Artificial Intelligence in Ophthalmology: Fundus-Based Multi-Classification Deep Learning for Glaucoma Detection

Zeina Moammer Alabido
Student ID: g202423020
Supervised by Dr. Muzammil Behzad
King Fahd University of Petroleum and Minerals, Saudi Arabia
muzammil.behzad@kfupm.edu.sa

October 2025

Abstract

Eyes are important part of the human body, which guides humans to see, and be conscious about what is going on around them, be safe from danger situations, and able to do the daily tasks easily, without the need to have an assistant from others. However, humans might feel some of the visual symptoms, such as pain, blurring view, seeing cobwebs of floating particles, eyesight that is distorted making straight lines appear wavy, the side vision is flawed, or vision loss [1]. Glaucoma is one of the well-known Retinal diseases that happen due to factors that lead to irreversible sight loss. It occurs when the fluids of the eye do not drain correctly, consequently, that leads to eyeball pressure. There is an optic nerve connected in between their eye and the brain, which will be affected by this pressure. This kind of disease is not treated and tackled, the vessels within the eye tissue would be burst. Approximately 3.6 million people might be affected with blindness because of glaucoma. This is a serious problem the field of medicine always tries to find the most intelligent ways to discover it in its early stages. Artificial intelligence and machine learning models would revolutionize the early detection of glaucoma. In this research, deep learning models would be tested to see their ability to early detect glaucoma. Also, we will see how Artificial Neural Networks effectively work on high dimensional data such as the glaucoma patients' images [2].

1 Introduction

Eyes are important part of the human body, which guides humans to see, and be conscious about what is going on around them, be safe from danger situations, and able to do the daily tasks easily, without the need to have an assistant from others. However, humans might feel some of the visual symptoms, such as pain, blurring view, seeing cobwebs of floating particles, eyesight that is distorted making straight lines appear wavy, the side vision is flawed, or vision loss . Glaucoma is one of the well-known Retinal diseases that happen due to factors that lead to irreversible sight loss. It occurs when the fluids of the eye do not drain correctly, consequently, that leads to eyeball pressure. There is an optic nerve connected in between their eye and the brain, which will be affected by this pressure. This kind of disease is not treated and tackled, the vessels within the eye tissue would be burst. Approximately 3.6 million people might be affected with blindness because of glaucoma. This is a serious problem the field of medicine always tries to find the most intelligent ways to discover it in its early stages. Artificial intelligence and machine learning models would revolutionize the early detection of glaucoma. In this research, deep learning models would be tested to see their ability to early detect glaucoma. Also, we will see how Artificial Neural Networks effectively work on high dimensional data such as the glaucoma patients' images [2].

1.1 Types of Glaucoma

- **Primary Open-Angle Glaucoma:** In this type, the channels that work on fluids draining are blocked. It might send fewer warnings about having glaucoma, but it is better if it has been detected early [2].
- **Acute Angle-Closure Glaucoma:** 2- Acute angle-closure Glaucoma: this type causes symptoms that need urgent attention, such as a medicine to lower the eye pressure sometimes used before the surgery [2].
- **Secondary Glaucoma:** this type of Glaucoma might be caused because of an injury, taking certain medicine, or an illness, for example, steroids [2].

this type of Glaucoma might be caused because of an injury, taking certain medicine, or an illness, for example, steroids [2].

1.2 Problem Statement

The problem of Glaucoma is a significant medical health problem that should be addressed carefully. Nowadays, the Artificial Intelligence world has become demandable in the field of medicine, since it assists the doctors in the early discovery of a serious medical problem.

However, the data becomes an obstacle in this field, such as it is not available publicly, it is limited to specific region, etc. Moreover, patients sometimes might not accept the idea of having their image has been used for medical analysis to be used for research purposes, so the result is, small, and simple dataset. Also, deep learning models require high dimensional data, because the DL models are too complex, and it introduces non-linearity, so it requires more complex datasets.

1.3 Objectives

- To develop a model that able to classify the Glaucoma Retinal Diseases using Fundus Images and leverage the beauty of Multiclassification.
- Generalize the results of the model to other datasets (if available).
- If the model successes, we should generalize that the model will work well on other retinal disease classifications

1.4 Scope

The dataset has been used in this study it has been collected from Kim's Eye Hospital, Yeongdeungpo-gu, Seoul, South Korea. Unfortunately, the data set availability is extremely limited, despite some of them are publicly available to be used, but very few ones.

2 Literature Review

Ali et al [3]. presented a paper titled "A Hybrid Convolutional Neural Network Model for Automatic Diabetic Retinopathy Classification from Fundus Images". Their goal was to build a model to detect diabetic retinopathy from fundus images using Convolutional Neural Networks (CNN). The model was hybrid CNN, since they have used ResNet50 and Inceptionv3 as a feature extraction method, and they have been abbreviated as IR-CNN. To start their research, they have used datasets of Diabetic patients' retinal images. The dataset contains the following categories: 25,806 No DR, 2440 Mild, 5291 Moderate, 5291 Severe 5291, and 5291 PDR images. They have done preprocessing steps such as histogram equalization and intensity normalization. They have split the data into 80% train set, 20% test set, 0.001 learning rate. The training was over 100 epoch. They have used Adam optimizer, and the mini batch during the training with a batch size of 32. They have resized the image to 256 by 256. They have tested three models, CNN, ResNet50, Inceptionv3, and the proposed IR-CNN hybrid model. To evaluate the performances, accuracy, sensitivity, specificity, precision, and f1-score were used. The Hybrid CNN model outperforms the other models with an accuracy equals 96.85%. ResNet50 scored 90.65%, and Inceptionv3 was the least with an accuracy of 87.18%.

Sathishkumar et al[4]. presented a paper about an AI-driven Glaucoma Detection using Deep Learning techniques. They have used a private dataset to conduct their experiment from a hospital. The dataset contains glaucoma, and nonglaucoma pictures. They have done several preprocessing such as resize the images, histogram equalization, Gaussian bar to remove the noise in the images, using rotation and reflection to apply data augmentation, and playing with the image brightness. They have suggested using CNN and comparing it to other models in the previous studies such as ANN, Efficient Net, SVM, LightGBM, and VGGNet. As an activation function, they have used sigmoid, since they are predicting glaucoma, or nonglaucoma, it is a binary classification. They have run the experiment 15 epochs, to assess the model performance, accuracy, precision, recall, and f1-score where used. Consequently, CNN was the highest among the previous models with an accuracy equals to 97.5%.

Zedan et al[5]. Presented a paper titled " Deep Learning-based Model Benchmarking of Glaucoma Segmentation Using a Novel Ibn Al Haitham Fundus Image Dataset". To start their study, they have worked using Ibn Al-Haitham dataset, which consist of 3000 high-resolution fundus images, it was collected form Iraqi hospital named Ibn al-Haytham Eye Teaching Hospital, which include Glaucoma and normal images. The data has been collected from patients aged between 5 to 80 years old. Distorted images have been filtered out. As a preprocessing step, they have applied the GNU Image processing Program known as GIMP-2, has been applied as an annotation process, and apply mask to annotate an optic disc or cup. Consequently, the images are converted into black and white; white represents the optic cup or disc, the rest is black as a background. Also, they have been resized the image to be 512 by 512 and applying Contrast-Limited Equalization (CLAHE) to enhance the image contrast. Also, applying the median filtering to smooth the image and reduce the noise. They have used the models U-Net, SegNet, FCDenseNet, and PSPNet. To evaluate accuracy, accuracy, specificity, and sensitivity, were used. SegNet was outperformed by the other models, with an accuracy equal to 98.1% and 97.4% in the case of optic dis, and optic cup, respectively.

Zhou et [6]. presented a paper titled "A foundation model for generalizable disease detection from retinal images". To start their experiment, they have used unlabeled dataset contains 1.6 million retinal images, which contains fundus images, and OCT images from MEH-MIDAS (Moorfields Diabetic imAge dataSet) and OCT, which is available online. The Images have been preprocessed as using AutoMorph to cut parts in the image and keep the retina part. Also, all images have been resized to 256 by 256, and the images for training resized to 224 by 224. Moreover, OCT photos have been processed by extracting and resizing the middle slices. Other data augmentation techniques such as random crops, random horizontal flip, and normalization have been done. Then, they have developed the RETFound, which is a self-supervised pretrained using MAE, which is a mask encoder strategy. The MAE is a ViT-large, which is a vision transformer large consist of 24 transformer blocks, the embedding size is 1024, and the decoder is ViT small, which is 8 blocks, and 512 embeds. The patch size is 16 by 16. The output f encoder feed into MLP Classifier with a batch size equals to 16, 50 epochs, and scheduling the learning rates values. This model has been compared with other previous models and found that it outperforms other models with an AUROC equals to 0.943 for APTOS Images in predicting the Diabetic retinopathy. Also, the model has been generalized to other diseases detection, such as heart failure AUROC 0.794, ischaemic stroke 0.754, myocardial infarction 0.737., and OCT AUROC 0.799

2.1 Previous Work Summary

Table 1 summarizes the reviewed studies and their results.

Table 1: Summary of Related Works

No.	Reference Title	Dataset	Models	Evaluation	Result
	Hybrid CNN for Diabetic Retinopathy	Private Dataset	ResNet50, InceptionV3, IR-CNN	Accuracy, F1-score	96.85%
	AI-Driven Glaucoma Detection	Private Dataset	CNN	Accuracy, Recall, Precision	97.5%
	Glaucoma Segmentation (Ibn Al-Haitham Dataset)	Ibn Al-Haitham	U-Net, SegNet	Accuracy, Specificity	98.1%
	Foundation Model for Disease Detection	1.6M Retinal Images	RETFound + MLP	AUROC	0.943

2.2 Gap Analysis

- Dataset Limitation: Most of the papers found have been using private datasets, which are not publicly available
- It is limited to a specific scope. In other words, most of the papers have collected datasets from their own region. Consequently, the models generalize to all types of images in different countries.
- Most of the paper focuses on classification, rather than trying more advanced techniques such as segmentation, object detection, or even image captioning.

- Universal Model: Do these models work to classify or detect other retinal diseases? This is one of the limitations in the previous papers

The limitation behind all the previous studies is how they can grant that the tested models on specified dataset will work for other datasets from other countries. Also, these developed models will work on other types of diseases related to retina. When it comes to medicine, taking care of people, we should think about how to make our model enhanced, strong, and propose strong architecture to predict reliable results. Not only for the purpose of research, but it should also care about a high accuracy for the medical field as well.

2.3 Research Questions

- RQ1: What are the best Neural Networks models that able to classify the normal, early, or advanced glaucoma cases using multi-classification model?
- RQ2: What are the best Neural Networks models that able to classify the normal, early, or advanced glaucoma cases using multi-classification model?

2.4 Hypothesis

- Hypothesis 1: This research will build a strong, and robust deep learning model to classify glaucoma cases.
- Hypothesis 2: Apply the needed steps for data preparation, processing, and augmentation to prepare the data for models training and testing.

3 Proposed Methodology

3.1 Existing Model and Challenges

3.1.1 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) is an enhanced version of Artificial Neural Networks (ANNs) [7], it has an automatic feature extraction technique from grid matrix dataset, such as images. CNNs are used mostly in Computer Vision Applications. CNNs are consist of input layer, pooling layer, and fully connected dense layer, ended with an output layer to give the predicted label. What makes CNNs challenging is that it is requires a lot of resources for training the model, which is computationally expensive, and it is prone to overfit if the dataset is not big enough, which requires large amount of data [8].

3.1.2 MobileNetV2

MobileNetV2 is a pretrained model, and a powerful CNN architecture has been established for light-weight employment such as mobile devices, or embedded vision applications [9] What makes MobileNetV2 interesting is that it introduces the concept of inverted residual blocks, and bottlenecks, which results in high accuracy. Also, it uses ReLU6 [10], which is an activation function that clips the output at 6, preventing numerical instability and making the model more suitable.

3.2 Model Architectures

Custom CNN: Includes three convolutional blocks (Conv2D + ReLU + BatchNorm + MaxPooling), followed by a dense layer with dropout and a softmax output for multiclass classification Fig.1.

MobileNetV2: A pretrained lightweight model using ImageNet weights, followed by global average pooling, dense layers, dropout, and softmax output Fig.2.

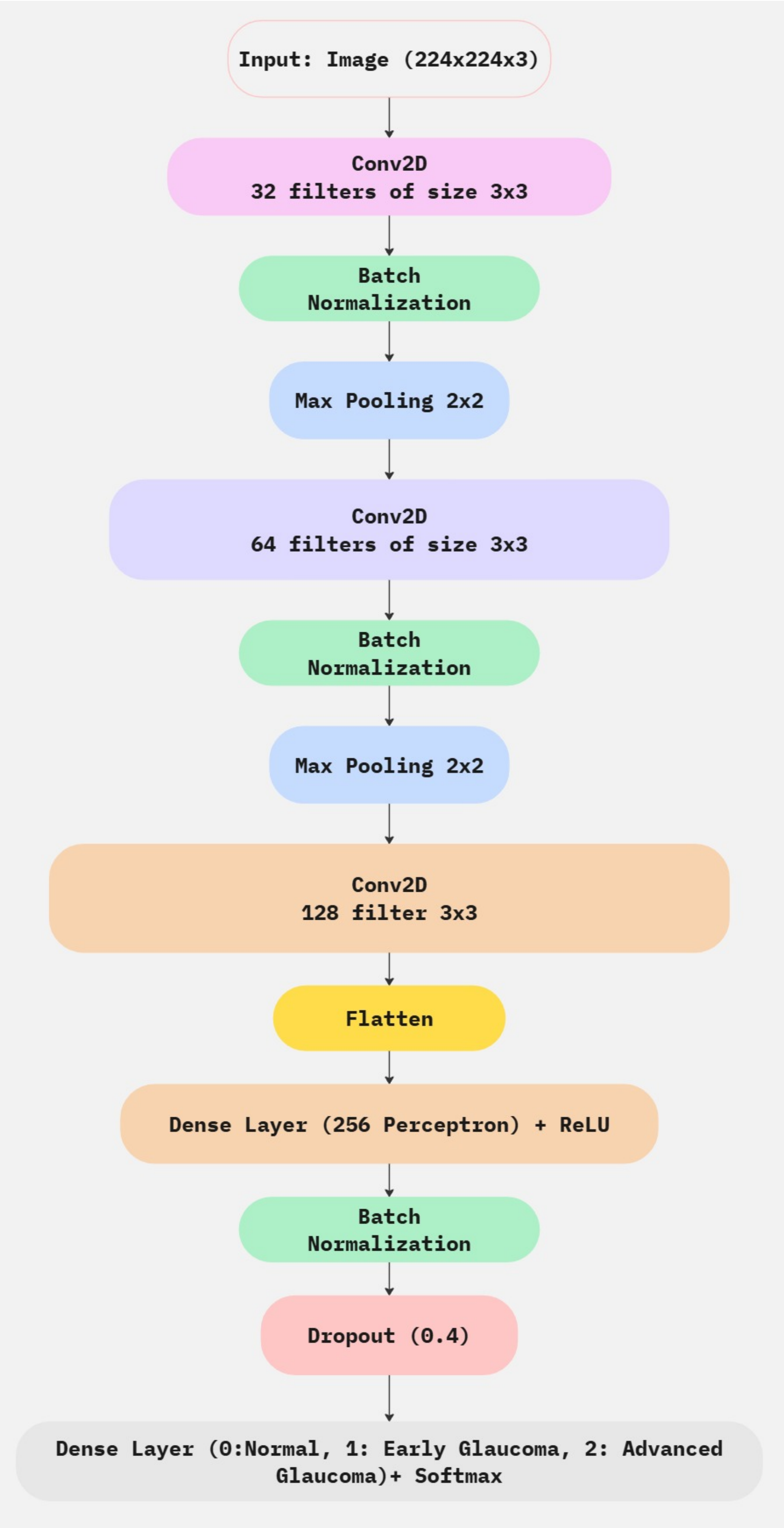

Figure 1: Proposed Architecture of Custom CNN.

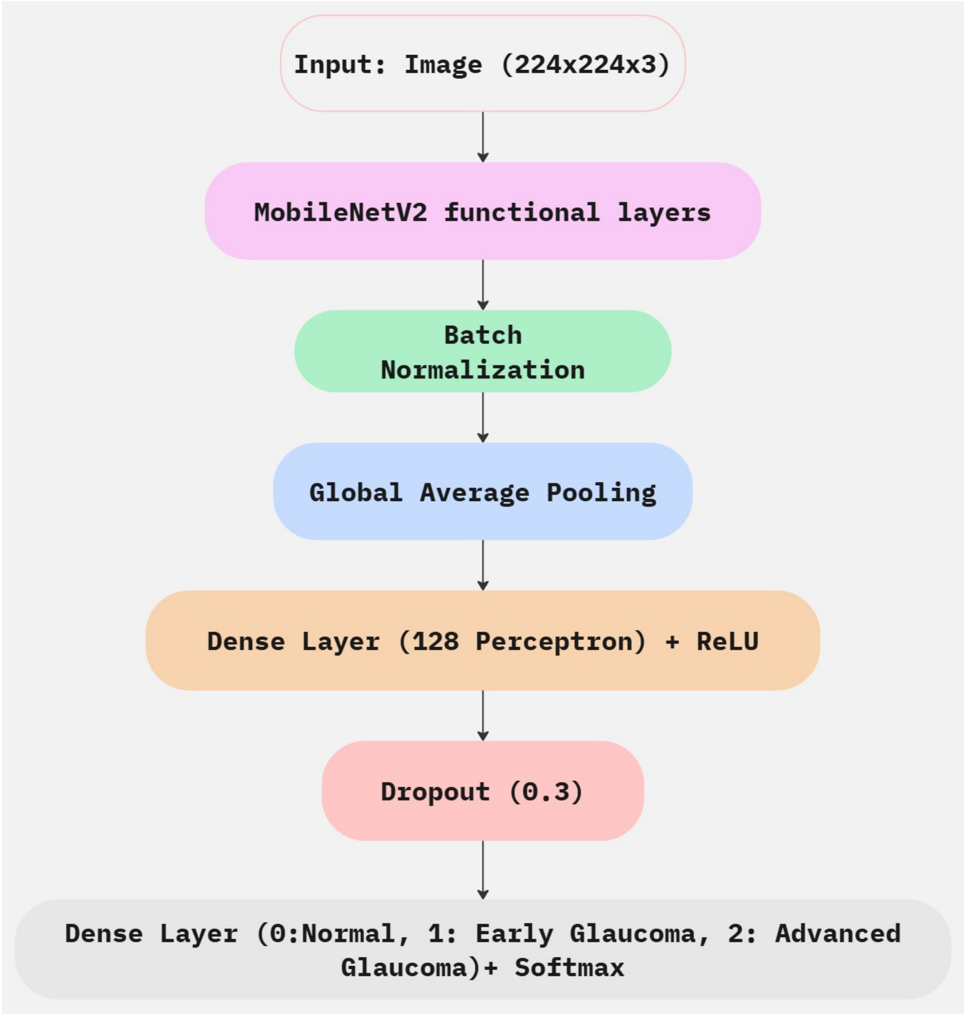

Figure 2: Proposed Architecture of Custom MobileNetV2.

3.3 Proposed Enahncements

3.3.1 Custome CNN:

The proposed CNN used for this project is as follows: Conv2D + ReLU: it consists of 32 filters, and the size of each filter is 3×3, over the input image of size 224×224×3. The output of this layer is going through ReLU activation function to introduce non-linearity and scale the data. Next, followed by Batch Normalization, which normalize the activations within each batch, by calculating the meaning and the variance of data, and then normalize [11]. Batch normalization followed by Max Pooling2D, which works on reducing the dimensionality by reducing the spatial size of feature map. In this case, the Max Pooling applied with the size (2,2). The next layer is another Conv2D + ReLU with 64 filters, each 3×3, which will start learning more hidden patterns, and more complex features, followed by Batch Normalization and Max Pooling2D. Next is Conv2D + ReLU with 128 filters of size 3×3, start to go deep inside the network and reach higher level of learning, followed by Batch Normalization, and Max Pooling2D. These 3 Conv2D layers followed by Flattening layer, which converts the 3D feature map to 1D vector, preparing them for the next layer, which is the Dense Layer. The Dense Layer consists of 256 perceptron with ReLU activation, to learn non-linear combinations. After that, Dropout with 0.4 has been applied to avoid overfitting, followed by Batch Normalization. The Last layer, which gives the prediction of multi-classification using SoftMax. See Fig.1.

3.3.2 Custome MobileNetV2:

This is a CNN model from TensorFlow Kera Applications. Basically, the weights of "ImageNet" dataset have been loaded, the top fully connected layers of MobileNetV2 have been freeze, because we will build our own. We have used the feature extractor of mobileNetV2 with the custom layers that have been added. The base of MobileNetV2 is followed by a batch normalization, with a global average pooling2D. This is followed by a dense layer that has a 128 + ReLU, with a dropout equals to 0.3. The last layer is a dense layer to produce the output using SoftMax function. See 2.

3.4 Algorithm and Implementation

3.4.1 Dataset Collection and Preprocessing

Dataset: *Machine Learn for Glaucoma* from Harvard Dataverse [12]. Dataset has been downloaded from Harvard Dataverse [12] and uploaded on drive to be able to apply them for training. The data resized to (224 x 224 x 3) and has been visualized to see the images of what they look like. The dataset contains 1544 Fundus images, as shown in Fig.3.

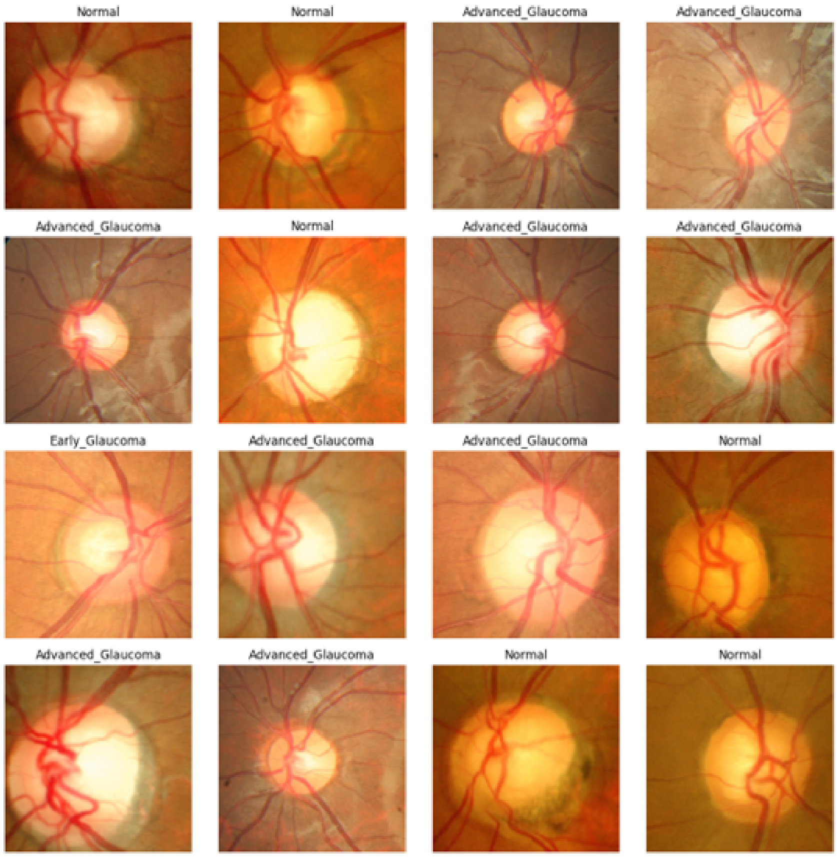

Figure 3: Dataset Images of size 224x224, which consist of 1544 images.

3.4.2 Data Preprocessing and Augmentation

Checking for class imbalance [13] is important to balance the data and avoid overfitting results. For that, data augmentation techniques applied during the training to enhance the training and avoid overfitting. Also, dividing the dataset into 80% training data, and 20% testing data. To ensure getting a good accuracy, High Performance NVIDIA GeForce RTX 4046 Laptop GPU, and to ensure faster model implementation. The dataset used in this study is named "Machine learn for glaucoma" available on Harvard Dataverse[12]. This dataset consists of 1544 images, and the dataset contains 3 separate files: Normal, Early Glaucoma, and Advanced Glaucoma. The dataset is mainly designed to apply deep neural network for the purpose of predicting the glaucoma and classify the images.

The preprocessing and data augmentation [14] done on the images of Normal, Early Glaucoma, Advanced Glaucoma are:

- Image Resize: The images resized both trains, and test dataset to: (224,224,3).
- Rotation Range: applying random rotation of the image by 20 degrees (clockwise or counterclockwise).
- Width Shift Range: By shifting the image horizontally by 20% of the image width.

- Height Shift Range: shift the image vertically by 20% of the image height.
- Shear Range: Applying shear transformation by slanting the image along one axis.
- Zoom Tange: randomly apply zooming in or out up to 15%.
- Horizontal Flip: randomly flip the images horizontally.
- Horizontal Flip: randomly flip the images horizontally.

Label-encoding: The three labels represented by the dataset have been converted into numbers, each representing a type of classification (0: Normal, 1: Early Glaucoma, 2: Advanced Glaucoma) [15]. Once the data augmentation has been applied, the images reach to 28,011 Image. See Fig.4 for the enhancement. In Fig.5 shows the division of data into tarin and test after data augmentation

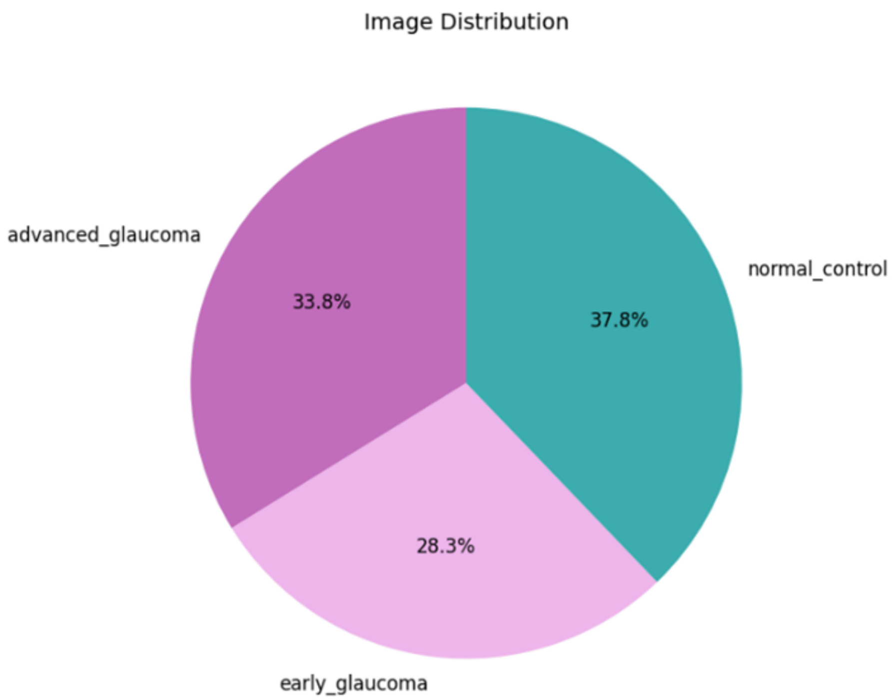

Figure 4: Class Balancing after data augmentation application.

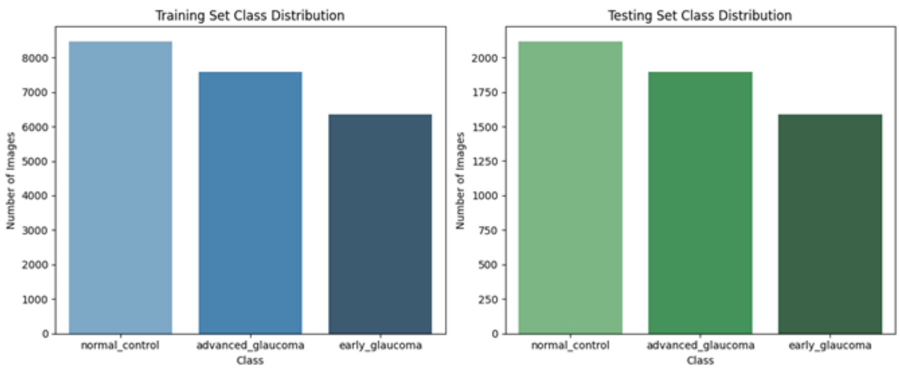

Figure 5: Dataset Splitting into train 80%, and test 20%.

3.4.3 Activation Functions

The Rectified Linear Unit (ReLU) activation function [16] is defined as:

$$f(x) = \max(0, x) \tag{1}$$

or equivalently, in piecewise form:

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases} \tag{2}$$

where x is the input to the neuron. The function returns x if $x > 0$, otherwise it returns 0.

The SoftMax activation function for a multi-class classifier is defined as:

$$\text{SoftMax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3}$$

where:

- z_i is the logit (the output of the previous layer for class i),
- K is the total number of classes,
- e^{z_i} represents the exponential of the logit,
- $\sum_{j=1}^{K} e^{z_j}$ is the normalization term (sum of exponentials across all classes).

3.4.4 Training Setup

- Optimizer: Adam (learning rate = $1e^{-5}$)
- Loss: Categorical Cross-Entropy
- Epochs: 20, 40, 80

- Batch Size: 32, 64
- Hardware: NVIDIA GeForce RTX 4046 GPU

3.4.5 Loss Functions and Optimization

Categorical Cross Entropy Loss Function: categorical cross entropy loss function [17] is used to calculate the loss by the models they are classifying, and it is used when we have more than two classes to classify (multi-classification problems). As an evaluation Metrex for the quality of the models, Categorical Cross Entrop Loss [18] has been used to assess how neural networks performs good on data.

The categorical cross-entropy loss function is defined as:

$$L(Y, \hat{Y}) = -\sum_{i=1}^{C} Y_i \log(\hat{Y}_i) \tag{4}$$

where:

- $L(Y, \hat{Y})$: the categorical cross-entropy loss.
- Y_i : the true label (e.g., 0: Normal, 1: Early Glaucoma, 2: Advanced Glaucoma).
- $\hat{Y}_i$: the predicted probability for class i.
- C : the total number of classes.

Optimization

Mini-Batch Gradient Descent:
it is a version of the traditional gradient decent [19] algorithm, which is used to optimize the models' parameters, such as the weights and biases. By dividing the dataset of training into mini batches of smaller sizes. This technique is more frequent than the usual mini bath, because we are using sample data for parameter update. In our experiment, 32 and 64 sizes have been implemented for both CNN, and MobileNetV2.

Adam Optimizer:
is an optimization technique used to adjust the learning rate during the training, it combines between the advantages of Root Mean Square Propagation (RMSProp), and the momentum [20].

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\frac{\partial L}{\partial w_t} \tag{5}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\left(\frac{\partial L}{\partial w_t}\right)^2 \tag{6}$$

Bias correction is then applied to compensate for initialization effects:

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{7}$$

Finally, the weight update rule is given by:

$$w_{t+1} = w_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{8}$$

where:

- m_t : first moment (mean) estimate,
- v_t : second moment (variance) estimate,
- β_1, β_2 : exponential decay rates for the moment estimates,
- α : learning rate,
- ϵ : a small constant to prevent division by zero,
- w_t : model weight at iteration t.

Performance Metric: Accuracy

Accuracy has been used as a primary performance metric to evaluate the effectiveness of both CNN and MobileNetV2 models. It measures how often the model's predictions are correct by dividing the number of correct predictions by the total number of predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

where:

- TP : True Positives (correctly predicted positive cases),
- TN : True Negatives (correctly predicted negative cases),
- FP : False Positives (incorrectly predicted positive cases),
- FN : False Negatives (incorrectly predicted negative cases).

3.5 Experimental Setup Environment

For the experiment, custom CNN, and MobileNetV2 have been applied. Adam optimizer has been used for weights updates. Also, minibatch has been applied for better model interpretation, and as a lightweight approach for updating gradients used in weight update. The models have been running for 20, 40, and 80 epochs. The core setup for these types of multi-classification, dealing with a huge number of images, requires a strong and robust machine to be used. Also, selecting a code editor to run the models. Device Name: Victus by HP Gaming Laptop. CPU: AMD Ryzen 7 7840HS w/ Radeon 780M Graphics (3.80 GHz). Memory size of 32.0 GB. NVIDIA GPU: High Performance NVIDIA GeForce RTX 4046 Laptop GPU. Visual Studio Code with Python 3.13.8 kernel has been used to run the code

4 Results and Evaluation

4.1 Performance Metrics

Accuracy and Loss values for both models are shown in Table 2.

Table 2: Results of CNN and MobileNetV2 on Glaucoma Images

Model	Epochs	Test Accuracy (%)	Test Loss
Custom-CNN	20	84.23	0.4641
	40	84.29	0.4829
	80	85.22	0.5503
MobileNetV2	20	73.56	0.6382
	40	78.81	0.5342
	80	83.72	0.4145

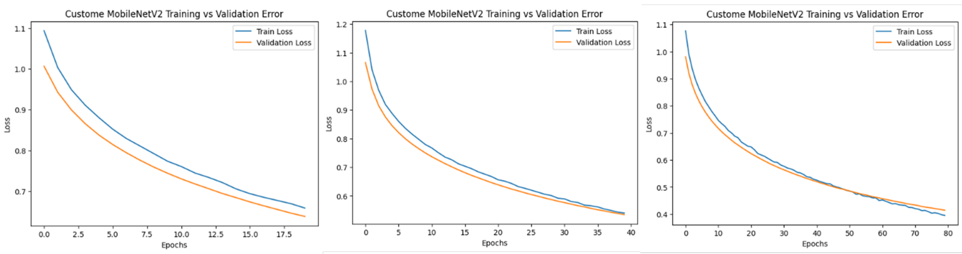

Figure 6: Train/Test Loss Curve for MobileNetV2

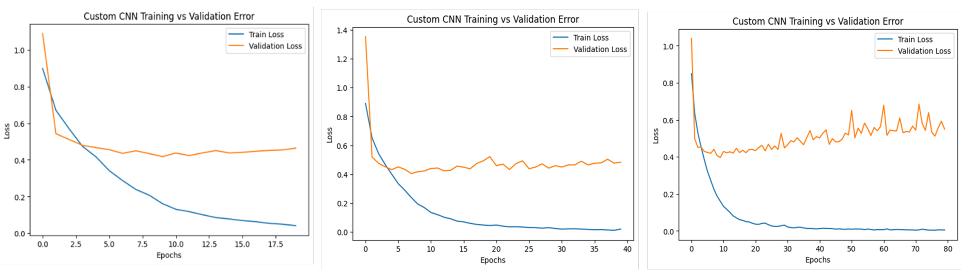

Figure 7: Train/Test Loss Curve for Custom CNN

As we shown in table1, the results of MobileNetV2 was showing a good performance by this lightweight model, after 80 epochs applied, the accuracy was equals to 83.72%

percent, which is showing a good performance by MobileNetV2, however, we can reach higher than this if we improve the models compilation steps, such as modifying the learning rate carefully, or by increasing the number of epochs, will grant reaching a much higher accuracy. The reason behind agreeing on MobileNetV2 has a good performance is that the flow of error for both training and testing was somehow the same, and there is no indication of overfitting sign found in the plot for all epochs trials. Moreover, the model indicates better generalization on the unseen data. In addition, what shows that the model is showing a great performance is that the accuracy of training increases slightly each time we increase the epoch, and the loss decreases approximately by 1% each time. See Fig.6

The above images in Fig.7, are showing how Custom-CNN behaves on the images. the accuracy of the model on test data was increasing silightly, and the accuracy on 80 epochs it was 85.22%, which is higher than the highest accuracy of MobileNetV2 by 1%. However, as we increase the epochs, the loss during the prediction is increasing, while the training loss is decreasing. This let us investigate that this is a sign of overfitting in the future if we increase the epoch. From the plots of each epoch, we can infer how the model test loss is not stable if the epoch starts from 20+ epochs and above. CNN models in its nature are complex deep networks, which require a huge amount of fundus images to train them, and get a higher accuracy. Basically, we can conclude that CNN model gets a higher accuracy than mobileNetV2, but it might be overfit and have a worser performance if we increase epochs, and due to its complex architecture. Also, it is not generalizing well on unseen data. On the other hand, mobileNetV2 was performing better every epoch, since we have increased accuracy, and lower testing errors. It has a better generalization on the unseen data.
See the Blow-the-Live prediction for both MobileNetV2 fig.8, and CNN fig.9 applied.

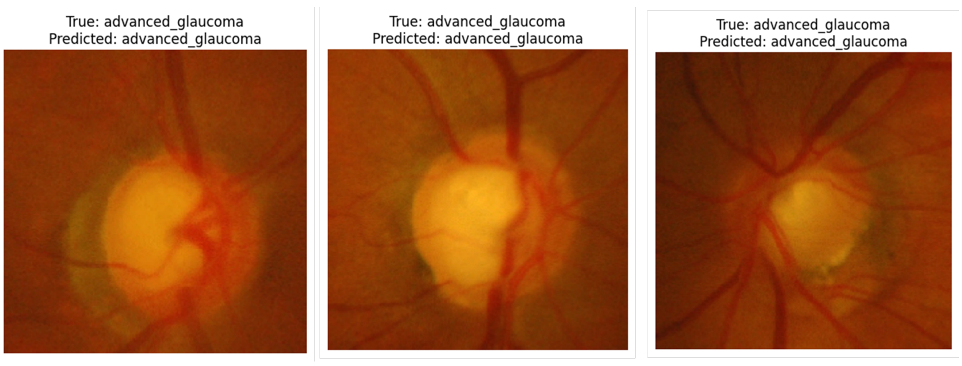

Figure 8: MobileNetV2 live predictions for 20, 40, and 80 epochs.

Figure 9: CNN live predictions for 20, 40, and 80 epochs.

4.2 Ablation Study

The studies have been done in this area of Glaucoma detection and classification of images in minor. For our work, we have used CNN, which is customized, and MobileNetV2, which has not been seen in any previous works, and we have got considerably good results. Also, the idea of applying for multi-classification is not seen in most of pervious works, so this area needs to be investigated more in future.

5 Extended Contributions

The studies have been done in this area of Glaucoma detection and classification of images in minor. For our work, we have used CNN, which is customized, and MobileNetV2, which has not been seen in any previous works, and we have got considerably good results. Also, the idea of applying for multi-classification is not seen in most of pervious works, so this area needs to be investigated more in future.

6 Discussion and Conclusion

Using deep learning neural networks in the field of medicine and healthcare will contribute dramatically to changing the way of diagnosis diseases, especially the retinal diseases such as the glaucoma cases, which it is not easy to build a strong, and accurate diagnosis model. This research tackled the problem of building a multi-classification deep network for predicting glaucoma cases, in order to predict normal, early glaucoma, or advanced glaucoma. We have introduced custom-CNN, and MobileNetV2. The custom-CNN has scored after 80 epochs 85.22%. However, the testing loss was increasing each time we increase the number of epochs, compared to the train loss was decreasing, this is a sign of overfitting might happen if we increase the epochs, and it is not generalizing well on unseen data. However, MobileNetV2 was showing better performance, and generalized

well on the unseen data. As the testing accuracy increases, the test loss decreases gradually with the train loss.
As a future contribution to this work, we will work on improving the current models for Custom-CNN and MobileNetV2, by improving the internal architecture, enhancing the gradient updates, and using other optimizers. Also, improving the hyperparameters, and suggesting new additional features, and applying new deep networks. Moreover, future research should focus on developing a generalizable model to classify and predict more retinal diseases

7 References

References

[1] C. J. Flaxel, R. A. Adelman, S. T. Bailey, A. Fawzi, J. I. Lim, and G.-S. Ying, “Age-related macular degeneration preferred practice pattern®,” *Ophthalmology*, vol. 127, no. 1, pp. P1–P65, Jan. 2020.

[2] Sightsavers, “Glaucoma: causes, symptoms and treatment — sightsavers,” https://www.sightsavers.org/eye-conditions/glaucoma/, accessed: Sep. 17, 2025.

[3] G. Ali, A. Dastgir, M. W. Iqbal, M. Anwar, and M. Faheem, “A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 11, pp. 341–350, 2023.

[4] B. R. Sathishkumar, S. Sudevan, M. Vignesh, and R. V. Kumar, “Ai-driven glaucoma detection using deep learning,” in *2025 International Conference on Sensors and Related Networks (SENNET) Special Focus on Digital Healthcare*, Jul. 2025, pp. 1–6.

[5] M. J. M. Zedan, M. A. Zulkifley, B. A. Saeed, and S. R. Abdani, “Deep learning-based model benchmarking of glaucoma segmentation using a novel ibn al-haitham fundus image dataset,” in *2024 International Conference on Smart Computing, IoT and Machine Learning (SIML 2024)*, 2024, pp. 244–249.

[6] Y. Zhou *et al.*, “A foundation model for generalizable disease detection from retinal images,” *Nature*, vol. 622, no. 7981, pp. 156–163, Oct. 2023.

[7] “Artificial neural network - an overview — sciencedirect topics,” https://www.sciencedirect.com/topics/earth-and-planetary-sciences/artificial-neural-network, accessed: Oct. 25, 2025.

[8] "Introduction to convolution neural network - geeksforgeeks," https://www.geeksforgeeks.org/machine-learning/introduction-convolution-neural-network/, accessed: Oct. 08, 2025.

[9] "What is mobilenet v2? - geeksforgeeks," https://www.geeksforgeeks.org/computer-vision/what-is-mobilenet-v2/, accessed: Oct. 08, 2025.

[10] P. Kalgaonkar and M. El-Sharkawy, "Condensenextv2: Light-weight modern image classifier utilizing self-querying augmentation policies," *Journal of Low Power Electronics and Applications*, vol. 12, no. 1, Mar. 2022.

[11] "What is batch normalization in deep learning? - geeksforgeeks," https://www.geeksforgeeks.org/deep-learning/what-is-batch-normalization-in-deep-learning/, accessed: Oct. 08, 2025.

[12] "Machine learn for glaucoma - harvard dataverse," https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/1YRRAC, accessed: Oct. 08, 2025.

[13] "Class distribution - an overview — sciencedirect topics," https://www.sciencedirect.com/topics/computer-science/class-distribution, accessed: Oct. 25, 2025.

[14] S. Kumar, P. Asiamah, O. Jolaoso, and U. Esiowu, "Enhancing image classification with augmentation: Data augmentation techniques for improved image classification," https://arxiv.org/pdf/2502.18691v1, feb. 2025, Accessed: Oct. 25, 2025.

[15] M. Kumar and V. Bhardwaj, "Evaluating label encoding and preprocessing techniques for breast cancer prediction using machine learning algorithms," *International Journal of Computational Intelligence Systems*, vol. 18, no. 1, pp. 1–35, Dec. 2025.

[16] "Relu activation function in deep learning - geeksforgeeks," https://www.geeksforgeeks.org/deep-learning/relu-activation-function-in-deep-learning/, accessed: Oct. 10, 2025.

[17] M. Kwon, S. Jeong, J. Choi, G. Lee, and K. Jhang, "Smooth categorical cross-entropy: An improved approach to overcome the limitations of cce," in *Proceedings of the IEEE International Conference on Big Data and Smart Computing (BIGCOMP)*, no. 2025, 2025, pp. 96–102.

[18] "Categorical cross-entropy in multi-class classification - geeksforgeeks," https://www.geeksforgeeks.org/deep-learning/categorical-cross-entropy-in-multi-class-classification/, accessed: Oct. 08, 2025.

[19] "Mini-batch gradient descent in deep learning - geeksforgeeks," https://www.geeksforgeeks.org/deep-learning/mini-batch-gradient-descent-in-deep-learning/, accessed: Oct. 13, 2025.

[20] Z. Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, Jan. 2019.