

# Automated detection of Glaucoma Using SegNet and Ensemble Learning



By

Ayesha Naeem

(Registration No: 00000329222)

Department of Computer Science

School of Electrical Engineering and Computer Sciences

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)

# Automated detection of Glaucoma Using SegNet and Ensemble Learning



By

Ayesha Naeem

(Registration No: 00000329222)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in

Information Technology

Supervisor: Dr. Farzana Jabeen

Co Supervisor : Dr. Mehvish Rashid

School of Electrical Engineering and Computer Sciences

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)


## THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis entitled "Automated detection of Glaucoma Using SegNet and Ensemble Learning" written by Ayesha Naeem, (Registration No 00000329222), of SEECs has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.


Signature: \_\_\_\_\_  \_\_\_\_\_

Name of Advisor: \_\_\_\_\_ **Dr Farzana Jabeen** \_\_\_\_\_

Date: \_\_\_\_\_ **26-Jun-2024** \_\_\_\_\_

HoD/Associate Dean: \_\_\_\_\_  \_\_\_\_\_

Date: \_\_\_\_\_ **26-Jun-2024** \_\_\_\_\_

Signature (Dean/Principal): \_\_\_\_\_  \_\_\_\_\_

Date: \_\_\_\_\_ **26-Jun-2024** \_\_\_\_\_

FORM TH-4

# National University of Sciences & Technology

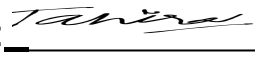
## MASTER THESIS WORK

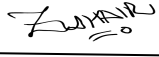
We hereby recommend that the dissertation prepared under our supervision by: (Student Name & Reg. #) Ayesha Naeem [00000329222]


Titled: Automated detection of Glaucoma Using SegNet and Ensemble Learning

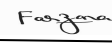
be accepted in partial fulfillment of the requirements for the award of Master of Science (Information Technology) degree.

### Examination Committee Members

1. Name: Tahira Lashari Signature:   
31-Jul-2024 9:59 AM

2. Name: Zuhair Zafar Signature:   
31-Jul-2024 9:59 AM

3. Name: Mehvish Rashid Signature:   
31-Jul-2024 9:59 AM

Supervisor's name: Farzana Jabeen Signature:   
31-Jul-2024 3:21 PM



Arham Muslim  
HoD / Associate Dean

02-August-2024

Date

### COUNTERSIGNED

05-August-2024

Date



Muhammad Ajmal Khan  
Principal

## Approval

It is certified that the contents and form of the thesis entitled "Automated detection of Glaucoma Using SegNet and Ensemble Learning" submitted by Ayesha Naeem have been found satisfactory for the requirement of the degree

Advisor : Dr Farzana Jabeen

Signature: Farzana

Date: 26-Jun-2024

Committee Member 1:Dr. Tahira Lashari

Signature: Tahira

Date: 26-Jun-2024

Committee Member 2:Dr Zuhair Zafar

Signature: Zuhair

Date: 26-Jun-2024

Co-Advisor: Dr Mehvish Rashid


Signature: Mehvish

Date: 27-Jun-2024

## Certificate of Originality

I hereby declare that this submission titled "Automated detection of Glaucoma Using SegNet and Ensemble Learning" is my own work. To the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics, which has been acknowledged. I also verified the originality of contents through plagiarism software.

Student Name:Ayesha Naeem

Student Signature:  \_\_\_\_\_

## AUTHOR'S DECLARATION

I Ayesha Naeem hereby state that my MS thesis titled “Automated Detection of Glaucoma Using SegNet and Ensemble Learning” is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Student Signature: \_\_\_\_\_ 

Name: Ayesha Naeem

Date: 12-8-2024

### **Certificate for Plagiarism**

It is certified that PhD/M.Phil/MS Thesis Titled "Automated detection of Glaucoma Using SegNet and Ensemble Learning" by Ayesha Naeem has been examined by us. We undertake the follows:

- a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.
- b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.
- c. There is no fabrication of data or results which have been compiled/analyzed.
- d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.
- e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

#### **Name & Signature of Supervisor**

Dr Farzana Jabeen

Signature : Farzana



## DEDICATION

I dedicate this thesis to my supporting parents along with my siblings whose unwavering support and love kept me going through every phase of life.

## **ACKNOWLEDGEMENTS**

Paying my utmost gratitude to Almighty Allah who provided me the strength during my vulnerable time of life, when He Almighty showered His countless blessings and supported me, no matter how much I thank Him, it shall never be enough! Dedicated to all of my family members who stayed by my side, never relinquishing their belief in me when I doubted my own abilities. Finally, I would like to thank my supervisor Dr. Farzana Jabeen whose utmost support and encouragement helped me in this journey.

# Contents

<b>LIST OF TABLES</b>	<b>V</b>
<b>LIST OF FIGURES</b>	<b>VI</b>
<b>ABSTRACT</b>	<b>VII</b>
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Traditional methods for detecting glaucoma . . . . .	2
1.2.1 Types of Eye Tests . . . . .	2
1.2.2 Problems with Traditional Glaucoma Detection . . . . .	3
1.3 Motivation . . . . .	4
1.4 Purpose of Research . . . . .	4
1.5 Thesis Composition . . . . .	4
<b>2 LITERATURE REVIEW</b>	<b>6</b>
2.1 Glaucoma Identification by Machine Learning . . . . .	6
2.2 Detection Using Neural Networks . . . . .	11
2.3 Detection Using Segmentation and Transfer Learning . . . . .	11
2.3.1 Summary of Existing Literature . . . . .	14
<b>3 MATERIALS AND METHODS</b>	<b>16</b>
3.1 Dataset . . . . .	16
3.2 Proposed Methodology . . . . .	16
3.2.1 Dataset Acquisition . . . . .	17
3.2.2 Pre-processing . . . . .	17
3.2.3 Segmentation Module . . . . .	17
3.2.3.1 Segmentation . . . . .	18

3.2.3.2	Segmentation Module of Proposed Work . . . . .	18
3.2.3.3	Architecture of Generalized SegNet . . . . .	19
3.2.4	Step-by-step Algorithm . . . . .	20
3.2.5	Classification Module . . . . .	22
3.2.5.1	CNN Structure . . . . .	22
3.2.5.2	VGG-16 Framework . . . . .	24
3.2.5.3	Algorithm for Classification . . . . .	26
3.2.6	Step-by-Step Algorithm . . . . .	26
3.2.7	Mathematical Representation . . . . .	27
<b>4</b>	<b>RESULTS</b>	<b>28</b>
4.1	Segmentation Results . . . . .	28
4.1.1	IOU(Intersection Over Union) . . . . .	28
4.2	Classification Results for Generalized CNN and VGG-16 . . . . .	29
4.2.1	Training Accuracy . . . . .	29
4.2.2	Validation Accuracy . . . . .	29
4.2.3	Training Accuracy and Validation Accuracy for generalized CNN	30
4.2.4	Training Loss . . . . .	30
4.2.5	Validation Loss . . . . .	30
4.2.6	Training and Validation Loss for generalized CNN . . . . .	31
4.2.7	Training and Validation Accuracy of VGG-16 Model . . . . .	31
4.2.7.1	Training Accuracy . . . . .	31
4.2.7.2	Validation Accuracy . . . . .	31
4.2.8	Training and Validation Loss of VGG-16 . . . . .	32
4.2.8.1	Training Loss . . . . .	32
4.2.8.2	Validation Loss . . . . .	32
4.3	Final Results(Ensemble Learning) . . . . .	33
4.3.1	Training Accuracy . . . . .	33
4.3.2	Validation Accuracy . . . . .	33
4.3.3	Training Loss . . . . .	34
4.3.4	Validation Loss . . . . .	34
4.4	Confusion Matrix Analysis . . . . .	35
4.4.1	True Negative(TN) . . . . .	35
4.4.2	False Positive(FP) . . . . .	35
4.4.3	False Negative(FN) . . . . .	36
4.4.4	True Positive(TP) . . . . .	36

4.5	Evaluation Matrices . . . . .	36
4.5.1	Accuracy . . . . .	37
4.5.2	Precision(Positive Predicted Value) . . . . .	37
4.5.3	Recall(Sensitivity or True Positive Rate) . . . . .	38
4.5.4	F1 Score(Harmonic Mean of Precision and Recall) . . . . .	38
4.5.5	Specificity(True Negative Rate) . . . . .	38
4.5.6	Comparison with Advanced Approaches . . . . .	39
4.5.7	Obtained Results . . . . .	40
<b>5</b>	<b>CONCLUSION AND FUTURE DIRECTIONS</b>	<b>41</b>
5.1	Conclusion . . . . .	41
5.2	Future Direction . . . . .	41
	<b>BIBLIOGRAPHY</b>	<b>43</b>

# List of Tables

2.1	Comparison of various methods for glaucoma detection . . . . .	14
3.1	Architecture of Generalized SegNet . . . . .	20
3.2	CNN Architecture Details . . . . .	23
3.3	Layer-wise Architecture of the VGG16 Model . . . . .	25
4.1	Comparison with other approaches . . . . .	40
4.2	Performance comparison of different models . . . . .	40

# List of Figures

3.1	Proposed Methodology . . . . .	17
4.1	Training and Validation IoU . . . . .	29
4.2	Training and Validation Accuracy(CNN) . . . . .	30
4.3	Training and Validation loss(CNN) . . . . .	31
4.4	Training and Validation Accuracy(VGG-16) . . . . .	32
4.5	Training and Validation Loss(VGG-16) . . . . .	33
4.6	Training and Validation Accuracy of Ensemble Model . . . . .	34
4.7	Training and Validation Loss of Ensemble Model . . . . .	35
4.8	Confusion Matrix . . . . .	36

# Abstract

Glaucoma is a neurodegenerative eye disease which occurs due to the elevated IOP (Interocular Pressure) that builds in the eye. The aqueous humour fluid in the chamber of eye presses the visual nerve due to which the condition of optic nerve deteriorates causing permanent blindness. Traditional methods such as gonioscopy, tonometry, dilated eye exam for the diagnosis is time consuming, expensive and require multiple visits to the hospital. Due to this, detection at early stage is compromised. With the advent in deep learning due to its pattern recognition capabilities, it is quite commonly used in early prediction of ocular diseases, glaucoma being one of them. This research focuses on developing a new methodology through which glaucoma can be detected using coloured fundus retinal images. The system consists of two basic modules: first segmenting the optic disc using a segmentation model i.e. customised Segnet being followed by a classification architecture that consists of an ensemble of VGG-16 and a customized CNN. The results have shown that the proposed model provides improved accuracy along with better sensitivity, specificity as compared to other existing architectures giving better generalization results in less time using less memory and computationally intensive system aids in early detection.

**Keywords:** Classification, Glaucoma Detection, Optic disc, Segmentation, VGG-16



# Chapter 1

## INTRODUCTION

Glaucoma is described as a category of gradual visual nerve degeneration eye condition caused by atrophy of visual nerve cells present in retina and degradation caused to RFNL (Retinal Nerve Fibre Layer) that results in permanent changes in the anatomy of ONH (Optical Nerve Head). Glaucoma can be classified into two major kinds. One is primary glaucoma and the second is secondary type of glaucoma. Based upon the anatomy of the condition they are further subdivided into Open Angle Glaucoma and Angle Closure Glaucoma. In first type of glaucoma there are mostly no identifiable symptoms seen in the patient whereas in secondary glaucoma there are several factors that can be identified that includes abnormal range of Inter Ocular Pressure (IOP) as well as there is visible damage seen in Optic Nerve Head (ONH). It is the leaving point of retinal nerve cells present inside the fundus of the eye. Primary open angle glaucoma is the type of chronic simple glaucoma. Secondary-open angle as well as normal tension glaucoma (NTG) is also included in its forms. In POAG optic nerve degrades over time due to increase in IOP. The cause of NTG is degradation of optic nerve fibres despite having normal IOP whereas in secondary-open angle glaucoma can be caused by both of these reasons. The opening between the cornea and the iris of the eye is responsible for the cause of Angle-closure glaucoma. The number of kinds of it is two. One is known as primary angle-closure glaucoma which is considered as an acute type because of abrupt increase in IOP and chronic due to slow increase in IOP which cannot be identified at first and secondary-angle closure glaucoma. Angle-closure glaucoma poses a higher risk of blindness as compared to open-angle glaucoma even though it is less prevalent.

### 1.1 Background

In the world, a significant reason of lifelong blindness that happens to patients is Glaucoma[1]. A survey was carried out by World Health Organization (WHO) which

revealed that the percentage at which glaucoma happens is 3.54% for people aged between 40 to 80 with people over 40 being high risk to developing glaucoma[2]. Glaucoma is mostly susceptible in patients whose age is above 60, those who have family history of glaucoma, diabetes, raised blood pressure, near and far sightedness or individuals using steroid medication for a long time or occurrence of any eye injury in the past can lead to glaucoma damage.111.8 million person in the world population are at the risk of contracting glaucoma by the year 2040[3]. One of the concerning reasons is that this ocular disease is asymptomatic in nature means its symptoms are not really visible in patients unless it has been progressed to a later stage.[4]This is the reason glaucoma is also being called as sneak thief of vision. Glaucoma is also responsible for the increasing health economy burden of the evolving countries as by time it is likely that the population of the country will grow old and this will result in the number of people that will contract this disease. Detection of glaucoma further complicates this issue as there are fewer glaucoma specialists as compared to the increasing ratio of patients. Therefore it is a very crucial task that glaucoma detection must be made better by incorporating such practices that will aid to decrease the health care stress that is caused by glaucoma.[4]

## 1.2 Traditional methods for detecting glaucoma

In order to rule out that a patient has glaucoma or not, a lot of clinical methods have been used in the initial stages that are:

### 1.2.1 Types of Eye Tests

The ophthalmologist performs one or more of these tests in order to detect glaucoma.

#### **Gonioscopy:**

Gonioscopy is known as the gold standard for the identification of type of glaucoma. It is the part of evaluating glaucomatous eyes including other procedures. It is performed using a hand-held convex lens or prism for looking at the back part of the eye.[5]

#### **Ophthalmoscopy:**

It is a test that is used to inspect any fluctuations in physical form and colour that occurs in optic nerve. If the IOP is higher than normal range and there is some abnormality seen in the optic nerve, more tests are advised by practitioners.[6]

#### **Tonometry:**

Tonometry is the standard test for the measurement of interocular pressure (IOP) of the eye. IOP indicates the pressure of ocular fluid that is present in eye. Normal range of IOP is between 11 mm Hg to 21 mm Hg. Pressure higher than this is the indicator of glaucoma.[7]

#### **Optical Coherence Tomography:**

It is a non-surgical eye examination test which makes use of interferometry with low coherence beam obtaining high resolution 2D images of internal structure of tissue. CT has been used in vitro for the imaging of retinal structure and coronary artery.[8]

**Slit lamp examination:**

The slit lamp exam is a procedure for diagnosing the overall condition of the eyes used by eye specialists. It helps in evaluating the eye anatomy thus giving a hint of any ocular disease including glaucoma and cataracts.[9]

**Dilated eye exam:**

In this method the ophthalmologist first dilates the pupil by using specific drops. This expands the pupil of the eye and the ophthalmologist can easily determine certain features of the eye namely retinal structure, optic nerve head, its size, hue and also the structure of blood vessels. Upon observing any physical damage, it can be assessed that whether a patient has glaucoma or not.[10]

**Perimetry:**

It is a standard test that is carried out for the assessment of eyes visual field. This visual field test checks the peripheral vision of the eyes and rules out any blind spots in the peripheral vision. This test plays a pivotal role in detecting early glaucoma as in this ocular pathology peripheral vision is affected first.[11]

## 1.2.2 Problems with Traditional Glaucoma Detection

There are some problems that are associated with detection of Glaucoma using that are:

**Increase in Socio-Economic Burden**

It increases the socio-economic burden, augmenting societal load. Glaucoma management also has significant impact on medical costs in the society provided that the patients require medication throughout their life. Thus, this implies the lasting impact of glaucoma care on the society, highlighting the need to relieve the medical infrastructure as well as patients by introducing glaucoma management practices that are efficient in the long run.

**Time Consuming**

All of these methods require manual assessment that is time taking and requires multiple visits to the ophthalmologist. Therefore, the need for automated detection of glaucoma at early stages is necessary.

## 1.3 Motivation

Deep learning has been profoundly used for pattern recognition. Through years deep learning has shown worthwhile progress in the area of artificial intelligence. It exhibits notable success in various domains including healthcare informatics, computer vision and image analysis. Deep learning has also shown advancements in medical image analysis by identifying underlying patterns responsible for disease detection [12]. Due to this reason, in the area of ophthalmology there is quite common inclusion of deep learning techniques specifically in glaucoma detection at early stages. Therefore, the need for automated detection of glaucoma at early stages is necessary.

## 1.4 Purpose of Research

Although there are methods available for diagnosing glaucoma but they are computationally expensive and memory intensive. Using private datasets requires external validation. Existing segmentation architectures requires training on large number of annotated images. Existing architectures shows promising results but none have been used in clinical settings. The main purpose of the proposed work is listed below:

- Implementing State of the art Feature Extraction to enhance diagnoses Glaucoma Detection
- A resource efficient methodology for predicting Glaucoma
- Ensemble Learning based glaucoma Classification to enhance predictions
- Portability to available datasets in medical organizations
- Facilitating Ophthalmologists to detect and prescribe for glaucoma in early stages

## 1.5 Thesis Composition

The total number of chapters included in this study are 5. Chapter 1 introduces glaucoma and its types and the features that are responsible for the occurrence of the disease. Its socio-economic impact and the traditional methods that helps in identification of glaucoma and the necessity for early glaucoma prediction. In Chapter 2, an extensive literature is covered that elaborates on different methods of machine learning that have been employed to predict glaucoma, deep learning as well as segmentation techniques that are given for the proposed problem. In Chapter 3 details about the

proposed methodology is presented, the segmentation module and the classification module are discussed in detail. Moreover, the algorithms for these modules are also given. Chapter 4. includes the results of the classification and segmentation modules including the evaluation metrics. Finally, the conclusion, limitations along with future direction is presented in Chapter 5.

# Chapter 2

## LITERATURE REVIEW

Numerous studies have been conducted to develop systems that are helpful in identification and classification of glaucoma, employing a diverse range of research methodologies and computational techniques. Glaucoma can also be detected using electronic health record of patients when they visit the ophthalmologist for normal eye checkup. The electronic health record consists of patient's age, gender, any previous eye disorder, eye sight information and diseases that can cause glaucoma such as high blood pressure, hypertension and auto-immune diseases. There are various imaging modalities that exists in ophthalmology that are used in order to detect glaucoma. Two commonly used modalities are OCT (Optical Coherence Tomography) and the other is Fundus Imaging. It is also known as fundoscopy. One of the most used in research in fundoscopy because it is not that much expensive and most of the datasets that are available for glaucoma detection consists of fundus images. The posterior part of eye can be seen by examining the retinal fundus of eye. It shows clear representation of optical disc including optic cup, blood vessels, ONH, neuro-retinal rim that is the width measured between the disc of fundus and optic cup. The existing literature that uses the retinal fundus photos in order to identify glaucoma from them is given below:

### 2.1 Glaucoma Identification by Machine Learning

Most of the traditional based Machine learning models focus on separating disc of the fundus as well as cup that in present inside the disc as well as blood vessels to analyse any deformity in ONH of the fundus picture in order to detect glaucoma. In these methods those physical features are humanly crafted by experts and based on those decisions are made.

A detection method for glaucoma was presented that made use of the fundus images [13]. Early assessment and treatment are effectively possible and will aid doctors to make timely decision and hence perform medical treatment that is essential with the

help of presented method. The technique introduced is also inexpensive and is feasible for every patient. The features that are present in the fundus image of the eye and can be responsible for glaucoma is optical disc that consists of optical cup together with blood veins and capillaries. In the presented method, identification of glaucoma with the help of ratio of cup and disc that is enlarged in case of glaucoma and it is also analysed by the alignment of vessels in the retinal image. The mean as well as highest intensity value pixels in the grey scale image are employed in extraction of optic disc and optic cup. These are found using histogram technique. Thus, the boundaries are obtained and are used for constructing the circle that fits appropriately and in this way the radius of the optical disc as well as cup is achieved. After the computation of CDR, the normal image can be distinguished from the image having glaucoma. If the CDR of image is greater than the specific threshold, it means the image is the part of positive instance that is having glaucoma. Contrary to it, the image belongs to the negative class i.e. the photograph is healthy. Glaucoma is predicted by the placement of the vessels in the image that is obtained with help of system. The system extracts the blood vessels and through their orientation glaucoma is identified. The presented system was able to achieve specificity of 95% and sensitivity of 82%.

A novel technique presented that identifies glaucoma incorporating Variational mode decomposition (VMD) of retinal fundus images. The decomposition mode was used in a repetitive manner until desired features were obtained [14]. The retinal fundus image was broken down into various parts or components in order to obtain the attributes. Several features were extracted through the layers produced by decomposition process that includes Kapoor entropy, Renyi entropy and extraction of various recurring geometric patterns that are represented by fractal dimensions. They employed signal processing in order to obtain the important information although the signal being unstable. This method is non-recursive. VMD has certain advantage over EMD method and can be used in place of EMD as it is more resilient to capture noise in an image, and performs more better. The distinguishable attributes that are necessary for the purpose of classification were opted by ReliefF algorithm. These particular features are then used as input to the support vector machine in order to perform classification. Eye specialist can make more robust decisions that whether a patient has glaucoma or not with the help of this system. This study incorporated 488 retinal images in which 244 were normal images and 244 images were glaucomatous. On the fundus images, contrast-limited-adaptive-histogram-equalization (CLAHE) was employed for the reduction of uneven brightness that also improved the vast range of retinal fundus images. The accuracy achieved was 95.19% by using 3-fold cross-validation. Similarly, 10-fold cross-verification resulted in accuracy of 94.79% but the classification accuracy of the method may decrease if employed on a big dataset.

A system was developed in order to detect the disease using features presented on bar graph of the retinal fundus images[15]. An integration of magnitude and phase features was used in this study for the purpose of detection of glaucoma as their com-

bination provides necessary information regarding frequency. The computation of their histogram features was also done. These features or attributes were extracted by employing Local binary patterns and an algorithm named Duagman's algorithm. Gabor Filter is used for edge detection in an image that purely depends upon the occurrence and alignment of the specific image. The optical disc was extracted by the implementation of a 2D Gabor filter on retinal fundus image. Every pixel location of the grey-scale image was contrasted with its adjacent pixel and converting it to a binary representation by employing Local binary patterns. In order to recognise glaucoma, the Euclidean space that was present in the middle of feature attributes was calculated. The evaluation metrics for measuring the efficiency of the proposed system were sensitivity, specificity as well as accuracy and they were compared with high order spectra features in order to analyse them effectively. The sensitivity, specificity and accuracy achieved by the detection system was 95.45%. This system was also able to make faster decisions and the processing time of the presented technique was also decreased as compared to other methods. They used a total of 44 fundus images. 22 images belonged to the disease class and the remaining 22 belonged to non-disease. An implanting technique was introduced in order to obtain an OD image that consists of specific region of localization. A level-set procedure was include in the presented segmentation technique[16].The opinion of the ophthalmologists was taken into consideration for the final computation of accuracy. Their combined opinion on the aggregate of images made the final decision. Evaluation was carried out using RIGA (retinal images-for-analysis-of-glaucoma) using 550 images. It achieved an accuracy of 83.9%. Similarly, another hybrid segmentation approach was produced by [17] in which the vessels in the fundus of the eye were extracted from retinal images making the diagnosing procedure more accurate. The algorithm introduced is the combination of morphological operations, transformation and an enhancement algorithm. A morphological technique is incorporated so that the vessels are clearly visible. The validation of the results was carried out by comparing the algorithm results with reference label of HRF dataset. The sensitivity, specificity and accuracy achieved was 94.11%,95.34% and 95.28% respectively.

For the classification of glaucoma ,a detection system in [18] was proposed by incorporating a statistical approach and used k-nearest neighbour (knn) algorithm. The first step was extracting the feature attribute responsible Feature extraction was performed and a subsequent step of feature selection was applied incorporating correlation procedure. Then for classification knn was used. 84 retinal images for the purpose of evaluation were utilized in this work. It had 41 positive cases. The negative cases(non-glaucoma) were 43. The achieved accuracy was 95.24%. A novel approach was presented in [19] in which quasi-bivariate variational mode (QB-VMD) was implemented. A reduction process was applied on fundus images resulting in sub-band images and consequently 70 features were extracted from those images and normalization of those images was carried out using ReliefF method. Dimensionality reduction was performed



using SVD (Singular Value Decomposition) and classification SVM. Three-fold cross-verification resulted in an accuracy of 85.94%. Ten-fold cross validation resulted in an accuracy of 86.13%. [20] presented a system using Discrete wavelet transforms and Empirical wavelet transform and a hybrid concatenation method for increasing accuracy was introduced in this work. 14 features were extracted from concatenation method and classified by SVM. Ten-fold cross validation resulted in 83.57%, 86.40% and 80.80% of accuracy, sensitivity and specificity respectively. As these methods required explicit feature extraction methodologies, so with neural networks better performance was seen.

A system was created in [21] for glaucoma detection known as Glaucoma Detection System (GDS) in which combination of classifiers was used. The classifiers that were incorporated were Random Forest as well as Naive Bayes for the detection purpose. The vertical optical cup consisting of optical disc were obtained from the retinal images that were the input. After that the particular region of interest was extracted. ROI can be used to identify between particular points that can help in pattern recognition. In this specific study, the ROI was developed from optic disc and optic cup. From that region a feature known as energy feature was captured and that energy feature was stored in a data repository. The classification was done using both the classifiers separately. In the last step, the combination of Naive Bayes and Random Forest algorithm classified that which images are having glaucoma. The Glaucoma Detection System for the purpose of assessment, 50 fundus photographs were used which consisted of both glaucoma and non-glaucoma images. The accuracy achieved using Naive Bayes and Random Forest was 93 % and 95% respectively and the combined accuracy achieved was 98%.

A system was developed in [22] by utilizing primary as well as secondary features for glaucoma classification of retinal fundus images. The proposed method presents a unique algorithm that uses the integration of different features. The combination of structural features that are represented by optic disc, optic cup, optic nerve head and CDR are included. The non-structural features on the other hand represent change in colour, pixel intensities based on the different texture that is present throughout the retinal fundus image. The accuracy for the detection of glaucoma was elevated by using this combination. They also developed a third class other than positive and negative classes if there was an ambiguity in the prediction made by using the hybrid combination of features. The classification was done using Principal Component Analysis in which the significant feature vectors were drawn out from the retinal fundus images. Those feature vectors were then used in classification process. Support Vector Machine distinguished between the positive and the negative class by choosing the best fitting hyperplane. The introduced method in this work was evaluated with the help of two database that were local. The first dataset consisted of 50 images in which 15 fundus images were that of glaucoma and 35 were normal images that is not having glaucoma. The second database consisted of 100 images in which 26 fundus images were glauco-

matous and the remaining 74 images were non-glaucomatous. All these images were previously annotated by eye specialists and the comparison between calculated CDR was made. The aim of this work was to ensure that the diseases cases from the rural economy which have no available glaucoma care can take advantage of the system by referring them. The proposed system exhibited a high level of sensitivity which is good in detection systems. The presented system had a sensitivity of 100% and the achieved specificity was 87%.

A detection system using physical features was proposed in [23] that mainly focused on the optic nerve head region for detection. The proposed system was trained on both positive and negative classes that is fundus images having glaucoma and those not having glaucoma. The classification module was trained with one dataset and to assess the system the other was used as a benchmark dataset. In the proposed system different optical markers for the need of glaucoma prediction in fundus image are extracted. These markers or indicators are CDR, any defect in the retinal nerve fibre layer, physical transformation of optical cup as well as disc that is the change of the colour, size or intensity or any other physical changes. These are represented in a feature vector that is used for the classification of glaucoma. The quantity of retinal fundus images used in this study were 2252 which were gathered from two eye hospitals located in India. The system proposed by them resulted in the sensitivity of 71.16% and specificity of 71.17%.

Glaucoma detection system was proposed in [24] that was developed using a combination of structural and attributes that represents pattern characteristics. The proposed system consisted of two basic subsystems that were based on the type of features being incorporated. The first system was called hybrid structural feature set and the second one was called hybrid texture feature set. The first module used support vector machine for classifying between glaucoma and non-glaucoma based upon the structural features. The structural features included the CDR computation that is the ratio of vertical optical cup to disc, shape of the cup etc. Similarly, the decision for classification by the second module was also done using support vectors based on textural characteristics of the eye that are consistency of certain patterns in the image, homogeneity that is how much the pixels are uniform in a particular space as well as their intensities. In this work the extraction of rim vessels was done for the computation of space in which vessels are present in rim region of fundus, mean, standard deviation and variance was computed. Using Grey-level matrix, other features such as depression of optic disc due to deterioration of fibres were achieved. Discrete wavelet transform (DWT), Mean grey level were also used for extraction of some features. Prior to classification image preprocessing such as scaling was done and optic disc was segmented using Otsu's thresholding. The final decision was done based upon the combination of the both the subsystems. The detection system was evaluated on a dataset that was publicly available named as Glaucoma Database. This data repository consisted of 100 images and the ophthalmologists labelled them. These labels were then used in

comparison with the results obtained by the presented system. The HTF module was able to achieve a sensitivity of 94% using super-pixels attribute and the best accuracy and specificity was obtained using Wavelet features that is 91% and 92% respectively. The accuracy achieved was 94%, 92% specificity and 96% sensitivity of HSF module.

## 2.2 Detection Using Neural Networks

A deep learning architecture was introduced in [25] consisting of 6 layers consisting of 4 convolutional and two fully-connected layers. Dropout mechanism was used for the neurons having probability 0.5. Data augmentation techniques were employed to increase accuracy. The datasets on which experimentation was performed are ORIGA and SCES which consists of retinal fundus images. The obtained accuracy from this method was 83.1% and 88.7% for both datasets respectively.

Convolutional Neural Network was deployed in [26] for recognition of glaucoma. In order to compute ratio of vertical cup and disc the method of Sparse Differentiation Constrained Coding (SDC) was implemented that facilitated in the optical cup location. Using retinal photos to segment the parts of fundus such as optic cup as well as vertical disc, a technique known as super-pixel classification was utilized that is used to partition a particular region in the given image. Important features in images were with-drawn by implementing CNN from the given image. For the purpose of classification between glaucoma and non-glaucoma support vector machine was implemented. This study was evaluated on a private and public dataset that are PSGIMSR and HRF dataset. The sensitivity obtained by system was 86.58% and the specificity was 95.21% by using PSGIMR dataset whereas the sensitivity and specificity using HRF dataset was 100% and 97.04% respectively.

## 2.3 Detection Using Segmentation and Transfer Learning

Another approach that gained popularity in medical image analysis was the use of automated segmentation techniques such as U-net, AlexNet, MobileNet. With the advent of U-net, the segmentation approaches for performing medical image segmentation increased drastically. A diagnostic tool was proposed in [27] for glaucoma detection using extraction of disc as well as cup features in which retinal fundus images were included. For the segmentation of optical disc U-net was used in the first step followed by using a pre trained mobileNet architecture for classification between glaucoma and healthy eyes. The experimentation was performed on RIM-ONE and DRISHTI datasets. Their

generalized U-NET achieved specificity and sensitivity of 93.0% and 76.0%. The integral area obtained for mobileNet was 0.93. A 3D augmentation algorithm in [36] was developed in order to achieve a topographic map of the Optical Nerve Head (ONH) from 2D retinal images. Those 3D maps that gave a detailed representation of all the physical structure present in the fundus of the eye were then used to train AlexNet and VGG-16 for classification of glaucoma. The results showed that training using 3D augmented images yielded better performance accuracy of 94.3% as compared to 2D fundus images.

In [29], the authors proposed a study using enhanced U-net for segmentation in which for the segmentation of optic cup transfer learning methodology was combined along with U-Net. DenseNet-201 was implemented for feature extraction and thus classification was carried out. The accuracy obtained during training was 98.82%. Similarly, the accuracy achieved while testing was 96.90%. In [30] a less computationally intensive system was presented for glaucoma detection. For segmentation purpose YouOnlyLookOnce (YOLO) architecture was implemented and for classification MobileNet architecture was implemented. The proposed system achieved accuracy was 94.4% and the F1-score was 97.3%.

In [31], a system classifying glaucoma using U-net for image segmentation and for classification 3 supervised machine learning algorithms were used. The first step was the input image validation that was done using LeNet architecture because of its less time consuming ability. This made sure that only fundus images is being used as input to the system. The features that were used for the classification were CDR also known as vertical cup ratio, ISNT region (Inferior, Superior, Nasal, Temporal) that are located in NRR (Neuro-retinal rim) and blood vessels present in disc as well cup depression region. Optical cup was segmented by implementing U-net model that produces a binary image. ROI is obtained in this way and support vector machine is employed for classification, a deep network that consists of starting layer, 2 secret layers as well as a final layer and an adaboost classifier. The final decision of the classification was made on the ground of voting. 666 retinal fundus images used from 5 datasets that are available to public that are RIM-ONE, DRISHTI-GS, DRIONS-DB, JISEC, DRIVE and one private dataset obtained from a hospital. The accuracy obtained was 98.8%. The specificity of the presented system was 99%.

The authors in [32] developed a hybrid approach for the screening of glaucoma using deep learning. A combination of CNN, recurrent neural network (RNN) was developed for the purpose of detecting glaucoma. A deep convolutional network was used in combination of RNN-LSTM that can be also called as a technique of transfer learning. Preprocessing of the fundus images was carried out before segmentation. A technique was incorporated which hides the vessels present in the focal area of fundus. For preprocessing gaussian smoothing filter was used to reduce the unnecessary distortion from fundus image. After that optic cup segmentation was done using an algorithm named Modified Level Set. Feature extraction was performed to extract features that

were area of the disc, cup area, retinal nerve fibre layer thickness, shape, optical disc colour, cup colour. These features were then fed to the classifier to perform classification. The spatial and temporal features are obtained by using the combination of VGG16, ResNet 50 and LSTM were employed for the purpose of classifying glaucoma. The training plus validation was performed by including DRISHTI The total number of images used in this study were 661 out of which 224 images were non-glaucoma and 137 fundus images were of glaucoma. The purposed model was able to achieve 98.21% training accuracy and 96.34% of test accuracy.

A detection system that used Generative Adversarial Network(GAN) was presented in [33].The first step was image collecting .For this purpose RIM-ONE and DRISHTI dataset were used. 455 images of fundus are in the RIM-ONE dataset. 200 images in the dataset are those having glaucoma and the remaining 200 images are not having glaucoma.101 images of retinal fundus are in Drishti dataset. The number of images having glaucoma is 70 and those not having glaucoma are 31 images. Hole filling method was employed that was a preprocessing step. Then a generative adversarial network was trained for segmenting the optic disc from retinal image of fundus. The architecture of GANs is such that it consists of two neural networks. One is called generator that is used to receive the input image and also generates new data. The other neural network is known as discriminator that is used to differentiate between the generated samples and the actual data. Over the time both neural networks learn to perform better. In this work, a conditional GAN (cGAN) was employed in order to obtain desired OD image and for the input retinal fundus image was utilized. The generator employed was a CNN U-net and for the discriminator path PatchGan was used. The texture attributes were extracted by deploying index of taxonomic diversity. Classification was carried out by implementing Multiple -layer Perceptron (MLP), Sequence Minimal Optimization (SMO) including random forest algorithm. The accuracy attained was 77.9% for glaucoma classification.

A glaucoma detection system in [34] was deployed that consisted of two phase approach. The dataset used in this study was ORIGA that consists of retinal fundus images. The localization process extracted optic disc implementing Yolo-v4. A thresholding technique was incorporated to help in the process of OD localization supported in identifying region of interest. To smoothen optic disc boundary the implementation of morphological opening took place. Canny edge detector was used in order to discover the edges of the extracted optical disc. The obtained optic disc was used as an input to the ResNet-101 architecture in order to classify between positive and negative cases. They proposed a novel semi-automatic ground truth generation method that is used to provide accurate labels for the training of Yolo-v4 architecture. The accuracy achieved by using the system was 88.5%. Transfer Learning technique was used in [35] incorporating retinal images of fundus for the glaucoma diagnosis in an efficient manner. Diabetic retinopathy is another disease and the system is made to detect both eye disease. Retrained AlexNet was used for the recognition of glaucoma. This study

included datasets that are available to the public and some were accessed upon request. The achieved validation accuracy was different for the proposed model as network was trained to include the dataset. NetTransfer1, NetTransfer2, NetTransfer3, NetTransfer4 and NetTransfer5 achieved a validation accuracy of 94.3% ,91.8% ,89.7% ,93.1% and 92.1% respectively.

**Table 2.1.** Comparison of various methods for glaucoma detection

Reference	Architecture	Dataset	Segmentation	Accuracy
[16]	Level-set method	RIGA	Novel algorithm	83.9%
[17]	MSVE	HRF	-	95.28%
[18]	KNN	Private dataset	-	95.24%
[19]	SVD, SVM	RIM-ONE	-	85.94%, 86.13%
[27]	CNN with 6 layers	ORIGA, SECS	-	83.1%, 88.7%
[27]	mobileNet	RIMONE, DRISHTI	U-net	93%
[36]	AlexNet, VGG16	DRIONS-DB, HRF, RIM-ONE, DRISHTI	-	94.3%
[29]	DenseNet-201	ORIGA	Enhanced U-net	98.82%, 96.9%
[30]	mobileNet	LAG, ACRIMA, DRISHTI, HRF, RIMONE	YOLO	97.3%

### 2.3.1 Summary of Existing Literature

According to a thorough examination of the literature, table 2.1 compares many efforts that have already been made based on characteristics that are important for classification of glaucoma implementing machine learning, deep learning and pattern recognition using fundus images. Many studies include detection systems with segmentation, but most of these detection systems requires computationally and memory intensive hardware in order to deploy them and hence cannot be deployed in real world environment. Some of the systems have used the datasets that are private and are not publicly available due to which their generalization capabilities can be questioned. The proposed

system incorporates a generalized segmentation module that segments the optical disc and then the classification module classifies it as glaucoma or normal image. Generalization ability of the presented system is well on the unseen data, produces faster results and is less computationally intensive.

# Chapter 3

## MATERIALS AND METHODS

### 3.1 Dataset

An Online Retinal Fundus Image Database for Glaucoma Analysis and Research (ORIGA) was made a part of this study to train, test and evaluate the system[37]. A study was carried out in Singapore that took place in an eye care facility and the data of various images of retinal fundus were collected. It was developed in order to facilitate the researchers for their segmentation algorithms and is publicly available. It consists of 650 coloured fundus images that are labelled by specialists from Singapore Eye Institute. All the images are captured using fundus camera by the eye specialists. In those 650 images, the images that are having glaucoma (positive class) are 168. The negative class or the images that are not having the disease are 482. Along with those fundus image, their annotations are also provided within the dataset that are carefully considered by experts in this field and further we reviewed this from our own eye specialist. Through augmentation techniques 1019 images were finally used in the classification step.

### 3.2 Proposed Methodology

The propose work consists of five steps for glaucoma classification using retinal fundus images. The acquisition of dataset is the first step, the second step is pre-processing. In the third step the already processed fundus retinal images are used as an input to the segmentation module that generates the masks of optic disc. In the fourth step the faulty segmented masks are dropped. In the final steps those segmented masks are given to the classification models and this module gives the conclusive judgement that the input image is classified as the positive class or whether should be put in negative class (not-glaucoma) Thus, in this work there are two basic modules: first is a segmentation module which takes resized fundus image as input and gives binary



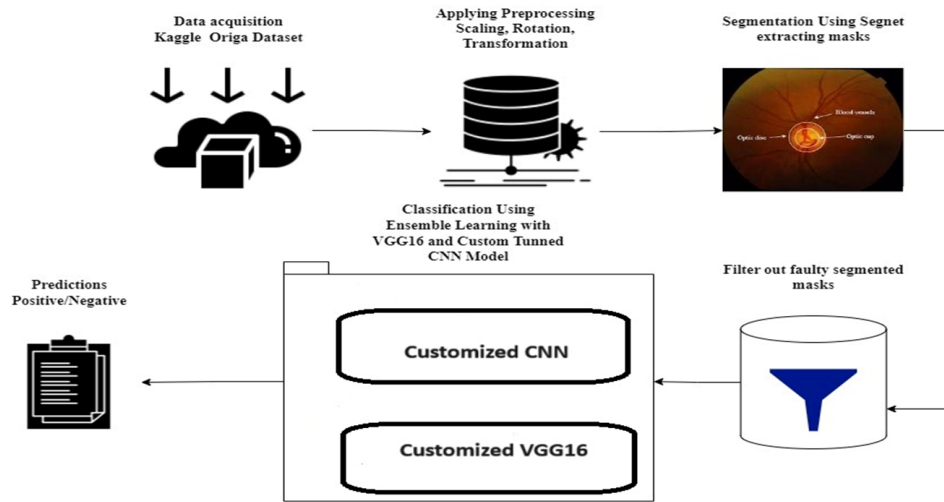


Figure 3.1: Proposed Methodology

segmented mask of optic disc, second is a classification module that takes those binary segmented masks as input and classifies them as glaucoma or non-glaucoma. Details about these steps is given below:

### 3.2.1 Dataset Acquisition

The first step is the acquisition of the dataset. The dataset used in this work was acquired from Kaggle.

### 3.2.2 Pre-processing

After acquiring the dataset in order to make it compatible with the system requirements and for better accuracy preprocessing techniques such as resizing of the image was applied. The input image was also resized from size of  $3072 \times 2042$  to  $224 \times 224$  in order for better segmentation performance. As the dataset set consisted of only 650 images, scaling technique was adopted, rotation of images, transformation of images was carried out to augment the given data in order to train the model better.

### 3.2.3 Segmentation Module

This section gives the details about the segmentation module used in this research work and also explains the algorithms.

### 3.2.3.1 Segmentation

Segmenting an image into various components is a part of image processing practiced in a variety of fields such as medical image processing, scene understanding, detection of and recognition of objects in images. Segmenting an image can be described as division of digital electronic image in different parts or segments so that the image is easier to understand and analyse. There are three basic kinds which can be used to represent segmentation of image that are described below:

**Instance Segmentation** This type of segmentation is used to differentiate between instances of the same class that is present in the image. It assigns different labels to each instance belonging to the same class. This allows to detect objects in a more precise manner.

**Panoptic Segmentation** Panoptic segmentation is basically a combination of both instance segmentation and semantic segmentation. It assigns labels to each pixel in an image, moreover the instances that belong to the similar category can also be differentiated by this segmentation. Thus, giving a combined outcome.

**Semantic Segmentation** In this type of segmentation, every pixel that is the part of the image is given a label through which its category of class is easily identified. The occurrences that are part of the same category of class cannot be identified by this type of segmentation.

### 3.2.3.2 Segmentation Module of Proposed Work

The segmentation is performed using SegNet. The basic architecture of SegNet was constructed with the objective of creating an algorithm that is able to perform segmentation implying a semantic method for segmenting. It can be described as deep neural model that is entirely convoluted. The construction of SegNet is such that it is made up by a layer that encodes the information. A decoder part of it upsamples. At the end is a classifying pixel-wise layer. The encoding block produces features map representation having decreased resolution and the decoder block converts those low-resolution feature maps back to input resolution. The decoder performs upsampling that is not linear by the use of pooling indices generated in the encoder block during its max-pooling operation. In this way decoder does not have to learn to upsample rather it uses those pooling indices. Thus saving inference time and utilizing little memory in contrast to various advanced architectures[38]. This work consists of a generalized SegNet specifically trained for the segmentation of optic disk from fundus image. As it consumes less memory and is not computationally intensive as compared to other segmentation architectures as U-net. It also has less trainable parameters as compared to other segmentation architectures. This section has the proposed segmentation module. The segmentation of the optical disc is the objective of this segmentation algorithm

for the given input fundus retinal image. The input is the original fundus image  $I$  of size  $3072 \times 2042$ . The size of the resized fundus image is  $s = x \times y$  where  $x$  and  $y$  are new height and width of fundus image.  $M$  is the segmented mask of optic disk.

### 3.2.3.3 Architecture of Generalized SegNet

The SegNet model is used for the purpose of semantic segmentation and it is constructed by incorporating an encoding block and a decoding block. Input image having dimensions of  $224 \times 224$  is passed to the input layer consisting of 3 colour channels (RGB). After that the encoder network begins with its first convolutional layer the output of which is 64 feature maps. Its subsequent layer is a maximum-pooling layer reducing the spatial dimensions by half producing an output of  $112 \times 112$  having 64 feature maps.

Following the maximum-pooling layer is the subsequent convolutional layer that produces 128 feature maps and it has 73,856 parameters. After that for dimensionality reduction another max-pooling layer **encoder\_pool2** is introduced in the encoder block that results in dimension of  $56 \times 56$  having 128 feature maps. To give the network more depth, third convolutional layer **encoder\_conv3** is introduced into the encoder network having 256 feature maps and the number of parameters at this point are 295,168. Another maximum-pooling layer is added after third convolution named **encoder\_pool3** the purpose to which is to further decrease dimension to  $28 \times 28$ . Then in the encoding part, final layer that is a convoluted layer is added named **encoder\_conv4** that produces 512 feature map representations having 1,180,160 parameters followed by a max-pooling layer encoder-pool4 that outputs  $14 \times 14$  feature maps. After that the decoder network begins which consists of the upsampling layer **decoder\_upsample** that is used to reconstruct the spatial dimension to  $28 \times 28$ . This upsampling layer is followed by a convolutional layer **decoder\_conv4** having 512 filters that outputs 512 feature maps with ReLU activation function having 2,359,808 parameters.

After that, an upsampling layer is introduced in the architecture **decoder\_upsample3** and performs the upsampling by doubling the dimensions to  $56 \times 56$  that is followed by a convolutional layer **decoder\_conv3** that is used to reduce the depth having 256 filters and the total number of parameters becomes 1,179,904. Further upsampling is done by introducing **decoder\_upsample2** that performs the upscaling and thus doubles the spatial dimension to  $112 \times 112$ . A subsequent convolutional layer **decoder\_conv2** consists of 128 features resulting in 128 feature maps and 295,040 parameters followed by a ReLU activation function. Upsampling layer at the end in the decoder network is used to restore the original input dimension that was  $224 \times 224$  that has a subsequent Conv2D layer named **decoder\_conv1** having 64 filters that outputs 64 feature maps with total 73,792 parameters.

At last, in the architecture there is **Conv2D** output layer. The end result in form of the segmentation map is given by the last convolutional layer. It consists of 3 filters whose size is 3\*3. It incorporates linear activation function with the computation of parameters that results in 195 parameters. SegNet model is comprised of 5,461,715, parameters all of which are trainable, allowing the model to effectively learn and perform pixel-wise classification for semantic segmentation tasks.

The following table gives a summary of the input layers, its final shape as well as all the quantity of parameters that were present in layers.

**Table 3.1.** Architecture of Generalized SegNet

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
encoder_conv1	(None, 224, 224, 64)	1792
encoder_pool1	(None, 112, 112, 64)	0
encoder_conv2	(None, 112, 112, 128)	73856
encoder_pool2	(None, 56, 56, 128)	0
encoder_conv3	(None, 56, 56, 256)	295168
encoder_pool3	(None, 28, 28, 256)	0
encoder_conv4	(None, 28, 28, 512)	1180160
encoder_pool4	(None, 14, 14, 512)	0
decoder_upsample4	(None, 28, 28, 512)	0
decoder_conv4	(None, 28, 28, 512)	2359808
decoder_upsample3	(None, 56, 56, 512)	0
decoder_conv3 (Conv2D)	(None, 56, 56, 256)	1179904
decoder_upsample2	(None, 112, 112, 256)	0
decoder_conv2 (Conv2D)	(None, 112, 112, 128)	295040
decoder_upsample1	(None, 224, 224, 128)	0
decoder_conv1 (Conv2D)	(None, 224, 224, 64)	73792
output (Conv2D)	(None, 224, 224, 3)	195

### 3.2.4 Step-by-step Algorithm

The elaboration of the proposed algorithm for segmentation is given in this section. The initial task is the acquisition of the original image. After that the image is resized. In the third step an input is given to the segmenting module that is the resized image

and in the fourth step the segmented mask is acquired. The steps for the extraction of mask from optic disc are given below in detail:

### 1. Loading Original Image as Input:

Original fundus retinal image in this step is used as an input that is acquired from the dataset. Let us denote the original fundus image as  $I$  where the dimension of  $I$  is  $3072 \times 2042$ :

$$I \in \mathbb{R}^{3072 \times 2042} \quad (3.1)$$

### 2. Resizing Image:

The actual image undergoes a resizing operation.  $I$  is converted into desired dimensions  $x \times y$ . Let's denote resized images as  $I'$ :

$$I' = \text{resize}(I, (x, y)) \quad (3.2)$$

In the above equation,  $I' \in \mathbb{R}^{x \times y}$  represents the dimension and form of the new resized image.

### 3. Loading of Resized Image to SegNet:

The resized retinal image of fundus  $I'$  is made the input for SegNet to get the segmented mask of the optic disk:

$$M = \text{SegNet}(I') \quad (3.3)$$

where  $M$  is denoted as the segmented mask of the optic disk of the fundus image.

### 4. Output Segmented Mask:

The output of the segmentation module is the segmented optical disc mask. The optical disc segmented dimensions are given by:

$$M \in \mathbb{R}^{x \times y} \quad (3.4)$$

where  $x$  and  $y$  are the dimensions of the segmented optic disc.

### Mathematical Representation

In equation 4,  $I$  represents the original fundus image with dimensions  $i$  and  $j$  that are  $3072 \times 2042$ . In equation 5,  $I'$  is the resized image in which the new dimensions are  $x$  and  $y$ . In equation 6,  $M$  is the segmented optic disc mask obtained through SegNet. In equation 7, the dimension of the final output  $M$  is given.

$$I = \{I(i, j) \mid 1 \leq i \leq 3072, 1 \leq j \leq 2042\} \quad (3.5)$$

$$I' = \text{resize}(I, (x, y)) \Rightarrow I' = \{I'(i, j) \mid 1 \leq i \leq x, 1 \leq j \leq y\} \quad (3.6)$$

$$M = \text{SegNet}(I') \quad (3.7)$$

$$M = \{M(i, j) \mid 1 \leq i \leq x, 1 \leq j \leq y\} \quad (3.8)$$

### 3.2.5 Classification Module

The classification module in this proposed work consisted of an ensemble of a customized CNN and VGG-16.

#### 3.2.5.1 CNN Structure

The generalized CNN framework is developed for classification of glaucoma or not-glaucoma. The architecture consists of an input layer that is a convolutional layer used for the extraction of features, a series of convolutional blocks that are formed using Conv2D layer, after that Batch Normalization is introduced and a dropout layer to turn off the activations of some neurons in the network, followed by MaxPooling2D thereby reducing spatial dimensions by down sampling the feature maps but they do not lose the important information required for classification. At the end, a flatten layer and dense layer is incorporated and for regularization dropout mechanism is used. The final dense layer uses softmax activation for classification. A Conv2D layer is incorporated as an initial layer which consists of 1024 filters ,3\*3 kernel and a ReLU activation function. After that a MaxPooling2D layer of 2\*2 that lowers the dimensions by half. It is followed by BatchNormalization layer that is used for improvement in training stability and overall performance. In order to prevent overfitting, dropout layer is incorporated in the architecture.

After the input layer, there is a series of convolutional blocks that consists of the same layers as in the input layer each with their own specifications. Block 1 consists of Convo2D layer with 512 filters ,a maximum-pooling operation having pool size 2\*2 ,a layer of batchnorm and in the end a dropout layer .Block 2 consists of a Convo2D layer with 256 filters ,a maxpooling layer with pool size of 2\*2 followed by normalization process that is in batch and dropout layer.Block 3 consists of a Convo2D layer with 128 filters followed by a maxpooling , a batch normalization and a dropout layer .Block 4 consists of the same architecture with a Convo2D layer with 64 filters. Block 5 starts with a Convo2D layer 32 filters, a maxpooling layer with pool size of 2\*2, a batch normalization layer and a dropout layer to prevent overfitting. Block 6 consists of a Conv2D layer with 8 filters, a maxpooling layer having a pool size of 2\*2, a batch normalization layer and finally a dropout layer.

**Table 3.2.** CNN Architecture Details

Layer (type)	Output Shape	Param #
input (Conv2D)	(None, 512, 512, 1024)	28,672
max_pooling2d	(None, 256, 256, 1024)	0
batch_normalization	(None, 256, 256, 1024)	4,096
dropout (Dropout)	(None, 256, 256, 1024)	0
conv2d (Conv2D)	(None, 256, 256, 512)	4,719,104
max_pooling2d_1	(None, 128, 128, 512)	0
batch_normalization	(None, 128, 128, 512)	2,048
dropout_1 (Dropout)	(None, 128, 128, 512)	0
conv2d_1 (Conv2D)	(None, 128, 128, 256)	1,179,904
max_pooling2d_2	(None, 64, 64, 256)	0
batch_normalization	(None, 64, 64, 256)	1,024
dropout_2 (Dropout)	(None, 64, 64, 256)	0
conv2d_2 (Conv2D)	(None, 64, 64, 128)	295,040
max_pooling2d_3	(None, 32, 32, 128)	0
batch_normalization	(None, 32, 32, 128)	512
dropout_3 (Dropout)	(None, 32, 32, 128)	0
conv2d_3 (Conv2D)	(None, 32, 32, 64)	73,792
max_pooling2d_4	(None, 16, 16, 64)	0
batch_normalization	(None, 16, 16, 64)	256
dropout_4 (Dropout)	(None, 16, 16, 64)	0
conv2d_4 (Conv2D)	(None, 16, 16, 32)	18,464
max_pooling2d_5	(None, 8, 8, 32)	0
batch_normalization	(None, 8, 8, 32)	128
dropout_5 (Dropout)	(None, 8, 8, 32)	0
conv2d_5 (Conv2D)	(None, 8, 8, 8)	2,312
max_pooling2d_6	(None, 4, 4, 8)	0
batch_normalization	(None, 4, 4, 8)	32
dropout_6 (Dropout)	(None, 4, 4, 8)	0
flatten (Flatten)	(None, 128)	0
dense (Dense)	(None, 512)	66,048
dropout_7 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 256)	131,328
dropout_8 (Dropout)	(None, 256) <sup>23</sup>	0
class_out (Dense)	(None, 2)	514

After that a flatten layer is introduced in the architecture in order to flatten the output obtained from the last block of the convolutional series. This output is converted into one dimensional vector. This vector is used as an input for the fully connected layers that are later introduced in the architecture.

The fully connected layers in the architecture are those layers in which every neuron is in connection with every other neuron of the next layer and the previous layer. In the proposed architecture after the flatten layer first dense layer is incorporated that consists of 512 units followed by a dropout layer. Then a second dense layer is introduced which consists of 256 units and an activation function. After that a dropout layer is introduced for the purpose of regularization.

In the end an output layer is incorporated that is a dense layer which consists of two units that represents the total number of classes that is glaucoma or not glaucoma used for binary classification. The probability that every category will have is computed by a softmax activation. It is used for the normalization of output the output produced by the model.

The generalized CNN model consists of 6,523,274 total parameters. 6,519,226 parameters are trainable and the remaining 4,048 cannot be trained.

This architecture is designed to effectively handle large image inputs, gradually reducing their dimensions while learning hierarchical feature representations. The inclusion of dropout and batch normalization ensures that the model generalizes well to new data, mitigating overfitting and enhancing performance on unseen samples.

### **3.2.5.2 VGG-16 Framework**

The categorization of an image in its respective class is an important task and VGG-16 has the ability to do so. It is a neural model that is entirely convoluted. It adapts transfer learning technique due to which it produces better results. Moreover, its simplicity and the depth of the architecture makes it suitable for classification tasks. The details about tailored architecture of VGG-16 in order to diagnose glaucoma is given:

In VGG-16 model, total number of layers is 16. The initial 13 layers present in the model are convolutional. The remaining three layers are identified as entirely integrated. The shape of the input layer is  $224 \times 224 \times 3$ . The size of the input photograph is  $224 \times 224$  and 3 represents the numerical value of colour channels that is RGB.

After that the formation of framework is: 5 blocks of convoluted layers. Block 1 is made up of 2 convolutional layers that are fed an image as input to extract those features that are essential. Reducing the spatial size is the task of a subsequent maxpooling layer. Block 2 consists of 2 convolutional layers in the same way as block 1 that are followed by a maxpooling layer for dimension reduction. Block 3 consists of 3 convolutional layers for feature extraction and a maxpooling layer for reducing the spatial dimensions.



**Table 3.3.** Layer-wise Architecture of the VGG16 Model

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1,792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36,928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73,856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147,584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295,168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590,080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590,080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1,180,160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2,359,808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2,359,808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2,359,808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2,359,808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2,359,808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25,088)	0
fc1 (Dense)	(None, 4,096)	102,764,544
fc2 (Dense)	(None, 4,096)	16,781,312
predictions (Dense)	(None, 1,000)	4,097,000

Similarly, block 4 and block 5 also are constructed using 3 layers that are convoluted. The subsequent maxpooling operation is used to reduce dimensions without losing important information. Those convolutional blocks consist of two to three Conv2D layers, ReLU activation followed by MaxPooling2D layer. The filters in the convolutional layers increase from 64 to 512. A Flatten layer at the end of the architecture converts the 3D tensor into 1D, followed by two Densely connected layers having ReLU as an activation function and finally a layer that is dense that incorporates softmax operation so that the binary input image can be classified as positive or negative class. The total number of parameters used in this model are 138,357,544.

### 3.2.5.3 Algorithm for Classification

This section consists of the proposed classification module. The input is the segmented mask  $M$  of the optic disk of size  $x \times y$ . The generalized CNN model is represented as  $C_{\text{gen}}$  and VGG-16 model as  $C_{\text{vgg}}$ . The output is (0) Not Glaucoma or (1) Glaucoma.

### 3.2.6 Step-by-Step Algorithm

1. **Input the Segmented Mask to the Generalized CNN Model:**

$$P_{\text{gen}} = C_{\text{gen}}(M) \quad (3.9)$$

where  $P_{\text{gen}}$  is the probability output from the generalized CNN.

2. **Input the Segmented Mask to the VGG-16 Model:**

$$P_{\text{vgg}} = C_{\text{vgg}}(M) \quad (3.10)$$

where  $P_{\text{vgg}}$  is the output from the VGG-16 model.

3. **Combining Results:**

$$P_{\text{combined}} = \alpha P_{\text{gen}} + (1 - \alpha) P_{\text{vgg}} \quad (3.11)$$

where  $\alpha$  is the weight assigned to the customized CNN's output.

4. **Classification:**

$$\text{Class} = \begin{cases} 1 & \text{if } P_{\text{combined}} \geq \theta \\ 0 & \text{if } P_{\text{combined}} < \theta \end{cases} \quad (3.12)$$

where  $\theta$  is the threshold probability for the classification.

### 3.2.7 Mathematical Representation

1. **Segmented Mask Input:**

$$M \in \mathbb{R}^{x \times y} \quad (3.13)$$

The output of the segmentation module becomes the input of the classification module.

2. **Generalized CNN Classification:**

$$P_{\text{gen}} = C_{\text{gen}}(M) \quad (3.14)$$

where  $P_{\text{gen}} \in [0, 1]$ .

3. **VGG-16 Classification:**

$$P_{\text{vgg}} = C_{\text{vgg}}(M) \quad (3.15)$$

where  $P_{\text{vgg}} \in [0, 1]$ .

4. **Combining Results:**

$$P_{\text{combined}} = \alpha P_{\text{gen}} + (1 - \alpha) P_{\text{vgg}} \quad (3.16)$$

where  $\alpha \in [0, 1]$  is a weighting factor.

5. **Final Classification:**

$$\text{Class} = \begin{cases} 1 & \text{if } P_{\text{combined}} \geq \theta \\ 0 & \text{if } P_{\text{combined}} < \theta \end{cases} \quad (3.17)$$

where  $\theta$  is the classification threshold.

# Chapter 4

## RESULTS

The chapter discuss about the segmentation and classification modules results. For segmentation IoU (Intersection over Union) graph is represented for the training and validation of segmentation module. Similarly, the training and validation accuracy and training and validation loss for generalized CNN, VGG-16 and ensemble model is also depicted. The detail of these results is given below:

### 4.1 Segmentation Results

This section gives detailed explanation about the results obtained by performing segmentation of the optical disc of the fundus photograph.

#### 4.1.1 IOU(Intersection Over Union)

Intersection over Union (IoU) measures the performance of the model by evaluating accuracy of a segmentation algorithm. The region in the middle of actual label(reference) and the predicted label given by segmentation algorithm is computed by IOU. It is used to measure the area to of overlap between the actual label (ground truth) and the prediction made by segmentation algorithm. Figure 4.1 shows Intersection Over Union (IoU) scores for the training and validation of the segmentation carried out in the proposed work. The total number of epochs that are used for training are 10. It can be observed that training IoU jumps from 0.70 to 0.89 whereas the validation IoU increases from 0.65 to 0.85. This indicates that the performance of segmentation model is generally increasing as the epochs are increasing. The generalization performance is also good as it can be observed from the small distance between training and validation. It depicts that the model can learn faster about the patterns that are masked in the data as well as generalize well to new data as time increases. Epochs are plotted on the x-axis and on y-axis IoU is plotted.

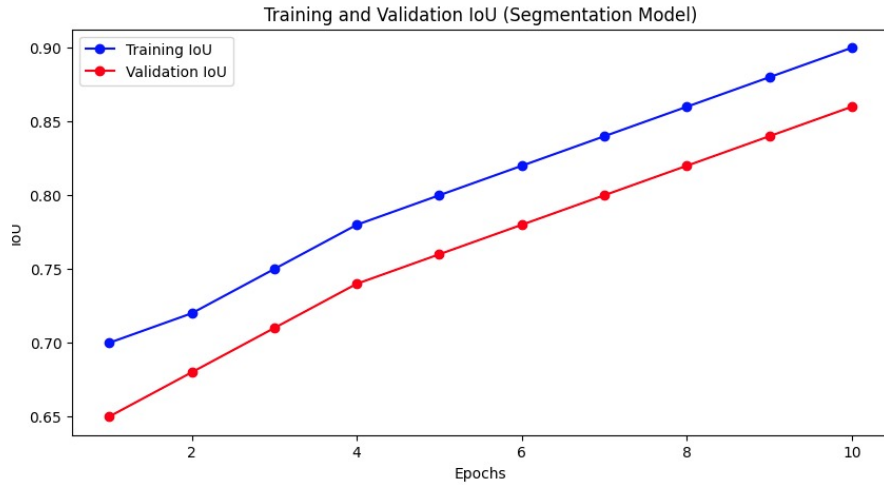


Figure 4.1: Training and Validation IoU

## 4.2 Classification Results for Generalized CNN and VGG-16

This section represents the detailed results that were obtained through the classification performed by generalized CNN and VGG-16 models.

### 4.2.1 Training Accuracy

The accuracy of the dataset used for training purpose showing accurate number of estimations given by the model. It illustrates the way in which the model is being able to fit to the training data in an adequate manner. In case, the model shows elevated accuracy of training data but that of testing is low, the means the model might be prone to overfitting.

### 4.2.2 Validation Accuracy

Validation accuracy measure the quality of a neural network classifier. This accuracy represents the ability of the model to correctly classify the data which the model has not seen before as the model is not trained on validation data. But this does not imply that if the model has elevated accuracy on validation data then it cannot make mistakes of miss classification on the data never seen before.

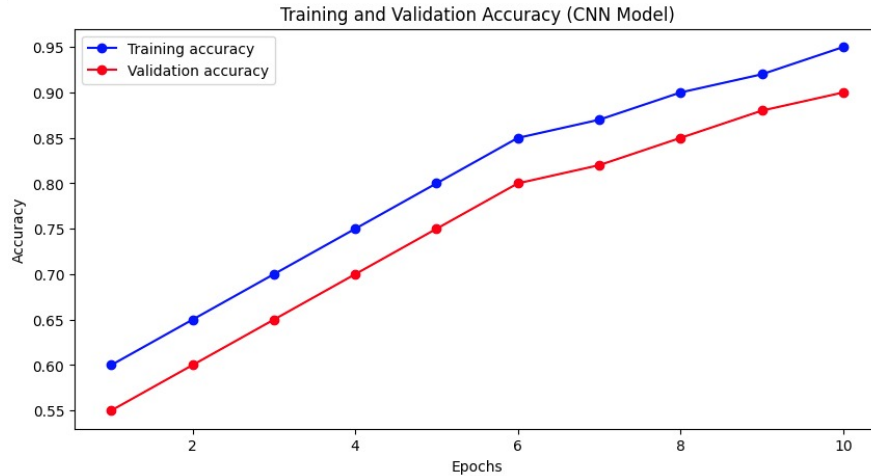


Figure 4.2: Training and Validation Accuracy(CNN)

### 4.2.3 Training Accuracy and Validation Accuracy for generalized CNN

Fig. 4.2 shows training and validation accuracy for the generalized CNN. The training accuracy can be seen at 0.60 gradually rising to 0.93. Similarly, validation accuracy can be seen elevating from 0.55 to 0.88 respectively over 10 epochs. The graph shows minute over-fitting depicting better generalization performance over unseen data. As the number of epochs increases both the curves moves upward showing better learning behaviour and increased performance of model. The model achieves highest accuracy in classifying glaucoma from segmented optic disc at the 10th epoch. The plot shows the accuracy on y-axis of a generalized CNN model over 10 epochs on x-axis for both sets of validation as well as training data.

### 4.2.4 Training Loss

The training loss is an evaluation metric for machine learning model. The proficiency shown by the data utilized for training as it effectively adjusts to the model.

### 4.2.5 Validation Loss

Validation loss quantifies generalization quality depicted by the algorithms. It shows that how well the models fits to the new data that is being given to the model.

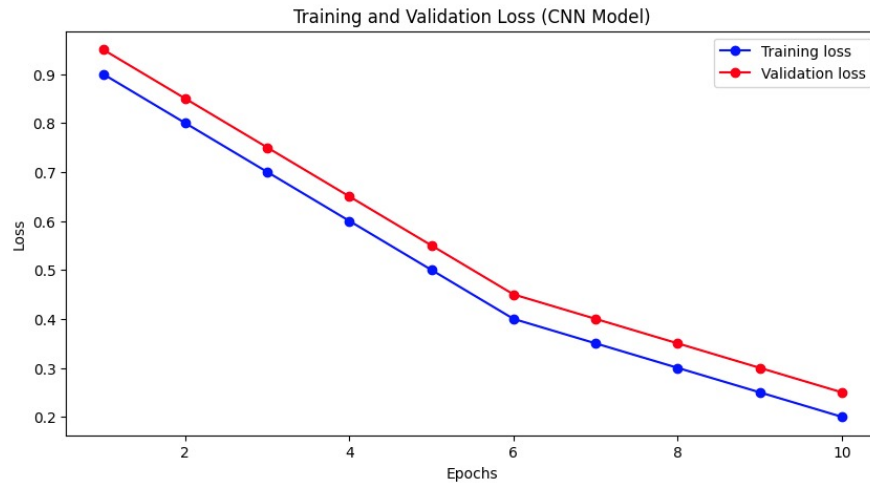


Figure 4.3: Training and Validation loss(CNN)

### 4.2.6 Training and Validation Loss for generalized CNN

The validity and training loss occurring while training generalized CNN is depicted in Fig.4.3. The training loss decreases consistently from 0.9 to 0.2 over 10 epochs. The graph depicts the validation loss also dropping down from 1.0 to 0.1. As both of the losses are decreasing in a parallel way, this is the indication of overfitting not occurring on the algorithm and successfully learning the hidden patterns in data. Thus, the model can generalize well to the unseen data. X-axis depicts the epochs while y-axis depicts the values of loss.

### 4.2.7 Training and Validation Accuracy of VGG-16 Model

The training as well as validation performance of the VGG-16 model is shown in Fig. 4.4

#### 4.2.7.1 Training Accuracy

The blue curve indicates that the training accuracy consistently increases with each epoch, starting from about 0.75 and reaching around 0.925 by the 10th epoch. This trend suggests the good learning behaviour the model is showing as well as increasing efficiency of the model as the time increases.

#### 4.2.7.2 Validation Accuracy

The red curve also shows an increasing trend, though at a slightly lower rate compared to the training accuracy. It starts at around 0.75 and reaches about 0.85 by the 10th

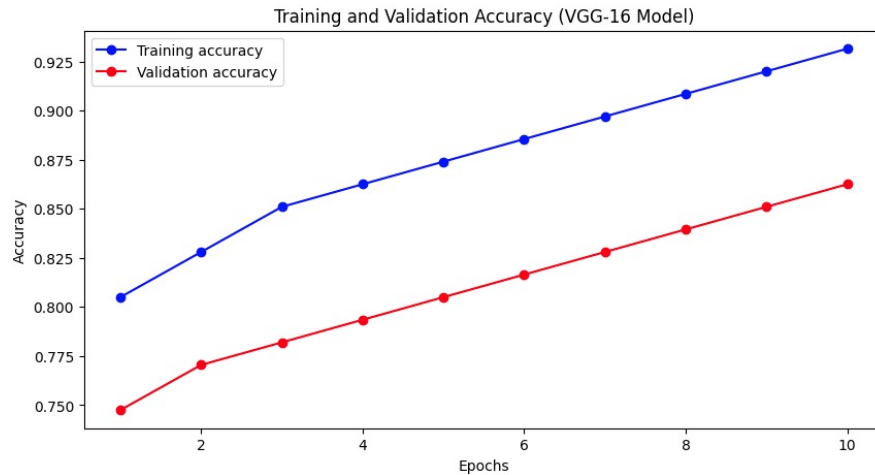


Figure 4.4: Training and Validation Accuracy(VGG-16)

epoch. This indicates that the model is also improving its performance on unseen data, but not as rapidly as on the training data. There is a rise in accuracy in case of training as well as when validation is performed which depicts that the VGG-16 algorithm is performing well on both datasets.

## 4.2.8 Training and Validation Loss of VGG-16

The training plus validation loss demonstrated by a VGG-16 model is illustrated in Fig. 4.5.

Epochs are plotted on x-axis that are 10. One cycle that occurs over the entire dataset used for training represents one single epoch. This axis represents the loss value of the model, ranging from 0.60 to 0.74. Loss represents the variation that occurs in the values of the estimation done by the algorithm verses the real reference values of data. Lower loss indicates better model performance.

### 4.2.8.1 Training Loss

The blue curve indicates that the training loss consistently decreases with each epoch, starting from around 0.70 and reaching approximately 0.60 by the 10th epoch. This information suggests the learning capacity exhibited by the model is increasing with the help the training data, as the loss is decreasing over time.

### 4.2.8.2 Validation Loss

The red curve also shows a decreasing trend, though at a slightly higher level compared to the training loss. It starts at around 0.74 and decreases to about 0.66 by the 10th



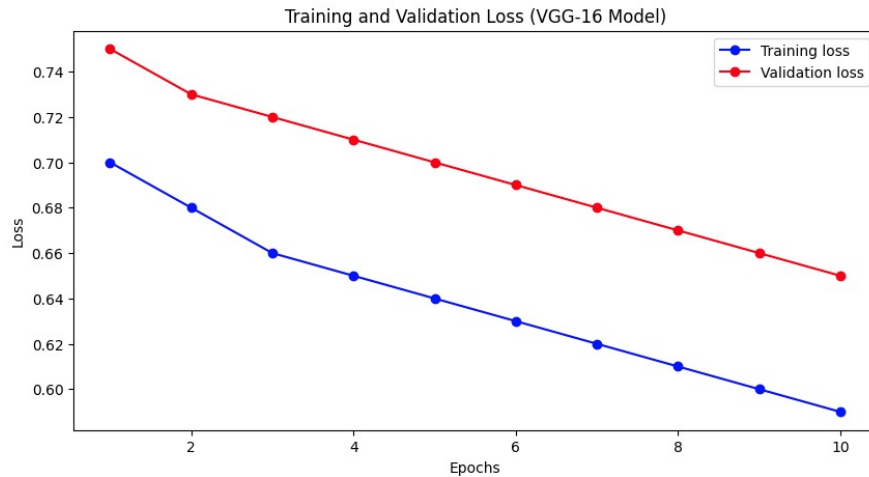


Figure 4.5: Training and Validation Loss(VGG-16)

epoch. This illustrates the model’s executes in a better manner on the dataset utilized for validation.

## 4.3 Final Results(Ensemble Learning)

This section represents the finals results that were obtained using the ensemble approach by combining the generalized CNN and VGG-16 models.

### 4.3.1 Training Accuracy

The blue curve indicates that the training accuracy increases with each epoch, although it shows some fluctuations. It starts around 0.60 and reaches approximately 0.90 by the 20th epoch. This suggests the good learning behaviour of ensemble model is from the data utilized for training as well as improving its performance as time increases. Epochs are plotted on x-axis. Model’s accuracy is plotted on y-axis.

### 4.3.2 Validation Accuracy

The red curve shows a more erratic but generally increasing trend. It starts around 0.60 and reaches about 0.90 by the 20th epoch, with noticeable fluctuations. This indicates model’s performance increasing on unseen data but experiences variations in accuracy at different points. Epochs are plotted on x-axis. Accuracy of models is plotted on y-axis.

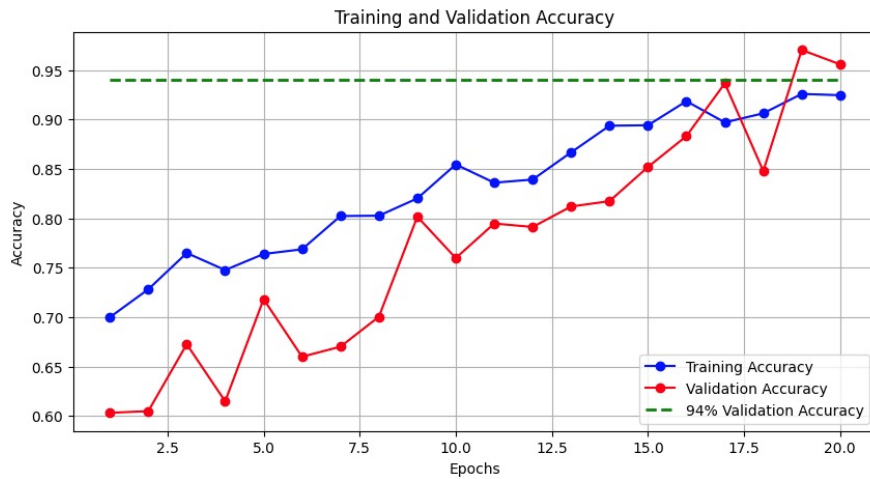


Figure 4.6: Training and Validation Accuracy of Ensemble Model

### 4.3.3 Training Loss

The blue curve indicates that the training loss consistently decreases with each epoch, starting from around 0.6 and reaching approximately 0.2 by the 20th epoch. The graph illustrates model's efficiency in acquiring information obtained through the data used for training as the loss is steadily decreasing over time.

### 4.3.4 Validation Loss

The red curve also shows a decreasing trend, starting from around 0.8 and decreasing to about 0.4 by the 20th epoch, with some fluctuations along the way. The plot pinpoints model's efficiency in gathering information obtained through the data used for validation but it experiences variations in loss at different points.

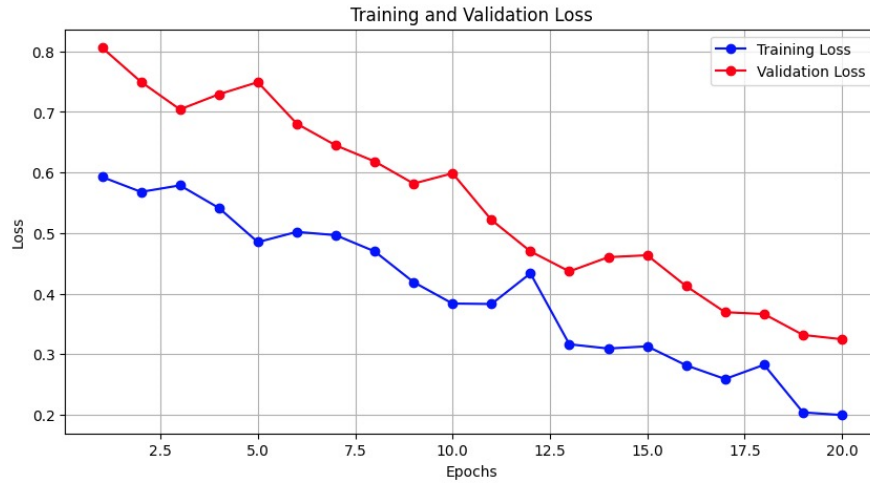


Figure 4.7: Training and Validation Loss of Ensemble Model

## 4.4 Confusion Matrix Analysis

The classification’s model efficiency can be seen with the help of a table or a matrix. This matrix is commonly called as confusion matrix and it displays the actual versus estimated values for classification. It gives the visual representation of the classification performed by the model. In short, it summarizes the classification information in a graphical form. It gives the comprehensive analysis of model’s performance by giving accurate representation of true positive, true negative, false positive and false negative values in the form of a table. For the given confusion matrix, the representation of the true labels are shown in rows, while the representation of the predicted values are shown in columns. Below is the detailed breakdown of the confusion matrix as well as its corresponding metrics:

### 4.4.1 True Negative(TN)

It is represented at the top-left corner of the table. It represents the number of cases when real value as well as the estimated value both are negative. It means that model correctly predicted non-glaucoma (0) when it was actually non-glaucoma. The value of true negative is 94 is this case.

### 4.4.2 False Positive(FP)

It is represented at the top-right corner of the table. It portrays the sum of instances in which the estimation made by model is of the positive class whereas in actual the

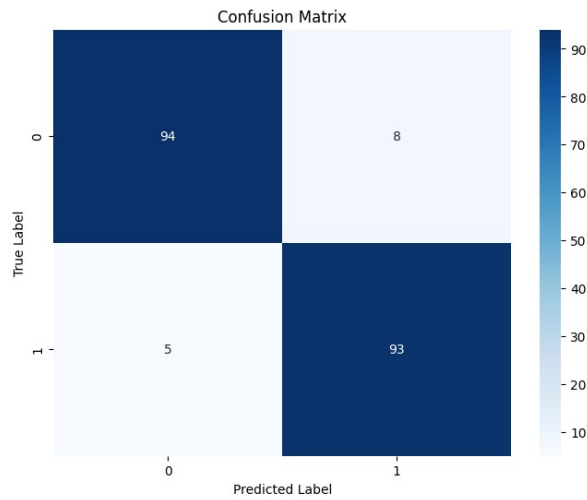


Figure 4.8: Confusion Matrix

value is negative. The model incorrectly predicted glaucoma (1) when it was actually non-glaucoma. The value of false positive is 8.

#### 4.4.3 False Negative(FN)

It is represented at the bottom-left corner of the table. It illustrates the sum of occurrences in which the estimation of the model is negative whereas the real value belongs to the positive class. The model incorrectly predicted non-glaucoma (0) when it was actually glaucoma. The total number of false negative cases is 5 in the given confusion matrix.

#### 4.4.4 True Positive(TP)

The depiction of true positive in the matrix is at the lower-most region. It is represented at the bottom-right corner of the table. It represents the case where the actual value was positive and the estimation of the model is also positive. The model correctly predicted glaucoma (1) when it was actually glaucoma. In the above matrix, the occurrences of true positive cases is 94 .

### 4.5 Evaluation Matrices

When solving classification problems, models make the prediction that to which class the specific data given belongs. The probability estimation is given for each data point or image in our case that whether the input image belongs to positive class that is

glaucoma or negative class that in not glaucoma. For deployment in real-world setting, it is necessary to test out their performance in form of some evaluation metrics. These metrics analyze the classification task performed in a specific scenario. Accuracy, precision, recall, F1 score and specificity are important performance measures used to evaluate and then compare the classification model. For this research evaluation was done using these matrices and their scores in detail is given below:

### 4.5.1 Accuracy

Accuracy is as an evaluation metric for classification tasks that quantifies the sum of correct predictions that is estimated in accordance of model. Accuracy of model can be obtained by taking the ratio of right predictions according to the model correct predictions which represents the summation of true positive and true negatives. An evaluation metric utilized commonly in machine learning framework is accuracy that gives the quantification of all the occurrences of true positives and true negatives compared to all positive and negative occurrences. It answers the question: how often does a model makes right decision? The accuracy achieved by our model is calculated below:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

$$\text{Accuracy} = \frac{93+94}{93+94+8+5} = \frac{187}{200} = 0.93$$

$$\text{Accuracy} = \mathbf{93.5\%}$$

### 4.5.2 Precision(Positive Predicted Value)

Precision can be defined as measure of quality of positive predictions. Precision can be given by the division of the true positive with the quantification of true positive combining false positive i.e. it is the ratio between the correctly identified positive cases verses all identified cases as positive. It can also be called as positive predicted value. In most of the cases, precision is used with recall that is another evaluation metric to correctly evaluate model's efficiency. Precision in the system and also if there is recall, a trade-off between the two is observed. There is a trade-off between precision and recall. When the precision of the model increases, the recall decreases. In the systems where false positives are valuable and costly, precision is given a higher consideration.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4.2)$$

$$\text{Precision} = \frac{93}{93+8} = \frac{93}{101} = 0.9208$$

$$\text{Precision} = \mathbf{92.08\%}$$

### 4.5.3 Recall(Sensitivity or True Positive Rate)

Recall can be seen as a quantitative measure. Recall prioritizes on representing all the relevant occurrences. The division of true positive by the summation of true positive along with false negative (positive occurrences in whole of the dataset) the value of recall can be achieved. It is also known as sensitivity. In most of the diagnosis systems a system with higher recall is preferred so that as many disease cases can be captured as possible even if it leads to some false positive cases. In cases where ignoring false negative leads to serious problems, recall is an important metric to consider.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4.3)$$

Recall=9393+5=93980.948

**Recall = 94.8%**

### 4.5.4 F1 Score(Harmonic Mean of Precision and Recall)

F1 score is another important evaluation metric that is mostly used for classification problems. It is combination of the model's precision along with recall in order to represent them as a single entity. It is commonly used in systems where both precision and recall are important and ignoring either of them will lead to serious consequences. In cases where imbalanced dataset set are used such as medical diagnosis where the instances in positive class are less as compared to instances in negative class, F1 score is used to represent the accuracy. In this way a common ground is kept among the precision of system and recall by representing the system's performance in a single number. It gives the harmonic mean of system's precision along with recall. Its formula and calculation for our model is given below:

$$F_1 = \frac{2.0 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.4)$$

F1 Score=2.0×0.9208×0.9480.9208+0.9482×0.87311.86880.9341

**F1 Score = 93.41%**

### 4.5.5 Specificity(True Negative Rate)

In response to it, when the model identifies the percentage of true negatives in the right manner it develops into a primary component of assessment. The ratio of the correctly identified eyes that are healthy to the actual quantity of eyes that really are healthy and normal. Nonetheless, it can be taken into account that the false positives can be created because of the instances that are actually negative but the diagnosis system predicted them as positive instances. This notion is implicit to this review.

These examples highlight the sensitivity of balance that is significant in maximizing the specificity of the model in a way that would not compromise sensitivity of the model. For the accurate detection of glaucoma the balance between these two evaluation metrics is critical. It acts as a trade off that needs to be balanced in order to provide relief to the patients.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4.5)$$

Specificity=94/94+8=941020.9216S

**Specificity = 92.16%**

This ensemble network of VGG16 and CNN shows high values of all of the evaluation metrics, indicating it performs well in classifying between glaucoma and non-glaucoma cases. A relatively low number of false positives and false negatives are depicted by the model, that is an important consideration in medical diagnoses to minimize misclassification.

#### 4.5.6 Comparison with Advanced Approaches

Many different advanced techniques were previously available for the diagnosis of glaucoma but each has own limitations. In [16] a segmentation method was proposed whose objective was the accurate segmentation of the retinal image of fundus to obtain optical disc as well as cup but the segmentation module employed for the extraction of the mentioned features still required the opinion of multiple ophthalmologists based upon which the final decision was made and accuracy of the segmentation module was calculated. In [17] a hybrid segmentation approach in which through fundus photo ,the withdrawal of the vessel in optic cup was made possible but the sensitivity decreased significantly of the detection system when it was tested on other datasets. Moreover, complex algorithms were used that increased the complexity of the detection system. Discrete and wavelet transform for the diagnosis of glaucoma with the help images of fundus were used but the specificity, sensitivity and accuracy was not that much as compared to other detection systems [20].In [27] a 2 step system was developed for glaucoma identification .First segmentation was done after that classification but the model required additional training that must be available publicly and also the data from private datasets so that the model can become more generalized .The authors in [36] proposed a method in which the 2D fundus image was converted to 3D fundus image which give far better visualization of optical disc including cup in the fundus image of eye .The algorithm developed a topographic map due to which the accuracy was increased but the creation of those 3D maps from 2D fundus images was done using complex algorithms that makes the diagnosis system complicated thus increasing the computational power and requiring more memory .A CNN based hybrid approach was developed in [39] for the detection of glaucoma but the detection system incurred

high cost because large quantity of data was used for the training of their model. The performance exhibited by our proposed work can be compared and is better in a sense to advanced techniques because the segmentation module used does not required much memory and computational efficiency. The generalizability of the proposed method was also good as it trained well on unseen data and the proposed model resulted in small cases of false positive of the system as well as false negatives that minimizes the misclassification of the detection system.

**Table 4.1.** Comparison with other approaches

Reference	Specificity %	Sensitivity %	Accuracy %
[16]	–	–	83.9
[17]	95.3	94.11	95.28
[20]	80.80	86.40	83.57
[27]	93.0	76.0	93.0
[36]	97.9	90.70	94.3
[39]	91.94	93.75	93.5
Proposed work	92.16	94.8	94

#### 4.5.7 Obtained Results

The metrics for evaluation of the generalized CNN, VGG-16 and their combined result were calculated. Their performance individually and combined is given below:

**Table 4.2.** Performance comparison of different models

Method	Accuracy	Precision	Recall	F1
<b>CNN</b>	91.0	91.0	91.70	0.91
<b>VGG-16</b>	85.0	86.0	82.0	81.0
<b>Ensemble</b>	94.0	92.0	95.81	93.0



# Chapter 5

## CONCLUSION AND FUTURE DIRECTIONS

### 5.1 Conclusion

An advanced method for detecting glaucoma was proposed with the main purpose of developing a system that is computationally and memory efficient but still delivers high diagnostic precision. The increased occurrence of glaucoma, which is one of the leading causes of irreversible blindness, necessitates the development of such advanced diagnostic systems that can be used in multiple clinical settings without huge computational resources requirements.

The proposed model is optimized for performance and efficiency. Essentially, it relies on a modified version of SegNet designed specifically to segment an optic disc from fundus images of eyes. In diagnosing glaucoma, this step is crucial because it enables accurate detection of optic disc as one key area needed to assess damage due to glaucoma.

After segmentation stage, model incorporates dual classification module for improving reliability and robustness in diagnosis. This classification module consists in a customized Convolutional Neural Network (CNN) and well-known VGG-16 architecture. The custom CNN is built to be lightweight and efficient so that the overall system remains compact.

### 5.2 Future Direction

In the future, we plan to make our proposed model much better. We'll do this by fine-tuning the ensemble model to tackle overfitting issues. These problems can sometimes cause inconsistent accuracy across different datasets and test scenarios. Overfitting is

still a common issue in machine learning models when they're dealing with tricky tasks like medical image classification. By cutting down on overfitting, we hope to make the model better at generalizing. This means it should work on many different types of data.

To do this, we'll use several cutting-edge methods. First, we'll look at tweaking the learning rate. This rate is key to how the model comes together during training; getting it just right can lead to steadier learning and stop the model from being too touchy about certain parts of the training data. Also, we'll check out ways to stop . This helps us halt training when the model starts doing worse on test data, which keeps it from getting too cozy with the training data.

What's more, we're going to use data augmentation techniques more often. Data augmentation is a good way to make the training dataset bigger and more varied by changing the input images in different ways, like rotating, flipping, and resizing them. This method not makes the model stronger but also helps prevent overfitting by showing the model many different kinds of data situations while it's learning. One more big improvement we're thinking about is adding skip connections to the SegNet setup. Skip connections, which work well in models like U-Net and ResNet, let the model keep important basic info from earlier layers and mix it with more complex features in later layers. This blending of info can make the segmentation results more accurate, which in turn makes the overall classification better. Adding skip connections is key to spot glaucoma , as it can help the model pick up on small details that are crucial to diagnose the disease when it's just starting.

What's more, we're going to take a closer look at how well our segmentation module stacks up against other top-notch segmentation designs. We'll pit our tweaked SegNet against cutting-edge models like U-Net, DeepLab, and Mask R-CNN to get a better handle on its performance and spot areas we can make even better. This side-by-side comparison will help us fine-tune our approach making sure our model doesn't just meet but goes beyond what the current big players in medical image segmentation can do. To wrap up, our next steps will zero in on beefing up the model's toughness and precision. We'll do this by mixing and matching fine-tuning tricks making smart upgrades to its architecture, and putting it through its paces against the best methods out there. All this hard work will lead to a more dependable and powerful tool to catch glaucoma , which in turn will help prevent blindness and boost how well patients do.

# Bibliography

- [1] R. N. Weinreb, T. Aung, and F. A. Medeiros, "The Pathophysiology and Treatment of Glaucoma: A Review," *JAMA*, vol. 311, no. 18, pp. 1901-1911, 2014.
- [2] M. M. C. V. S. Mary, E. B. Rajsingh, and G. R. Naik, "Retinal Fundus Image Analysis for Diagnosis of Glaucoma: A Comprehensive Survey," *IEEE Access*, vol. 4, pp. 4327-4354, 2016.
- [3] K. Allison, D. Patel, and O. Alabi, "Epidemiology of Glaucoma: The Past, Present, and Predictions for the Future," *Investigative Ophthalmology and Visual Science*, vol. 60, no. 1, pp. 1-5, 2021.
- [4] M. T. Leite, L. M. Sakata, and F. A. Medeiros, "Managing glaucoma in developing countries," *Arq. Bras. Oftalmol.*, vol. 74, no. 1, pp. 83-84, 2011.
- [5] R. Thomas, S. G. Thomas, and G. Chandrashekar, "Gonioscopy," *Indian J. Ophthalmol.*, vol. 49, no. 4, pp. 247-251, 2001.
- [6] V. Biousse, B. B. Bruce, and N. J. Newman, "Ophthalmoscopy in the 21st century: The 2017 H. Houston Merritt Lecture," *Neurology*, vol. 90, no. 2, pp. 1-11, 2018.
- [7] P. Brusini, M. L. Salvetat, and M. Zeppieri, "How to Measure Intraocular Pressure: An Updated Review of Various Tonometers," *Journal of Clinical Medicine*, vol. 10, no. 17, pp. 1-21, 2021.
- [8] D. Huang, et al., "Optical Coherence Tomography," *Science*, vol. 254, pp. 1178-1181, 1991.
- [9] M.-M. Gellrich, "The fundus slit lamp," *SpringerPlus*, vol. 4, no. 1, pp. 1-8, 2015.
- [10] R. Thomas, K. Loibl, and R. Parikh, "Evaluation of a glaucoma patient," *Indian Journal of Ophthalmology*, vol. 59, no. S1, pp. S43-S52, 2011.
- [11] G. E. Trope, E. M. Eizenman, and E. Coyle, "Eye movement perimetry in glaucoma," *Canadian Journal of Ophthalmology*, vol. 26, no. 3, pp. 133-138, 1991.

- [12] X. Wang, "Deep Learning in Object Recognition, Detection, and Segmentation," *Foundations and Trends in Signal Processing*, vol. 8, no. 4, pp. 217-382, 2016.
- [13] S. Atheesan and S. Yashothara, "Automatic glaucoma detection by using fundus-copic images," in *Proc. 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, India, 2016, pp. 1453-1458.
- [14] S. Maheshwari, et al., "Iterative variational mode decomposition based automated detection of glaucoma using fundus images," *Biomedical Signal Processing and Control*, vol. 68, pp. 102590, 2021.
- [15] K. Sakthivel and R. Narayanan, "An automated detection of glaucoma using histogram features," *International journal of ophthalmology*, vol. 8, no. 1, pp. 194-200, 2015.
- [16] A. Almazroa, et al., "Optic disc segmentation for glaucoma screening system using fundus images," *BioMed Research International*, vol. 2016, pp. 1-10, 2016.
- [17] R. Sundaram, et al., "Extraction of Blood Vessels in Fundus Images of Retina through Hybrid Segmentation Approach," *Mathematics*, vol. 7, no. 2, pp. 169-181, 2019.
- [18] A. Septiarini, et al., "Automatic Glaucoma Detection Method Applying a Statistical Approach to Fundus Images," *Healthcare Informatics Research*, vol. 24, no. 1, pp. 53-60, 2018.
- [19] D. K. Agrawal, B. S. Kirar, and R. B. Pachori, "Automated glaucoma detection using quasi-bivariate variational mode decomposition from fundus images," *IET Image Processing*, vol. 13, no. 13, pp. 2401-2408, 2019.
- [20] B. S. Kirar and D. K. Agrawal, "Computer aided diagnosis of glaucoma using discrete and empirical wavelet transform from fundus images," *IET Image Processing*, vol. 13, no. 1, pp. 73-82, 2019.
- [21] M. Sreedhar and B. Radhika, "GLAUCOMA DETECTION SYSTEM ON THE BASIS COMBINING NB and RF CLASSIFIERS," *Global Journal of Novel Research in Applied Sciences (NRAS)*, vol. 1, no. 2, pp. 37-41, 2023.
- [22] A. A. Salam, et al., "Automated detection of glaucoma using structural and non structural features," *SpringerPlus*, vol. 5, no. 1, pp. 1-13, 2016.
- [23] L. Chakrabarty, et al., "Automated Detection of Glaucoma From Topographic Features of the Optic Nerve Head in Color Fundus Photographs," *Journal of Glaucoma*, vol. 25, no. 7, pp. 590-597, 2016.

- [24] T. Khalil, et al., “Improved automated detection of glaucoma from fundus image using hybrid structural and textural features,” *IET Image Processing*, vol. 11, no. 9, pp. 693-700, 2017.
- [25] X. Chen, et al., “Glaucoma detection based on deep convolutional neural network,” in *Proc. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 2015, pp. 7156-7159.
- [26] V. P, et al., “Glaucoma Detection using Convolution Neural Networks,” in *Proc. 2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2023, pp. 324-329.
- [27] J. Civit-Masot, et al., “Dual Machine-Learning System to Aid Glaucoma Diagnosis Using Disc and Cup Feature Extraction,” *IEEE Access*, vol. 8, pp. 127519-127529, 2020.
- [28] P. Wang, et al., “3D augmented fundus images for identifying glaucoma via transferred convolutional neural networks,” *International Ophthalmology*, vol. 41, no. 6, pp. 2065-2072, 2021.
- [29] R. Kashyap, et al., “Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model,” *Healthcare*, vol. 10, no. 12, pp. 1-12, 2022.
- [30] S. Saha, J. Vignarajan, and S. Frost, “A fast and fully automated system for glaucoma detection using color fundus photographs,” *Scientific Reports*, vol. 13, no. 1, pp. 18408, 2023.
- [31] R. Shinde, “Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms,” *Intelligence-Based Medicine*, vol. 5, pp. 100038, 2021.
- [32] S. S and D. V. Babu, “Retinal Glaucoma Detection from Digital Fundus Images using Deep Learning Approach,” in *Proc. 2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2023, pp. 315-321.
- [33] T. R. V. Bisneto, A. O. de Carvalho Filho, and D. M. V. Magalhães, “Generative adversarial network and texture features applied to automatic glaucoma detection,” *Applied Soft Computing*, vol. 90, pp. 106165, 2020.
- [34] H. Sheraz, T. Shehryar, and Z. A. Khan, “Two stage-network: Automatic localization of Optic Disc (OD) and classification of glaucoma in fundus images using deep learning techniques,” *Multimedia Tools and Applications*, vol. 83, no. 6, pp. 17921-17939, 2024.

- [35] I. A. Arias-Serrano, et al., “Artificial intelligence based glaucoma and diabetic retinopathy detection using MATLAB - retrained AlexNet convolutional neural network,” *Scientific Data*, vol. 8, no. 1, pp. 1-12, 2021.
- [36] P. Wang, et al., “3D augmented fundus images for identifying glaucoma via transferred convolutional neural networks,” *International Ophthalmology*, vol. 41, no. 6, pp. 2065-2072, 2021.
- [37] Z. Zhang, et al., “ORIGA(-light): an online retinal fundus image database for glaucoma analysis and research,” *Scientific Data*, vol. 4, pp. 170032, 2017.
- [38] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, 2017.
- [39] C. Oguz, T. Aydin, and M. Yaganoglu, “A CNN-based hybrid model to detect glaucoma disease,” *Multimedia Tools and Applications*, vol. 83, no. 6, pp. 17921-17939, 2024.