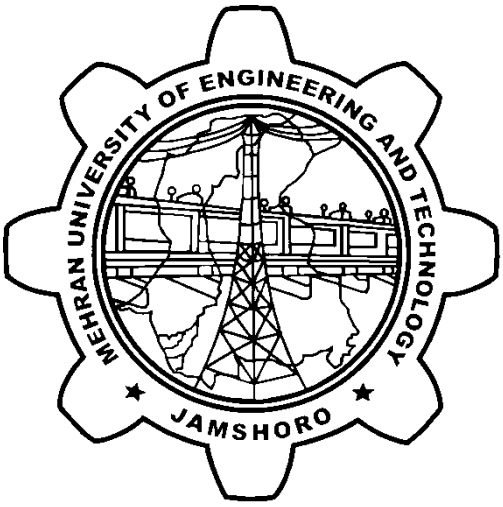
AUTOMATED DETECTION OF DIABETIC RETINOPATHY USING SMARTPHONE-BASED PHOTOGRAPHY



A thesis submitted by

Hina Lilaram (GL) (18CS74)
Sarmad Talpur (18CS80)
Sharjeel Ali (18CS122)

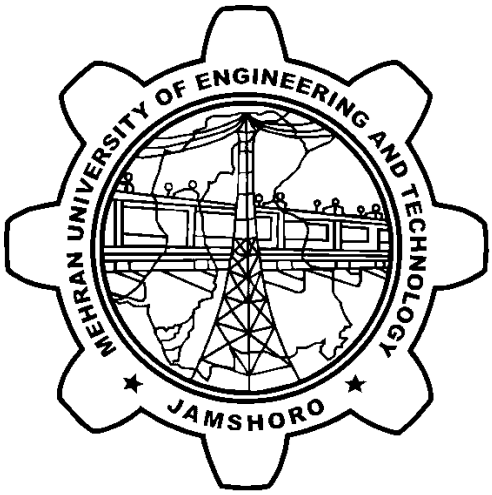
Supervised by

Dr. Irfan Ali Bhacho

Submitted in the partial fulfillment of the requirements for the degree of Bachelor of Engineering in Computer Systems

Faculty of Electrical, Electronics & Computer Engineering
MEHRAN UNIVERSITY OF ENGINEERING & TECHNOLOGY, JAMSHORO

October 2022



CERTIFICATE

This is to certify that "**Project/Thesis Report on, Automated Detection of Diabetic Retinopathy using Smartphone-Based Technology"** is submitted in partial fulfillment of the requirement for the degree of Bachelor of Computer Systems Engineering by the following students:

Hina Lilaram (GL) (18CS74)
Sarmad Talpur (18CS80)
Sharjeel Ali (18CS122)

_____

Supervisor

Dr. Irfan Ali Bhacho

_____

(Chairman, Department of Electronic Engineering)

Date: _____

ACKNOWLEDGEMENT

Words cannot explain how grateful we are to our supervisor and professor Dr. Irfan Ali Bhacho and chairman Dr. Shahnawaz Talpur for their tremendous understanding and feedback.

In addition, we could not have made this voyage without the rich knowledge and experience of Dr. Khalid Iqbal Talpur (Director, Sindh Institute of Ophthalmology and Visual Sciences, Hyderabad). Aside from that, our project would not have been possible without the co-operation of doctors and the IT department of SIOVS.

We are also appreciative of the technical assistance, and support we received from our team work and other fellow mates from 18-Batch.

Moreover, we received enormous amount of support and availability to discuss technical and academic matters with the faculty of Computer Systems Engineering Department, Mehran University of Engineering and Technology Jamshoro.

ABSTRACT

According to recent findings in 2019, there has been a dramatic increase in the number of diabetics in Pakistan. One in every three diabetics is prone to develop a visual impairment called “Diabetic Retinopathy (DR)” that can cause permanent blindness if left untreated in its early stages. Diabetics are required to go through an annual eye examination that involves taking images of the eye using an expensive fundus camera that is only available in specific urban hospitals. Additionally, there is a lack of ophthalmologists. There are only 1,500 ophthalmologists in Pakistan, which means one ophthalmologist for just under 100,000 people, out of which about 80% provide their services in urban locations, leaving out a vast majority of eye patients unattended that are living in remote and rural regions of the country. We aim to develop a smartphone-based cost-effective handheld AI-integrated product to detect visual Impairment and generate reports of the patient on the same day with a minor intervention by an eye specialist. This project focuses on the detection of DR using a 20D (20 Diopter) Lens and a smartphone camera to capture fundus images that are further classified and compared against different deep learning models. The model with the best accuracy is integrated into a mobile application for the development of a cost-effective and portable method for DR screening in rural areas. This project aims to combat three major Sustainable Development Goals (SDGs): Good health and well-being; Affordable and clean energy and Industry Innovation and Infrastructure respectively.

Table of Contents

1. INTRODUCTION 1

1.1 Diabetes 1

Type 1 diabetes 1

Type 2 diabetes 2

Gestational diabetes 2

1.2 Diabetic Retinopathy 2

1.2.1 Background retinopathy 4

1.2.2 Diabetic maculopathy 4

1.2.3 Proliferative retinopathy 5

1.2.4 Vitreous hemorrhages 6

1.2.5 Retinal Detachment 6

1.3 Symptoms 6

1.4 Surveys and Reports 7

1.5 Existing solutions 8

1.6 Proposed solution 8

1.7 Organization of the thesis 9

2. LITERATURE REVIEW 10

2.1 Imaging Modalities 10

2.1.1 Single field imaging 10

2.1.2 Wide field imaging 10

2.1.3 Optical Coherence Topography (OCT) 11

2.1.4 OCT-angiography (OCTA) 11

2.1.5 Smartphone based systems 11

2.2 Processing methods 12

2.2.1 Image Processing-Based DR Detection 12

2.2.2 Traditional Machine Learning Algorithms 12

2.2.3 Neural Networks, Deep Learning and Transfer Learning 13
3. DESIGN AND METHODOLOGY 14
3.1 Dilation 14
3.2 Device Design 15
3.2.1 20D lens 15
3.2.2 3D Printed Scope 16
3.2.3 Smartphone 16
3.3 Working of the Device 17
3.4 Dataset Collection 18
3.4.1 Dataset 18
3.4.2 Preprocessing 18
3.5 Algorithms ..... **Error! Bookmark not defined.**
3.5.1 Convolutional Neural Network (CNN) 19
3.5.2 Experimental Setup 24
3.5.3 VGG16 25
3.5.4 Resnet50 27
3.5.5 Custom CNN Model 30
3.6 Mobile application 32
3.6.1 Image sharing Flow 32
3.6.2 Automated Report Generation Flow 36
4. RESULTS AND DISCUSSION 38
4.1 VGG16 38
4.2 ResNet50 40
4.3 Custom Model 41
4.4 Comparative Analysis 42
5. CONCLUSION AND FUTURE RECOMMENDATIONS 43
5.1 Our Research Outcomes 43
5.2 Future Recommendations 43
6. REFERENCES 44

List of Figures

Figure 1.1- Comparing Normal and DR infected Eye 2
Figure 1.2- Classification of Stages of DR 3
Figure 1.3- Macular Edema Stages 5
Figure 1.4- Eye structures and lesion types crucial to the diagnosis of DR. 7
Figure 1.5- Retina images from with different DR levels, (a) normal, (b) mild, (c) moderate, (d) severe, and (e) proliferative 7
Figure 1.6- Estimated prevalence of diabetes in the age groups 20–79 in millions (IDF Atlas 2019) 8
Figure 3.1- Dilation of Eye 14
Figure 3.2- Complete Device 15
Figure 3.3- Horizontal view of a 20D lens 15
Figure 3.4- Volk Aspheric 20D Lens 15
Figure 3.5- 3D Printed Scope 16
Figure 3.6- Image Acquisition Process 17
Figure 3.7- Before Preprocessing 18
Figure 3.8- After Preprocessing 19
Figure 3.9- Basic CNN Architecture 20
Figure 3.10- Basic CNN Architecture (with layers) 23
Figure 3.11- Fine tuned VGG16 Architecture 25
Figure 3.12- Image Preprocessing 26
Figure 3.13- Resnet 50 Architecture 28
Figure 3.14- Skip connection in Resnet 50 29
Figure 3.15- Data Preparation and Preprocessing - Resnet50 29
Figure 3.16- Custom Model Architecture 30
Figure 3.17- Image Preprocessing- Custom CNN Model 31
Figure 3.18- Application User Flow 32
Figure 3.19- Login and Home Screen 33
Figure 3.20- Image Selection and Questionnaire Screen 34
Figure 3.21- Consultant Selection and Confirmation Screen 35
Figure 3.22- Select images and enter patient data for AI report 36
Figure 3.23- AI-Generated DR Report 37
Figure 4.1 Training and Validation Accuracy/Loss 38

Figure 4.2- ROC Curve - VGG16 39
Figure 4.3 Training and Validation Accuracy/Loss 40
Figure 4.4- ROC Curve- ResNet50 40
Figure 4.5 Training and Validation Accuracy/Loss 41
Figure 4.6- ROC Curve- Custom CNN Model 42

List of Tables

Table 3.1- Cloud Instance Configurations 25
Table 4.1- Comparative Analysis 42

List of Abbreviations

DR	Diabetic Retinopathy
CFP	Colour fundus photography
FOV	Field of View
ETDRS	Early Treatment Diabetic Retinopathy Study (ETDRS).

Chapter-1

INTRODUCTION

According to WHO (World Health Organization) diabetes is the ninth leading cause of death worldwide. Prevalence of diabetes has been rapidly increasing in low and middle-income countries than in high-income countries [1]. Diabetic retinopathy (DR) is a complication of diabetes, caused by high blood sugar levels damaging the back of the eye (retina). DR damages the retina and leads to arbitrary growth of blood vessels. It can make the vessels clot or burst as well as depicted in Figure 1.1. DR can cause blindness if left undiagnosed and untreated.

It is one of the leading causes of blindness worldwide accounting for more than 3.9 million. [2] However, it takes several years for diabetic retinopathy to reach a stage where it could threaten the sight of a diabetic patient. Most patients remain asymptomatic until they reach advanced stages of DR. The available treatments for visual impairment caused by diabetes have been more effective at slowing visual loss than reversing visual impairment. Hence, early detection of DR before the permanent loss of vision is crucial to ensure better patient health [3].

1.1 Diabetes

Diabetes, commonly known as diabetes mellitus (DM), is a chronic condition that develops when the pancreas either produces insufficient insulin or when the body cannot properly utilize the insulin that it generates. A hormone called insulin controls blood sugar. Uncontrolled diabetes frequently results in hyperglycemia, or elevated blood sugar, which over time causes substantial harm to many different bodily systems, including the neurons and blood vessels. In 2019, 1.5 million deaths occurred due to diabetes of which 48% were adults aged less than 70 years. One in every three patients with diabetes mellitus develop DR [3]. The three major types of diabetes mellitus are:

Type 1 diabetes stems from the pancreas' inability to generate adequate insulin. Previously, this kind was known as "insulin-dependent diabetes mellitus" or "juvenile

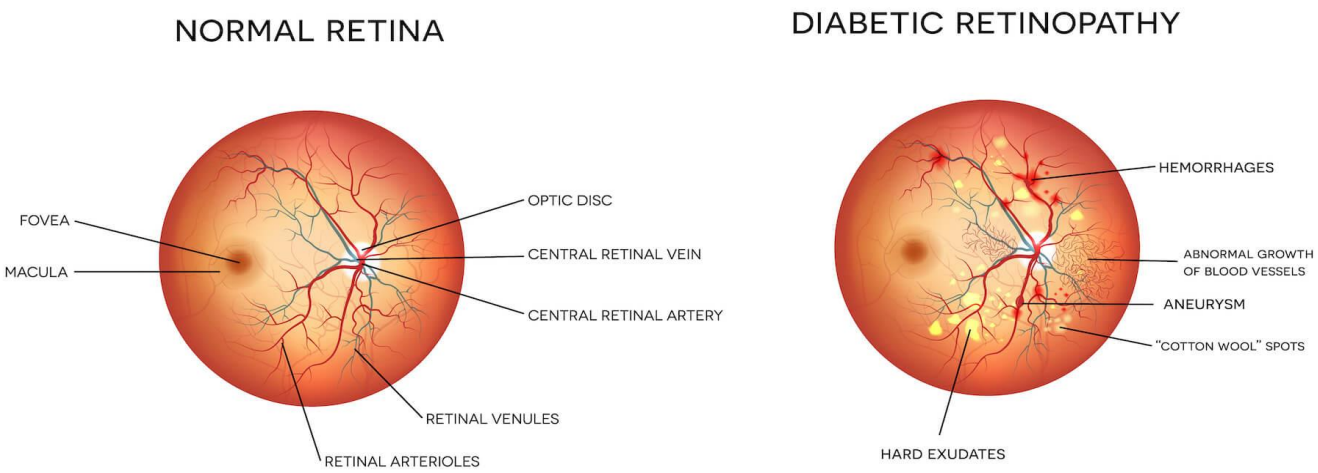


Figure 1.1- Comparing Normal and DR infected Eye

diabetes." While Type 1 diabetes often first manifests in infancy or adolescence, it can also strike adults.

Type 2 diabetes This form was formerly known as "non-insulin-dependent diabetes mellitus" or "adult-onset diabetes." Type 2 diabetes is more common in older adults, but a significant rise in the prevalence of obesity among children has led to more cases of type 2 diabetes in younger people. The most common type 2 diabetes is insulin resistance, a condition in which cells fail to respond to insulin properly.

Gestational diabetes occurs in pregnant women who have never had diabetes have excessive blood sugar levels. It is the third major kind. After delivery, blood sugar levels in women with gestational diabetes often recover to normal. The chance of acquiring type 2 diabetes later in life is increased for women who had gestational diabetes during pregnancy.

1.2 Diabetic Retinopathy

Alterations in blood glucose level or diabetes results in changes in the retinal blood vessels, which in turn results in the development of a vision threatening eye disease referred to as

"Diabetic Retinopathy" (DR). The retinal blood vessels occasionally develop swelling and leakage of blood vessels called macular edema. On occasion, the surface of the retina will develop aberrant blood vessels. These blood vessels can get enlarged and may lead to blood clots and leakages in the back of the eye. Without treatment and comprehensive annual examination, this disease can develop itself significantly from mild retinopathy to impaired vision and can even result in permanent blindness. Other eye conditions such as glaucoma and cataracts are highly likely to be developed with individuals suffering from diabetes compared to the rest of the healthy population. [4]

We have used four different classification types for different stages of DR:

1. Mild
2. Moderate
3. Sever
4. Proliferative

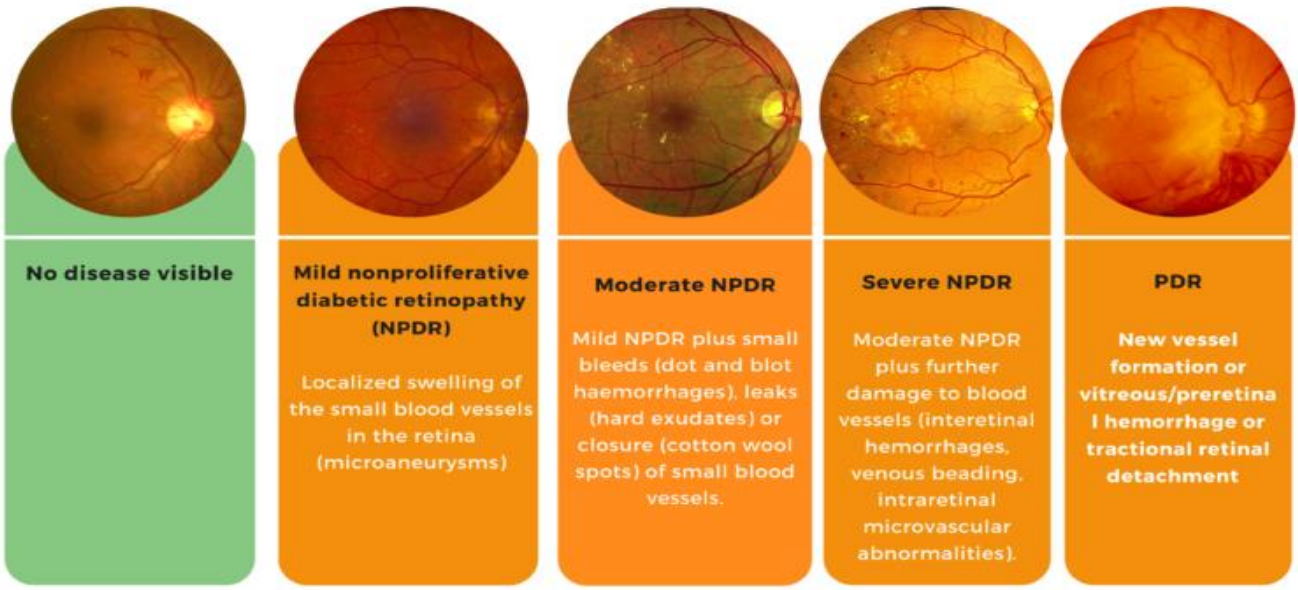


Figure 1.2- Classification of Stages of DR

Diabetic retinopathy includes three different types:

1. Background retinopathy
2. Diabetic maculopathy
3. Proliferative retinopathy

1.2.1 Background retinopathy

Background retinopathy is likely to develop if your retina has acquired microaneurysms. Microaneurysms occur when the capillaries—very small blood vessels—that supply the retina expand. There probably will not be any visual issues if there are only a few microaneurysms present. However, your vision is more likely to be in risk if the retinopathy's severity might increase noticeably. Typically, a retinal screening check, which involves taking a picture of your retina, is the only way to detect the symptoms of background retinopathy. For those who have diabetes, retinopathy in some form is frequent. Patients with type 1 diabetes have a higher incidence of retinopathy than those with type 2 diabetes because they have had the disease for a lot longer than those with type 2 diabetes. Treatment must start right away (within a year) if retinopathy progresses to other kinds such maculopathy or proliferative retinopathy in order to preserve the patient's vision.

1.2.2 Diabetic maculopathy

Our central vision is aided by the macula, which is a component of the eye. When the macula has some sort of injury, it is called diabetic maculopathy. Diabetic macular oedema, in which blood vessels close to the retina leak fluid or protein onto the macula, is one such cause of macular degeneration. Clinically Significant Macular Oedema is the medical name for a disorder where blood fat deposits called exudates (which form when blood vessels leak) causing the retina to harden and become noticeably big and near to the fovea (CSMO).

One of the signs of diabetic maculopathy is blurry central vision. This may be noticed by:

- Having trouble reading
- Spotting faces in your line of sight.

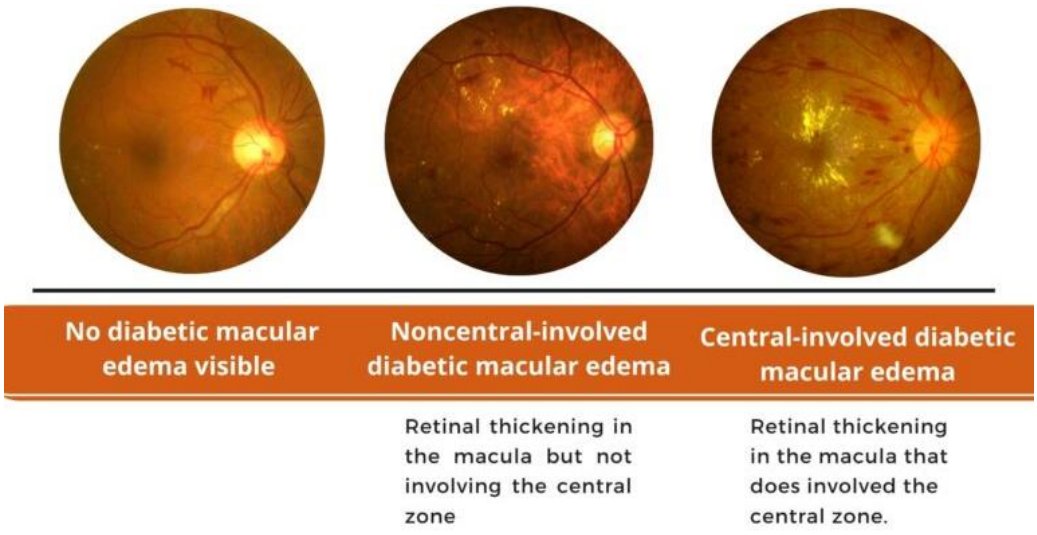


Figure 1.3- Macular Edema Stages

1.2.3 Proliferative retinopathy

A more advanced form of retinopathy known as proliferative retinopathy causes the formation of new, weak blood vessels on the retina in an effort to improve blood flow. The body tries to save its retina through proliferative retinopathy; however, it frequently results in scarring and can separate the retina, leaving the patient blind. Attending routine retinal screening tests is crucial since proliferative retinopathy can manifest without any symptoms.

A hemorrhage or detached retina may cause symptoms such as:

- The sudden emergence of floaters (dots, specks, or streaks) in your field of vision
- Suddenly developing floaters (dots, specks, or streaks) in your vision
- Cobweb-like effect throughout your field of vision
- Obvious object distortion
- Vision loss

If enough blood vessels on the retina are broken, your body will react by generating a growth hormone called Vascular Endothelial Cell Growth Factor (VEGF). The development of new blood vessels is promoted by growth hormone. However, since these new blood vessels are so weak and prone to leakage, visual problems may develop.

Proliferative retinopathy can lead to a number of issues, including:

- Vitreous hemorrhage
- Retinal detachment

1.2.4 Vitreous hemorrhages

Proliferative retinopathy may cause new blood vessels to grow from the retina and enter the vitreous humour (the gel within the eye). Your field of vision may become obscured by cobwebs when these blood vessels bleed as a result of the vitreous humour decreasing, making it harder to see. A vitreous hemorrhage's blood will vanish, but proper and prompt examnation is still required to prevent further issues.

1.2.5 Retinal Detachment

The retina may become detached as a result of bleeding into the vitreous humour, which might shrink as a result. Suddenly finding floaters in your field of vision is one sign of retinal detachment. Vision may become affected if the retina has an abnormal shape.

1.3 Symptoms

Early signs of retinopathy are rare. The signs of diabetic retinopathy often appear gradually because elevated blood sugar levels damage the retina's blood vessels. The first retinal indications can only be seen during a retinopathy screening exam. Retinopathy can develop over many years or decades. One or more of the following indicators may be present in established retinopathy in addition to the standard symptoms of retinopathy:

- Loss of vision.
- Sudden changes in vision
- Blurred, distorted, or floaters in vision
- Detection of dark spots or patches
- A decrease in night vision

- Sudden changes in vision

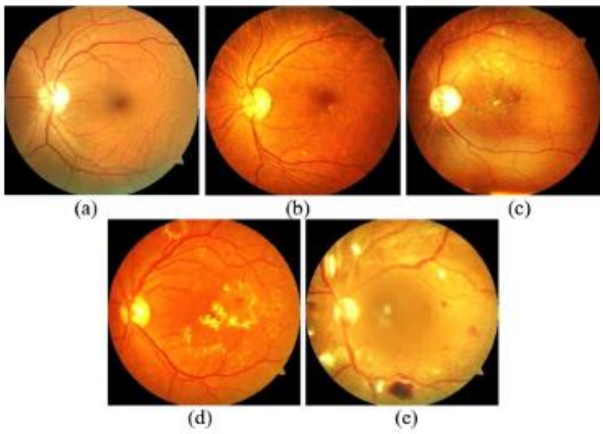


Figure 1.5- Retina images from with different DR levels, (a) normal, (b) mild, (c) moderate, (d) severe, and (e) proliferative

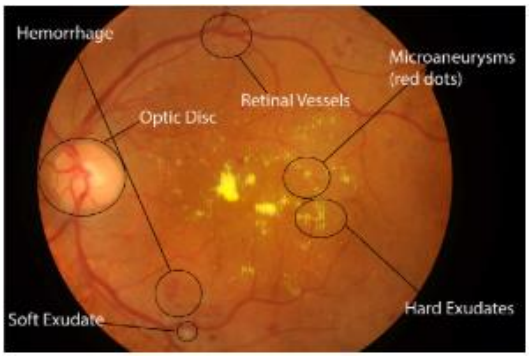


Figure 1.4- Eye structures and lesion types crucial to the diagnosis of DR.

1.4 Surveys and Reports

There are nearly 2.2 billion people suffering from distant or near vision impairment globally. About half of these, or a minimum of one billion people, involve visual damage that either might have been avoided or is still unaddressed and untreated. [2]

Population based studies demonstrate that by the age 20 years following the onset of the disease, 60% of individuals with type II diabetes and virtually all of those with type I diabetes will acquire DR in their life span. [5]

According to recent findings in 2019 there has been dramatic increase in the number of diabetics in Pakistan. Over 19 million adults are at risk of life-threatening complications caused by diabetes. [6]

According to statistics from the World Health Organization, there were 422 million diabetics worldwide in 2014, and by 2030, that number is expected to rise to 552 million [7]. According to the US Department of Health and Human Services' National Diabetes Statistics Report [8], the estimated number of Americans with diabetes in 2020 will be 30.5 million (10.5 percent), with 7.3 million remaining undiagnosed across all age categories. Diabetes increases a person's chance of developing eye conditions such diabetic retinopathy, diabetic macular edoema, and glaucoma.

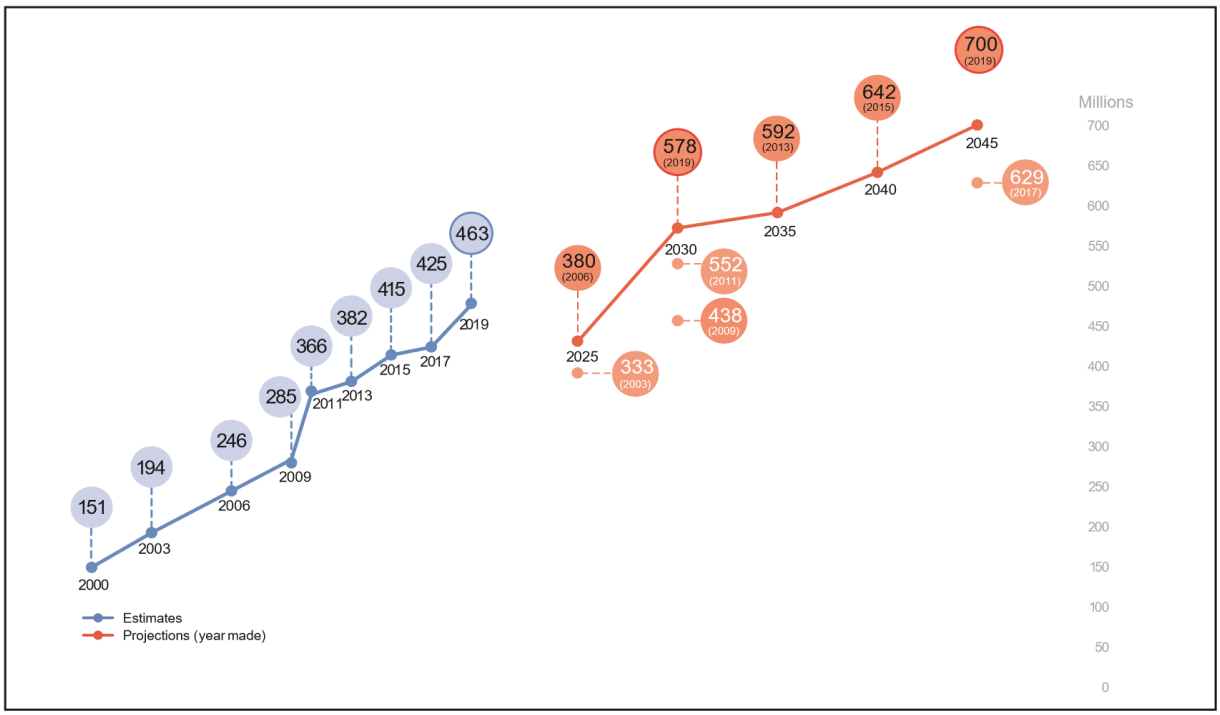


Figure 1.6- Estimated prevalence of diabetes in the age groups 20–79 in millions (IDF Atlas 2019)

1.5 Existing solutions

Diabetic retinopathy screening is required to identify referable patients that require prompt comprehensive ocular examination and treatment to prevent irreversible vision loss. However, many underdeveloped nations, including Pakistan, do not have enough funding for widespread screening programs. The utilization of telemedicine, portable imaging equipment, and emerging artificial intelligence-based technologies that enable individualized prediction model implementation are transforming screening tactics and increasing screening's cost-effectiveness.

1.6 Proposed solution

In this study, we detect DR using a 20D (Diopter) ophthalmic Lens and a smartphone-based camera to capture fundus images of 150 OPT patients at the SIOVS hospital. These images are used to test deep learning model based on CNN, VGG16. A dataset of 10,000 fundus images acquired from a fundus camera operating at the hospital is used to train the model and 150 images captured from a 20D lens are used to test the model. The results are classified into two categories namely (DR and no DR) that are compared against the three

deep learning models. The model with the best accuracy is integrated into a mobile application for the development of a cost-effective and portable method for DR screening in rural areas.

1.7 Organization of the thesis

The remaining portions of the document are arranged as follows. A brief overview of earlier techniques for detecting diabetic retinopathy is included in Section II. The phases of the suggested approach and dataset preparation are shown in Section III. The findings of the suggested system are presented in Section IV, along with a comparison of the suggested model with earlier research. The research is concluded in Section V, which also covers potential directions.

Chapter-2

LITERATURE REVIEW

Yearly screenings are necessary for diabetic patients since early DR diagnosis might assist to prevent blindness. However, ophthalmologists must perform physical examinations, which takes time, and there are not enough specialists to match the demand for screening. The idea of autonomous detection of retinal exudates, towards diagnosis and tracking the success of a patient's treatment, is intriguing given the limits of human screening. This issue has been addressed in a number of ways.

2.1 Imaging Modalities

2.1.1 Single field imaging

The optic nerve and macula are visible in single-field fundus imaging within a 20–50-degree field of vision. Seven standard 30-degree ophthalmic images have been standardized by 1991, from the Early Treatment Diabetic Retinopathy Study and have been the industry standard for DR diagnosis (ETDRS) [9], which are used in traditional methods. Montage photographs demonstrate success in DR detection and categorization, although collecting is labor-intensive and time-consuming in actual use.

2.1.2 Wide field imaging

The development of contemporary technology has enabled field-of-view (FOV) up to 200 degrees, which has significantly advanced retinal imaging. WF color imaging, autofluorescence, fluorescein angiography, and indocyanine green angiography are examples of WFI modalities. Over fifty percent of the time, mostly peripheral lesions occur outside of the typical ETDRS fields; hence, lesions that aid in the detection of DR, such as microaneurysms (MAs), hemorrhages, venous beading, and microvascular abnormalities, may go unreported. Optical coherence tomography (OCT)

2.1.3 Optical Coherence Topography (OCT)

Advanced optical sectioning abilities, unrestricted by spatial or depth resolutions, are illustrated by the OCT methodology. It allows for very accurate noninvasive examination for diabetic retinopathy, enabling patients to get treatment before experiencing irreparable visual loss [10]. This method analyses cross-sections of tissues by employing low-coherence light and observing the reflections that it creates off tissue surfaces.

2.1.4 OCT-angiography (OCTA)

DR detection and monitoring may benefit from several OCT-angiography (OCTA) measurements. To standardize these measurements, investigations must first develop criteria that restrict the influence of other variables including age, comorbidities, and the length of diabetes. A non-invasive alternative to fluorescein angiography for Diabetic Retinopathy detection is provided by the OCT-angiography examination, due to which it generates a lot of interest for research opportunities and examination.

2.1.5 Smartphone based systems

The price of many current devices has decreased as a result of advances in imaging technology. The study in [10] examined the various types of fundus cameras that are employed globally for the detection of diabetic retinopathy.

The US Food and Drug Administration has already given the green light to many smart phone-based fundus camera devices for retinal imaging; some of these devices also support AI-based image processing. It is hoped that when these cheap fundus cameras become more widely available, it would improve access to care and lessen the burden of eye illnesses worldwide.

Despite all the recent developments and improvements in ocular imaging, there are still specific open issues that the ophthalmic community should consider and resolve. One such issue is the curriculum and training for aspiring ophthalmologists [11].

2.2 Processing methods

2.2.1 Image Processing-Based DR Detection

Matching filtering (MF) and bilinear top-hat screening are used in the study in [11], and the outcomes are then supplied to a region-growing algorithm. Other publications, such as wavelet transform in [12] and [13] and radon transform in [5], convert image-based characteristics to a different space. Many of these algorithms have produced pleasing outcomes while still being computationally economical, however, straightforward architectures such as these are prone to have functional problems. Such as, many fundamental features like microaneurysms and haemorrhages can be falsely identified with intensity-based algorithms.

2.2.2 Traditional Machine Learning Algorithms

SVMs were used by [14] and [15] to find blood vessels and locate PDR instances. A study in [16] focusses on finding microaneurysms (MAs) by implementing an SVM based ensemble model structure, multi-model medium and Gaussian Mixture Models (GMM). Study in [17] and [18] proposed a method where SVM assigned a grade of DR and tracked changes in red lesions. Although this approach allows for rapid class separation and complicated feature representations, it still involves some human interaction that might lead to mistakes in difficult jobs.

A likelihood distribution normalization and the K Nearest Neighbors (KNN) algorithms were implemented by [19]. The study in [13] generates splattered partitions after dividing fundus images, it selects the best features using a filter and a wrap-per sequence, then implements a KNN model to classify possible hemorrhages. KNN performs well across tasks as a general-purpose classifier, although it may have problems with computing efficiency and generalizability. Therefore, training on new data across various population applications or retaining extremely big training data would be required for distinct patient groups.

Based on RFs, [20] and [21] detect binary DR. RFs can also distinguish between lesions and non-lesions thanks to candidate morphology, according to [12].This technique

can provide obvious class distinctions, but it involves manual work and may have problems with less defined class boundaries (i.e. mild from mod- erate).

Particle swarm optimization (PSO) was used by [22] to extract characteristics from retinal pictures, which were then input to a NN to categorise DR.

2.2.3 Neural Networks, Deep Learning and Transfer Learning

The AlexNet model was altered by [23] to classify retinal pictures according to the degree of DR. For assessing DR in retinal pictures, [17] studied existing deep learning CNN frameworks and architectures such as VGGNet, ResNet, GoogleNet and AlexNet,).

Chapter-3

DESIGN AND METHODOLOGY

In this chapter we have discussed the steps and processes to develop a mobile application, integrated with an artificial intelligence model, to detect DR. Section 1 discusses details about the device design that was used to collect the dataset, as the data to train our model based on 20D lens previously did not exist publicly, therefore, we obtained it by capturing images of OPD patients with the support of Sindh Institute of Ophthalmology and Visual Sciences hospital, Hyderabad (SIOVS), details are discussed in section 2 of this chapter. The data collected was labelled and graded with the help of ophthalmologists at the supporting hospital. The algorithms used are discussed in section 3 and finally, the development of the mobile application, integrated with the model is described in the 4rth section.

3.1 Dilation

The patient's eye is fully dilated using dilating drops prescribed by the ophthalmologists before capturing the image in order for the fundus to be visible as depicted in figure 3.1.

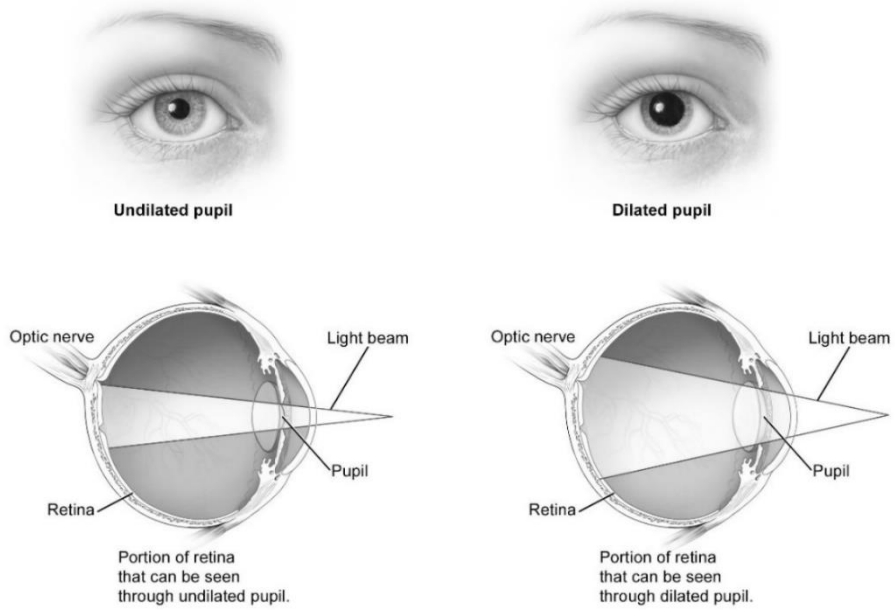


Figure 3.1- Dilation of Eye

3.2 Device Design

This device is a hand held fundoscope that can be used by a trained ophthalmologist or anyone with minimal training to capture images. A complete device is depicted in figure 3.2.

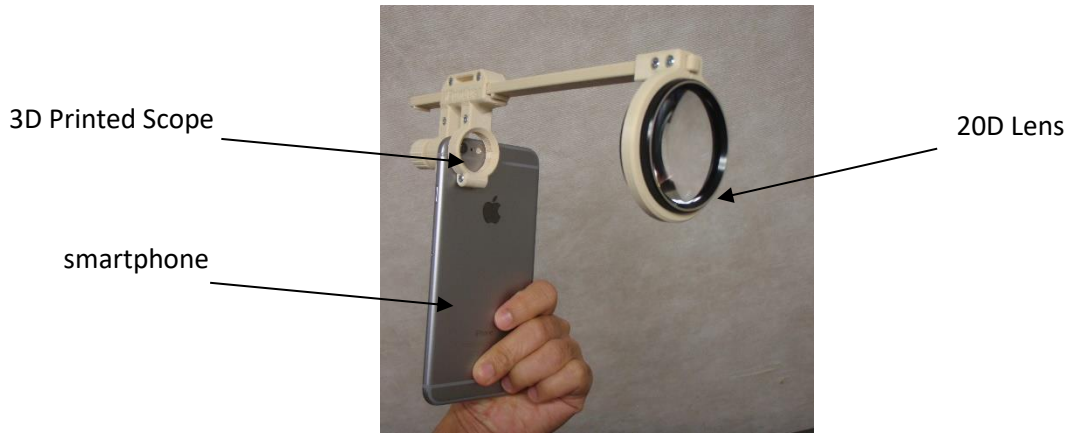


Figure 3.2- Complete Device

3.2.1 20D lens

The 20D lens is used for general examination of the fundus. This provides a high-resolution image of the retina in the OPD or the operating room. A 45-degree field of view helps in visualization up to the mid peripheral region.

Figure 3.4- Volk Aspheric 20D Lens

Figure 3.3- Horizontal view of a 20D lens

3.2.2 3D Printed Scope

The 3D scope is adapted from the Odocs eye care (https://odocseyes.co.nz/pages/3d-print) open source 3D printing source. It is the world's first open source retinal imaging adaptor called the OphthalmicDocs Fundus. It is used to perform indirect ophthalmoscopy using a smartphone and a 20D lens. This device converts a smartphone into a fundus camera that can capture images of the retina.

3.2.3 Smartphone

Images of the fundus were taken using an Android smartphone with OS Version 6 or above. The smart phone is attached to one end of the 3D scope and the other end is fixed with a 20D lens. We opened the camera (preferably in video mode with flash/torch option on) and focused the retina to the center of the camera to capture a clear image or record a video from which the image was extracted later on.

Figure 3.5- 3D Printed Scope

3.3 Working of the Device

1. To get rid of any debris, thoroughly clean the 20D lens and the smartphone camera lens before the process. The slots at the ends of the 3D scope gadget were used to hold a smartphone with a Volk 20D lens.
2. The gadget is held around 3-5 cm in front of the patient's eye. Once the fundus of the eye is visible and clear in the camera, the image is captured or a video is recorded from the smartphone screen, the flash option is activated throughout this process for a consistent light source.
3. The gadget was slightly tilted or moved back and forth to do this, and an infant depressor was used to stabilize the patient's eyes. Additionally, good mydriasis reduced reflections.
4. After the recording was complete, the high-quality snapshots that were obtained were extracted from the video frames using the smartphone's screenshot capability.
5. To classify the snapshot photos using DR, the photographs were then uploaded to a smartphone application.

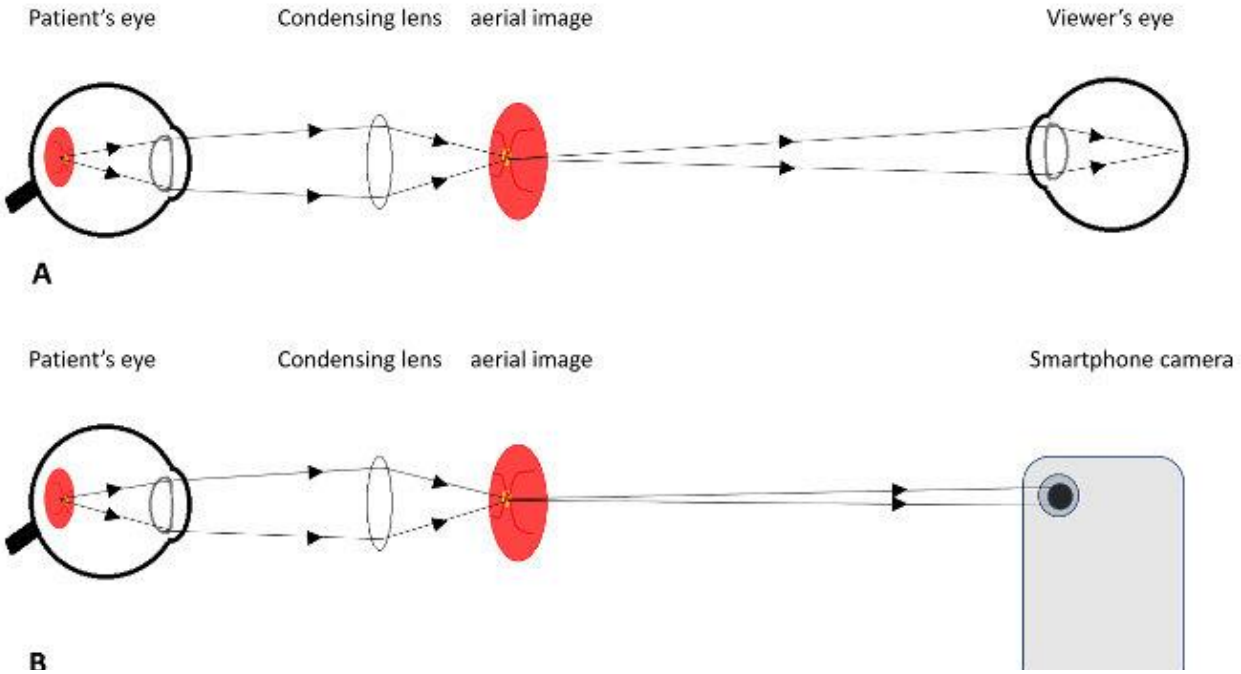


Figure 3.6- Image Acquisition Process

3.4 Dataset Collection

3.4.1 Dataset

We collected the fundus image data from the Sindh Institute of Ophthalmology and Visual Sciences (SIOVS) Hyderabad hospital and merged it with an open source Kaggle competition data for detection of diabetic retinopathy. Finally, we were able to create a dataset of 35,122 total images divided into two classes: DR and No DR. 9,313 images were DR fundus and 25,809 images were No DR fundus.

3.4.2 Preprocessing

Since the open-source data was collected from different sources it had a lot of difference in the images. Some were zoomed in and some zoomed out. The ratio of fundus of an eye to image had to be set for all images. Hence, we created a Python script to accomplish the preprocessing of all images and cut the extra black background of all the images and set the standard resolution for all images, i.e., 300*300.

Figure 3.7- Before Preprocessing

Figure 3.8- After Preprocessing

We separated 790 images from the data for validation. Later we divided the rest of the data into 80% training dataset and 20% validation dataset.

3.5 Convolutional Neural Network (CNN)

As the focus of this research is based on the classification of diabetic retinopathy images, we used Convolutional Neural Network algorithms, which is a Deep Learning algorithm. Reason to focus on CNNs is that they compete dominant in object detection and image classification problems.

3.5.1 How CNNs Operates

CNNs are a subset of neural networks that take their cues from how the brains of both humans and animals process information and recognize patterns and pictures. Therefore, neurons with learnable weights and biases are used by CNNs.

3.5.2 CNN Architecture

Input, hidden, and output layers are the layers that make up CNN models. Layers are a group of neurons that receive several inputs and process the information to produce an output.

Each neuron in the input layer receives values from the picture pixel(s). The processing of the inputs, which includes feature selection, is handled by hidden layers. In the end, the output layer has as many neurons as the classes or targets that the user needs to forecast. The output class is thought to be the neuron with a high likelihood of producing an output.

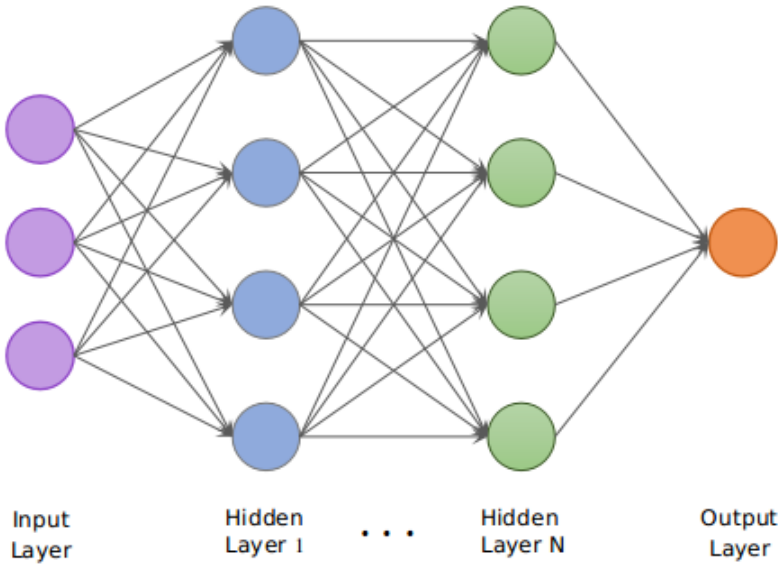


Figure 3.9- Basic CNN Architecture

3.6 CNN Layers

The CNN is made up of three different kinds of layers: fully-connected (FC), pooling, and convolutional layers. A CNN architecture is created when these layers are layered. The dropout layer and the activation function, which are defined below, are two additional crucial factors in addition to these three layers.

3.6.1 Convolutional Layer

The different characteristics from the input photos are extracted from the first layer, which is used. The input picture and a filter with the specific size M x M are convolutional mathematically in this layer. The model filter passed across each dataset image and the filter M x M is applied accordingly to specific regions of the image.

An a resultant image, providing details about the image corners and edges is obtained, which is also known as a Feature Map. This Feature Map is then used as an input image for further layers, in order to learn more features and produce optimal training accuracy. The outcome is later shared with the following layer by the CNN's convolutional Layer.

The result is known as the Feature map, and it provides details about the image, including its corners and edges. This feature map is later supplied to further layers to teach them more features from the input picture.

Once the convolution operation has been applied to the input, CNN's convolution layer transfers the output to the following layer. The spatial link between the pixels is preserved thanks to convolutional layers of CNN.

3.6.2 Pooling Layer

A Pooling Layer often comes after a Convolutional Layer. This layer's main goal is to lower the convolved feature map's size in order to save on computational expenses. This is done individually on each feature map and by reducing the links between layers. There are several sorts of pooling operations, depending on the mechanism utilized. In essence, it provides a summary of the characteristics produced by a convolution layer.

The convolutional and the FC layers are interlinked. Sum Pooling calculates the components' combined sum in the specified segment. Average pooling is used to calculate the average of the components in an image segment of a specific size. Feature Map is a used to obtain important components and assets for Max Pooling.

3.6.3 Fully Connected Layer

The fully connested layer is placed before the final output layer of the CNN model structure. To link the neurons between two layers, the Fully Connected (FC) layer, which also includes weights and biases, is utilized.

The process above feeds a flattened image to the FC layer. After passing it through a few more FC layer levels, the input image is used to calculate some standard mathematical operations and functions. These two layers act together to perform categorization which in turn reduces human effort by developing an intelligent CNN architecture.

3.7 Dropout

When the model performs exceptionally well on the training data, however, fails to produce similar accuracy on the new data in the prediction process then the model has been overfitted. This phenomenon generally happens when all major features of an input image are connected to the FC layer.

A dropout layer, which reduces the size of the model by removing a few neurons from the neural network during training, is used to solve this issue.

3.8 Activation Functions

The activation function is one of the most crucial elements of the CNN model. They are employed to discover and approximation any type of continuous and complicated link between network variables. In layman's terms, it determines which model information should shoot ahead and which should not at the network's end.

The network gains nonlinearity as a result. The ReLU, Softmax, tanH, and Sigmoid functions are a few examples of regularly used activation functions. Each of these operations has a particular use. Sigmoid and softmax functions are recommended for a CNN model for binary classification, while softmax is typically employed for multi-class classification. To put it simply, activation functions in a CNN model decide whether or not to activate a neuron. It determines through mathematical processes if the input to the work is significant or not.

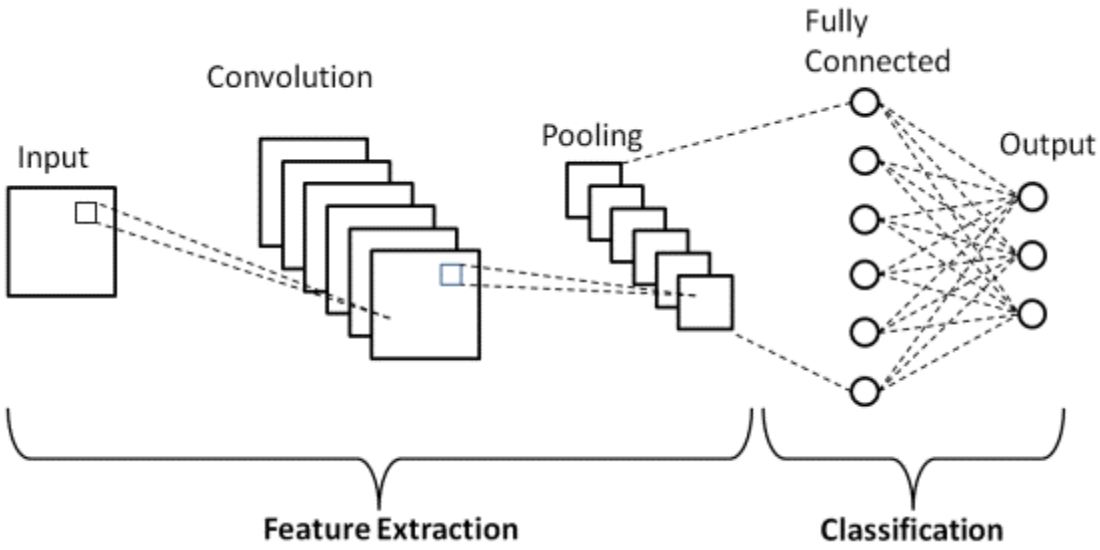


Figure 3.10- Basic CNN Architecture (with layers)

3.9 Training Process of CNN

CNN algorithms are trained using the following two steps: Forward and Backward Propagation.

3.9.1 Forward Propagation

Numbers are used to feed images into the input layer. These numbers represent the pixel intensities in the picture. A few mathematical operations are performed on these values by the neurons in the hidden layers.

There are some parameter values that are randomly initialised in order to carry out these mathematical processes. Following these calculations at the hidden layer, the output layer sends the result, which creates the final prediction.

3.9.2 Backward Propagation

The next step is to compare the output with the real value once the output has been created. The values of the parameters are changed based on the final result and how close or distant this is from the real value (error). With the revised parameter values, the forward propagation process is restarted, and fresh outputs are produced.

3.10 Experimental Setup

To perform the experiments for the classification of Diabetic Retinopathy using Deep learning approach, we have used AWS cloud computing Machine Learning service and Google Colab Pro configuration as an experimental setup.

Accelerated computing machine instances in AWS and Google Colab Pro subscriptions were selected to perform experiments as depicted in Table 1.

Table 3.1- Cloud Instance Configurations

Service	Machine	Memory	Pricing
AWS	ml.g4dn.xlarge	16GB 4 cores	$0.7364/hour
Google Colab Pro	K80 and P100	32GB	$9.99/month

3.11 Algorithms

We used three different deep learning algorithms along with transfer learning to train our model. ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is an annual worldwide deep learning competition that evaluates algorithms developed on the ImageNet dataset that has a variety of classes, for image classification and object detection. The algorithms used below are the top-performing algorithms from ILSVRC. These algorithms can outperform almost every dataset with a bit of fine-tuning.

3.11.1 VGG16

VGG16 is a 16-layer CNN algorithm that won ImageNet competition on classification of images and detection of objects in 2014. VGG16 consists of 16 layers having weights, out of which 13 are convolution layers, and 3 are Dense layers. 5 Max Pooling Layers are used that do not hold weights.

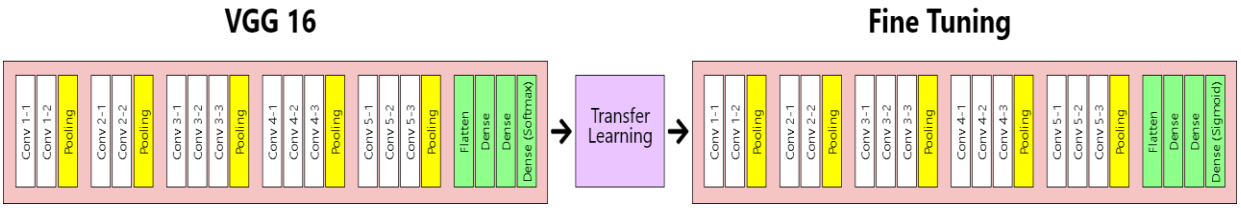


Figure 3.11- Fine tuned VGG16 Architecture

VGG16 model accepts a tensor of size 224*224 with 3 RGB channels.

Conv-1 Layers have filter size of 64, Conv-2 filter size is 128, Conv-3 having 256 filters and Conv-4 and Conv-5 consists of 512 filters.

We prepared the dataset in image data generator and used data augmentation to increase the data size as well as performed preprocessing on images for better results on the algorithm used.

We downloaded the VGG16 model and fine tuned it. We removed the default output Dense layer and added a new Dense output layer to predict from 2 classes (DR or No DR). We used ‘sigmoid’ activation function for output layer.

We compiled the model using ‘Adam’ optimizer with default learning rate i.e., 0.001 and ‘binary_crossentropy’ loss function.

We used early stopping callback on the training process by monitoring validation accuracy, which means if validation accuracy doesn’t seem to improve after certain number of epochs the training process will end. We also used the restoration of weights property on early stopping to use the weights from the epoch performing best on validation data.

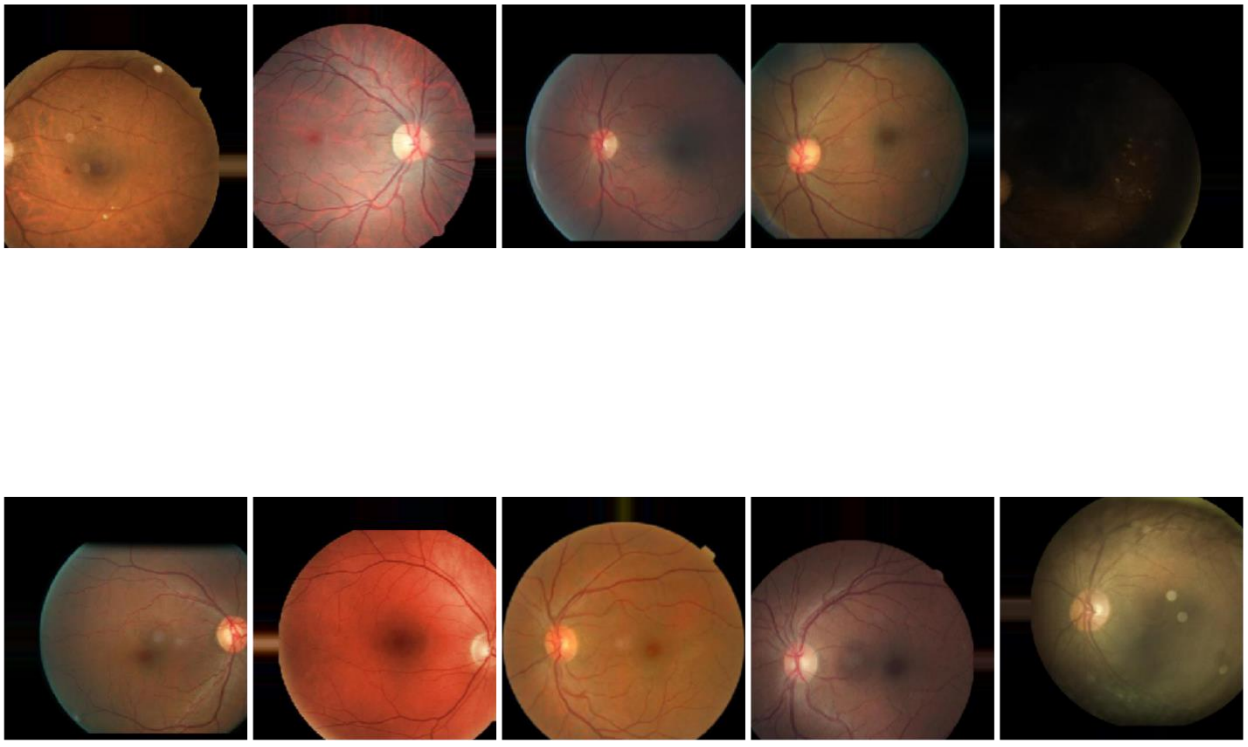


Figure 3.12- Image Preprocessing

3.11.2 Resnet50

A convolutional neural network with 50 layers is called ResNet-50. A common neural network that serves as the foundation for many computer vision applications is called ResNet, short for Residual Networks. The main innovation with ResNet was that it enabled us to train very deep neural networks with more than 150 layers.

A significant drawback of convolutional neural networks is the "Vanishing Gradient Problem." Gradient value greatly reduces during back propagation, therefore weights scarcely change at all. ResNet is employed to get around this. It employs "SKIP CONNECTION."

The following component can be found in the Resnet 50 architecture:

- The first convolutional layer is given by 64 distinct kernels having a stride and size of value 2. Along with a kernel size of 7 * 7.
- Max pooling was observed when the size and stride value was kept to 2.
- A 1 * 1 with 64 kernel is used in the below mentioned convolution layer. Along with it, a 1 * 1 with 256 kernel is included.

- Following that, a kernel of 1 * 1 128 is shown, then one of 3 * 3 128 and ultimately one of 1 * 1 512. This procedure was iterated for a total of four times on 12 layers respectively.
- The next kernels are 1 * 1 256, 3 * 3 256, and 1 * 1 1024, and this is done six times, giving us a total of 18 layers.
- After that, a 1 * 1 512 kernel was followed by two more layers of 3 * 3 512 and 1 * 1 2048, which was then repeated three times to give us a total of nine layers.
- Next, we do an average pool, finishing with a fully linked layer made up of 1000 nodes and a softmax function to give us one layer.

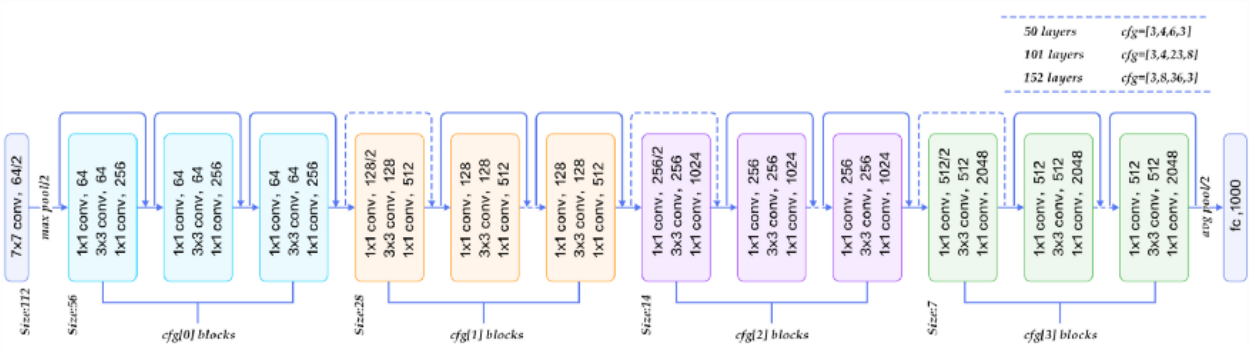


Figure 3.13- Resnet 50 Architecture

Skip connection was initially offered by ResNet. The skip connection is depicted in the diagram below. Convolution layers are being stacked in the left-hand figure one on top of the other. the right, we continue to build convolution layers as previously, but this time we

additionally include the original input in the convolution block's output. It's known as skipping the connection.

Figure 3.14- Skip connection in Resnet 50

The input tensor used by Resnet50 has a dimension of 224*224. Thus, we created the dataset by employing an image data generator, preprocessing the photos to reduce their size, and then adding data augmentation by rotating, flipping, zooming in, and adjusting the width and height of the images.

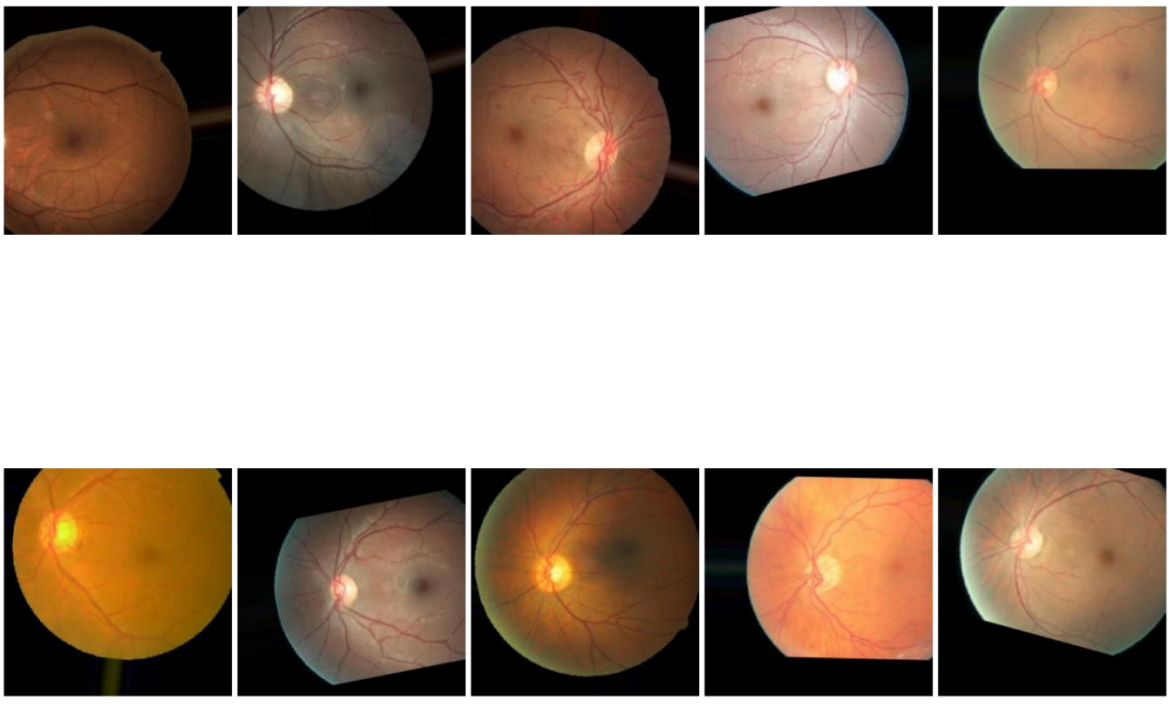


Figure 3.15- Data Preparation and Preprocessing - Resnet50

In order to address the issue of DR categorization, we applied transfer learning on the downloaded Resnet50 model and fine-tuned it to add a few extra layers.

We added the next layers, in that order:

- Flatten
- Dropout layer
- Dense layer
- Dropout
- Dense output layer

We made our finely adjusted Resnet model to learn on the last block while also freezing the default weights of the Resnet50 model trained on the ImageNet dataset.

We built the model using an Adam optimizer with a learning rate of 0.0001 and a binary cross-entropy loss function.

To recover the best weights and to stop the training process after 12 epochs, we employed Early Stopping Callback. We used a batch size of 64 for both the training and validation datasets for doing the training.

3.11.3 Custom CNN Model

A deep learning model with 16 layers is created; out of which seven are convolutional layers, three are pooling layers, the first layer is a flattening layer, second layer is dropout layers and the third layer is dense layer.

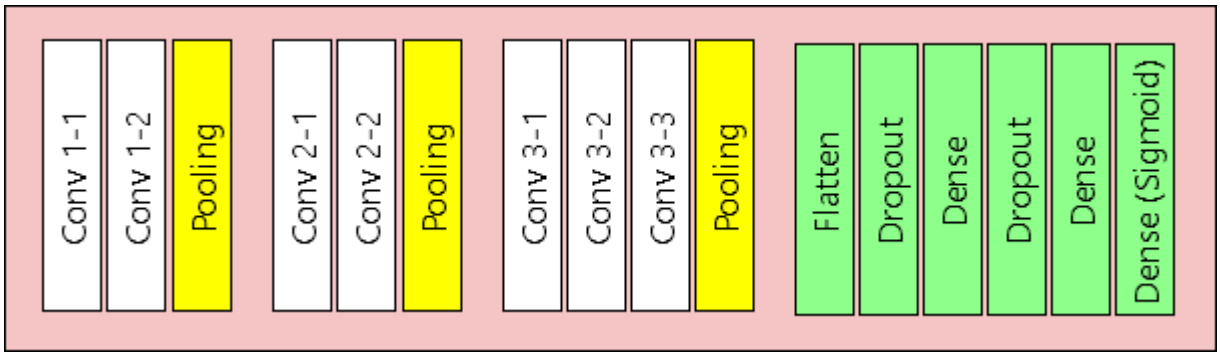


Figure 3.16- Custom Model Architecture

The model is directed to accept an input image of size 224 * 224. According to the Figure 3.16, Conv 1 layers use 64 filter size with kernel size 3. Conv 2 uses 128 filters with 3 * 3 kernel size. Conv 3 uses filter size of 256 with kernel size 3. For all pooling layers, which we max pooling layers, a pool size and stride size of value 2 was used. Dropout layers are used with a factor of 0.3. The last dense layer is used for classification and it uses sigmoid activation function. All the Convolutional layers are used with same padding i.e., the image size won't reduce due to convolution operation.

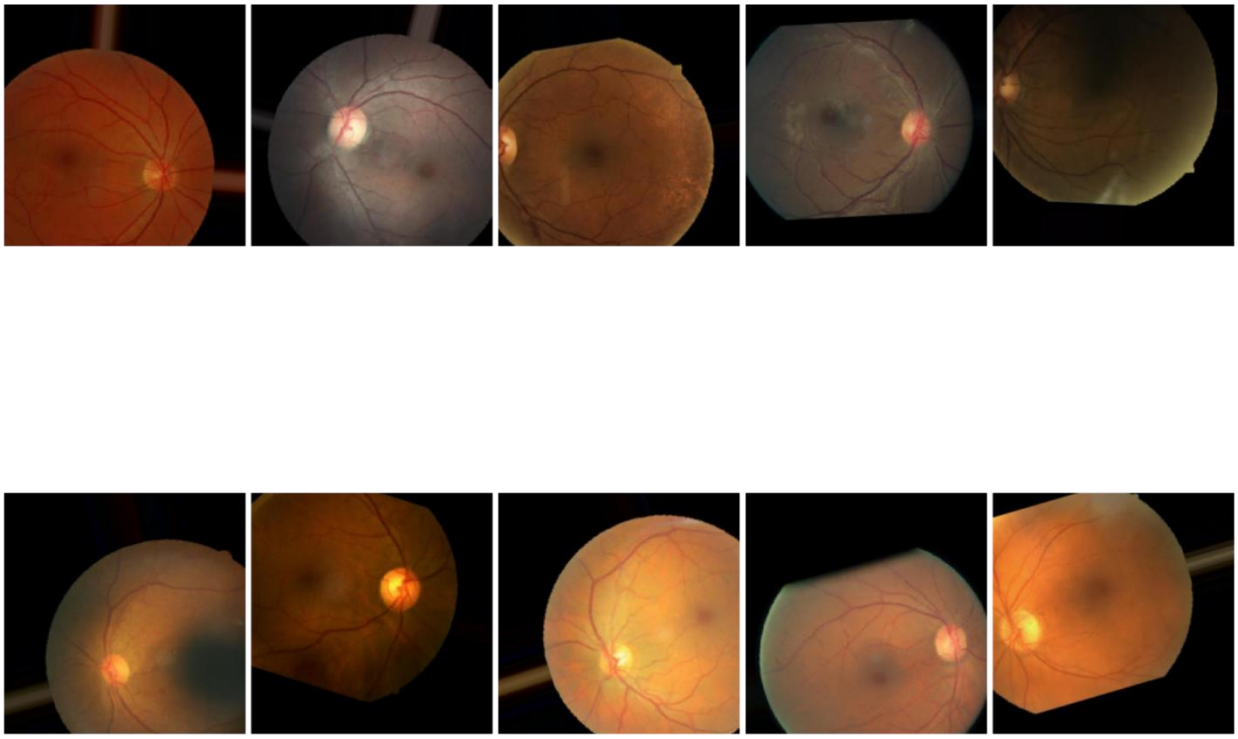


Figure 3.17- Image Preprocessing- Custom CNN Model

We prepared the dataset using image data generator. We used image data augmentation on train and validation data. We compiled the model using 'adam' optimizer. We used 'binary_crossentropy' loss function. For efficient training we used early stopping by checking the improvement in validation accuracy and also defined to restore the best weights. We trained the model on training data and validated on validation dataset using 200 epochs and early stopping callback.

3.12 Mobile application

An Android phone with OS Version 6 or higher may be used to take wide-angle digital color photos of the retina (fundus) using this mobile application built on the Android OS. This application is a hybrid device made up of a smartphone app and a 3D-printed scope with an indirect ophthalmoscopic lens attached. The objective of this application is to capture pictures of the retina of the eye for general viewing and patient education. With the aid of this application, the consultant may easily and swiftly take fundus photos for patient education and examination. It helps doctors and patients have talks about how the condition

Figure 3.18- Application User Flow

is progressing and the best course of action.

3.12.1 Image sharing Flow

1. The user will login the application using their unique credentials. The user will select camera to capture images and will also provide access to select images from gallery.

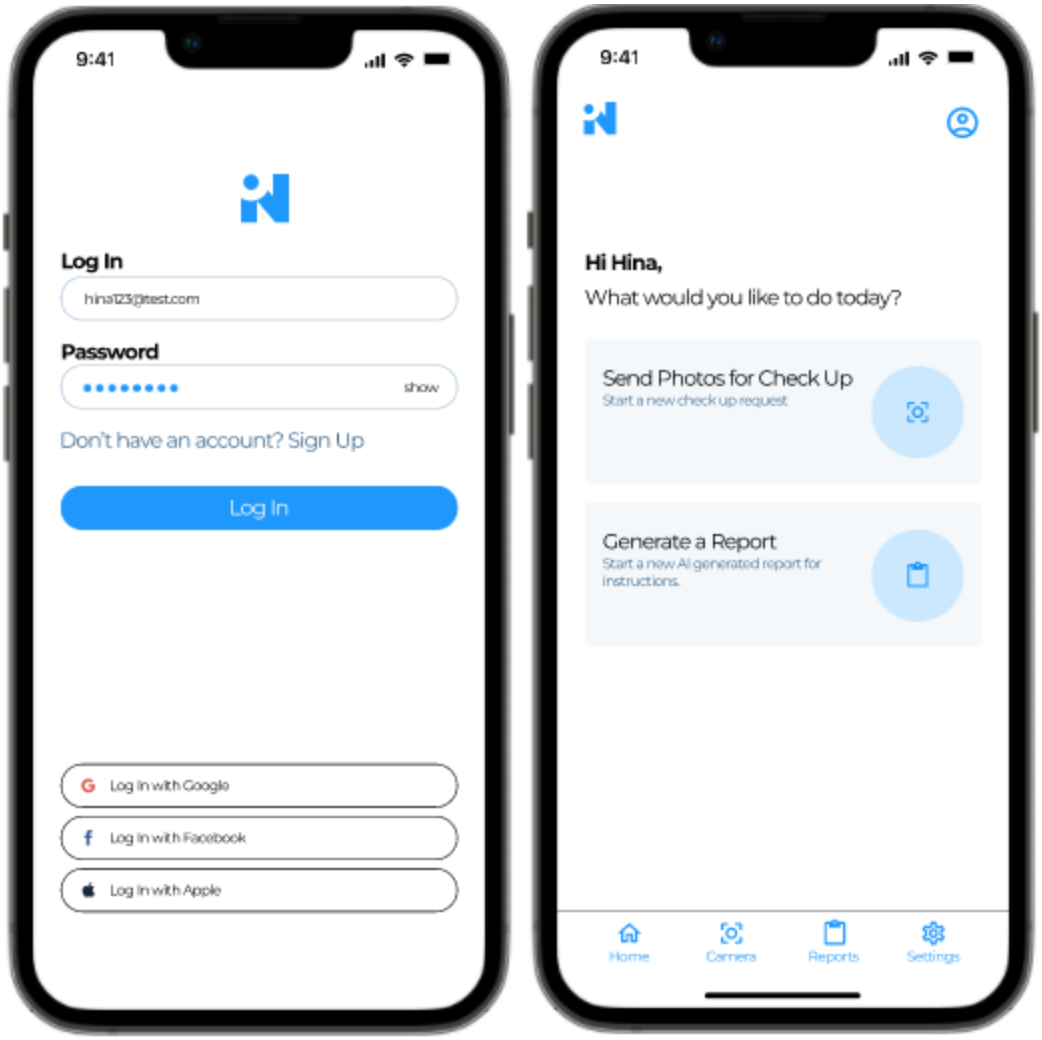


Figure 3.19- Login and Home Screen

2. The user will then attach the 3D scope with the smartphone and adjust the camera position so that the fundus of patient is at the center of the camera.

3. After capturing multiple images, the user will then select clear fundus images for further sharing.

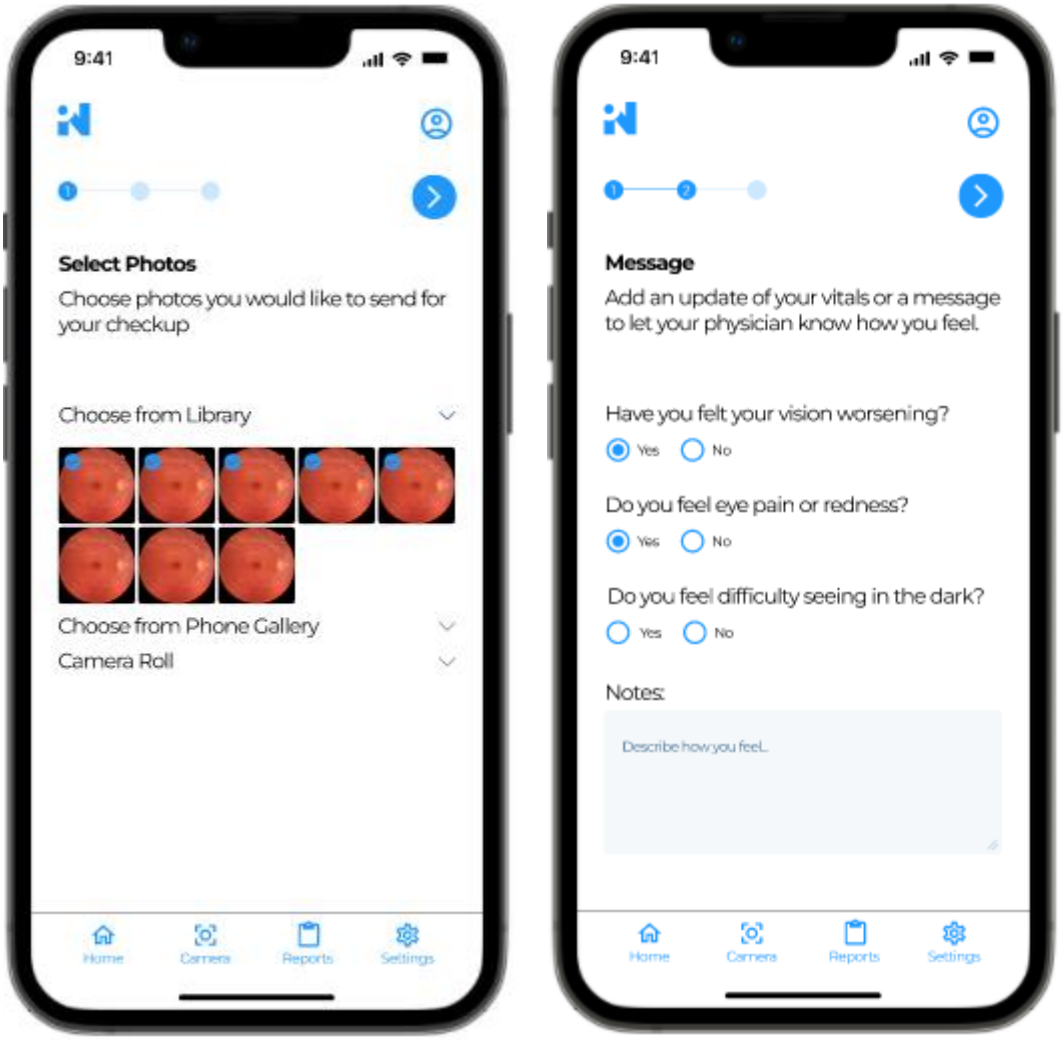


Figure 3.20- Image Selection and Questionnaire Screen

4. The user will then answer some questions regarding patient eye history and examination.

5. The user will the select the consultant and share the examination data.

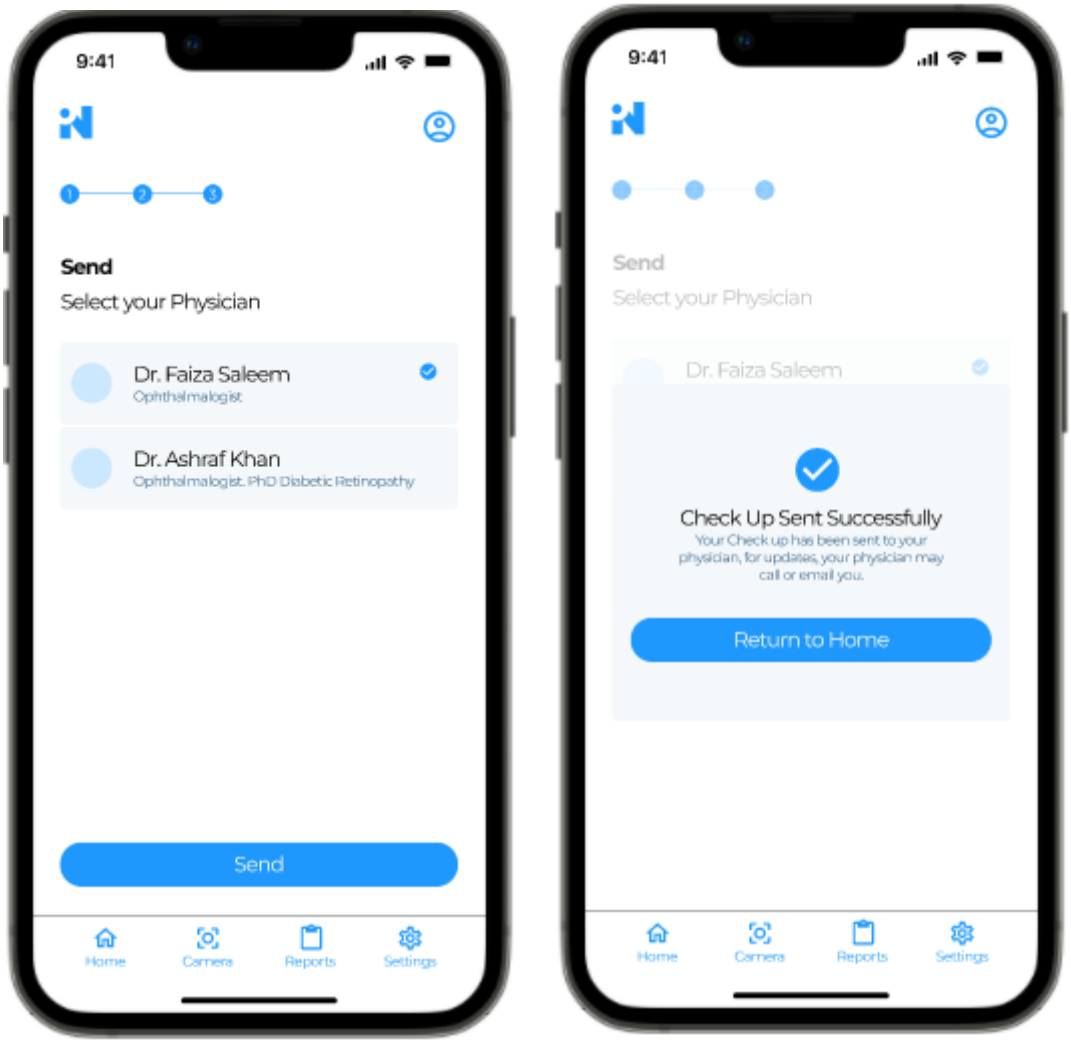


Figure 3.21- Consultant Selection and Confirmation Screen

3.12.2 Automated Report Generation Flow

1. For generating an AI report, the user will select generate a report option from the Home Screen. Figure 3.19.
2. The user will open camera or select images from gallery and upload them to the application. Figure 3.22.

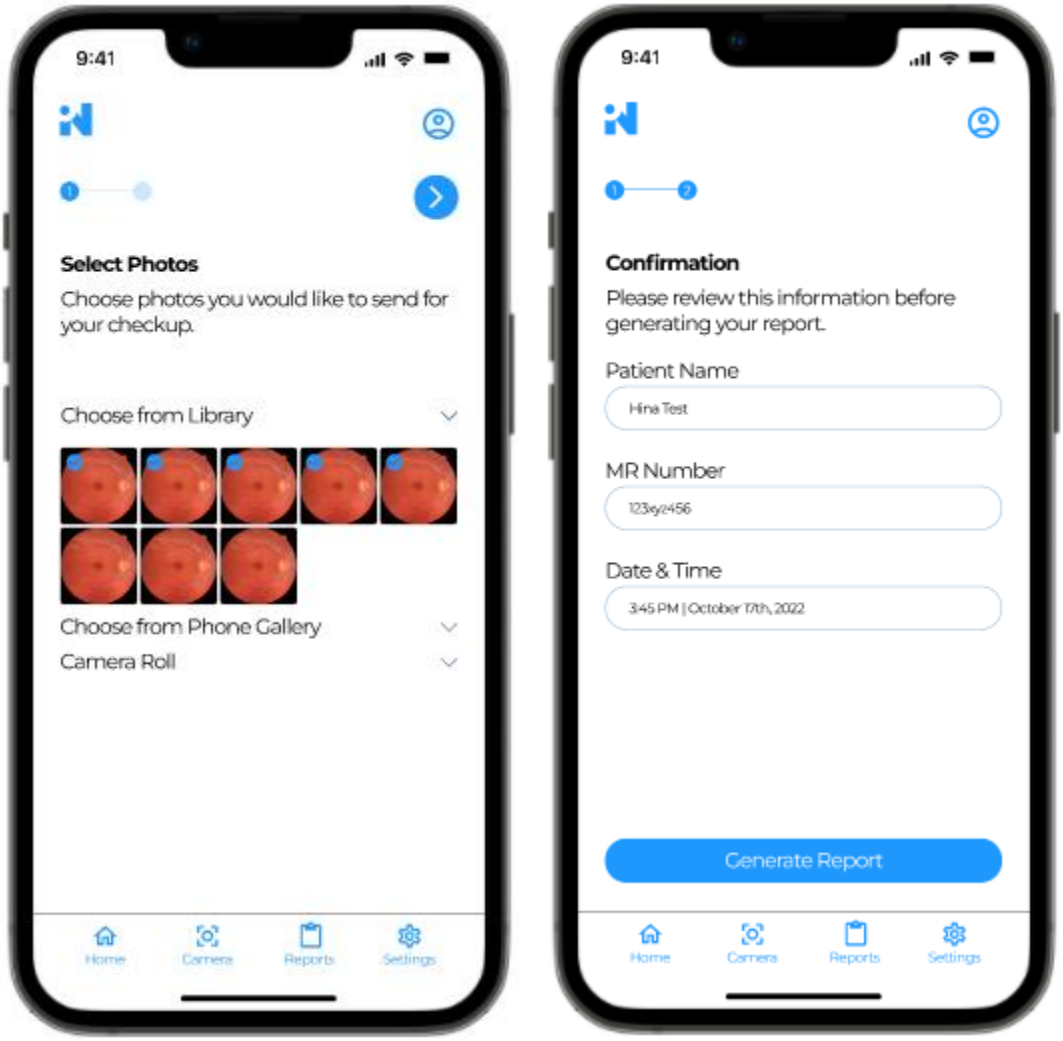


Figure 3.22- Select images and enter patient data for AI report

3. Then the user will enter patient information for patient record and click on generate report. Figure 3.23.

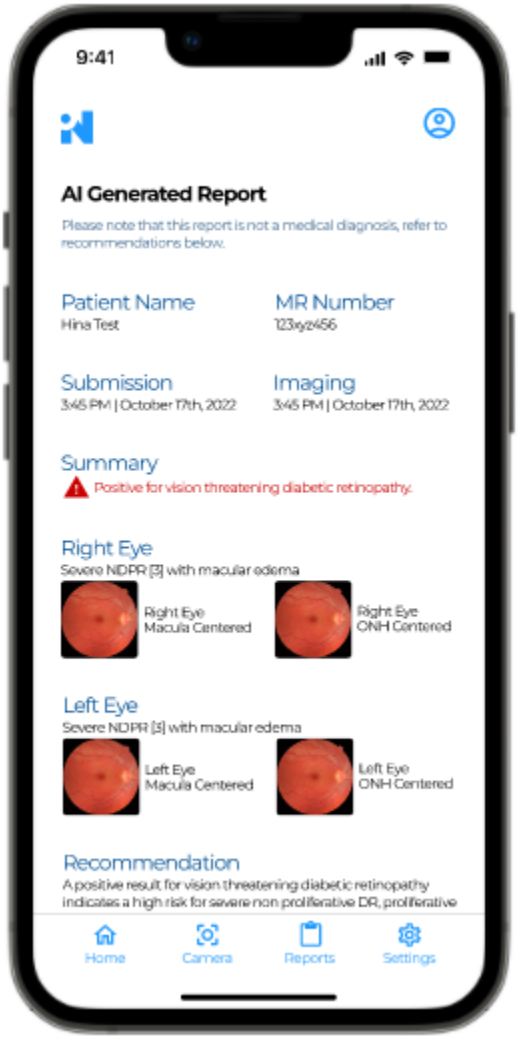


Figure 3.23- AI-Generated DR Report

Chapter-4

RESULTS AND DISCUSSION

After performing various deep learning frameworks mentioned in Chapter 3 on image dataset of Diabetic Retinopathy, the obtained results are reviewed in this section for further comparative analysis.

4.1 VGG16

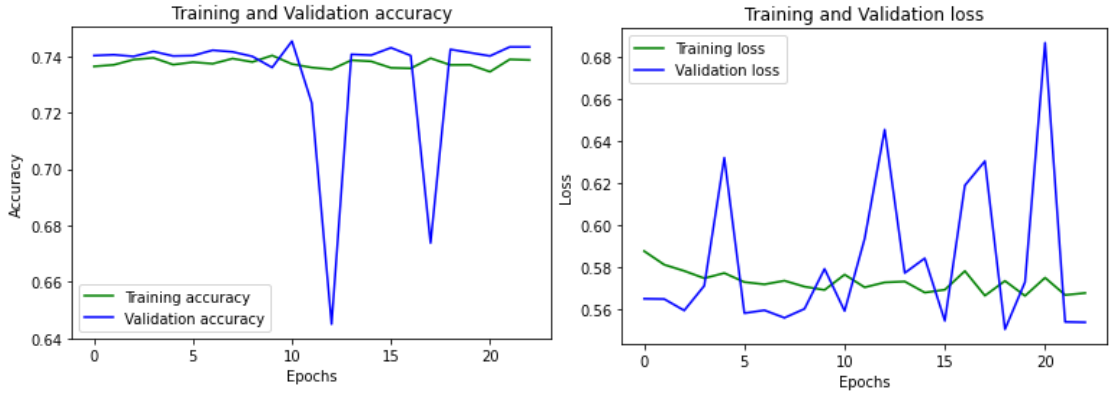


Figure 4.1 Training and Validation Accuracy/Loss

The VGG16 Model was trained on 100 epochs. On approaching 23rd epoch the model validation accuracy remained constant and did not improve. Therefore, training process was stopped due to early stopping callback. Configurations and weights were restored from optimal epoch i.e. epoch 11.

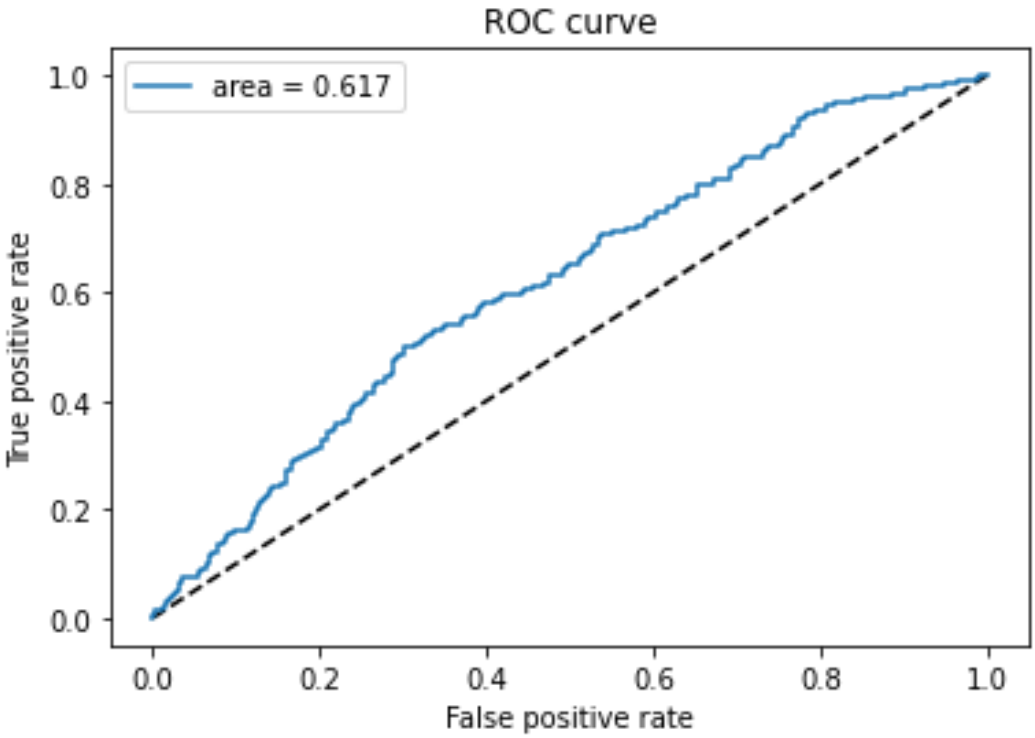


Figure 4.2- ROC Curve - VGG16

Area Under the Curve (AUC) refers for the level or measurement of separability, and Reciever Operating Characteristics (ROC) is a probability curve.

True Positive Rate (TPR) or recall or sensitivity is plotted against False Positive Rate (FPR) or specificity on the ROC curve, with FPR on the x-axis and TPR on the y-axis.

4.2 ResNet50

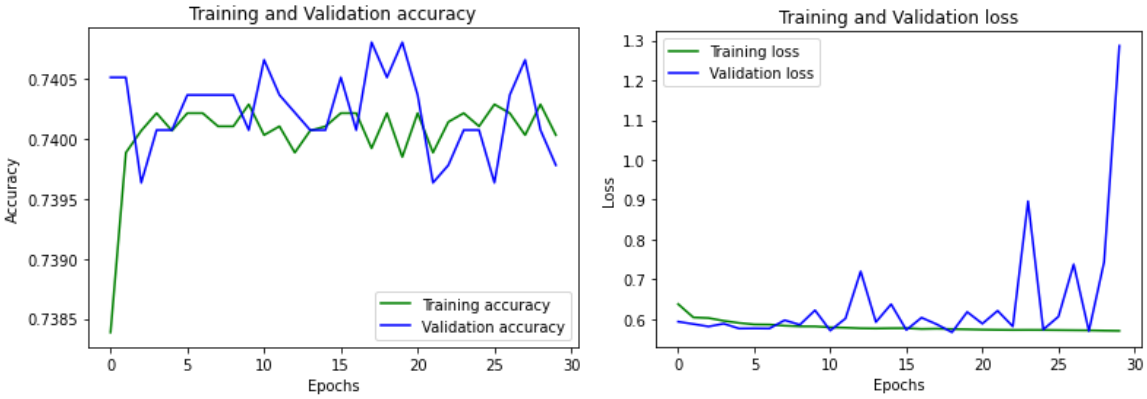


Figure 4.3 Training and Validation Accuracy/Loss

After 30 epochs model training stopped with a callback of early stopping. Validation accuracy didn't improve for a few epochs and callback interrupted the training. The best model weights were restored from epoch 18.

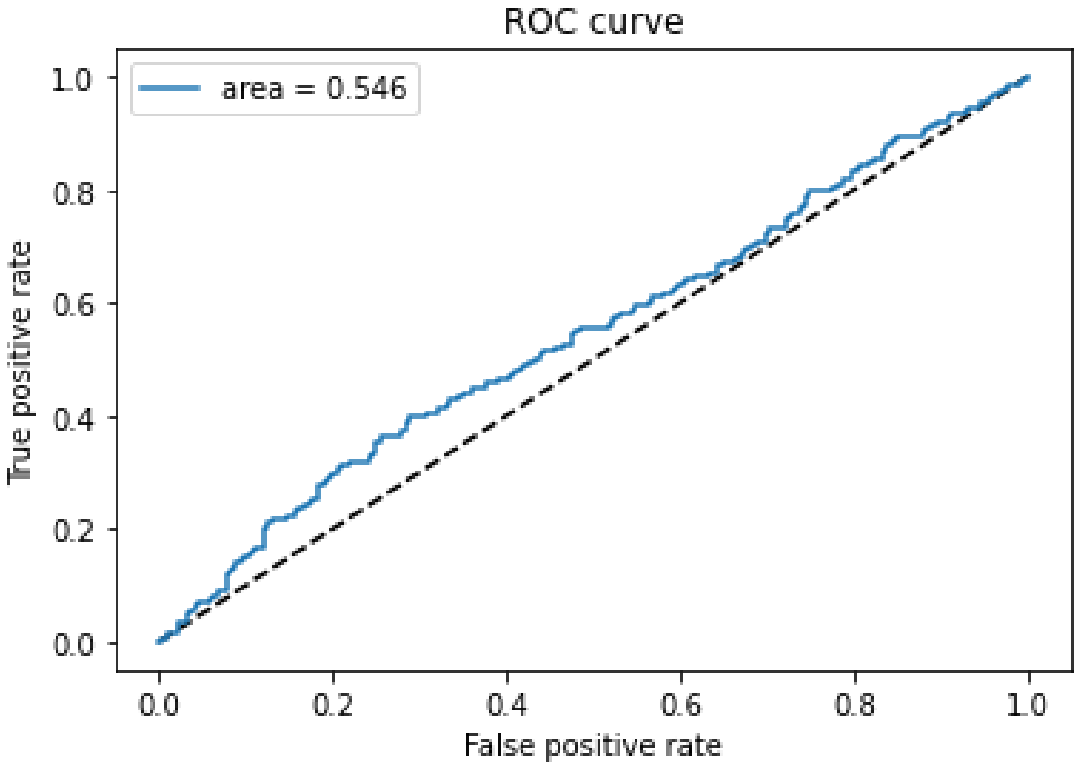


Figure 4.4- ROC Curve- ResNet50

4.3 Custom Model

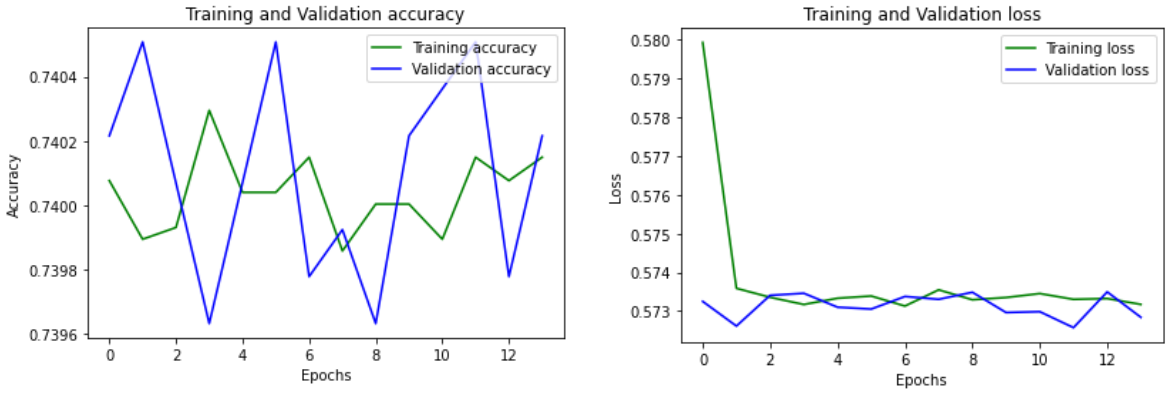


Figure 4.5 Training and Validation Accuracy/Loss

On approaching 14th epoch Model training occurred successively till 14 epochs. After 14th epoch, model training wmodel was force stopped due to early stopping callback on approaching highest validation accuracy. Additionally, the model weights from epoch 2 due to optimal performatce.

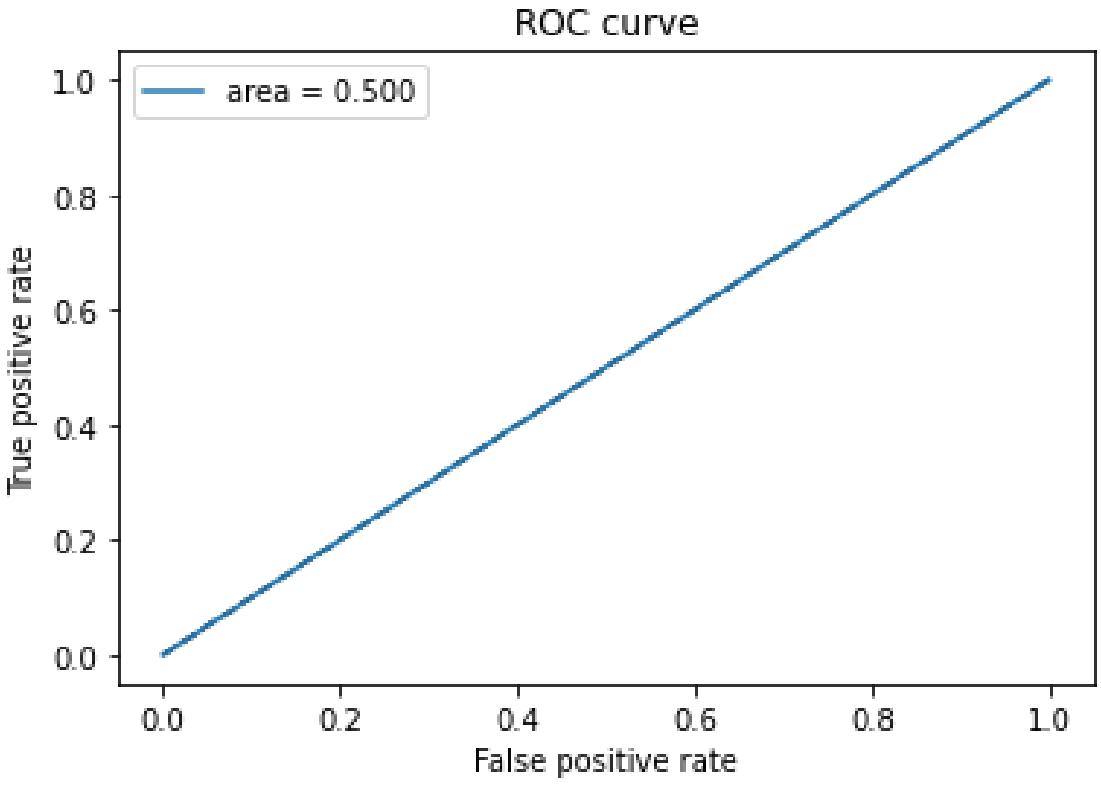


Figure 4.6- ROC Curve- Custom CNN Model

4.4 Comparative Analysis

After comparing results of all the trained deep learning models, VGG16 outperformed other model by obtaining highest validation accuracy i.e., 74.53% as well as lowest validation loss of 55.94%. Thus, VGG16 performed best on the dataset provided and is deployed in the smartphone application.

Table 4.1- Comparative Analysis

Model	**Training Accuracy**	**Training Loss**	**Validation Accuracy**	**Validation Loss**	**AUC**
VGG16	0.7372	0.5767	0.7453	0.5594	0.6174
ResNet50	0.7399	0.5762	0.7408	0.5872	0.5464
Custom CNN Model	0.7399	0.5736	0.7405	0.5726	0.5

Chapter-5

CONCLUSION AND FUTURE RECOMMENDATIONS

After researching the existing systems and approaches for the detection of Diabetic Retinopathy using Smartphone Based photography with 20D lens or Fundus on a phone Photography, we conclude that our proposed DR detection technique is successful in detecting Diabetic Retinopathy. Our proposed method classifies the DR disease in two classes, namely DR (Patient has Diabetic Retinopathy) and NO DR (Patient is healthy).

5.1 Our Research Outcomes

A Publicly available Kaggle dataset EyePACS is merged with the dataset obtained from the SIOVS hospital. Three components area of exudates, the area of blood vessels, and the area of the microaneurysm have been used to classify the samples. We have divided the conditions into two categories: DR and NO DR. With the VGG16 model, our highest obtained accuracy is 74.53% respectively. The metrics that we operated on in this work are contrasted with previous efforts.

5.2 Future Recommendations

A larger dataset with direct ophthalmic 20D images needs to be constructed for training data using deep learning frameworks in order to obtain optimum accuracy. Dataset images need to be labelled using expert opinion and must have consistent and high resolution. Additionally, dataset image frames must be kept identical. With more good resolution training data obtained from 20D lens the accuracy can be increased and an efficient and robust system can be developed to diagnose Diabetic Retinopathy.

REFERENCES

[1] "Diabetes." https://www.who.int/news-room/fact-sheets/detail/diabetes (accessed Aug. 09, 2022).

[2] "Blindness and vision impairment." https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (accessed Aug. 04, 2022).

[3] C. H. Tan, B. M. Kyaw, H. Smith, C. S. Tan, and L. T. Car, "Use of smartphones to detect diabetic retinopathy: Scoping review and meta-analysis of diagnostic test accuracy studies," *J. Med. Internet Res.*, vol. 22, no. 5, 2020, doi: 10.2196/16658.

[4] "British Medical Journal," *British Medical Journal*, vol. 1, no. 2509. pp. 295–301, 1909, doi: 10.1136/bmj.1.2509.295.

[5] S. Stolte and R. Fang, "A survey on medical image analysis in diabetic retinopathy," *Med. Image Anal.*, vol. 64, p. 101742, 2020, doi: 10.1016/j.media.2020.101742.

[6] IDF, "Media Hub | World Diabetes Day," pp. 1–3, 2021, [Online]. Available: https://worlddiabetesday.org/media/.

[7] R. E. Hacisoftaoglu, M. Karakaya, and A. B. Sallam, "Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems," *Pattern Recognit. Lett.*, vol. 135, pp. 409–417, 2020, doi: 10.1016/j.patrec.2020.04.009.

[8] US Department of Health and Human Services, "National Diabetes Statistics Report, 2020," *Natl. Diabetes Stat. Rep.*, p. 2, 2020.

[9] E. Treatment and D. Retinopathy, "Grading Diabetic Retinopathy from Stereoscopic Color Fundus Photographs—An Extension of the Modified Airlie House Classification: ETDRS Report Number 10," *Ophthalmology*, vol. 98, no. 5, pp. 786–806, 1991, doi: 10.1016/S0161-6420(13)38012-9.

[10] R. Rajalakshmi, V. Prathiba, S. Arulmalar, and M. Usha, "Review of retinal cameras for global coverage of diabetic retinopathy screening," *Eye*, vol. 35, no. 1, pp. 162–172, 2021, doi: 10.1038/s41433-020-01262-7.

[11] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, no. 4, pp. 284–302, 1996, doi: 10.1006/cbmr.1996.0021.

[12] X. Zhang *et al.*, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Med. Image Anal.*, vol. 18, no. 7, pp. 1026–1043, 2014, doi: 10.1016/j.media.2014.05.004.

[13] L. Tang, M. Niemeijer, J. M. Reinhardt, M. K. Garvin, and M. D. Abramoff, "Splat feature classification with application to retinal hemorrhage detection in fundus images," *IEEE Trans. Med. Imaging*, vol. 32, no. 2, pp. 364–375, 2013, doi:

10.1109/TMI.2012.2227119.

[14] R. A. Welikala *et al.*, “Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification,” *Comput. Methods Programs Biomed.*, vol. 114, no. 3, pp. 247–261, 2014, doi: 10.1016/j.cmpb.2014.02.010.

[15] K. A. Goatman, A. D. Fleming, S. Philip, G. J. Williams, J. A. Olson, and P. F. Sharp, “Disc Using Retinal Photographs,” vol. 30, no. 4, pp. 972–979, 2011.

[16] A. Soni and A. Rai, “A Novel Approach for the Early Recognition of Diabetic Retinopathy using Machine Learning,” *2021 Int. Conf. Comput. Commun. Informatics, ICCCI 2021*, 2021, doi: 10.1109/ICCCI50826.2021.9402566.

[17] W. Zhang *et al.*, “Automated identification and grading system of diabetic retinopathy using deep neural networks,” *Knowledge-Based Syst.*, vol. 175, pp. 12–25, 2019, doi: 10.1016/j.knosys.2019.03.016.

[18] K. M. Adal, P. G. Van Etten, J. P. Martinez, K. W. Rouwen, K. A. Vermeer, and L. J. Van Vliet, “An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1382–1390, 2018, doi: 10.1109/TBME.2017.2752701.

[19] M. Niemeijer, M. D. Abràmoff, and B. Van Ginneken, “Information fusion for diabetic retinopathy CAD in digital color fundus photographs,” *IEEE Trans. Med. Imaging*, vol. 28, no. 5, pp. 775–785, 2009, doi: 10.1109/TMI.2008.2012029.

[20] R. Casanova, S. Saldana, E. Y. Chew, R. P. Danis, C. M. Greven, and W. T. Ambrosius, “Application of random forests methods to diabetic retinopathy classification analyses,” *PLoS One*, vol. 9, no. 6, pp. 1–8, 2014, doi: 10.1371/journal.pone.0098587.

[21] S. Sanromà, A. Moreno, A. Valls, P. Romero, S. De La Riva, and R. Sagarra, “Assessment of diabetic retinopathy risk with random forests,” *ESANN 2016 - 24th Eur. Symp. Artif. Neural Networks*, no. April, pp. 313–318, 2016.

[22] A. Herliana, T. Arifin, S. Susanti, and A. B. Hikmah, “Feature Selection of Diabetic Retinopathy Disease Using Particle Swarm Optimization and Neural Network,” *2018 6th Int. Conf. Cyber IT Serv. Manag. CITSM 2018*, no. Citsm, pp. 2016–2019, 2019, doi: 10.1109/CITSM.2018.8674295.

[23] T. Shanthi and R. S. Sabeenian, “Modified Alexnet architecture for classification of diabetic retinopathy images,” *Comput. Electr. Eng.*, vol. 76, pp. 56–64, 2019, doi: 10.1016/j.compeleceng.2019.03.004.