Glaucoma Detection and Progression Prediction using Deep Learning

Debayan Biswas
MSCDAD_C
National College of Ireland
Dublin, Ireland
x22242821@student.ncirl.ie

Ishita Kundu
MSCDAD_C
National College of Ireland
Dublin, Ireland
x22242091@student.ncirl.ie

Pinaki Pani
MSCDAD_C
National College of Ireland
Dublin, Ireland
x23112573@student.ncirl.ie

***Abstract*—This study outlines the process of creating an advanced deep-learning model to predict and detect the development of eye disease Glaucoma. Glaucoma if left untreated may lead to irreversible eye damage leading to permanent blindness. Traditional diagnosis methods are costly and error-prone. The deep-learning model can explain detection and prediction which will help medical technicians identify and manage the disease effectively. The model uses techniques to examine retinal fundus images, capturing valuable diagnostic insights from the back of the eye. The project aims to deliver healthcare officials an improved diagnostic tool for glaucoma, allowing early disease detection and intervention, and improving patient lives in fighting glaucoma.**

Keywords*— *glaucoma, retinal fundus, deep-learning, early detection, MobileNet

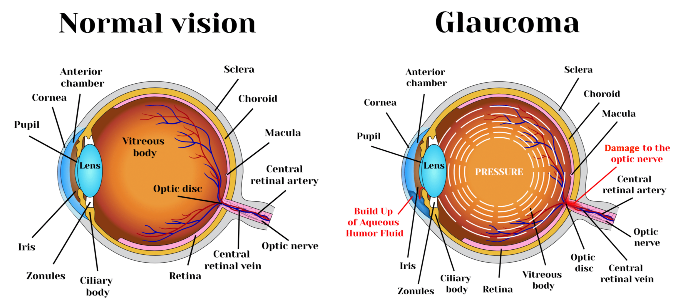


Fig. 1: Normal Vision vs Glaucoma
Source: Adapted from [2]

I. Introduction

Glaucoma is the primary contributor to worldwide irreversible blindness and is a challenge for the healthcare sector to tackle glaucoma as it is difficult to diagnose initially [1]. Early diagnosis of glaucoma is necessary for preserving eyesight and improving patient life. This project uses an advanced deep-learning method to detect glaucoma early and predict its progression over time.

The main motivation that led to the idea of the project is the need to enhance early diagnosis and detection of glaucoma. As cases of glaucoma are rising globally each year, there is an urgent need for a solution that is advanced enough to enable early detection. This would help to provide timely treatment to the affected individuals. Figure 1 shows the cross-sectional difference between a healthy and a glaucoma eye. The glaucoma eye shows increased intraocular pressure, optic nerve damage, and fluid accumulation inside the eye.

The project relies upon retinal fundus images which are specialized eye images and are the main components to train the model for prediction and detection of disease. The images are obtained ethically from trusted repositories. The aim is to build a powerful tool that can help clinicians diagnose the disease early and predict its progression using retinal fundus images.

The project holds significant business value and has the potential to make patient's lives better. Early glaucoma detection will not only preserve eyesight but will also lower the socio-economic burden on an individual and the healthcare system globally. This approach promises to be cost-effective and cost-saving for healthcare systems globally. Moreover, by increasing the prediction accuracy and its progress, the project could lead to a major achievement in the research and treatment of glaucoma in the future. This signifies a healthier life and better results for the patients.

II. Hypothesis

The main hypothesis behind the project is that by using advanced machine learning techniques on the retinal fundus images, one can accurately predict the progression and spread of glaucoma. The suggested machine learning model will display strong capabilities in the accurate identification and classification of cases of glaucoma and predict the development of the disease over time with a level of accuracy. These identifying and predictive capabilities are important as they allow timely intervention which can eventually prevent vision loss and improve patient lives. By using the information extracted from the retinal fundus images, the hypothesis aims to state that early detection and prediction of the disease can benefit the affected individuals in terms of patient care and therefore reduce the financial burden related to eye problems globally. The goal is to verify the hypothesis and uncover new knowledge that has the potential to change the healthcare industry and positively impact public health worldwide.

III. Literature Review

The main task of this project is to detect glaucoma early using deep-learning techniques. Glaucoma is an eye disease that can lead to irreversible vision loss and can be identified with the help of retinal fundus images. The research in [3], mentions that in glaucoma the optic nerve is affected which causes the intraocular fluid pressure to increase within the eye. Other symptoms besides this can also help to identify glaucoma through the fundus images. Disc segmentation, the thickness of the retinal nerve fiber layer (RNFL), and Cups-to-Disc ratio measurements are a few such features of the eye that can be used to identify glaucoma in fundus images as discussed in the study [4]. These features that help to identify glaucomatous eyes can be identified with the help of machine-learning methods as well as deep-learning techniques.

Deep-learning methods can be used to classify retinus images. Convolutional Neural Networks (CNN) are being used here as one such technique. In the study [5], authors state that CNNs reduce the number of operations by using convolution on pixel patches of adjacent pixels. They further mentioned that classic artificial neural networks are outperformed by CNNs when used in pattern recognition within images. Also, [6] discussed how the features of layer pooling in CNN can downscale an image which creates accurate information from large image data.

This project uses advanced data pre-processing and augmentation techniques. The research in [7], states how pre-processing techniques such as resizing, data transformation, normalization, and feature extraction can be helpful to save computational resources or increase the flexibility of training an image dataset over a pre-trained model. Besides image pre-processing techniques, hyperparameter tuning, model interpretability, and the use of evaluation metrics are also important. The technologies mentioned above play an important role in increasing the effectiveness, accuracy, and interpretability of the model that has been devised. The techniques are explained elaborately by [8]. Transformations like zooming, rotation, flipping, and color shifting have been used to create variation in the raw images which helps in expanding dataset quality and ultimately improves the generative ability of the model. Both data processing and augmentation fall under image pre-processing and these methods help to improve the clarity of characteristics in the retinal fundus images.

CNN works as a base for this project. Transfer learning techniques in the paper [9] leveraging CNNs improve the performance of the deep-learning models. Pre-trained CNN models like RestNet, VGG, or MobileNet aid in boosting the capabilities of glaucoma detection by understanding the varied range of image data even when the dataset is limited. The study [10] discusses the usage of these models for the classification where they achieve an accuracy of 80% with ResNet-50 and 74% with VGG-16.

All the above mentioned research papers and their different approaches for the task shows that it is worth researching further into the topic with pre-trained deep neural networks.

IV. Methodology

The project uses retinal fundus image data from a public dataset called EyePACS and the images are in colored format having a dimension of 512 x 512. The RGB images of the retinal fundus can provide crucial information. The crucial information fetched from the retina images allows to check the retinal nerve fiber layer (RNFL) and Cup-to-Disc ratio(CDR) within the retinas. The RNFL and CDR information are fetched with the help of edge detection and brightness information along with pattern recognition through filters in CNN. The filters in CNN identify the patterns and help extract the relevant features for this task. Before the models are trained to fetch the relevant features the dataset is prepared for the task. The project follows a Knowledge Discovery in Databases (KDD) based approach to discover patterns, connections, and insights from the dataset. The steps consist of data selection, followed by data pre-processing, after which the data is transformed into the required format for the model training thus helping to do data interpretation, and knowledge gathering, which is demonstrated in Figure 2.

A. Data Collection

The first step of KDD flow is data collection. Through data collection, KDD aims to discover valuable information hidden within the dataset. A dataset of retinal fundus images is initially fetched using Python libraries that are downloaded from the source. The data fetched are already divided into train and test sets.

B. Data Preprocessing

The next step is image preprocessing which includes various preprocessing techniques to enhance feature extraction and improve model performance. In the beginning, the images are pre-processed using multiple techniques. The images are resized initially, along with flipping the picture both vertically

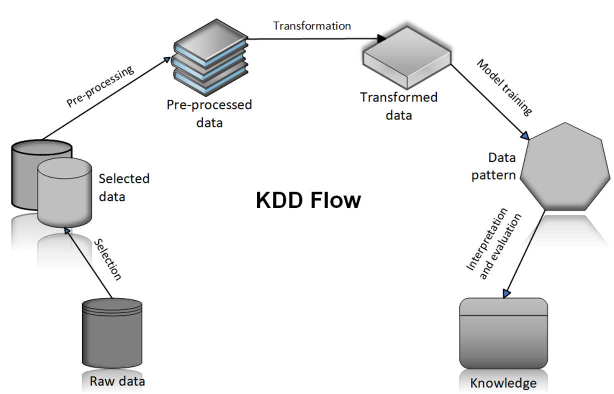


Fig. 2: KDD Flow of the project

and horizontally. Following this, the images are then rotated and undergo batch normalization. Once the pre-processing is completed the dataset is ready for the models to train on. A sample of the preprocessed eye images is shown in Figure 3.

C. Model Training

The next step is model training, where the CNN pre-trained models are identified that could be used for the task. Given the research done in the field, there are multiple pre-trained models out of which MobileNet and InceptionResNet models are trained on the large ImageNet dataset. This project uses MobileNetV3 as the transfer learning model to classify the data. Once the model is finalized then the project uses MobileNet as the classifier and the top layers are turned off. The learnable parameters used here are of Imagenet. While training the alpha value used here is 0.1. The model also uses regularization techniques. This project uses Lasso (L1) regularization. While training the models both regular and spatial dropouts are included. Along with this the model also includes early stopping to avoid any overfitting in case the model experiences such a scenario while training.

D. Model Evaluation

The trained model is evaluated on the test dataset to estimate the performance of the model. Performance metrics like accuracy, precision, recall, and F1-score are there to determine the effectiveness of prediction. Initially, the model training is done with 25 epochs. However, upon further training of the model, it is noticed that hyperparameter tuning can lead to better results. Here the tuning helps the model to converge at around 12 epochs. The hyperparameters like lambda values and alpha values are changed. The optimizer chosen for the model is Adam Optimizer and the loss function used is binary cross-entropy. Further, the evaluation of the model is also performed with the help of a confusion matrix that shows the predictions made on the test dataset.

V. Results

The results of the implemented CNN model using MobileNetV3 displayed robust performance. The model shows a test accuracy of around 91% in categorizing and classifying fundus images. Figure 4 depicts the results obtained from the project. The test loss achieved here is 0.33. The evaluation metrics, precision, recall, and F1 score provide a further understanding of the model performance. The precision value achieved is around 89% while the recall value is high at around 93%. The area under the curve (AUC) shows an efficiency of 96%.

Fig. 3: Glaucoma pre-processed images

```
1000/1000 [==============================] - 14s 14ms/step -
loss: 0.3365 - binary_accuracy: 0.9090 - auc: 0.9654 - preci
sion: 0.8925 - recall: 0.9300
Test Loss: 0.33652451634407043
Test Accuracy: 0.9089999794960022
Test AUC: 0.9653699398040771
Test Precision: 0.8925144076347351
Test Recall: 0.9300000071525574
1000/1000 [==============================] - 14s 14ms/step
```

Fig. 4: Summary result

A confusion matrix is represented by the model results in Figure 5. It is a visual summary of the performance of a classification model using positive and negative instances. The confusion matrix shows a matrix of True positive (TP) in 465 instances & True negative (TN) in 444 instances, False positive (FP) in 55 instances, and False negative (FN) in 34 instances. The matrix represents the number of correct and incorrect predictions by the model.

Data augmentation techniques helped to improve model generalization by expanding the training dataset. This overall helps to improve model performance. The model performance is also improved by the hyperparameter tuning done within the model. The tuning done here is the usage of optimizers and also the lambda values used for regularization. Evaluation metrics such as precision, recall, and F1 score provided additional insights into model performance and the respective trade-offs between true positive and false positive rates. The model accuracy and model loss are depicted in the Figure 6.

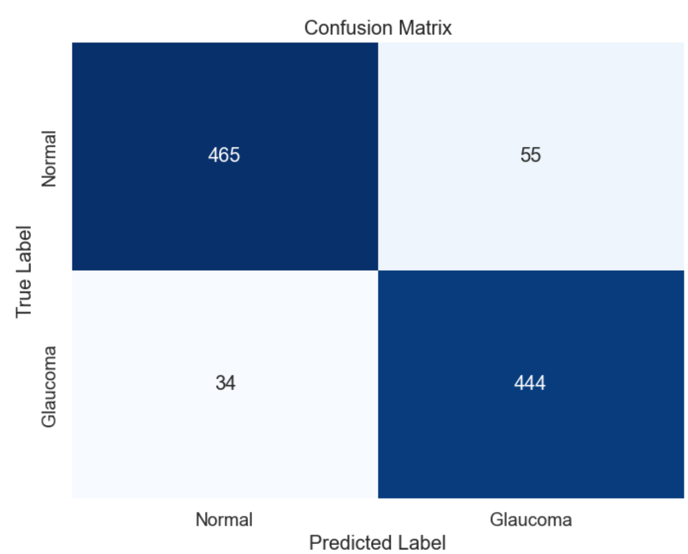
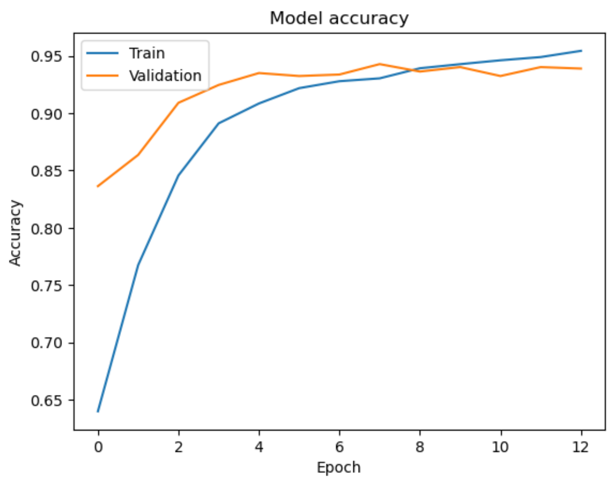
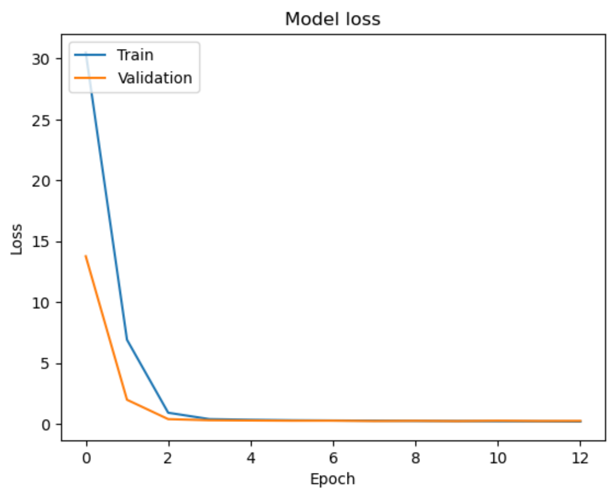


Fig. 5: Confusion Matrix



(a)



(b)

Fig. 6: Model accuracy and loss

VI. Interpretation of the Result

A. Quantitative Results

The results obtained from the model depict a complete evaluation of the model's performance. The test metrics obtained provide in-depth details of the model. The test loss of 0.336 shows the model's accurate prediction capability with a minor difference between the actual and the predicted result. The test accuracy metric is the measure of correctly classified instances out of all test instances. The high value of 0.91 depicts that the model properly identifies 91% of the instances present. The higher the test accuracy, the better the model performs in real-world applications. The Area Under the Curve (AUC) is the model performance metric that shows class discrimination. The higher the AUC value, the better the model is capable of differentiating between the classes present. The Test AUC score of 0.97 represents the model's strong discrimination capability. The two measures of precision and recall indicate the model's capability to identify positive instances. Precision is the fraction of true positive instances to the sum of true positive and false positive instances while Recall is the proportion of true positive and sum of true positives and false negatives. The high value of Test precision and recall at 0.89 and 0.93 suggests the model's capability of minimizing false positive instances and accurately identifying positive instances. The high value of TP and TN with a low value of FP and FN depicts robust model performance. In summary, the confusion matrix along with precision, recall, and accuracy suggests that the model has high precision and accuracy in accurately identifying positive and negative instances that are important for glaucoma prediction.

B. Business Value Qualitative Interpretation

This research on glaucoma prediction and detection has significant business value that can be achieved by improving early diagnosis and treatment of glaucoma. Accurately predicting glaucoma early will allow healthcare professionals to effectively devise treatment plans to prevent further vision loss. This will not only boost the healthcare domain but also reduce the socio-economic impact linked to glaucoma by reducing healthcare expenses. The process of early detection of glaucoma using machine learning is better in terms of accuracy and prediction capabilities than traditional diagnosis techniques. It has the potential to detect the disease early with fewer resources thus leading to lesser infrastructural cost. The application of machine learning will reshape the eye-care sector. The application of machine learning in medical science is just an entry point to the field of medical innovation and much more can be achieved with further use of advanced technology that will eventually reduce the socio-economic burden related to healthcare thus making people's life better.

VII. Conclusion & Future Scope

To conclude, the project highlights the importance of glaucoma detection and the immense potential of the application of deep learning in this domain. The model generated can identify and predict the progression of glaucoma with better efficiency than traditional methods. This is a state-of-the-art technology in image classification that can contribute immensely to the medical diagnosis of glaucoma as well.

There are several scopes for exploration and improvement in this field as well as the model used in this project. Advanced models based on both the fundus and the Optical Coherence Tomography images can be developed to diagnose glaucoma at an early stage using multiple different modes of imaging. The multimodal imaging approach can involve multiple datasets combined and used to train a model that can be further useful. The models can be even further tuned as well with larger batch sizes and better computational resources so that results can be achieved faster. Besides this, Vision Transformers are new technologies that could be implemented as they are making huge progress in pattern recognition and feature extraction from images. Besides this collaboration with clinicians can shed further light into the domain and even exploring other eye diseases could turn out fruitful.

ACKNOWLEDGEMENT

The project owes its completion to the support, advice, guidance, and critical feedback provided by Professor John Kelly, whose insights and constructive ideas were crucial throughout the process.

REFERENCES

[1] I. V. Wagner, M. W. Stewart, and S. Dorairaj, "Updates on the diagnosis and management of glaucoma," *Mayo Clinic proceedings. Innovations, quality outcomes*, vol. 6, pp. 618–635, 12 2022.

[2] I. Pharmaceuticals, "What is Glaucoma? Learn about Glaucoma — Glaucoma Resources," 4 2024. Available: https://www.inmedpharma.com/learn/what-is-glaucoma/.

[3] R. Krishnan, V. Sekhar, J. Sidharth, G. S, and G. P. Gopakumar, "Glaucoma detection from retinal fundus images," pp. 0628–0631, 07 2020.

[4] J. Carrillo, L. Bautista, J. Villamizar, J. Rueda, M. Sanchez, and D. rueda, "Glaucoma detection using fundus images of the eye," in *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, pp. 1–4, 2019.

[5] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *CoRR*, vol. abs/1511.08458, 2015.

[6] A. Semma, S. Lazrak, Y. Hannad, M. Boukhani, and Y. El Kettani, "Writer identification: The effect of image resizing on cnn performance," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVI-4/W5-2021, pp. 501–507, 2021.

[7] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data," *Sec. Sustainable Energy Systems*, vol. 9, March 2021.

[8] T. Hope, Y. Resheff, and I. Lieder, *Learning TensorFlow: A Guide to Building Deep Learning Systems*. O'Reilly Media, 2017.

[9] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," *Technologies*, vol. 11, no. 2, 2023.

[10] P. Jibhakate, S. Gole, P. Yeskar, N. Rangwani, A. Vyas, and K. Dhote, "Early glaucoma detection using machine learning algorithms of vgg-16 and resnet-50," in *2022 IEEE Region 10 Symposium (TENSYMP)*, pp. 1–5, 2022.