | Multiscale Average Pooling Net |
| MD | Mean Deviation |
| ML | Machine Learning |
| OC | Optic Cup |
| OCT | Optical Tomography |
| OCT | Optical Coherence Tomography |
| OD | Optic Disc |
| ONH | Optic Nerve Head |
| PACG | Primary Angle-Closure Glaucoma |
| POAG | Primary Open-Angle Glaucoma |
| PSD | Pattern Standard Deviation |
| RNFL | Retinal Nerve Fibre Layer |
| ROC | Receiver Operating Characteristic |
| ROI | Region of Interest |
| SDOCT | Spectral-Domain Optical Coherence Tomography |
| SGD | Stochastic Gradient Descent |
| SMOTE | Synthetic Minority Over-Sampling Technique |
| VAE | Variational Autoencoder |

ABSTRACT

The recent emergence of deep learning (DL) algorithms has shifted the way in which glaucoma imagery is now analysed. Automated feature extraction using Convolutional neural networks (CNN) have become the most prominent algorithm for image analysis, surpassing manual methodologies. Normally, DL algorithms including CNN require large amounts of data for training. In the medical imaging data is scarce, this is now being alleviated with Transfer learning (TL).

INTRODUCTION

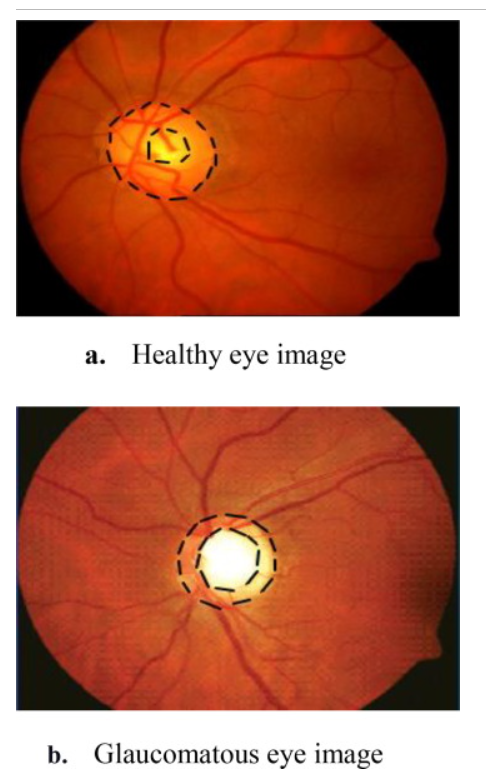
Glaucoma can lead to permanent blindness and is currently an incurable eye disease. The origin and progression of glaucoma is uncertain. There is often no indication of glaucoma in its early stages, making it difficult to diagnose.

Figure 1 & 2: healthy eye vs glaucomatous eye

Age, race, and genetics all play a significant role in glaucoma, with those over 60, a family history, and being of African descent more at risk. A thicker cornea and myopia can also lead to glaucoma. (1)

Glaucoma is either primary or secondary. Primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG) occur in the early stages of the disease. Secondary glaucoma results from trauma, inflammation, tumours, or by taking certain medications, for example corticosteroids. (2)

It is estimated that 40 – 75% of glaucoma patients have normal intraocular pressure. Normal pressure ranges from 10 – 21 mmHg. Some patients develop glaucoma damage at lower pressure levels while others have minor damage even at higher pressure levels. (3)

Glaucoma usually presents with optic disc cupping, whereby the optic disc (or cup) becomes larger. The ratio of the optic cup (OC) to the optic disc (OD) is used when diagnosing glaucoma, with a ratio of > 0.6 being significant. (4) The cup-to-disc-ratio (CDR) is used by experts to detect glaucoma, the manual process can however be time-consuming and labour intensive as well as subjective. Automated systems of segmenting the OD and OC and diagnosing glaucoma are common with deep learning (DL) algorithms. (5)

Convolutional neural network (CNN) models are now the most common techniques for the classification of fundus images. Fundus imagery are specifically for the interior of the eye, whereas retinal images capture the back of the eye. In contrast, an Optical tomography (OCT) image, is a cross-section of the retina. (6)

Project aim

Investigation of automated segmentation and classification of fundus image using several DL CNN architectures for glaucoma prediction. DL techniques save time and computational resources as well as address data scarcity.

Medical image analysis enhancement by way of Transfer learning (TL) together with ImageNet pre-trained CNN models to classify images.

Stakeholders

Individuals from the higher risk categories swill benefit from this project.

Clinicians will benefit from this project to reduce their workload and assist in the screening and diagnosing of glaucoma in patients.

Researchers and developers within the glaucoma domain will also benefit from this research.

related works

AI-Bander et al. used a method based on DenseNet and a Fully Convolutional Network (FCN). The method, a U-shaped architecture results in pixel-based classification of fundus imagery.(4)

Shoukat et al use a Contrast Limited Adaptive Histogram Equalization (CLAHE) method to enhance fundus images. (8). A median filter is applied during pre-processing as a noise reduction method. Thereafter, three pre-trained Convolutional Neural Network (CNN) architectures, VGG19, ResNet50 and EfficientNetB7 are used to classify glaucoma.

Alghamdi and Abdel-Mottaleb use three CNN models with different learning methods. (9) The models use TL methods on both labelled and unlabelled data and outperformed two experienced opthalmologists.

Latif et al use the ODGNet DL model in a two-phased approach using TL and pre-trained models for automated OD glaucoma classification of fundus images. (10)

Sallam et al. investigated early glaucoma detection using TL from pre-trained CNN models. Nine different models were used in the study and were subject to forward and backward propagation for optimal performance. (11)

Kim et al investigated the automation of OD and OC segmentation methods using DL. Their method locates the Region of Interest (ROI), adds a mask, thereafter segmenting the OC and OD from the ROI using the Multiscale Average Pooling Net (MAPNet) algorithm. (11)

Akter et al. analyzed multiple features using Logistic Regression (LR) and DL algorithms using CNN. The method uses a cross-section to separate the optic nerve head (ONH) from optical coherence tomography (OCT) images. (12)

Shortcomings and gaps in previous work

Lack of reference standard for diagnosing glaucoma

The lack of spectral-domain optical coherence tomography imagery (SDOCT) and other clinical information make replication of research more difficult. (14)

Non-inclusive features

Using only structure or function alone when training a ML classifier is not enough to detect glaucoma accurately. Clinical factors such as intraocular pressure (IOP) and corneal thickness have been shown to enhance diagnostic accuracy. (1)

Not considering glaucoma variations

Variations in cup-to-disk ratio to detect the disease severity is an area of research that is lacking in identifying glaucoma. (7)

Lack of diversity in data

DL models require large amounts of data to effectively train an algorithm, these need to be maintained and need to be replicated. Unlabelled retinal imagery would also improve the CNN models, as would a more representative sample of the Black and Hispanic populations. (16)

Data

The following public datasets are proposed to be used in this study:

|  |  |  |  |
| --- | --- | --- | --- |
| **DATASET** | **Total** | **Glaucoma** | **Normal** |
| PAPILA | 488 | 155 | 333 |
| DRIONS-DB  (Carmona et al, 2008) | 110 | No label | No label |
| DRISHTI-GS1  (Sivaswamy et al, 2014) | 101 | 70 | 31 |
| RIM-ONEV1 | 485 | 313 | 172 |
| RIM-ONEV2 |  |  |  |
| RIM-ONEV3  (17) | 159 | 74 | 85 |
| HRF  (Khohler et al, 2013) | 45 | 27 | 18 |
| ORIGA-light  (Zhang et al, 2010) | 650 | 168 | 482 |
| ACRIMA | 705 | 396 | 309 |
| REFUGE | 1200 | 120 | 1080 |
| RIGA | 750 | No label | No label |
| G1020 | 1020 | 296 | 724 |
| CHASE-DB1 – 14 children left and right eyes | 28 |  |  |

Table 1: Public glaucoma datasets

Methods

There are two phases in this project, firstly automated segmentation of the OD from fundus imagery, followed by TL-based ensemble learning with pre-trained CNN models to classify glaucoma.

Pre-processing

Fundus images will be pre-processed using CLAHE and Median Filter.

Advantages of using the CLAHE method include sharper contrast in the ROI, as well as enhanced edge definition The non-linear median filter assists with noise reduction and computes the value of the output pixel by the median neighbouring pixel.

Segmentation

OC segmentation is performed using the U-Net platform, a very common approach in medical image segmentation. Greyscale images are the result, they are cropped around the OD to segment the OD and OC regions.

Chart, box and whisker chart

Description automatically generated

Figure 2: U-Net architecture for OD and OC segmentation

Data split

Segmented images containing the ROI are split into 70 % training, 15% validation and 15% testing images, with a resolution of 512 x 512 pixels.

Data Augmentation

Data augmentation is used to create multiple variations of the images in the training dataset. This includes the use of horizontal flip, vertical flip, image rotation, image mirroring and shape deformation.

Proposed CNN models with Transfer Learning

Images are classified by a four-layer CNN using InceptionV3, ResNet50, VGG19 and Xception architectures.

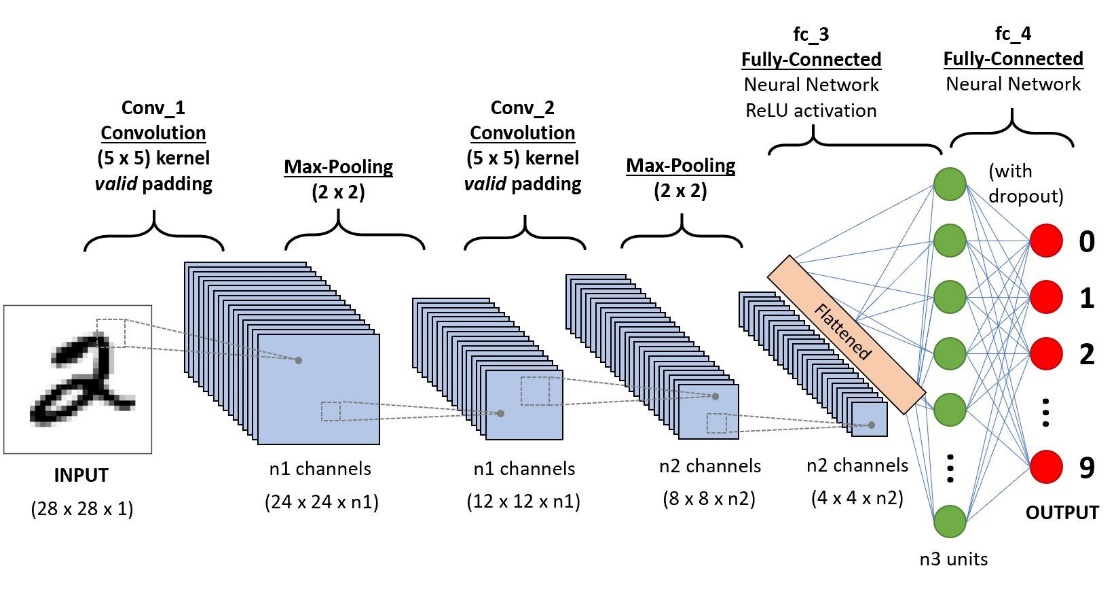


Figure 3: CNN architecture

VGGNet19:

The Visual geometry group (VGG) is, trained on the ImageNet database, consists of 19 convolutional layers. As the network gets deeper, the image is compressed, and the convolution kernels increase.

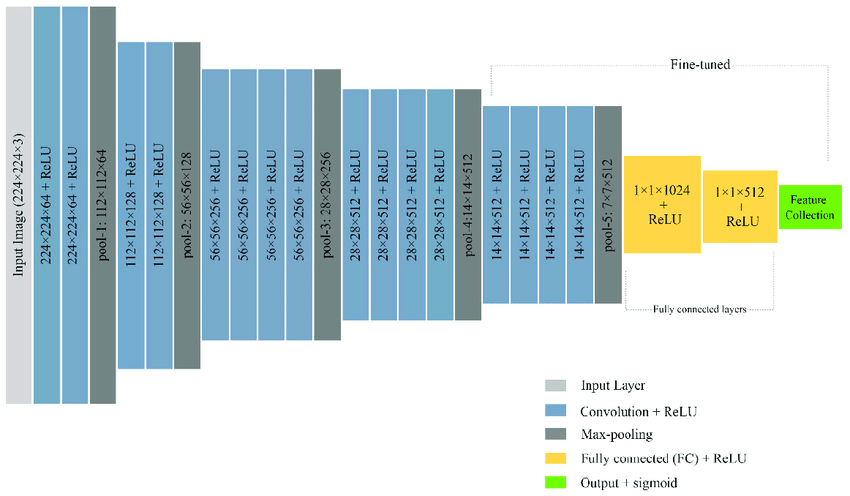


Figure 4: VGG-Net 19 Model Architecture

ResNet-50:

ResNet-50 is a 50-layer convolutional network. The network has 2 FC layers and 16 residual layers, with each residual layer containing 3 convolutional layers. The network improves the performance in the upper layer and connecting features from one layer to the next, avoiding degradation.

A picture containing text

Description automatically generated

Figure 5: ResNet50 Model Architecture

GoogLeNet (InceptionV3):

This architecture is a popular model used in TL with the last layer capable of being retrained. The architecture can be applied to smaller datasets efficiently with high classification accuracy. The architecture is trained with the ImageNet dataset and has 48 layers.

Diagram

Description automatically generated

Figure 6: Inception V3 Model Architecture

Xception (extreme inception):

Xception models although more costly to train, have the advantage of lower compute time and higher efficiency.

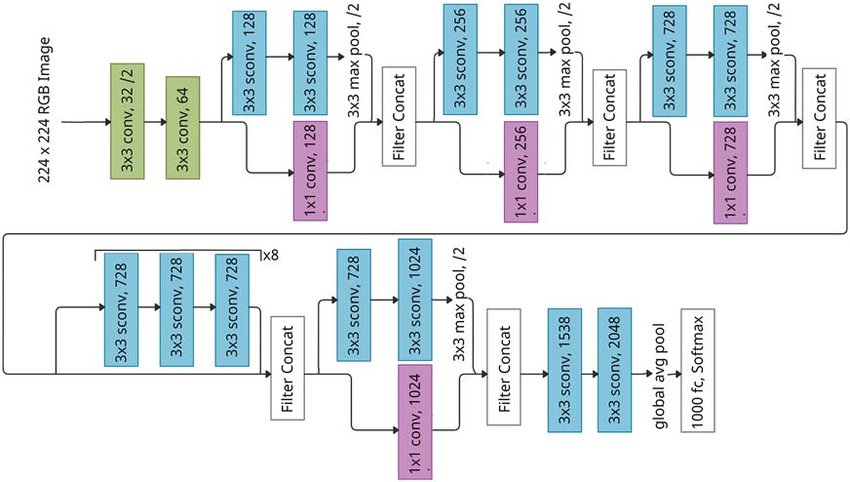


Figure 7: Xception Model Architecture

Ensemble learning:

Ensemble learning combines all four of the CNN architectures into one classifier.

Transfer learning:

Pre-trained models with pre-trained weights from the non-medical labelled ImageNet dataset are used to classify new fundus images to diagnose glaucoma.

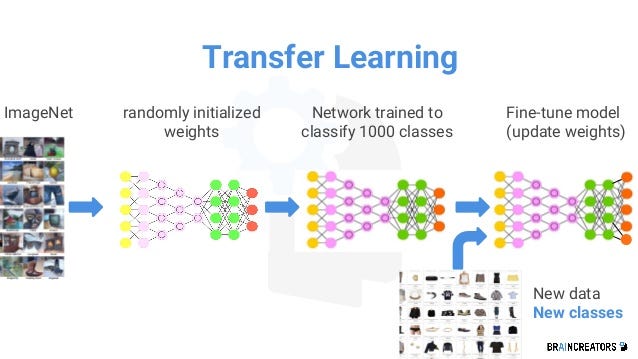


Figure 8: Transfer Learning Model Architecture

Evaluation and performance

Area Under Curve (AUC), accuracy, sensitivity, specificity, precision, F1 score, Receiver Operating Characteristic (ROC) curve are the important metrics to consider when evaluating classification models.

Precision-recall and F1 scores are found in a confusion matrix and are linked to true and false classifications. The true positive rate or sensitivity is known as recall, whereas precision is the positive predictive value. The frequency at which the CNN accurately predicts glaucoma is measured by sensitivity. F1-score are a test of image accuracy, using a combination of accuracy and recall obtaining the score.

The accuracy score measures classifier efficiency, it is a ratio between the correctly identified predictions and the total number of predictions a classifier makes.

ROC and AUC metrics are often used to determine how a model can distinguish between classes. The higher the AUC, the better the model. With these two metrics, there is a trade-off between sensitivity and specificity.

Work plan

Using the iterative Knowledge Discovery in Database (KDD) process, the following steps are required to implement this project:

1. Gather the data together.
2. Pre-process the data.
3. Segmentation of the optic nerve using U-Net CNN architecture.
4. Create four individual CNN learning architectures to the cropped images.
5. Create the ensemble model from the individual CNN architectures.
6. Validation and testing.
7. Interpret and evaluate the models using various performance metrics.
8. Report the results.

Risk assessment

Class imbalance

Class imbalances may lead to models having poorer predictive capabilities. The Synthetic Minority Over-Sampling Technique (SMOTE) can solve this problem by generating new synthetic images of the minority class instead of creating copies of existing samples.

Biased datasets

Because of biased datasets, the predictive power of current machine learning algorithms is limited to certain ethnic groups and may perform poorly for under-represented groups, such as the hispanic and black populations. Algorithms trained on balanced and representative datasets perform better. (18)

Size of datasets

Transfer learning-based have been shown to have higher performance on datasets with over 1000 images. Training of CNN usually requires larger datasets. Smaller datasets may lead to less data variability which in turn can reduce an algorithms performance.

The need for image synthesis techniques such as Variational Autoencoder (VAE) and Deep Convolutional Generative Adversarial Network (DCGAN) is likely to increase a model’s performance.

Computational resources

AI-based frameworks are often hindered by computational efficiency, with most methods having limitations. In addition, these models usually require more parameters to ensure proper training, this adds to the need for higher computational resources. (18)

Expected results

Based on similar research, metrics for the VGG19 architecture should be the highest, followed by InceptionV3, ResNet50 then Xception respectively. The methodology incorporates ImageNet trained CNNs together with the ADAM optimizer and 10-fold cross validation.

An ensemble learning classifier should reduce the error rate by using a combination of output predictions of each CNN.

The accuracy of these CNN models is expected to be anywhere from 89% to close to 100%. Precision, F1-score, sensitivity and specificity, AUC scores are expected from 0.89 to 1.

Conclusions

It is essential that glaucoma be diagnosed and treated as early as possible to avoid damage to the optic nerve that leads to permanent vision loss. VGG19, ResNet50 and EfficientNetB7 have been shown to be good performers of image classification. Deep learning models can predict glaucoma development before disease onset with reasonable accuracy.

Ensemble learning has been found to have higher metrics. Each network can be fine-tuned independently and then combined in an average ensemble setting to obtain a glaucoma score for each image.

Results from other studies show that performance is high when generalising to new datasets, though it was found that in many cases that data insufficiency affected performance. Transfer learning was successful in several studies; based on these results, models used in this project could successfully applied to other retinal diseases.

To improve a DL models performance, it is evident that there is a need for balanced and representative datasets and could also be a good topic for future research.

# References

1. **glaucoma.org.** glaucoma.org. [Online] 2022. https://glaucoma.org/are-you-at-risk-for-glaucoma/?print=print.

2. *Machine learning applied to retinal image processing for glaucoma detection: review and pespective.* **D. M. S. Barros, J. C.C. Moura, C. R. Freire, A. C. Taleb, R. A. M. Valentim, and P. S. G. Morais.** 2020, BioMedical Engineering Online, Vol. 19.

3. **R. Kashyap, R. Nair, S. Machinathu P. Gangadharan, M. Botto-Tobar, S. Farooq, and A. Rizwan.** Glaucoma detection and classification using improved U-Net deep learnng model. *Healthcare.* December 9, 2022, p. 19.

4. *Dense Fully Convolutional Segmentation of the Optic Disc and Cup in Colour Fundus for Glaucoma Diagnosis.* **B. Al-Bander, B. M. Williams, W. Al-Nuaimy, M. A. Al-Taee, H. Pratt and Y. Zheng.** 20738994, 2018, Symmetry, p. 17.

5. *A noel optic disc and optic cup segmentation technique to diagnose glaucoma using deep learning concolutional neural network over retinal fundus images.* **H.N. Veena. A. Muruganandham, T. Senthil Kumaran,.** 8, 2022, Journal of King Saud University - computer and information science, Vol. 34, pp. 6187 - 6198.

6. *Classification of glaucoma in retinal images using EfficientnetB4 deep learning model.* **A. Geetha, and N. B. Prakash.** 3, 2021, Computer Systems Science & Engineering, Vol. 43, p. 15.

7. *Optic Disc Boundary and Vessel Origin Segmentation of Fundus Images.* **S. Roychowdhury, D. D. Koozekanani, S. N. Kuchinka, K. K. Parhi.** 6, 2016, IEEE Journal of Biomedical and Health Informatics, Vol. 20, pp. 1562 - 1574.

8. *An Automated Deep Learning Approach to Diagnose Glaucoma using Retinal Fundus Images.* **A. Shoukat, S. Akbar, S. Al E Hassan, A. Rehman, N. Ayesha.** Islamabad : 2021 International Conference on Frontiers of Information Technology (FIT), 2021. pp. 120-125.

9. *A compariative study of deep learning models for diagnosing glaucome from fundus images.* **Abdel-Mottaleb, M. Alghamdi and M.** 2021, IEEE, Vol. 9, pp. 23894-23906.

10. *ODGNet: a deep learning model for automated optic disc localization and glaucoma classification using fundus images.* **J. Latif, S. Tu, C. Xiao, S. Ur Rehman, A. Imran and Y. Latif.** 2022, SN Applied Science, Vol. 4.

11. *Identifying those at risk of glaucoma: a deep learning approach for optic disc and cup segmentation and their boundary analysis.* **J. Kim, Loc Tran, T. Peto, E. Y Chew.** 1063, 2022, Diagnostics, Vol. 12.

12. *Glaucoma diagnosis using multi-feature analysis and a deep learning technique.* **N. Akter, J. Fletcher, S.Perry, M. P. Simunovic, N. Briggs and M. Roy.** 2022, Scientific Reports, Vol. 12.

13. *A comparitive study of deep learning models for diagnosing glaucoma from fundus images.* **M. Alghamdi, M. Abdel-Mottaleb.** 2021, IEEE Xplore, Vol. 9, pp. 23894-23906.

14. *A review of deep learning for screening, diagnosis, and detection of glaucoma progression.* **Atalie C. Thompson, Alessandro A. Jammal, Felipe A. Medeiros.** 2, July 2020, Translational vision science and technology, Vol. 9.

15. **E. B. Mariottoni, A. A. Jammal, S.l I. Berchuck, I. M. Tavares, F. A. Medeiros.** *An objective structural and functional reference standard for diagnostic studies in glaucoma.* 2020.

16. **Leonard, C.** What lies ahead for AI in glaucoma? [Online] 2022. https://www.reviewofophthalmology.com/article/what-lies-ahead-for-ai-in-glaucoma.

17. **Medical Image Analysis Group (MIAG), Universidad de La Laguna (ULL).** RIM-ONE-DL. [Online] http://medimrg.webs.ull.es/research/retinal-imaging/rim-one/.

18. *Assisting glaucoma screening process using feature excitation and information aggregation techniques in retinal fundus images.* **A. Raza, S. Adnan, M. Ishaq, H. Seok Kim, R. Ali Naqvi and S. Lee.** 257, 2023, Mathematics, Vol. 11.

19. *Optic disc and cup segmentations for glaucoma assessment using cup-to-disc-ratio.* **A. Neto, J. Camera, S. Oliveira, A. Cláudia, A. Cunha.** 2022, Procedia Computer Science, Vol. 196, pp. 485-492.

20. *Machine learning on glaucoma: the missing point.* **AlRyalat, S. Aldeen.** 2021, Eye, Vol. 35, pp. 2456-2457.

21. **J. Mandell, B.S. and B. Lin, M.D, A. Quan, M.D.** eyeguru.org. [Online] https://eyeguru.org/blog/evaluating-cup-and-disc/#:~:text=As%20the%20rims%20become%20thinner,%2FD%20ratio%20of%20%3E0.6..

22. *Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening.* **J. Cheng, J. Liu, Y. Xu, F. Yin, D. Wing Kee Wong, N. Tan, D. Tao, C. Cheng.** 6, 2013, IEEE Transactions on Medical Imaging, Vol. 32, pp. 1019 - 1032.

23. **A. Pinto, Y. Diaz.** *teleoptometria.com.* Department of electronic engineering, Universitat Politecnica de Valencia. 2019. PHD Thesis.

24. *Glaucoma and machine learning: a call for increased diversity in data.* **S. Sekimitsu, N. Zebardast.** 4, 2021, American academy of opthalmology, Vol. 4, pp. 339 -.