Smartphone Based Detection of Retinal Abnormalities

Final Year Project Report

Presented

by

Rohaid Ahmed Mirza
CIIT/FA19-BCE-006/ISB

Shakir Khan
CIIT/ FA19-BCE-008/ISB

Mohammad Zain Zubair Pirzada
CIIT/ FA19-BCE-027/ISB

In Partial Fulfillment

of the Requirement for the Degree of

Bachelors of Science in Computer Engineering

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

COMSATS UNIVERSITY ISLAMABAD

July 2023

Declaration

We hereby declare that this project neither as a whole nor as a part has been copied from any source. It is further declared that we have developed this project and the accompanying report entirely based on our personal efforts made under the sincere guidance of our supervisor. No portion of the work presented in this report has been submitted in support of any other degree or qualification of this or any other University or Institute of learning, if found we shall stand responsible.

Signature: _____________
Rohaid Ahmed Mirza

Signature: _____________
Shakir Khan

Signature: _____________
Mohammad Zain Zubair Pirzada

COMSATS UNIVERSITY ISLAMABAD
July 2023

Smartphone Based Detection of Retinal Abnormalities

An Undergraduate Final Year Project Report submitted to the Department of

ELECTRICAL AND COMPUTER ENGINEERING

As a Partial Fulfillment for the award of Degree

Bachelor of Science in Computer Engineering

By

Name	Registration Number
Rohaid Ahmed Mirza	CIIT/FA19-BCE-006/ISB
Shakir Khan	CIIT/FA19-BCE-008/ISB
Mohammad Zain Zubair Pirzada	CIIT/FA19-BCE-027/ISB

Supervised by

Dr. Syed Saud Naqvi
Associate Professor,
Department of Electrical and Computer Engineering
CU Islamabad

Co-Supervised by

Dr. Haroon Ahmed
Assistant Professor,
Department of Electrical and Computer Engineering
CU Islamabad

COMSATS UNIVERSITY ISLAMABAD

July 2023

Final Approval

This Project Titled
Smartphone Based Detection of Retinal Abnormalities

Submitted for the Degree of
Bachelor of Science in Computer Engineering

by

Name	**Registration Number**
Rohaid Ahmed	CIIT/FA19-BCE-006/ISB
Shakir Khan	CIIT/FA19-BCE-008/ISB
Mohammad Zain Zubair Pirzada	CIIT/FA19-BCE-027/ISB

has been approved for
COMSATS UNIVERSITY ISLAMABAD

Supervisor
Dr. Syed Saud Naqvi
Associate Professor

Co-Supervisor
Dr. Haroon Ahmed
Assistant Professor

Internal Examiner-1
Dr. M. Awais Javed
Associate Professor

Internal Examiner-2
Dr. Bakhtiar Ali
Associate Professor

External Examiner
Dr. Shabbir Majeed
Professor
EED, UET Taxila

Head
Department of Electrical and Computer Engineering

Dedication

This work is dedicated to all the individuals who have supported and encouraged us throughout our academic journey. To our families, who have always been our source of inspiration.

Your unwavering love and support have been our guiding light, and we could not have achieved this milestone without you.

To our friends and colleagues, who have shared their knowledge, insights, and experiences with us, and have challenged us to think critically and creatively. Your friendship has made this journey more rewarding.

And finally, to the faculty and staff at **“Comsats University Islamabad”**, who have provided us with the tools, resources, and guidance to excel in our field of study. Your dedication to teaching and mentorship has had a profound impact on our personal and professional growth, and we are grateful for the opportunities you have provided us.

This work is a testament to your collective efforts, and we hope to honor your legacy by making a positive impact in our field of study.

Acknowledgements

We would like to express our sincere gratitude to everyone who has contributed to the completion of this project work.

First and foremost, we would like to thank our supervisor **“Dr. Saud Naqvi”** and co-supervisor **“Dr Haroon Ahmed”**, for providing us with valuable guidance, support, and encouragement throughout the research process. Your expertise and insight have been invaluable, and we are grateful for the time and effort you have dedicated to my project.

We would also like to thank the faculty and staff at **“Comsats University Islamabad”**, who have provided us with the resources, facilities, and academic support necessary to complete this work. The opportunities and challenges that we have encountered during our studies have helped us grow both academically and personally.

Lastly, we would like to acknowledge the influence of our role models, whose dedication and commitment to their work have inspired us to pursue our academic goals. We hope to continue to learn from your example and make a positive impact in our field of study.

Thank you all for your invaluable contributions.

Rohaid Ahmed Mirza
Shakir Khan
Mohammad Zain Zubair Pirzada

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

Diabetic Retinopathy (DR)
Field of View (FOV)
Artificial Intelligence (AI)
International Diabetes Federation (IDF)
Category Attention Block (CAB)
Global Attention Block (GAB)
Area under the curve (AUC)
Generalized estimating equation (GEE)
User Interface (UI)
Contrast Limited Histogram Equalization (CLAHE)
Naturalness Preserved Enhancement Algorithm (NPEA)
Convolutional Neural Network (CNNs)
Graphical Processing Unit (GPU)
Stochastic Gradient Descent (SGD)
Adaptive Moment Estimation (Adam)
Root Mean Square Propagation (RMSProp)
Mean Squared Error (MSE)

Abstract

Diabetic Retinopathy (DR), also known as diabetic retinopathy, represents a retinal ailment, which arises as a common complication resulting from diabetes mellitus. The impairment of retinal blood vessels can lead to vision loss, as fluid flows from vessels to the retina. In the early stages, DR often lacks noticeable symptoms but can progress to severe consequences, including complete blindness, if untreated. Detecting DR promptly through regular eye exams plays a vital role in mitigating the severity of this condition. However, the utilization of fundus cameras for diagnosis remains limited due to their hefty cost, large size, and weight. Recent technological advancements have yielded a compact, low-power retinal imaging system that integrates with smartphones, providing a more affordable option for DR screening. Nonetheless, the image quality adversely impacts the accuracy of DR detection. Hence, this study investigates the suitability of smartphone-based portable systems for identifying DR during general health screenings. Drawing upon the findings, the study introduces automated DR detection algorithms utilizing deep learning frameworks specifically designed for retinal images captured on smartphones. Recognizing the demand for an independent diagnostic tool accessible to non-experts, this project proposes and implements diverse classifiers for the detection and classification of DR. Moreover, the proposed method is integrated into a mobile-based classification smartphone application, furnishing an online automated system for diagnosing DR. This research contributes to the advancement of smartphone-based DR screening and detection, thereby facilitating enhanced healthcare access and reducing the global burden of DR.

Chapter 1:
1. Introduction

With the increase in population comes an increase in the use of technology and its number. The number of smartphones in use is increasing day by day with the population. Smartphones are easily accessible. They have reached almost every household. With its increased use, every other person tends to have everything done on hand by the smartphone.

Smartphones have high processing speed, that is why they are now attracted by every other industry for its products' control and use. The field of medicine has also been tilted towards smartphones for their services. From the doctors' appointments to consultancy everything is just a touch away.

Artificial Intelligence (AI) is the study of algorithms that provide machines, human thinking capabilities for the tasks that normally require human brains. Its use has significantly increased in the last couple of years. Every field has adopted AI in daily routine tasks or production. AI is also in the eye of the medical field for detection and diagnosis of any abnormality present.

Eyes are very crucial and sensitive parts of the body. They are the organs that provide vision. They develop diseases like every organ. If eye diseases are not detected timely, the risk of permanent damage significantly increases, and vision can be lost.

1.1 Eye

Eyes are one of the most sensitive organs of the human body. They provide the sight to the living organisms. They receive, process, and send visual information to the brain. Eye is present in protective bone socket called the orbit. The main part of eye that is truly sensitive for sight is retina. The structure of retina is explained below.

1.1.1 The Structure of Retina

In the eye, a specialized and sensitive tissue is present, its job is to detect light and is known as retina. Found in the inner surface of eye, it is responsible for sight. Three layers compose the retina, each with its specialized function. The outermost layer is the photoreceptor layer. It has two types of cells: cones and rods.

Rods are responsible for sensing light in low-light environments and are more sensitive to motion, while cones are responsible for color vision and work best in bright light conditions.

The bipolar cell layer receives information from the photoreceptor layer and transmits it to the ganglion cell layer. The ganglion cell layer is the third and final layer of the retina, consisting of specialized cells that relay signals from the retina to the brain through the optic nerve.

The red circular structure illustrated in figure 1.1 is referred to as the macula. It is responsible for focusing light entering the eye through cornea and lens, giving a detailed vision. The macula is accountable for central vision while the rest of the retina provides peripheral vision. It is essential to maintain a dry macula for unclouded vision. The fovea, situated in the center of the macula, is the most significant part of the retina for visual acuity.

The optic disk visible as a white area in the figure 1.1, also known as the optic nerve head, is another crucial structure within the retina. It is where the optic nerve leaves the eye, facilitating the transmission of visual data to the brain in electrical impulses. However, the optic disk does not contain rods or cones, and thus represents a blind spot in the visual field. The anatomical structure of the retina is fundamental to its function, and damage or disruption to any of the layers can lead to vision loss or other visual impairments.

Comprehending the anatomy of the retina is essential for developing diagnostic and treatment strategies for retinal diseases, including diabetic retinopathy, macular degeneration and retinal detachment.

To carry out efficient research, vision experts and eye doctors need to possess a thorough grasp of the retina's structure. Acquiring profound comprehension in this vital area is crucial for their work, guaranteeing effective execution of their studies. Hence, it becomes imperative for field experts to dedicate time and energy to attaining an in-depth understanding of this significant element of the eye. This understanding is pivotal, as it serves as the basis for comprehending the way in which visual information is processed by the eye and the potential effects of retinal diseases on one's vision.

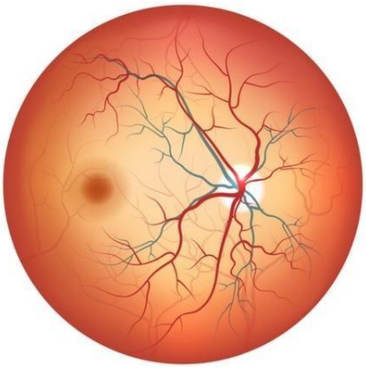


Figure 1.1 Retinal Image [3]

1.2 Diabetic Retinopathy

Diabetic Retinopathy (DR) is a significant disease and cannot be ignored in any case. People develop diabetes with increasing age. The case when diabetes becomes so intense it starts affecting other organs like kidneys, heart, and eyes. When diabetes starts affecting the eye, it is referred to as Diabetic Retinopathy or DR. It damages the blood vessels of light sensitive tissue at the back of the eye also known as retina. If DR is not detected on time, it can result in eye loss.

Diabetes, a chronic metabolic condition, displays elevated blood glucose levels due to diminished insulin function. Insulin, a hormone produced by the pancreas, regulates glucose levels. The disease encompasses two primary types: Type 1, or insulin-dependent diabetes, caused by the immune system attacking insulin-producing cells, resulting in reduced insulin production; and Type 2, were insulin resistance or insufficient production hampers glucose regulation.

This form of diabetes is often linked to lifestyle factors like obesity, physical inactivity, and an unhealthy diet. Diabetes can lead to acute and chronic complications. Acute complications manifest as low or high blood sugar levels, causing symptoms like sweating and altered consciousness. Chronic complications affect various organs, including the heart, kidneys, and eyes.

Diabetic retinopathy, a common diabetes-related complication, damages the retina's blood vessels responsible for light sensing and reaction. Elevated blood glucose levels harm these vessels, leading to leakage or blockage and potential vision loss or blindness if untreated. Glaucoma and cataracts are also eye problems associated with diabetes.

According to recent data from the International Diabetes Federation (IDF), Pakistan has the highest prevalence of diabetes worldwide. One out of every four adults in the country is affected by the condition. The IDF report reveals a 70% increase in diagnosed cases since 2019, with a current total of 33 million individuals affected. In terms of individuals affected, China and India top the charts with 140 million and 74 million people respectively. Pakistan, in the year 2021, saw diabetes claim the lives of 400,000 individuals.

This has caused the country to rank third globally in terms of the number of people affected by this condition. As per the International Diabetes Federation's (IDF) report, the number of individuals diagnosed with diabetes across the globe has risen from 108 million in 1980 to 463 million in 2019[1].

The IDF projects that this number will continue to rise, reaching 578 million by 2030 and 700 million by 2045, with an increase of 51% [1].

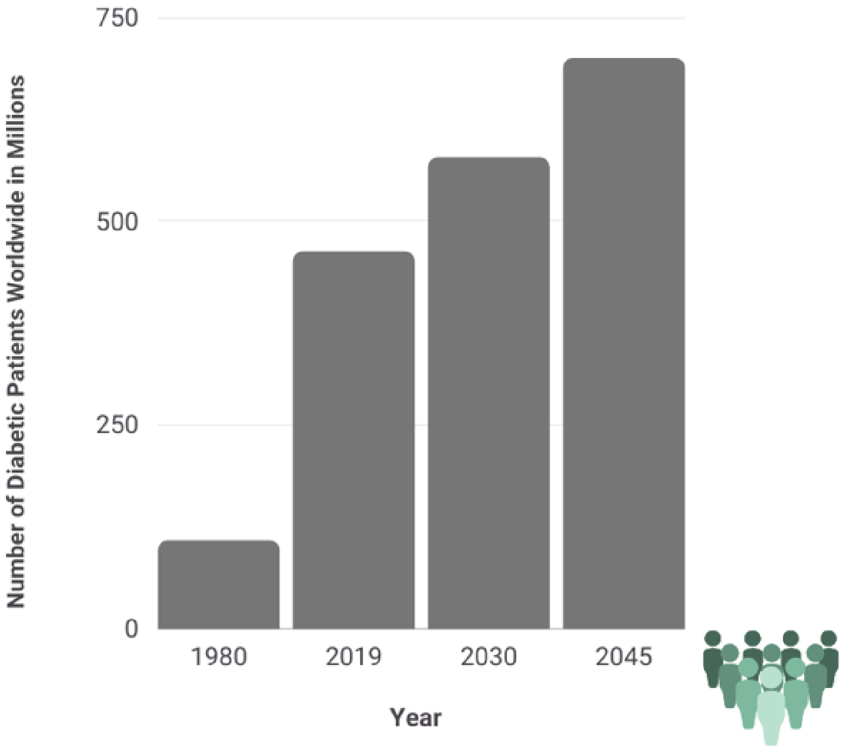


Figure 1.2: Number of Diabetic Patients Worldwide [1]

Diabetic Retinopathy (DR) is a retinal disease and one of the most common complications caused by Diabetes Mellitus. Vision loss can occur due to the damage of blood vessels in the retina, resulting in fluid flow from the blood vessels to the retina. The disease is often asymptomatic in the initial stages and can lead to complications, such as complete blindness if left untreated.

It is estimated that about one-third of people with diabetes have some degree of retinopathy.

The study conducted a systematic review based on IDF Atlas 2019 data up to March 2020. In 2020, it was approximated that approximately 103 million adults globally and regionally suffered from diabetic retinopathy. It is predicted that this figure will rise to about 160 million by the year 2045 [2].

The signs of diabetic retinopathy may differ based on how serious the condition is. During the initial phases, there might not be any evident indications and vision may remain normal. As the condition progresses, however, symptoms may include blurred or distorted vision, floaters, difficulty seeing at night, and eventually, vision loss.

1.2.1 Types of Diabetic Retinopathy

Diabetic retinopathy is conventionally diagnosed by ophthalmologists, who examine retinal images to identify signs of the disease. The process of diagnosis includes the observation of the retina for different irregularities like unusual blood vessels, inflammation, fat build-up, detachment of the retina, and abnormalities in the optic nerve. During the diagnosis of diabetic retinopathy, ophthalmologists also conduct vision tests for the blood vessels growth and scar tissues.

In a proposal by [3], a disease severity scale for diabetic retinopathy was suggested, which consists of five stages. The first stage is "No DR," as per name there is no damage to the retina caused by diabetes mellitus. The second phase of DR, named "Mild DR," is distinguished by only a small number of microaneurysms. As for the third stage, "Moderate DR," it can be detected through the occurrence of cotton wool spots, and several microaneurysms. The fourth stage is known as "Severe DR," where one can observe microvascular abnormalities and cotton wool spots. Finally, the last stage is called "Proliferative DR," and it is identified by the emergence of new retinal blood vessels, retinal detachment, and vitreous hemorrhages.

In 2020, a retrospective study was conducted during a medical camp in Muzaffargarh, Multan, Punjab, where 1150 individuals were screened. Among those screened, 522 (45%) were diagnosed with diabetes, with a higher prevalence observed in males (59.77%) compared to females (40.22%). Of the 522 individuals diagnosed with diabetes, 54.21% were tested for diabetes for the first time. Further investigation revealed that 124 patients were already suffering from diabetic retinopathy, while 398 individuals were newly diagnosed with diabetic retinopathy [3].

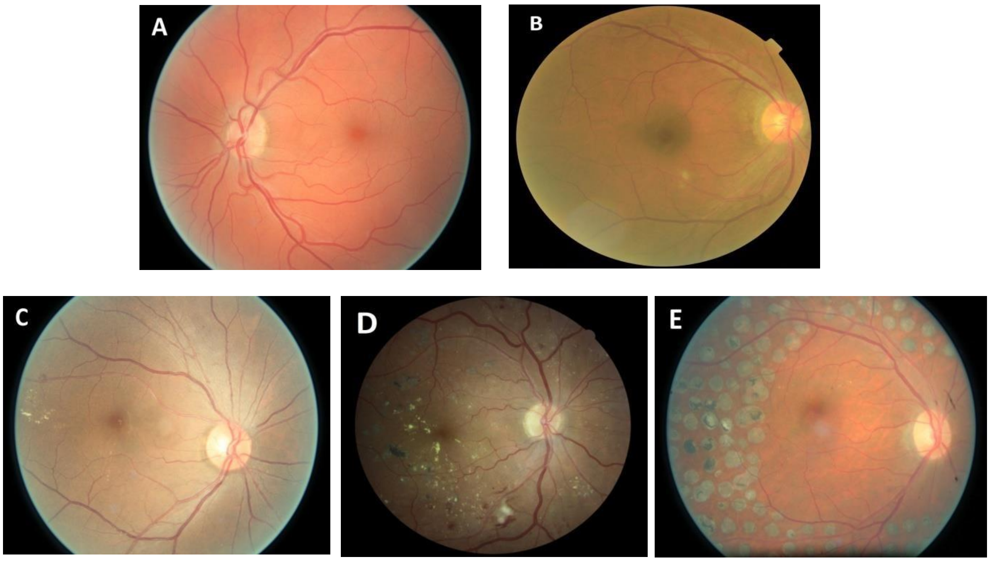


Figure 1.3: Referable vs Non-Referable; (A) No DR/Non-Referable; (B)(C)(D)(E) DR/Referable;

The approach to treat diabetic retinopathy is determined by the severity of the condition. During the initial stages, monitoring blood sugar and pressure along with regular eye check-ups may suffice. However, for more severe cases, interventions such as laser therapy or surgical procedures may be required to halt the development of abnormal blood vessels or damage to the retina.

1.2.2 Diabetic Retinopathy Detection

The diagnosis of diabetic retinopathy entails the installation of expensive fundus imaging devices and the engagement of trained professionals in healthcare centers. Various imaging techniques, including optical coherent tomography and colorful fundus photography are employed in the diagnosis of diabetic retinopathy [4]. Nevertheless, the development, deployment, and use of these techniques are accompanied by high costs and require trained professionals for their employment and utilization. Additionally, an ophthalmologist is needed to evaluate and diagnose the captured fundus image. Individuals with diabetes residing in homes, remote areas face difficulties in accessing early diagnosis and treatment because the equipment used for diagnosis is expensive and there is a shortage of ophthalmologists and medical facilities, it is difficult to provide sufficient access to eye care.

1.3 Automatic Diabetic Retinopathy Detection:

The dissertation presents a comprehensive review of a proposed method and approach for addressing the challenges in diagnosing and treating diabetic retinopathy (DR) using smartphones. The research explores the use of deep learning algorithms to develop an automatic DR detection system for retinal images captured on smartphones. The integration of this system into a mobile-based classification application allows individuals with diabetes to conveniently screen and detect DR without the need for expensive equipment or expert assistance. The study contributes to improving healthcare access by providing a cost-effective and accessible method for identifying and categorizing DR, especially in remote areas. The research also highlights the importance of performance metrics for evaluating the model's effectiveness, such as precision, recall, F1-score, and accuracy. Although the proposed solution has limitations, such as image quality constraints and reliance on smartphone-based imaging, future work suggests enhancing the deep learning models with larger datasets and advanced architectures and incorporating telemedicine capabilities for remote consultations. Overall, the research demonstrates the potential of smartphone-based systems in improving early detection and intervention for diabetic retinopathy.

1.4 United Nations Sustainable Development Goals (SDGs) and "Smartphone-Based Detection of Retinal Abnormalities"

1.4.1 Goal 3: Good Health and Well-being

The United Nations Sustainable Development Goals (SDGs) are a set of 17 global goals aimed at addressing various social, economic, and environmental challenges by the year 2030. Goal number 3, "Good Health and Well-being," specifically focuses on ensuring healthy lives and promoting well-being for all at all ages. This goal aligns perfectly with the project titled "Smartphone-Based Detection of Retinal Abnormalities," as it seeks to enhance healthcare access and improve the early detection of diabetic retinopathy (DR), a retinal ailment that can lead to severe consequences if untreated.

1.4.1.1 Enhancing Healthcare Access:

The project aims to address the limited access to retinal screenings due to the high cost and limited availability of fundus cameras. By utilizing smartphones as a portable retinal imaging system, the project provides a more affordable and accessible option for DR screening during general health screenings. This initiative directly contributes to the goal of ensuring healthy lives by enabling individuals, especially those in resource-constrained areas, to receive timely retinal examinations.

1.4.1.2 Early Detection of Retinal Abnormalities:

Diabetic retinopathy (DR) is a leading cause of vision loss and blindness, particularly among individuals with diabetes mellitus. The project's focus on the early detection of DR through regular eye exams aligns with the target of reducing non-communicable diseases under SDG 3. Timely detection allows for appropriate medical intervention, preventing the progression of DR to severe consequences, such as complete blindness.

1.4.1.3 Integration of Deep Learning Algorithms:

The project introduces automated DR detection algorithms that utilize deep learning frameworks specifically designed for retinal images captured on smartphones. By leveraging advancements in artificial intelligence, the project contributes to strengthening the capacity for early warning and risk reduction in health. The accurate detection and classification of DR enable healthcare professionals to identify individuals who require immediate medical attention, facilitating early treatment and management.

1.4.1.4 Online and Automated Diagnosis:

The proposed method is integrated into a mobile-based classification smartphone application, providing an online and automated system for diagnosing DR. This innovation aligns with the target of achieving universal health coverage, as it empowers non-experts to utilize a diagnostic tool that can detect retinal abnormalities. By enabling online diagnosis, the project addresses potential limitations in areas with limited internet connectivity, ensuring healthcare access for all.

1.4.1.5 Conclusion

The project "Smartphone-Based Detection of Retinal Abnormalities" effectively aligns with United Nations SDG 3, "Good Health and Well-being." By providing a more affordable and accessible option for DR screening, utilizing deep learning algorithms for accurate detection, and offering an automated diagnosis system, the project contributes to ensuring healthy lives and promoting well-being for all at all ages. This research initiative aims to reduce the global burden of DR by enhancing healthcare access and facilitating early detection, ultimately mitigating the severity of the condition and improving overall health outcomes.

1.4.2 Goal 9: Build Resilient Infrastructure, Promote Inclusive and Sustainable Industrialization, and Foster Innovation

The United Nations Sustainable Development Goals (SDGs) are a set of 17 global goals aimed at addressing various social, economic, and environmental challenges by the year 2030. Goal number 9, "Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation," focuses on promoting sustainable and inclusive industrial development, enhancing infrastructure, and fostering innovation. This goal closely aligns with the project titled "Smartphone-Based Detection of Retinal Abnormalities," as it leverages technological advancements and innovation to create a resilient and inclusive healthcare solution.

1.4.2.1 Technological Innovation:

The project incorporates recent technological advancements to develop a compact, low-power retinal imaging system that integrates with smartphones. By leveraging the capabilities of smartphones, the project fosters innovation in the field of retinal imaging, providing a more accessible and affordable option for the detection of retinal abnormalities. This initiative directly contributes to the target of fostering innovation by introducing a novel approach to retinal screening and diagnosis.

1.4.2.2 Resilient and Inclusive Healthcare Infrastructure:

Traditional retinal screening methods, such as the use of fundus cameras, face limitations due to their hefty cost, large size, and weight. By utilizing smartphones as a portable retinal imaging system, the project contributes to building resilient healthcare infrastructure that can be easily deployed in various settings, including resource-constrained areas. The accessibility and affordability of the smartphone-based system enables inclusive healthcare services, ensuring that individuals, regardless of their location or economic status, can receive retinal screenings.

1.4.2.3 Sustainable Industrial Development:

The project's focus on smartphone-based retinal imaging systems promotes sustainable industrial development in the healthcare sector. By providing an alternative to expensive and cumbersome equipment, the project encourages the development and production of more affordable and sustainable healthcare technologies. This aligns with the target of promoting sustainable industrialization by fostering the growth of industries that prioritize sustainability and accessibility.

1.4.2.4 ICT Access and Connectivity:

The integration of deep learning algorithms into the smartphone-based retinal imaging system requires access to information and communication technology (ICT). The project highlights the importance of increasing ICT access, particularly in the healthcare sector, to facilitate advanced diagnostic capabilities and improve healthcare outcomes. By utilizing smartphones and mobile-based applications, the project promotes the use of ICT for healthcare delivery, making diagnostic tools more accessible and widely available.

1.4.2.5 Conclusion

The project "Smartphone-Based Detection of Retinal Abnormalities" aligns with United Nations SDG 9, "Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation." By leveraging technological innovation, creating resilient healthcare infrastructure, promoting sustainable industrial development, and utilizing ICT for advanced diagnostics, the project contributes to achieving the goals of SDG 9. This research initiative not only improves the detection of retinal abnormalities but also fosters inclusive and accessible healthcare services, paving the way for sustainable and resilient healthcare systems worldwide.

Chapter 2:
2. Literature Review

This chapter provides an overview of smartphone-based fundus imaging faces challenges such as poor image quality and variations in lighting conditions, the literatures provide solutions to overcome these challenges. The second literature review proposes a deep learning model called CABNet for accurately grading retinal abnormalities, which addresses the challenges of imbalanced data distribution and small lesion identification through specific attention modules.

2.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in image classification and computer vision tasks, making them a popular choice for deep neural network applications. CNNs were initially introduced in the 1980s; however, they gained popularity in the 2010s with the advent of large datasets and powerful graphical processing units (GPUs) [50].

CNNs primary function is to replicate the functioning of the human brain's visual cortex, which has distinct layers of neurons that process visual information. CNNs, similarly, use layers to process image data through a series of mathematical operations such as convolution, pooling, and activation.

The convolution operation entails sliding a filter (also known as a kernel or feature detector) over the given image. During this process, multiplication element by element is carried out and the resulting products are summed up to produce a feature map. This process enables CNN to identify local patterns and features in the input image, such as edges, corners, and textures.

LeCun et al [50] were the pioneers who introduced the method of image recognition using convolution, where a small kernel size is applied, in their work. This approach was employed in the recognition of handwritten digits from input images.

The pooling operation is then utilized to down sample the feature maps by selecting the maximum value or the average value in a defined local region. This process is effective in reducing the size of the feature maps while still preserving the crucial features. The activation function is then applied to the output of the convolution and pooling operations to introduce nonlinearity into the model. This allows CNN to learn complex relationships between features and generate more accurate predictions.

After being processed through number of convolution and multiple pooling layers, an input image is classified into distinct categories using a fully connected layer. To enable the data to be transmitted from the convolution layer to the fully connected layer, a flattening layer is employed. The quantity of layers or nodes allocated to each layer is dependent on the prediction or classification task as well as the size of the dataset being used in a fully connected layer.

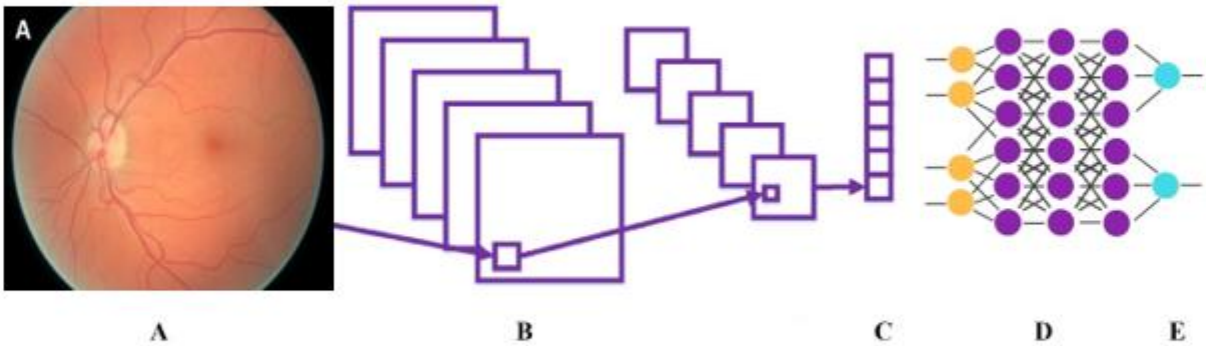


Figure 2.1: A typical convolutional neural network: (A) Input layer; (B) Convolution and pooling layers; (C) Flattening layer; (D) Fully connected layer; (E)Output layer.

2.1.1 Transfer Learning

The End-to-End learning technique involves training Convolutional Neural Networks (CNNs) from scratch, while Transfer learning refers to using pre-trained neural network models to classify new data that has not been seen before. In medical image classification, where there is a shortage of annotated training data, Transfer learning can be a useful tool [51, 52]. This technique has gained widespread acceptance due to the scarcity of labeled training data [53, 54]. Altaf et al. [55] recommended how to utilize transfer learning as a means of leveraging the capabilities of deep learning.

This involves bringing in an existing deep learning model, using it to extract features from the new dataset, and then creating a classifier based on those features. Only the weights of the new classifier are updated during training, while the pre-trained model's weights remain fixed. The resulting outputs are then calculated and stored as features. Transfer learning allows us to reuse the knowledge learned from one task and apply it to a new task, saving time and resources.

2.1.2 Transfer Learning Based Model

Transfer learning refers to using pre-trained neural network models to classify new data that has not been seen before. This technique leverages the knowledge learned from one task and applies it to a new task, saving time and resources. In the context of medical image classification, where annotated training data is often limited, transfer learning has proven to be a valuable tool. Transfer learning offers several advantages in medical image classification:

2.1.2.1 Utilizing Pre-trained Models:

By bringing in pre-trained deep learning models, we can leverage their learned representations of general image features. Models like Densenet-121, MobileNet1.0, MobileNetv2 etc. are trained on the ImageNet dataset, possess a strong foundation in recognizing visual patterns. These models capture a wide range of features applicable to various image classification tasks, including DR detection.

2.1.2.2 Reducing the Need for Annotated Data:

The scarcity of labeled training data in medical image classification poses a challenge. Transfer learning helps overcome this challenge by allowing us to extract relevant features from the pre-trained model and use them as inputs to a new classifier. Only the weights of the new classifier are updated during training, while the pre-trained model's weights remain fixed. This approach significantly reduces the amount of labeled data required for training, making it feasible to achieve accurate classification even with limited annotated data.

2.1.2.3 Enhancing Performance and Convergence:

The knowledge captured by pre-trained models helps initialize the network with meaningful weights. This initialization often leads to faster convergence during training, as the model already possesses a good understanding of general image features. By starting from a better initialization point, the model can focus on learning task-specific features relevant to DR classification, further improving performance.

In medical image classification, where there is a shortage of annotated training data, Transfer learning can be a useful tool [51, 52]. This technique has gained widespread acceptance due to the scarcity of labeled training data [53, 54]. Altaf et al. [55] recommended the use of Transfer learning to harness the power of deep learning.

2.1.3 Optimization Algorithms

The optimization functions used in deep neural network training are a crucial component, as they aim to minimize the difference between the predicted and actual targets. The first is known as Stochastic Gradient Descent (SGD), while the second is called Adaptive Moment Estimation (Adam), and the third is referred to as Root Mean Square Propagation (RMSProp) are among the optimization functions available, each with its own set of strengths and weaknesses.

The selection of an appropriate optimization function for a particular neural network and task is critical to achieving optimal results. When choosing an optimization function, one must take into account factors like how fast it converges, its stability, and its ability to avoid getting stuck in local minima.

Hence, it becomes crucial to conduct experiments with different optimization functions to identify the most suitable one for a given task. This might involve adjusting the hyperparameters of the optimization function, such as the learning rate and momentum, to enhance its performance. Ultimately, the selection of the right optimization function can have a significant impact on the performance of the neural network and should be done thoughtfully in any deep learning project.

2.1.4 Loss Function

In the realm of deep learning models, an integral aspect is the employment of a loss function to gauge the deviation between predicted and actual outputs. This variance is also identified as the error or cost function. The primary aim of a loss function is to reduce this discrepancy to heighten the model's accuracy. During the process of iterative training, loss functions are indispensable in assessing the efficiency of the neural network's weights and biases.

Several loss functions are employed in deep learning, namely, Cross-Entropy, Mean Squared Error (MSE), etc. The choice of an apt loss function is essential in obtaining optimal results, as distinct loss functions are better suited to different data types and models. It is crucial to consider the task type and output of the model when selecting a loss function. For example, regression tasks require different loss functions than classification tasks. Additionally, it is imperative to conduct experiments with a diverse range of loss functions and hyperparameters to determine the most suitable one for the specific model and task.

For multi-class issues, a widely used loss function is cross-entropy, which has proven to yield enhanced results in various experiments. The selection of an appropriate loss function can significantly influence the performance of a deep learning model. Therefore, it should be thoroughly considered in any deep learning project.

2.2 Detection of Retinal Abnormalities using Smartphone Captured Fundus Images

This proposed method gives a big picture of the diverse ways and approaches employed in ophthalmology to find retinal abnormalities and how well they work [5]. Spotting these issues early on is key in stopping blindness and using affordable smartphones for retinal imaging shows promise. Nevertheless, there are challenges like subpar image quality [6] and lighting differences that come with using smartphones to capture fundus images.

Spotting retinal abnormalities in the early stages is crucial to diagnose diseases such as diabetic retinopathy, age-related macular degeneration, and glaucoma. The usual screening techniques for these problems require specialized equipment and trained staff, which limits their availability. By using smartphones for fundus imaging, this problem could be addressed, making healthcare more accessible in underserved areas.

This assessment examines the current research status regarding the use of smartphone-captured fundus images to detect eye diseases. It covers aspects like image quality, field of view, and tele-ophthalmology. The potential for future research to enhance the accuracy of smartphone-based fundus cameras and tele-ophthalmology systems is also highlighted.

2.2.1 Quality and Field of View (FOV) of Smartphone Captures Retinal Images

Several studies have evaluated the fundus images quality taken using smartphones, and the results show that the quality is comparable to that of direct ophthalmoscope images. Russo et al [7] compared images taken with a 3.2-megapixel camera smartphone to those taken with a Topcon TRC-50DX camera. They found that the image quality of smartphone-captured images was comparable to that of Topcon camera images, with a sensitivity of 95% and specificity of 95%. Thus, smartphones can be considered a reliable tool for fundus imaging.

Different smartphone lenses offer varying FOVs, and the processing of the images should ensure that the retinal structures are accurately modeled for automatic ocular pathology detection methods to work effectively. Some studies propose the capture of multi-field fundus images, where several images are merged to provide a wider FOV. The studies reported that the images captured by the system showed clear optic disc and macula morphology and could be used for screening diabetic retinopathy. Following are the images captured with different FOV.

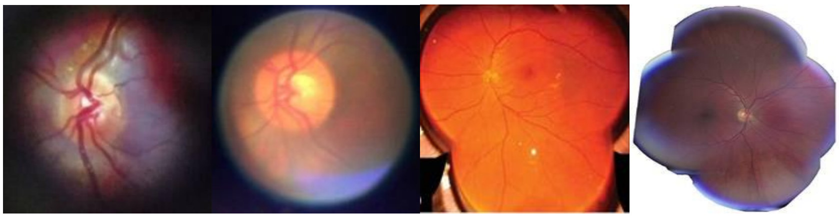


Figure 2.2: Smartphone-captured images with different FOV: (a) captured with D-EYE, (b) captured Welch Allyn Panoptic Ophthalmo-scope, (c) captured with ophthalmoscopy lens–based Smartphone fundus cameras [6], (d) captured with CellScope [7].

Numerous methods for preprocessing have been suggested to enhance the accuracy of detection. These techniques include using convolution filters, CLAHE, and Elloumi et al. [8] circular Hough transform method to adjust contrast. These techniques can help improve the accuracy of detection, making smartphones a more reliable tool for ocular disease detection.

2.2.2 Tele-Ophthalmology

Tele-ophthalmology has the potential to completely transform the way eye care is provided and expand its reach to underserved communities. Research conducted by Lam et al. delved into the exploration of using smartphone cameras to detect diabetic retinopathy through wide-field retinal images [9]. The researchers compared smartphone and single-field fundus cameras, finding that the smartphone camera yielded similar quality wide-field retinal images. This breakthrough paves the way for revolutionizing tele-ophthalmology and increasing access to eye care for remote populations. However, further research is needed to optimize the effectiveness of smartphone-captured fundus images in this emerging field.

2.2.3 Use of Smartphone Captured Retinal Images for Ocular Disease Detection

Smartphone images have proven handy for spotting various eye conditions like diabetic retinopathy (DR). Studies indicate their reliability as a screening tool. In a rural India study by Rajalakshmi et al. [10], they evaluated a smartphone-based retinal camera's accuracy in detecting DR. The findings showed a 92.7% sensitivity and 98.4% specificity in identifying diabetic retinopathy. These results suggest smartphones hold promise as an early detection and prevention tool for eye diseases.

Table 2.1: Characteristics and Performance of Ocular Diseases Detection from Smartphone Captured Retinal Images

Works	Pathologies	Number of images	Performance Evaluation	Capture Device
Toy et al. [11]	moderate non proliferative &Worse diabetic retinopathy	100	91% sensitivity 99% specificity 95% positive predictive value98% negative	Paxos Scope
Russo et al. [12]	Vertical Cup-to-DiscRatio (glaucoma)	107 eyes	95% sensitivity 95% specificity 63 % Kappa (k)	D-EYE
Ryan et al. [13]	DR	--	With 7-field mydriatic: Sensitivity 95% Specificity 94%Non-mydriatic: Sensitivity 95% Specificity 95%	--
Rajalakshmi et al. [14]	DR	--	Sensitivity 92.7 % Specificity 98.4 % Kappa (κ) = 0.90	(FOP) camera
Micheletti et al. [15]	DR	240 eyes	- -	D-Eye
Muiesan et al. [16]	hemorrages exudates papiledema	52 patients (104fundus images)	Hemorrages: 0.795 Exudates: 0.822 – 0.878 Papiledema: 0.815 – 0.804	D-Eye
thomas et al. [17]	glaucoma		specificity 79% sensibility 83%	
Jin et al. [18]	DR AMD	--	P>0.05	MiisDE C200

To make these methods more accessible, we must implement machine learning techniques on portable devices and optimize them. Deep learning methods have shown promising results in identifying diabetic retinopathy through smartphone-captured images.

Furthermore, ensuring patient privacy and security becomes crucial when employing smartphone-captured fundus images for tele-ophthalmology [19]. Patients' health data, including their fundus images, must be protected from unauthorized access, use, and disclosure. The use of secure transmission protocols and the implementation of encryption techniques can help ensure the privacy and security of patients' health data.

When using images captured by smartphones in tele-ophthalmology, it is crucial to also consider reducing the time it takes to process the images. This is because processing many images can be time-consuming and may cause delays in diagnosing and treating ocular diseases. However, incorporating efficient image processing techniques such as parallel processing [20] and cloud computing [21, 22, 23, 24] can significantly reduce the execution time and increase the efficiency of tele-ophthalmology systems.

2.3 Category Attention Block for Imbalanced Diabetic Retinopathy Grading

Diabetic Retinopathy (DR) is a retinal disease and one of the most common complications caused by Diabetes Mellitus. Vision loss can occur due to the damage of blood vessels in the retina. Based on the findings, the work proposes automatic DR detection methods using deep learning frameworks for smartphone-based retinal images. Due to the uneven distribution of labels in the dataset, accurate grading is challenging. The proposed method involves a deep learning model called CAB (Category Attention Block) for imbalanced DR grading. CABNet addresses the challenges of imbalanced data distribution and small lesion identification in DR grading by using two types of attention modules: Category Attention Block (CAB) for imbalanced DR data distributions and Global Attention Block (GAB) for capturing small lesion information [6].

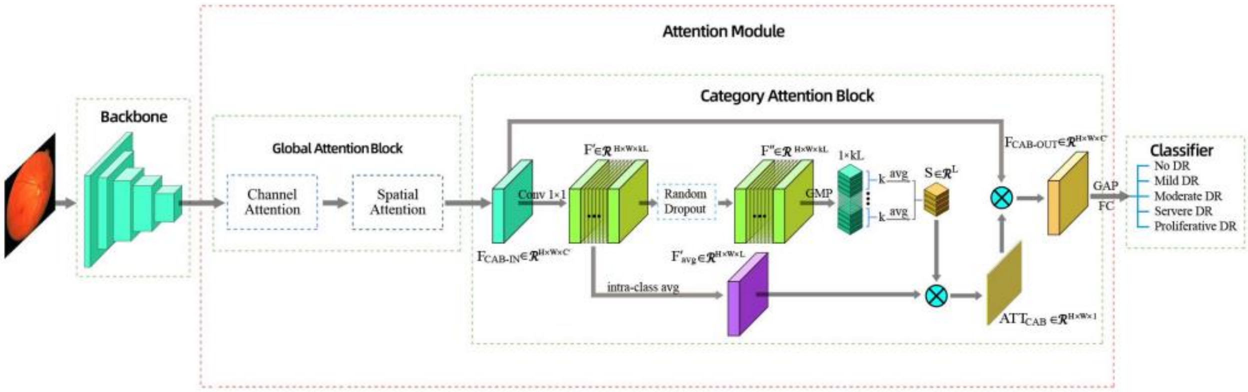


Figure 2.3: The overall structure of CABNET [6]

2.3.1 Methodology:

The proposed method involves a deep learning model called CABNet to evaluate the severity of diabetic retinopathy in cases where the dataset is imbalanced. The model architecture consists of a backbone network, followed by a category attention block, a global average pooling layer, and a fully connected layer.

CABNet relies on two crucial techniques: CAB and GAB. CAB extracts region-based features for each DR grade, while GAB utilizes class-specific attention maps to gather detailed information about small lesions in fundus images. By evaluating the spatial similarity between different regions, GAB calculates attention weights. The category attention block specifically targets categories with limited samples, highlighting the significance of training on imbalanced classes.

To conduct their experiments, the authors employed the Kaggle Diabetic Retinopathy Detection (KDRD) dataset, comprising 35,126 fundus images. The dataset exhibits an imbalance, with only 6% of the images belonging to the severe DR class. For their experiments, the proposed method employed MobileNetV2 and other network architectures as the backbone.

2.3.2 Results and Discussion

The CABNet model, suggested by the authors, was tested on three different datasets: DDR, Kaggle-2015, and Messidor. Additionally, a comparison was made with other advanced models like MobileNet and DenseNet-121. The authors utilized a range of evaluation metrics, such as Precision, Recall, F1-score, Area under the curve (AUC), quadratic weighted Kappa, and accuracy metrics, for both binary and multi-class DR grading.

The results of the experiment demonstrated that CABNet surpassed several cutting-edge models in terms of accuracy, F1-score, and AUC. Moreover, the authors conducted ablation studies to analyze the performance of the CABNet model. The results indicated that the suggested category attention block had a significant effect on the CABNet model's performance. The CABNet model's accuracy dropped substantially when the category attention block was not included.

The CABNet model proposed in the research shows promise for fine-grained grading of diabetic retinopathy. The experimental outcomes reveal that the model effectively tackles the challenges of imbalanced data distribution and identification of small lesions. The study's proposed attention blocks, namely CAB and GAB, can be incorporated into diverse deep architectures and trained end-to-end, making the model versatile and applicable.

The category attention block notably enhances the CABNet's performance, demonstrating the efficiency of the proposed approach in managing the imbalanced data distribution issue. The global attention block captures more details about small lesions. In the future, it would be beneficial to investigate the potential use of CABNet in other medical image analysis tasks, as well as to examine the interpretability of the attention mechanism.

2.4 Comparison of smartphone-based retinal imaging systems for diabetic retinopathy detection using deep learning

In this literature review, we discuss the materials and methods employed in a study that combined the RetinaScope device and the EyeArt system for DR screening. Furthermore, we present the study results and discuss their implications for the field.

2.4.1 Device Description

The RetinaScope, a portable retinal imaging device, was used in the study. Encased in a 3D-printed shell, the RetinaScope weighed approximately 310g and utilized deep red light diodes for illumination and imaging onto a smartphone camera. The device enabled operators to adjust focus, zoom, and exposure using touch and swipe gestures on an iPhone app connected via Bluetooth, enhancing image capture efficiency and patient comfort.

2.4.2 Results

2.4.2.1 Sensitivity and Specificity

Automated analysis utilizing the EyeArt system demonstrated promising results in terms of sensitivity and specificity. At the patient level, the system achieved a sensitivity of 87.0% and a specificity of 78.6%. At the eye-level, the sensitivity and specificity were 77.8% and 71.5% respectively. The human graders exhibited moderate inter-grader agreement, as indicated by a kappa value of 0.452 ± 0.334.

2.4.2.2 Discussion

The combination of the RetinaScope device and the EyeArt system showcased a promising approach for DR screening. The system demonstrated the necessary sensitivity for a DR screening tool, meeting the criteria set by the British Diabetic Association. Importantly, the study also highlighted the feasibility of utilizing the RetinaScope under non-ideal imaging conditions in conjunction with automated grading using EyeArt. The findings of this research support the integration of smartphone-based retinal imaging and automated grading systems for the early detection of diabetic retinopathy and related retinal diseases.

2.5 Comparison of Automated and Expert Human grading of Diabetic Retinopathy using Smartphone-based retinal photography

Several studies have explored the applications of smartphone-based retinal imaging platforms in medical diagnostics. These platforms have been employed in a range of experiments, with a focus on capturing synthetic retina images as well as actual fundus images for further examination. The devices that have been frequently used and examined in these studies include the iExaminer, D-Eye, Peek Retina, and iNview [26].

Studies have emphasized the need for an illumination source for retinal imaging due to the retina's dark nature. For instance, certain platforms such as D-Eye and iNview make use of the smartphone's flashlight, directing its light towards the retina. However, other systems like the iExaminer and Peek Retina come equipped with their integrated light sources. The image capture and storage in these scenarios are generally conducted using the smartphone's camera and internal memory.

2.5.1 Experimental Methodology Employed

The first phase of these studies often involves capturing images from a printed retina using various smartphone retinal imaging platforms. These images are then situated at the base of a synthetic eye model for further investigation. The subsequent phase typically includes acquiring retina images from the optimal distance for each device to maximize the field of view. This enables a comparative analysis of the image quality and field of view offered by each platform.

2.5.2 Image Quality Assessments in Retinal Imaging Platforms

Assessments are commonly conducted on the captured retina images using iExaminer, D-Eye, Peek Retina, and iNview. These images often present cases of normal retina, diabetic retinopathy, and proliferative diabetic retinopathy. Each image is then duplicated using the smartphone-based imaging systems for additional evaluation.

From various observations, it is found that platforms like the iExaminer can capture images of satisfactory quality, displaying visible optic nerve and blood vessels. Even signs of macula edema are discernible, indicating its potential in detecting diabetic retinopathy. Despite some platforms having uneven illumination, they can still showcase the macular edema, reinforcing their utility in detecting diabetic retinopathy. Certain systems provide a wider field of view, clear visibility of the optic nerve and vessels, making them suitable for diabetic retinopathy detection. Some systems offer the largest field of view, coupled with remarkable image quality and balanced illumination.

2.5.3 Field of View in Retinal Imaging Systems

The field of view varies across different systems, and while none of the systems can capture the entire retina, some provide a larger field of view than others. The importance of the field of view in the context of retinal imaging systems is often highlighted, particularly in relation to its impact on DR detection performance.

2.5.4 Conclusions from Previous Investigations

Existing studies primarily focus on evaluating the efficacy of smartphone-based retinal imaging systems such as the iExaminer, D-Eye, Peek Retina, and iNview. The results suggest that some systems provide a more extensive field of view and superior image quality. Furthermore, the correlation between the field of view and Diabetic Retinopathy detection performance has been noted, with some systems proving more effective than others. Despite the variations, these studies generally suggest that smartphone-based systems could potentially replace a direct ophthalmoscope. However, the field of view is frequently identified as a significant determinant of automatic DR detection accuracy.

2.6 CABNet over other proposed methods

The use of CABNet over other proposed literature reviews is justified based on its ability to address the challenges of imbalanced data distribution and small lesion identification in diabetic retinopathy (DR) grading. CABNet, a deep learning model, incorporates two types of attention modules: Category Attention Block (CAB) for imbalanced DR data distributions and Global Attention Block (GAB) for capturing small lesion information. This approach improves the accuracy, F1-score, and Area under the curve (AUC) of the model compared to other advanced models like MobileNet and DenseNet-121. The experimental outcomes demonstrate the effectiveness of CABNet in fine-grained grading of DR and its potential for managing imbalanced data distribution. The proposed attention blocks, CAB and GAB, can be incorporated into different deep architectures and trained end-to-end, making the model versatile and applicable to other medical image analysis tasks.

Chapter 3:
3. Software Tools

3.1 Development Environments

3.1.1 Google Colaboratory

Colab, the free cloud computing service from Google, utilizes Jupyter notebooks to enable the deployment of machine learning models. Its primary goal is to disseminate knowledge about machine learning and encourage research in this domain. Free of cost, Google extends this service to researchers and practitioners, granting them entry into a cloud-based realm with abundant GPU and RAM capabilities. By creating a Google account, one can effortlessly tap into this service.

The impressive computational capacity of Colab includes powerful GPU, TPU, and RAM resources, which are accessible for a set number of hours within a single period.

Consequently, it has become a favored option for the assessment of deep learning models. The infrastructure of this service is hosted on the Cloud platform provided by Google. Users can store their notebooks in Google Drive, which makes it easy to share notebooks and collaborate with others. Additionally, these notebooks can be uploaded or downloaded. Google Colab comes with many pre-installed libraries, including TensorFlow, PyTorch, and Keras, which makes it easy to get started with machine learning and deep learning [27].

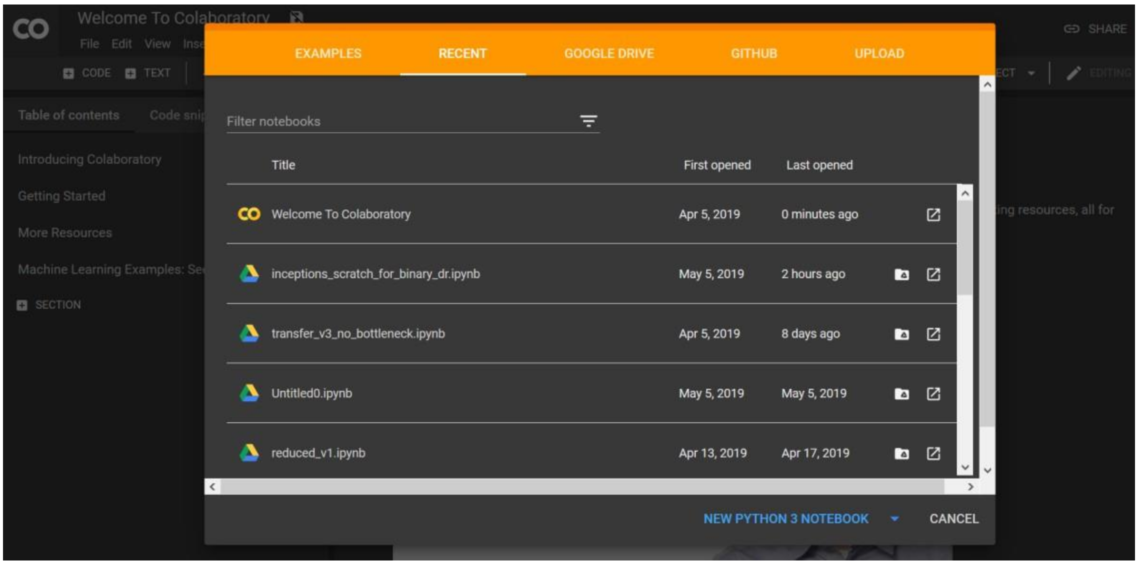


Figure 3.1: Google Colaboratory welcome screen

Google Colaboratory can achieve comparable performance with dedicated hardware [28]. Additionally, Colab has limitations due to the time limit of GPU, TPU utilization [28].

3.1.2 Jupyter Notebook

Jupyter Notebook is a freely available web-based tool that empowers individuals to create and distribute notebooks comprising executable code and graphical representations. It is a popular choice among data scientists and researchers who need an interactive environment for model training and other computational tasks The notebook encompasses a user-friendly interface that supports several programming languages, including Python.

Jupyter Notebook has garnered considerable popularity among data scientists and researchers owing to its versatility and ease of use. One of the key benefits of Jupyter Notebook is its real-time code execution ability, which allows users to experiment with their code and instantly visualize data. Additionally, it facilitates the integration of diverse libraries and tools, thereby adds flexibility to serve as a platform for data analysis and machine learning.

Moreover, due to the limited GPU and TPU available with Google Collab, Jupyter Notebook, equipped with dedicated hardware, was utilized for model training in this project. This deployment enabled the running of large-scale computations and the training of complex models, which would have been restricted with Google Collab's limited hardware resources.

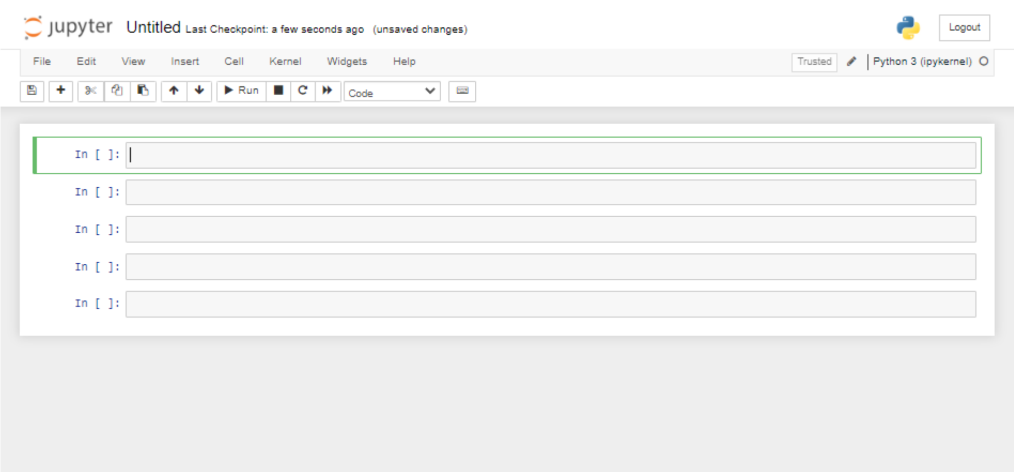


Figure 3.2: Jupyter Notebook Development Screen

3.1.3 Android Studio

In areas with limited access to professional ophthalmologists and medical diagnostic equipment, a low-cost and user-friendly smartphone-based diagnostic tool for diabetic retinopathy would be advantageous. The growing global acceptance of smartphones presents an opportunity to incorporate automatic identification and classification of abnormalities in the retina caused by diabetes, which could benefit those with restricted access to medical services. A software application is designed specifically for the detection of DR on a smartphone. "

Android Studio is a software tool developed by Google primarily for creating applications for the Android operating system. Developers are offered an extensive range of tools to aid in the creation, development, testing, and troubleshooting of Android applications.

Android Studio is commonly used to develop, test, and deploy a smartphone-based diagnostic application for identification and classification of retinal abnormalities, running on an Android operating system. A self-contained smartphone software can provide patients residing in homes or remote areas with easy access for DR identification without the need for network connectivity.

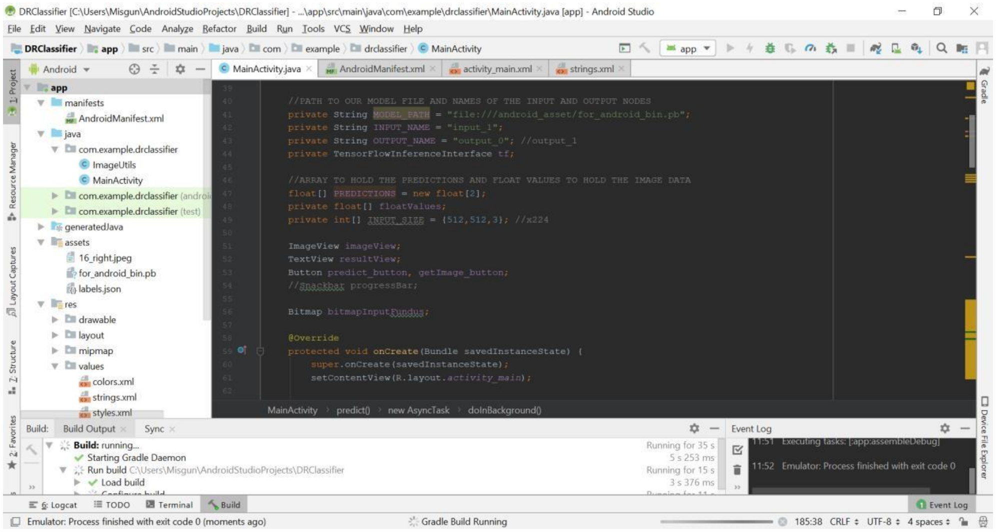


Figure 3.3: Android Studio Development Screen

3.1.4 MATLAB

MATLAB is a development environment and programming language that is widely used in engineering and scientific research. It is known for its powerful functions and tools that enable users to analyze, visualize, and model data. One of its notable strengths is its ability to manage large datasets and perform numerical computations quickly and efficiently.

MATLAB is known for managing massive data sets and conducting speedy numerical computations. Its potential can be harnessed for diverse image processing operations such as image enhancement and segmentation by making use of its built-in functions or custom scripts. For training models in the fields of machine learning and deep learning, MATLAB is a highly sought-after choice. It provides a variety of features for model creation and training, encompassing built-in algorithms and functions for neural networks, decision trees, and other methods. Furthermore, MATLAB comes equipped with tools to preprocess data, extract features, and visualize data, rendering it a comprehensive platform that supports end-to-end workflows for machine learning.

Research and development settings frequently make use of MATLAB, especially in fields like computer vision, medical imaging, and signal processing. Its adaptability and user-friendly interface make it a popular preference among both academic researchers and industry professionals.

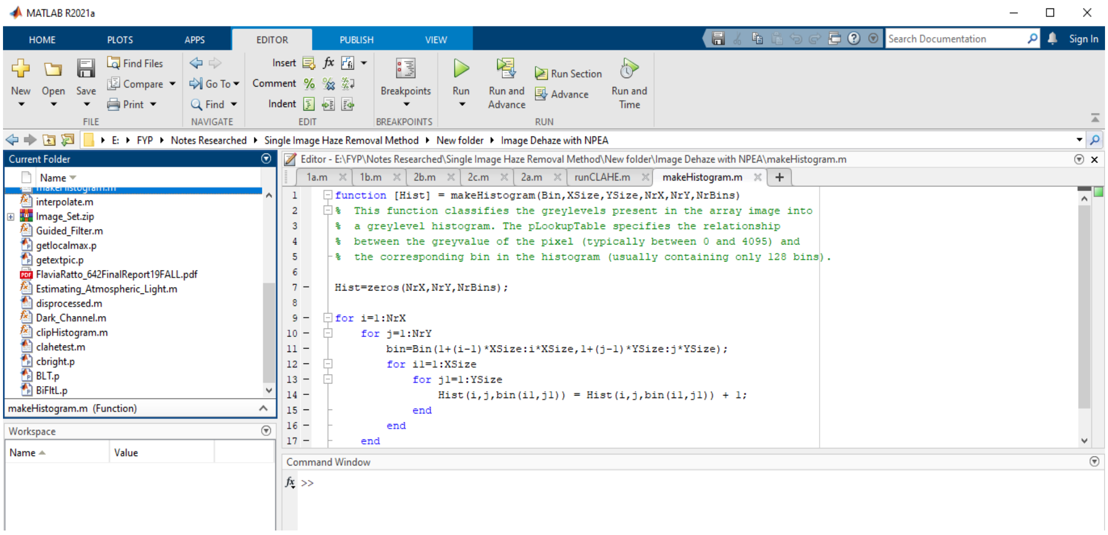


Figure 3.4: MATLAB Development Screen

3.2 Mobile Application Development

3.2.1 Flutter and Dart: Powering Seamless UI Development

Google's open-source User Interface (UI) software development kit (SDK), Flutter, is becoming a popular choice among developers. Its ability to enable creation of top-tier native interfaces for mobile, and desktop applications from one shared codebase sets it apart. Dart, another Google invention, is the programming language at the heart of Flutter.

3.2.2 Dart: Object-Oriented Programming for Enhanced Performance

Dart embodies a class-defined, object-oriented language that employs garbage collection and operates with a C-style syntax. The language's ability to compile effectively into native code is pivotal in augmenting the overall performance of the Flutter framework. Like JavaScript, Dart endorses a reactive framework which makes it particularly well-suited for developing UIs.

3.2.3 Flutter for Android App Interface Development

The ability to witness the impact of code changes in real-time is one of the defining features of Flutter. Known as the 'hot reload', this functionality eliminates the need to reboot the application for reviewing modifications. This accelerates the development process and enables faster iterations.

3.2.4 Interactive and Customizable UI

Flutter is designed to maintain UI design uniformity across diverse platforms. This enables developers to create apps with identical appearance and behavior on both Android and iOS. The resultant consistency saves valuable time and resources in the app development process.

A notable feature of Flutter is its array of customizable widgets. It offers an extensive assortment of pre-configured widgets, including scrolling, navigation, icons, and fonts, which can be employed to craft highly responsive and engaging applications.

3.2.5 Enhanced Performance

With Dart's ability to directly compile into native code, Flutter outperforms many other cross-platform frameworks in terms of performance. The absence of a JavaScript bridge, a common performance hurdle in many frameworks, enables Flutter to deliver an optimized, seamless user experience.

3.3 Web Application

Web application development has become increasingly important in the digital landscape, with a growing demand for interactive and responsive web experiences. This section focuses on utilizing React.js, a popular JavaScript library, for building robust and dynamic web applications.

3.3.1 React.js

React.js, an open-source JavaScript library developed by Facebook, has gained significant popularity among developers due to its ability to create reusable UI components. It follows a component-based architecture that enhances code reusability and maintainability, enabling efficient web application development.

3.3.1.1 Virtual DOM and Efficient Rendering

One of the key features of React.js is its virtual DOM (Document Object Model). By using a virtual representation of the DOM, React.js minimizes direct manipulation of the actual DOM, resulting in faster rendering and improved performance. This approach allows developers to efficiently update and modify the UI without reloading the entire web page.

3.3.1.2 Component-Based Architecture

React.js encourages the development of UI components as independent, reusable building blocks. These components can be composed to create complex user interfaces, facilitating a modular and scalable approach to web app development. This modular structure simplifies code organization, enhances reusability, and improves the overall maintainability of the project.

3.3.1.3 Responsive and Interactive User Interface

React.js enables the creation of highly responsive and interactive user interfaces. With its efficient rendering and component reusability, developers can build web applications that provide a seamless user experience across different devices and screen sizes. React.js also supports the integration of CSS frameworks and libraries, allowing for enhanced styling and customization options.

Chapter 4:
4. Methodology and Working

In this chapter, the project's overall functionality is presented, encompassing suggested methodologies and their outcomes. The work comprises techniques for preparing datasets, pre-processing, and why pre-processing is crucial for improved data visualization and accuracy during model training. The chapter also discusses various deep learning models, their architecture, and classification outcomes, as well as how these models are employed for training purposes. The transfer learning method is useful for classifying new data that has not been seen before, and it is also addressed. The proposed work employs deep learning models to classify diabetic retinopathy from fundus images. The first stage identifies whether the image is a retinal image, the second stage classifies it as referable or non-referable. The chapter also includes the integration of the two trained deep learning models into the smartphone and web application and the constraints or difficulties encountered during model deployment, as well as how model is integrated to the applications.

4.1 Retinal Image Dataset Description

The retinal fundus images dataset consists of high-resolution images captured using fundus photography cameras and optical coherence tomography machines for the screening of diabetic retinopathy. Clinics and healthcare sectors use telemedicine technology to upload the images to a centralized database for research purposes.

The retinal fundus images dataset is vital for developing and testing algorithms to diagnose and screen for eye diseases. These datasets may vary in terms of different classes, quality, and quantity of fundus images. Moreover, utilizing such datasets can also be advantageous in creating greater public awareness regarding eye health. In Table 4.1, an overview of the public available datasets is presented.

Table 4.1: Dataset for DR Detection

Dataset	Number of images	Resolution	Format	Annotation	Tasks
Messidor [29]	1200	1440x960, 2240x1488, 2304x1536	TIFF	Image level	- DR grading - Risk of DME
e-ophtha [30]	148 (MAs), 233 (Normalnon-MA) 47 (EXs), 35 (Normalnon-EX)	-	-	Pixel level	-MAs detection -EXs detection
Kaggle [31]	80000	-		Image level	-DR grading
DRIVE [32]	33 (Normal) 7 (Mild earlyDR stage)	584x565	JPEG	Pixel level	-Vessel's extraction

STARE [33]	400	605x700	-	Pixel level	-13 retinal diseases -Vessel's extraction -Optic nerve
DIARETDB1 [3**4**, 35]	5 (Normal) 84 (At least one NPDR sign)	1500x1152	-	Pixel level	-MAs
CHASE [36]	28	1280x960	-	Pixel level	-Vessel's extraction
DRiDB [37] -	50	-	BMP	Pixel level	-MAs -HMs -HEs -SEs -Vessel's extraction -OD -Macula
ORIGA [38]	482 (Normal) 168 (Glaucomatous)		-	Pixel level	-OD -Optic cup -Cup-to-Disc Ratio (CDR)
SCES [40]	1630 (Normal) 46 (Glaucomatous)			Pixel level	-Cup-to-Disc Ratio (CDR)
AREDS [41]	72000	-	-	Image level	-AMD stages
REVIEW[42]	16	-	-	Pixel level	-Vessel's extraction
EyePACS-1 [43]	10000			Image level	-Referable DR -MA
RIM-ONE [44]	18 (Normal) 12 (Early glaucoma) 14 (Moderate glaucoma) 14 (Deep glaucoma) 11 (Ocular hypertension)	-	-	Pixel level	-Optic nerve
DRISHI-GS [45]	31 (Normal) 70 (Glaucomatous)	2896x1944	-	Pixel level	-OD -segmentation -Optic cup (OC) segmentation
ARIA [46]	16 (Normal) 92 (AMD) 59 (DR)	768x576	-	Pixel level	-OD -Fovea location -Vessel extraction
DRIONS-DB [47]	110	600x400	-	Pixel level	-OD
SEED-DB [48]	192 (Normal) 43 (Glaucomatous)	3504x2336	-	Pixel level	-OD -Optic cup (OC)

4.2 EyePacs Dataset

The given dataset is a compilation of retinal images, which was made accessible from a fundus images classification competition on Kaggle [31].

EyePacs dataset comprises 35,126 fundus images, there are five stages labeled from zero to four [3], the first stage is "No DR" it contains 25812 fundus images, the second stage is "Mild DR" it contains 2442 fundus images, the third stage is "Moderate DR" it contains 5291 fundus images, the fourth stage is "Severe DR" it contains 873 fundus images, the final stage is "Proliferative DR" it contains 708 fundus images.

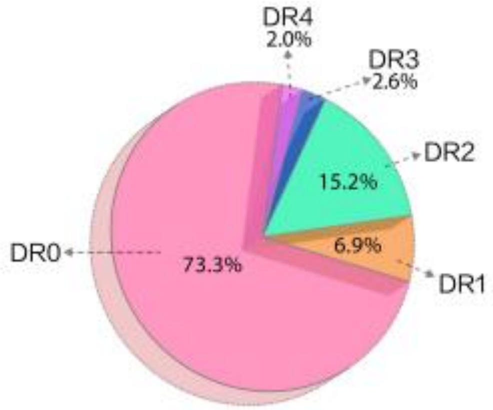


Figure 4.1: EyePacs Dataset Distribution [3]

The EyePacs fundus images dataset presents researchers with a valuable resource for devising and testing algorithms for the screening and diagnosis of diabetic retinopathy.

4.2.1 Kaggle Dataset Preparation

The fundus images were modified to have a width and height of 224pixels. The Kaggle dataset was then divided into three separate groups: the training set, the validation set, and the test set with distribution percentages of 60%, 20%, and 20% respectively, for binary and multiclass classifications. This dataset was used for model training and testing. Due to the uneven distribution of labels in the dataset, image augmentation was applied as it is necessary while applying deep learning algorithms for DR classification. The distribution of fundus images into training, validation and test directories can be observed in Figure 4.2

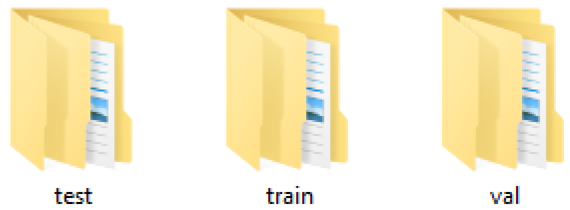


Figure 4.2: Fundus Images Distribution

Figure 4.2 illustrates the training, validation and test folders contain their perspective fundus images.

Figure 4.3: Arrangement of fundus Images in Respective Classes

Figure 4.3 illustrates that each fundus image has been allocated to its respective folder following its label.

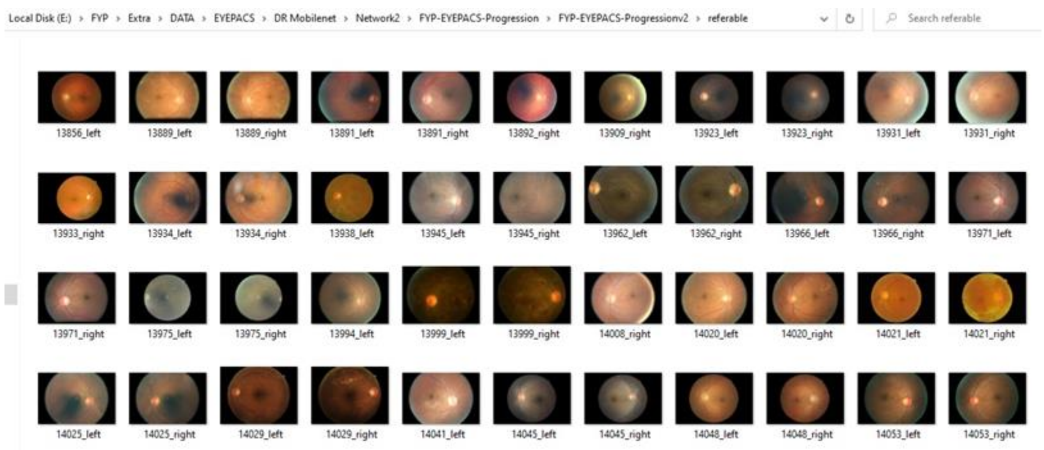


Figure 4.4: Fundus Images in referable Class

4.3 DDR Dataset

DDR dataset is assembled by Tao Li et. al. It is the second largest dataset. The dataset is a compilation of retinal images, it was made accessible from a fundus images classification competition on Kaggle [31]. The dataset was obtained from health facilities that use telemedicine technology and fundus photography cameras to screen for diabetic retinopathy.

Due to the uneven distribution of labels in the dataset, image augmentation is necessary while applying deep learning algorithms for DR classification. The fundus images dataset presents researchers with a valuable resource for devising and testing algorithms for the screening and diagnosis of diabetic retinopathy.

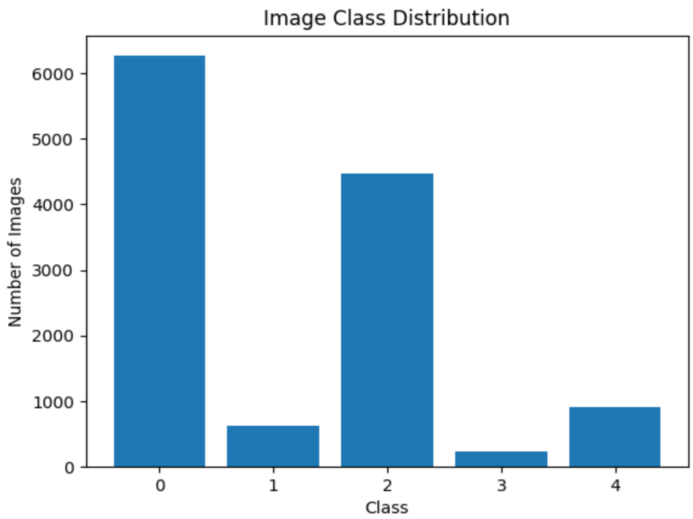


Figure 4.5: DR Dataset Distribution [31]

4.4 Image Preprocessing

The retinal images dataset [31] may pose a challenge for implementing classification algorithms or can lead to low contrast images that can affect the training of deep learning models due to the noise present in the images. This noise can impede the model's ability to extract the necessary features for effective diabetic retinopathy classification. To overcome this challenge, the selected fundus images underwent pre-processing using image pre-processing techniques. Preprocessing techniques considered for the project are described further below:

4.4.1 Contrast Limited Histogram Equalization (CLAHE)

One of the pre-processing techniques considered was Contrast Limited Adaptive Histogram Equalization (CLAHE), which is commonly used in image processing. It improves the contrast of the image and can also remove noise from images.

CLAHE overcomes this challenge by equalizing the image histogram locally in small regions, rather than the entire image. In CLAHE, the image is divided into tiles, and the histogram equalization process is applied to each tile separately to prevent over-amplification of noise in areas with little contrast and to prevent saturation in areas with high contrast.

4.4.2 Naturalness Preserved Enhancement Algorithm

We used a second pre-processing technique to address the issue of images captured in non-uniform illumination conditions. Such images often suffer from reduced contrast and less visible details due to uneven illumination distribution in the scene. To enhance the quality of these images, various image enhancement algorithms have been proposed. In this project, we presented a naturalness preserved enhancement algorithm (NPEA) that effectively enhances the visibility of details while maintaining the naturalness of the images. The NPEA algorithm consists of three key steps: correcting uneven lighting, improving contrast, and preserving naturalness [49].

4.4.3 Bens Graham Preprocessing Technique

The provided preprocessing technique, developed by Ben Graham, was used in a Kaggle competition for the detection of diabetic retinopathy. Its purpose is to prepare retinal images by enhancing relevant features, removing noise, and normalizing the images for effective classification. The technique involves cropping dark regions, converting images to a specific size, and applying an image enhancement process using Gaussian blur and linear combination [56]. These steps collectively improve the quality and suitability of the images for subsequent diabetic retinopathy classification tasks.

4.4.3.1 Steps in Bens Graham Preprocessing Technique

The technique involves the following steps:

4.4.3.1.1 Crop Image

This step involves cropping the black or dark regions around the retinal images to focus only on the relevant part that contains the retina. The process differs based on whether the image is grayscale or RGB:

4.4.3.1.1.1 For grayscale images

A binary mask is created, where pixels above a threshold value are set to True and the rest to False. The resulting cropped image contains only the region of interest without the dark areas.

4.4.3.1.1.2 For RGB Images

The image is converted to grayscale. The same binary mask creation process is applied. If the extracted region is empty (no part of the image exceeds the threshold), the original image is returned. Otherwise, the corresponding regions from each channel (R, G, B) are separately extracted using the binary mask, and the final cropped RGB image is created.

4.4.3.1.2 Preprocessing Images

This step is the main preprocessing pipeline that takes the file path of an image as input and returns the preprocessed image. The steps involved are as follows:

The image is read in BGR format. It is converted to RGB format. The previously defined **"Crop Image from Gray"** function is applied to remove dark regions. The image is resized to a specific size (IMG_SIZE) to ensure consistent dimensions for feeding into a neural network. An image enhancement technique is applied using the `cv2.addWeighted` function. It combines the original image and its Gaussian blur to enhance edges and features.

The resulting preprocessed image is ready for use as input in a diabetic retinopathy classifier. The technique incorporates parameter values, such as `tol` and `sigmaX`, which can be adjusted to fine-tune the preprocessing steps. For example, `tol` determines the threshold value for classifying pixels as dark, and `sigmaX` controls the amount of blur applied to the image. The application of the`cv2.addWeighted` function further enhances the image by sharpening edges and improving certain features, potentially benefiting the subsequent classification task of diabetic retinopathy [56].

Figure displays the outcome of applying image pre-processing techniques to sample fundus images.

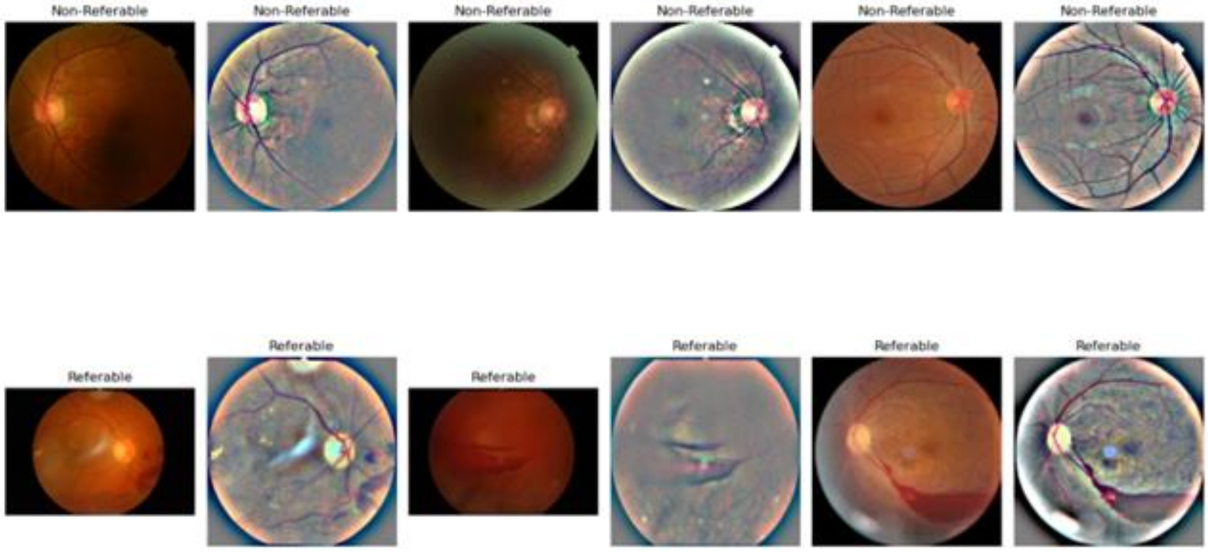


Figure 4.6: Images Before and After Applying Ben Graham's Preprocessing

4.4.4 Selection of Preprocessing Technique

We selected Bens Graham Preprocessing Technique [56] over the Naturalness Preserved Enhancement Algorithm (NPEA) [49] and CLAHE due to its superior efficiency during the preprocessing phase. This technique offers more precise and accurate detection of diabetic retinopathy.

4.5 Deep Learning

In recent years, Deep Learning has gained significant attention as a branch of machine learning that utilizes neural networks for the purpose of analyzing and forecasting patterns in data.

Deep learning models are composed of multiple layers of artificial neurons that work in tandem to analyze complex data patterns. Due to the emergence of large datasets and increased computing power, deep learning has found extensive use in many fields including computer vision, natural language processing, speech recognition, and healthcare.

In the healthcare industry, one of the most notable areas where deep learning is utilized is in the analysis of medical images, where deep learning models are utilized to detect and classify medical conditions from radiology images such as X-rays, CT scans, and MRI scans. For instance, in ophthalmology, deep learning models are developed to analyze fundus images to diagnose diabetic retinopathy, a common diabetes complication that could result in blindness if left untreated. Deep learning is expected to become an essential tool for solving complex problems and making breakthrough discoveries.

4.6 Diagnosis

This section provides a detailed overview of the various stages involved in the development of the system, including image acquisition, image preprocessing, and classification.

4.6.1 Image Acquisition

In the beginning, retinal images are acquired from a smartphone camera or an existing gallery. Using the camera on a smartphone, people can easily capture their retinas during regular health check-ups or exams. These images serve as the input for the subsequent stages of the diabetic retinopathy detection system.

4.6.2 Image Preprocessing

After acquiring the images, various adjustments are made to enhance their quality and extract important features for classification using Ben's Graham Preprocessing [56]. These adjustments include resizing, cropping out darker regions, noise reduction, and contrast enhancement. The primary objective is to improve the visibility and sharpness of retinal structures, enabling accurate analysis during the classification process.

4.6.3 Classification

Once the images are preprocessed, the phase of classification begins. The classification process continues in two steps:

1. Retinal Image Identification

   It identifies the image as retinal or non-retinal before it is put to disease classification. Non retinal images are of no use, so this step stops the users from entering the classification phase because it is just a waste of time.

2. Disease Classification

   After identifying the retinal image, the actual disease classification process begins. The image is classified as "Referable" diseased image or "non-Referable". Non referable means that the person is safe from the disease and does not need to visit the specialist.

Training networks are discussed in detail in the sections below.

4.7 System Architecture

The system architecture was designed in three phases to effectively process and analyze diabetic retinal images. The first phase involved acquiring the retinal images from a mobile phone or selecting them from the image gallery. The acquired images underwent image preprocessing, including steps such as noise removal, image resizing, and augmentation techniques.

Once the preprocessing phase was completed, the second phase focused on model training. The model was trained using the pre-processed images, with additional training performed using unseen data. Subsequently, the model underwent testing to evaluate its performance. Various metrics such as accuracy, losses, and verification techniques were employed to assess the effectiveness of the model.

In the third phase, the final model's performance was examined using evaluation criteria and metrics such as accuracy, specificity, and sensitivity. This evaluation helped determine the model's proficiency in diagnosing and classifying diabetic retinopathy. Once integrated into the app, the finalized model finds out the presence of diabetic retinopathy based on the provided retinal images as “Referable” and “Non-Referable”.

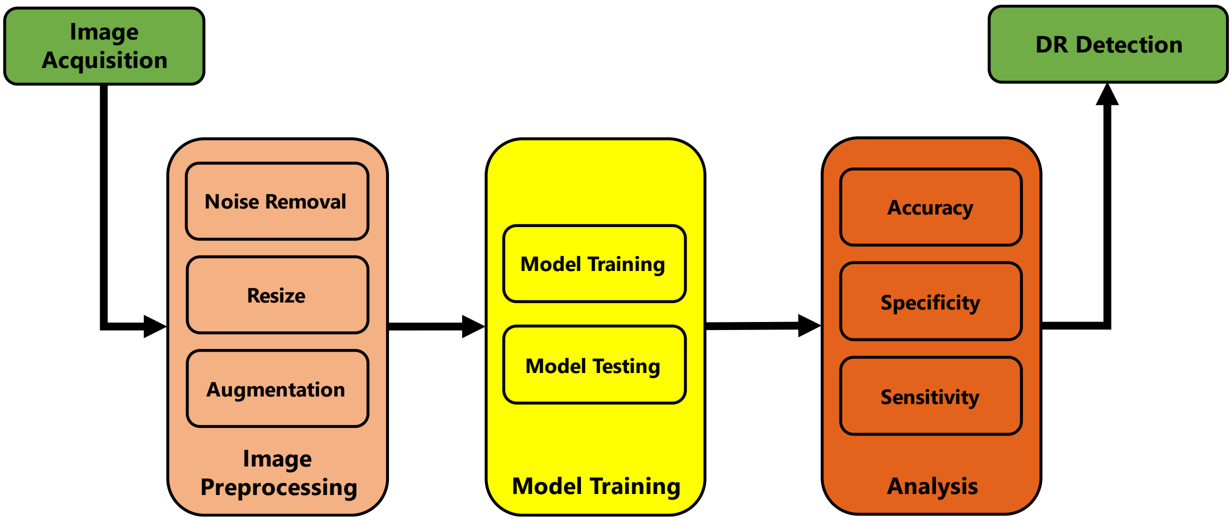


Figure 4.7: Process Flow Diagram

4.8 2-Stage Network

The detection of diabetic retinopathy involves the following two model network.

4.8.1 Network 1: Retinal Image Identification

In this project, the goal is to develop a classification model that can distinguish between retinal and non-retinal images. The dataset consists of two classes: "non-retinal" (real-world images) and "retinal" (images of the retina). The model is trained using a simple artificial neural network (ANN) architecture with specific layers and parameters. However, due to limitations in performance on images with complex features, an alternative approach using MobileNetV2 is explored.

4.8.1.1 Simple ANN Training

4.8.1.1.1 Dataset

The training dataset consists of 2800 images from the "non-retinal" class (labeled as 0) and 2800 images from the "retinal" class (labeled as 1). To ensure class balance, the dataset is carefully balanced.

4.8.1.1.2 Architecture

The simple ANN architecture comprises the following layers:

- Input Layer: Accepts the image input.
- Flatten Layer: Converts the 2D image data into a 1D feature vector.
- Dense Layer 1: Fully connected layer with a specified number of units/neurons.
- Batch Normalization Layer: Normalizes the activations of the previous layer.
- Dropout Layer: Regularizes the model to prevent overfitting.
- Dense Layer 2: Final fully connected layer with the same number of units as the number

of classes (2 in this case).

- Output Layer: Applies the SoftMax activation function for classification.

4.8.1.1.3 Training

4.8.1.1.3.1 Procedure

- Loss Function: Binary cross-entropy loss function is used to measure the model's performance.
- Optimizer: The Adam optimizer is utilized to optimize the model's weights.
- Activation Function: SoftMax activation is used in the output layer for class probabilities.

4.8.1.1.3.2 Process

- The model is trained using the training set, which contains both the "non-retinal" and "retinal" images with their respective labels.
- The training process is performed for a total of 30 epochs, with early stopping applied at 20 epochs if the validation loss does not improve.
- Early Stopping: Patience of 10 is set, meaning training is stopped if the validation loss does not improve for 10 consecutive epochs.

4.8.1.1.4 Evaluation

- After training, the model's performance is evaluated using the test dataset.
- The test dataset consists of 2400 images for each class, maintaining class balance.

4.8.1.1.5 Inference

- Single image inference is performed to assess the suitability of the model for real-time predictions.
- The simple ANN architecture performs well on some images but struggles with images containing complex features.

4.8.1.2 MobileNetV2 Training

4.8.1.2.1 Architecture

Considering the limitations of the simple ANN architecture, MobileNetV2 is chosen as a base model to enhance the model's performance.

- MobileNetV2 is a lightweight deep learning model architecture designed for mobile and embedded vision applications.
- A sequential model is created with MobileNetV2 as the base model and the previous simple ANN architecture added on top.

4.8.1.2.2 Freezing Layers

- To retain the features learned by MobileNetV2, the layers of the base model are frozen during training.
- Only the additional layers from the simple ANN architecture are trained.

4.8.1.2.3 Training

- The training process is conducted using the same dataset and class labels as in the previous training.
- The model is trained for 27 epochs, with early stopping applied at 17 epochs if the validation loss does not improve.

4.8.1.2.4 Learning Parameters

- The learning parameters used for this training are the same as those employed in the first training of the simple ANN architecture.

4.8.1.2.5 Inference

- Inference is performed on both simple and complex feature images using the saved model.
- The MobileNetV2-based model exhibits improved performance compared to the simple ANN architecture during real-time inference.
- Consequently, MobileNetV2 is chosen over the simple ANN architecture for deployment in the mobile and web app for classification purposes.

4.8.1.3 Conclusion

The project involves training a classification model to differentiate between retinal and non-retinal images. Initially, a simple ANN architecture is employed, but its performance is limited on images with complex features. To address this, MobileNetV2 is utilized as a base model, combined with the previous ANN architecture. Training is conducted, and the resulting model outperforms the simple ANN architecture during real-time inference. Hence, MobileNetV2 is selected for deployment in the mobile and web app for classification tasks.

4.8.2 Network 2: Disease Detection

If the image is recognized as a retinal image, the process moves on to the next stage. Here, the image undergoes evaluation to detect the presence of diabetic retinopathy. A deep learning model was trained on around 18,000 retinal images from the EyePacs dataset, which were categorized into two classes: "Referable" and "Non-Referable". The model's knowledge was then transferred using transfer learning to a new dataset called DDR, consisting of 12,000 images. This enabled the model to classify retinal images as either "DR" or "No DR" in the DDR dataset. This classification provides crucial information about the presence of the disease, facilitating timely medical intervention and treatment.

The entire process is designed to be executed on a smartphone during runtime, enabling image acquisition and classification to occur in real-time on the mobile and web supported device. This setup allows for convenient and efficient detection of diabetic retinopathy, providing valuable information for timely medical intervention and treatment.

4.8.3 Network-2 Classification Models for Diabetic Retinopathy Detection

The project proposes automatic DR detection methods for smartphone-based retinal images using deep learning frameworks. The proposed method involves the use of a deep learning model based on MobileNetV2 trained on a dataset of approximately 18,000 retinal images from the EyePacs dataset, classified into two classes: DR and No DR. Transfer learning was employed to leverage the preexisting knowledge of the model trained on the EyePacs dataset to classify retinal images from the DDR dataset, which comprised 12,000 images, into "DR" and "No DR".

To detect and classify diabetic retinopathy (DR) deep learning technique is implemented. The MobileNetv2 network model was chosen for training. The model is pretrained on the ImageNet dataset, providing them with a robust understanding of general image features. This section explores the effectiveness of MobileNetv2 for DR classification, particularly when used on smartphones.

4.8.3.1 Bens Graham Preprocessing Technique

The Bens Graham preprocessing technique is utilized to enhance retinal images for effective classification. It involves cropping dark regions, resizing images, and applying image enhancement techniques such as Gaussian blur and linear combination.

4.8.3.2 Training with MobileNetV2

MobileNetv2, pretrained using the vast ImageNet dataset, possesses an abundance of varied images. The model acquires broad image characteristics pertinent to DR classification. As a result, the necessity for arduous training with restricted DR-specific datasets is greatly diminished, expediting the models' convergence.

4.8.3.2.1 Model Compactness and Efficiency

MobileNetv2 stands out for its ability to be lightweight and highly efficient, making it an ideal choice for smartphones. Its architecture is simplified, featuring depth wise separable convolutions that lower computational complexity and memory requirements. As a result, resource-limited devices can perform real-time classification, detecting DR directly on the device without the need for external servers or internet connectivity. This approach ensures independence and accessibility.

4.8.3.2.2 Architecture

- MobileNetV2 is a lightweight deep learning architecture suitable for mobile and embedded vision applications.
- The additional ANN architecture, comprising five layers, is added on top of MobileNetV2 for classification purposes.

4.8.3.2.3 Dataset

- The training dataset consists of 9000 images for each class: "non-referable" (labeled as 0) and "referable" (labeled as 1).
- Class balance is ensured by carefully balancing the dataset.

4.8.3.2.4 Training

- The MobileNetV2 model is trained using the EyePacs dataset.
- The training is performed for a total of 15 epochs, with early stopping applied at 10 epochs if the validation loss does not improve.
- Early Stopping: Patience of 10 is set, meaning training is stopped if the validation loss does not improve for 10 consecutive epochs.

4.8.3.2.5 Learning Parameters

- Loss Function: Binary cross-entropy loss function is used to measure the model's

performance.
- Optimizer: The Adam optimizer is utilized to optimize the model's weights.
- Activation Function: Softmax activation is used in the output layer for class probabilities.

4.8.3.2.6 Evaluation

- After training, the model's performance is evaluated using the DDR dataset for testing and validation.
- The test and validation datasets consist of 2500 images for each class, maintaining class balance.
- Single image inference is performed to assess the model's suitability for real-time predictions.
- The MobileNetV2 model performs well on some images but struggles with images containing complex features.

4.8.3.3 Fine-tuning with Bens Graham Preprocessing and Transfer Learning

4.8.3.3.1 Fine-tuning

- The Bens Graham preprocessing technique is applied again to enhance the images in the training dataset.
- The previously saved MobileNetV2 model from the first training is loaded to leverage its knowledge on the new dataset.

4.8.3.3.2 Training Process

- The model is fine-tuned using the MobileNetV2 architecture as the base model.
- A sequential model is created with MobileNetV2 as the base model and the previous ANN architecture.
- The trainable layers of MobileNetV2 are set to "True" to learn complex features from the new dataset.
- The model is trained for 20 epochs, with early stopping applied at 10 epochs based on the validation loss.
- The learning parameters remain the same as in the first training phase.
- Inference is performed on both simple and complex feature images using the fine-tuned MobileNetV2 model.
- The fine-tuned MobileNetV2 model demonstrates improved performance during real-time inference.

4.8.3.3.3 Dataset and Evaluation

- The training dataset is changed to the DDR dataset, while the classes and labels remain the same.
- The training set consists of 3700 images for each class, maintaining class balance.
- The model's performance is evaluated using the DDR dataset for testing and validation.

4.8.3.4 Conclusion

The project involves training a classification model to differentiate between non-referable and referable images related to diabetic retinopathy. Bens Graham preprocessing technique is applied to enhance the images before classification. Initially, the MobileNetV2 architecture is used with an additional ANN architecture. However, due to limitations in handling complex features, a second training phase is conducted. The fine-tuned MobileNetV2 model, combined with the previous ANN architecture and Bens Graham preprocessing, performs better during real-time inference. Consequently, the fine-tuned MobileNetV2 model is selected for deployment in the mobile and web app for classification purposes.

4.9 Model Integration

The trained deep learning model involves merging it into a portable device like a smartphone. This merging process revolves around integrating the trained model into an application.

To integrate the model into a smartphone app, one must initially develop the app with suitable software libraries and tools. The app ought to incorporate a camera interface designed to capture fundus images, which can then undergo processing through the optimized deep learning model.

Once the app is developed, the saved model can be loaded and incorporated into the app's inference pipeline. Subsequently, the app can undertake real-time inferences on the acquired images, offering a diagnosis for diabetic retinopathy.

During the integration process, it is imperative to consider any restrictions imposed by the device's hardware or software, such as memory and processing capabilities. Hence, fine-tuning the optimized model becomes crucial to ensure its ability to perform inferences within the confines of the device.

4.9.1 Model Optimization

The process of optimization entails squeezing the saved model to shrink its size, while simultaneously enhancing its efficiency and performance. This method is crucial as deep learning models tend to be bulky and demand extensive resources, which may impede their usage on less capable gadgets like smartphones or embedded systems. Consequently, optimizing the model can boost its portability and enable its deployment on a broader array of devices. The project delves into quantization as an approach to reduce and streamline the trained model's complexity, without significantly compromising its accuracy. The process involves decreasing the weight and precision of the model, resulting in a decrease in its memory consumption and computational requirements.

4.9.1.1 Model Compression Techniques

In deep learning, this approach can be used to fine-tune a trained model for optimal performance on devices with limited resources.

Quantization entails the transformation of floating-point values into integers, leading to reduced memory consumption and faster inference due to fewer computations. The quantization process involves assigning integer values to the weights and activations of the saved model.

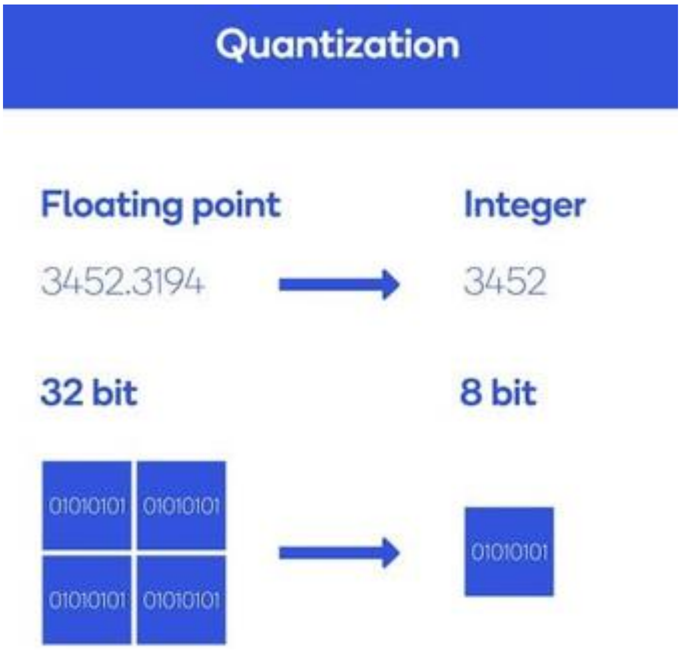


Figure 4.8: Quantization Process

4.9.1.2 TensorFlow Lite

Compressed model is saved in the *'tflite'* file format. It is then integrated to the application using libraries. Since, flutter was being used, the libraries required for TensorFlow Lite model integration are ***'tflite_flutter'*** and ***'tflite_flutter_helper'***. The two libraries make it easy for the application to accept the models.

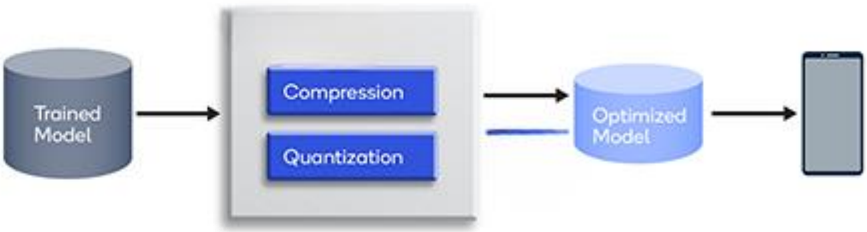


Figure 4.9: Quantization in model compression

4.9.2 Flask API

Flask is a popular web framework for building web applications and APIs in Python. It provides a simple and lightweight approach to developing server-side applications. In the context of the project, Flask is used to create an API that integrates the trained models and handles the classification requests from the app. The Flask API acts as a bridge between the front-end interface and the backend model deployment, allowing seamless communication and inference.

4.9.2.1 Flask for Model Integration

4.9.2.1.1 API Development

Flask is used to develop the API that connects the web app interface with the deployed models. The Flask application defines the necessary routes and functions to handle incoming requests.

4.9.2.1.2 Request Handling

When a user selects an image for classification, the web app sends a request to the Flask API. Flask receives the request and extracts the necessary data, such as the image file or image URL.

4.9.2.1.3 Model Inference

Flask then utilizes the deployed models to perform inference on the received image data. The models process the image and generate classification results based on their trained knowledge.

4.9.2.1.4 Deployment on Server

The trained models are deployed on server through Flask. The API generated by Flask helps transfer images to server and return the results after classification.

4.9.3 Selecting the Integration Technique

Two techniques are put forth for model integration in application. Both are efficient enough to be used but there are drawbacks for each of them.

Model deployed to server using the API generated by Flask cannot run unless an internet connection is available, but the performance is not compromised. The full original sized model without compression is deployed that works on its original and full performance.

TensorFlow Lite is compressed model that is, the size of the model is reduced for mobile application integration. The reduction in size also decreases the performance of the model such that complex problems such as our own cannot bear this reduced performance.

The reduced performance is not acceptable because the problem at hand is the sensitive organ of a living being and that cannot be left unattended. So, the server deployment of the model using API generated by Flask is being entertained because low performance cannot be tolerated.

4.9.4 Integration into a Smartphone Application

The suggested method is incorporated into a smartphone app built with Flutter for classification aims. This app provides a user-friendly interface that allows people with diabetes to easily take retinal pictures, process them, and quickly receive outcomes about the existence and seriousness of diabetic retinopathy. Additionally, it operates online and automates the diagnosis process. By leveraging the capabilities of smartphones, the system enables convenient and accessible screening and detection of DR, especially in remote areas with limited access to specialized healthcare centers.

4.9.4.1 Flutter App Interface Design

The integration of the proposed method into a mobile-based classification smartphone application involves designing an intuitive and user-friendly interface using the Flutter framework. The Flutter framework allows for the development of cross-platform applications, ensuring compatibility with both Android and iOS devices. Flutter provides an image picker package to access gallery or camera.

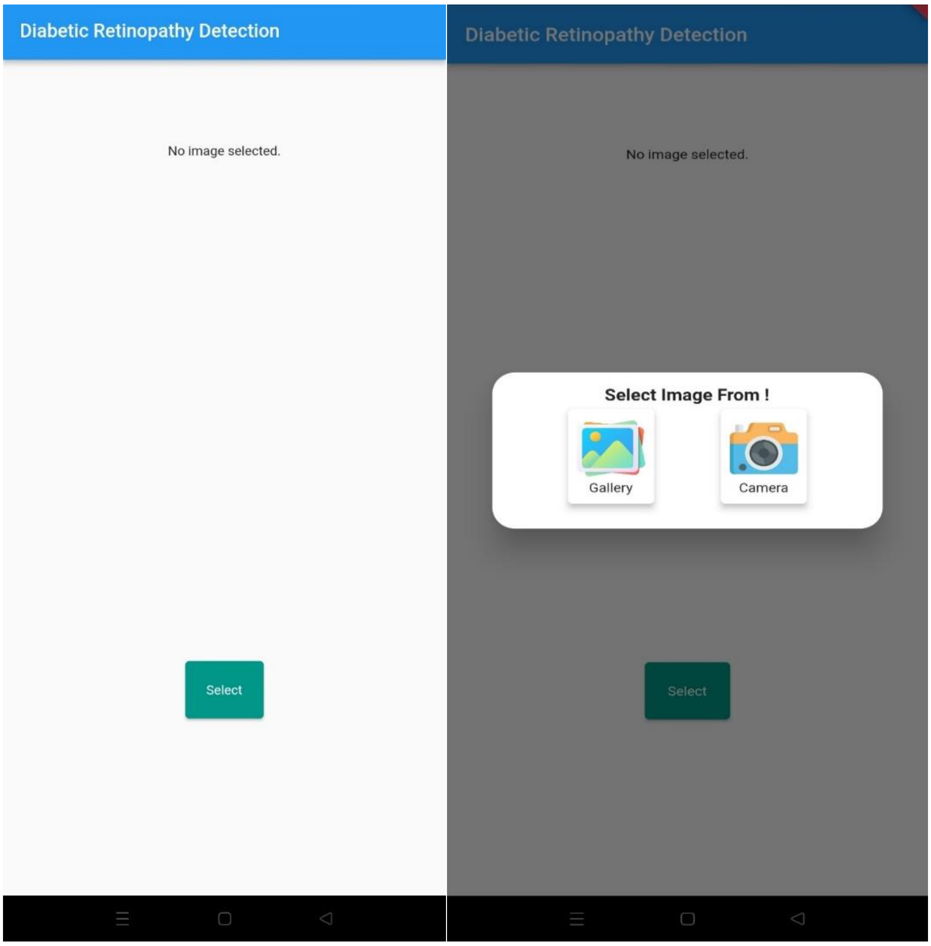


Figure 4.10: Application Home Screen

4.9.4.2 Classification Output Display

The selected image is passed through the trained model, which identifies whether it belongs to the retinal or non-retinal category. This initial classification step helps filter out irrelevant images, ensuring that only retinal images proceed to further analysis.

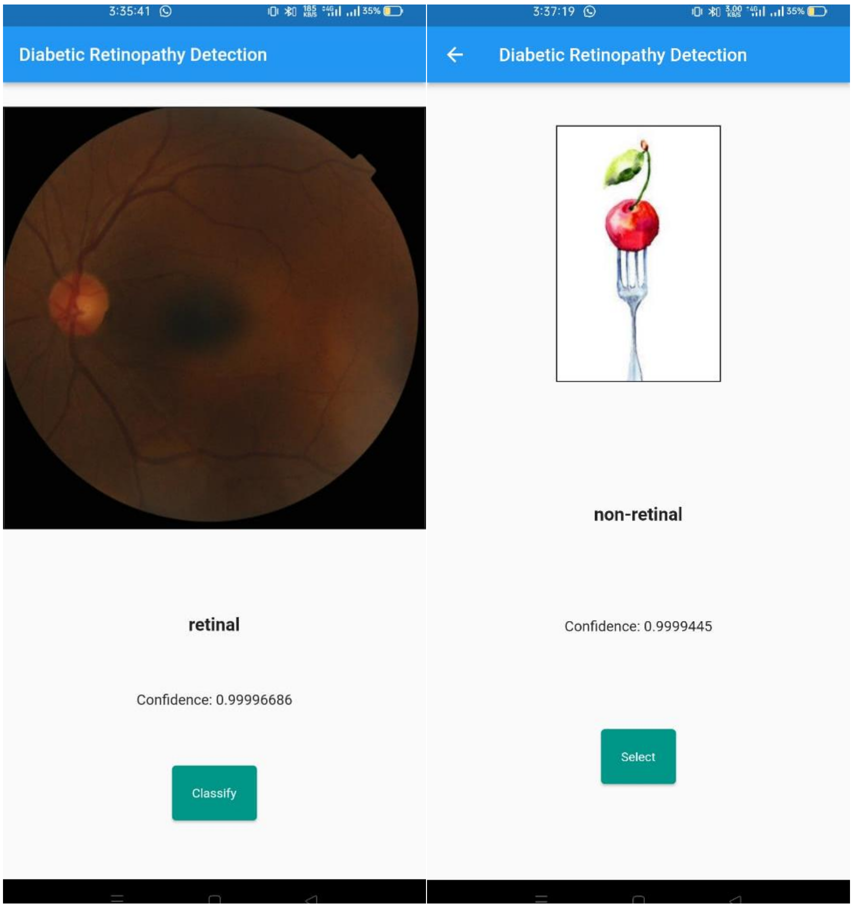


Figure 4.11: Retinal Image Classification

The presence of diabetic retinopathy is determined based on the second model deployed for classification process. This classification provides crucial information about the presence of the disease, facilitating timely medical intervention and treatment.

By leveraging the capabilities of trained model and the integration of the classification results into the app interface, users can conveniently and promptly assess the presence of diabetic retinopathy. This feature enhances the overall usability and effectiveness of the mobile application in supporting users' health management journey.

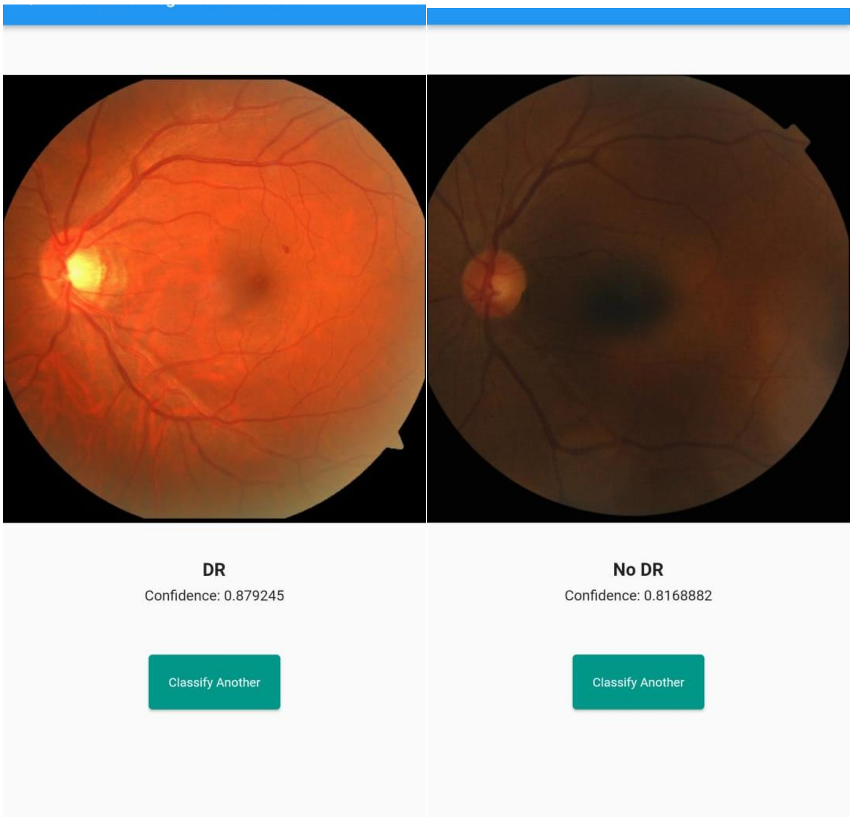


Figure 4.12: Disease Progression Level

4.9.5 Integration into a Web Application

After training both Network 1 and Network 2 models, the next step is to deploy and integrate them into a web app. The web app allows users to perform runtime inference on retinal and non-retinal images. The deployment involves setting up a server to host the models and handling user requests for image classification. The interface provides immediate feedback on the classification results, displaying whether the image is non-retinal or requiring further classification as referable or non-referable.

4.9.5.1 Web App Interface Design

The integration of the proposed method into a web-based classification application involves designing an intuitive and user-friendly interface using the React framework. The React framework allows for the development of web applications in JavaScript language.

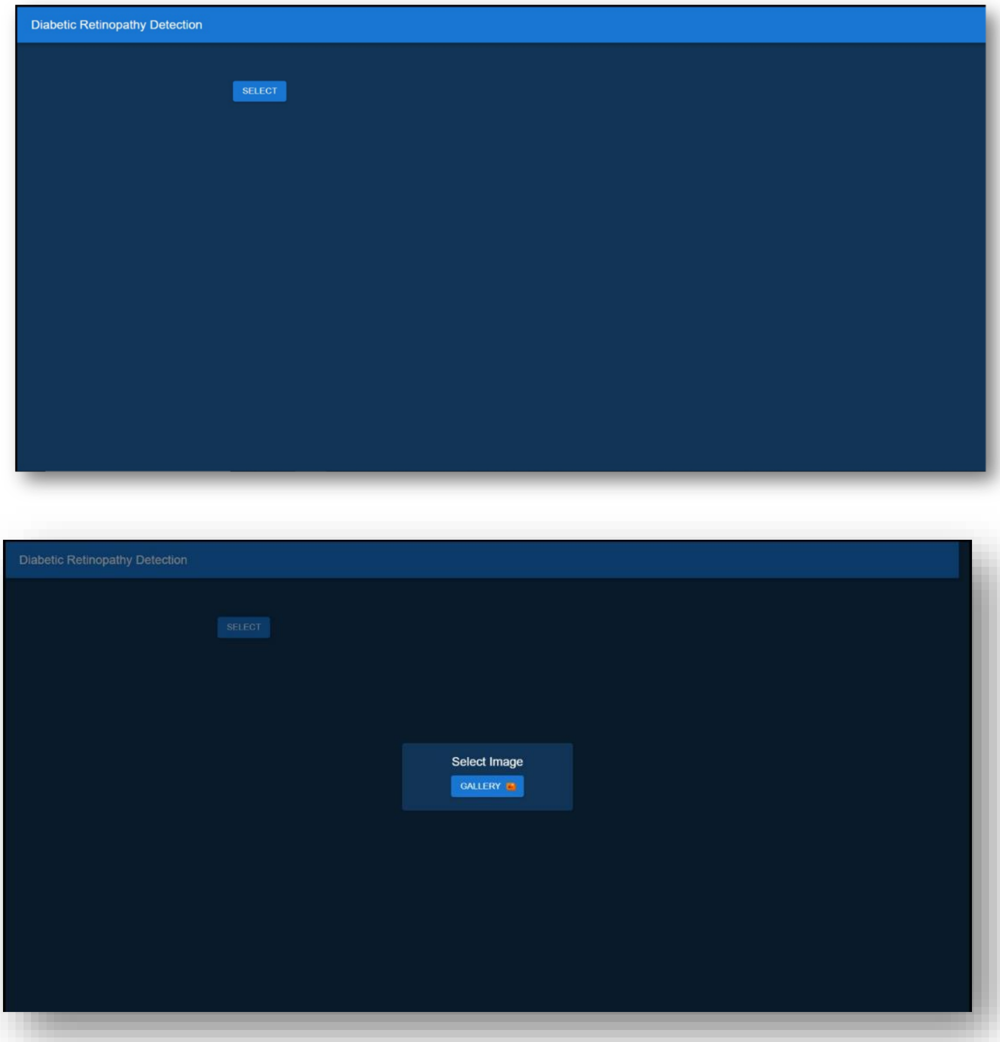


Figure 4.13: Application Home Screen

4.9.5.2 Classification Output Display

The selected image is passed through the trained model, deployed on server, using API, which identifies whether it belongs to the retinal or non-retinal category. This initial classification step helps filter out irrelevant images, ensuring that only retinal images proceed to further analysis.

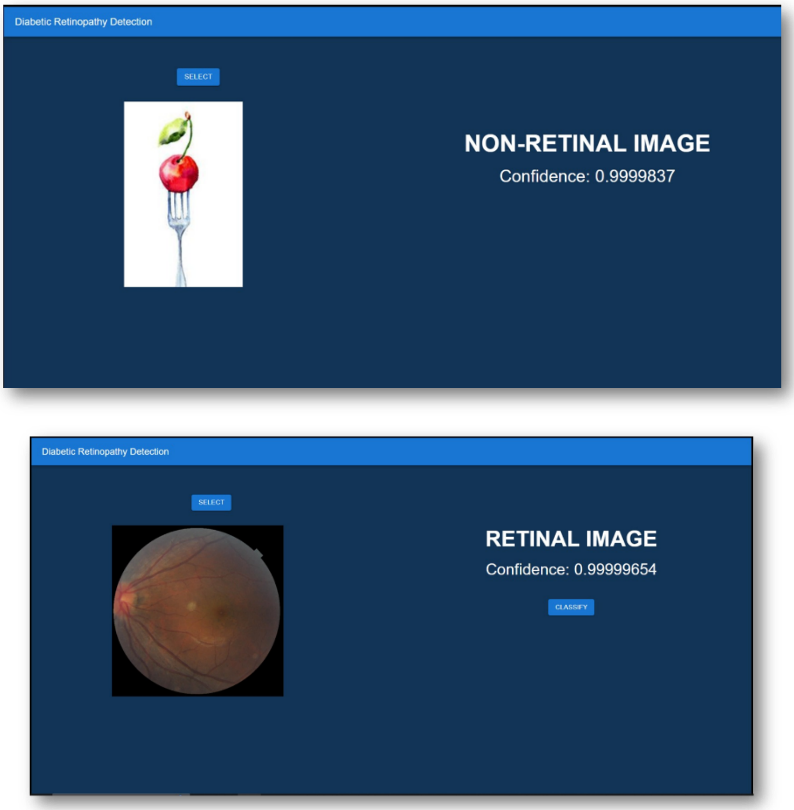
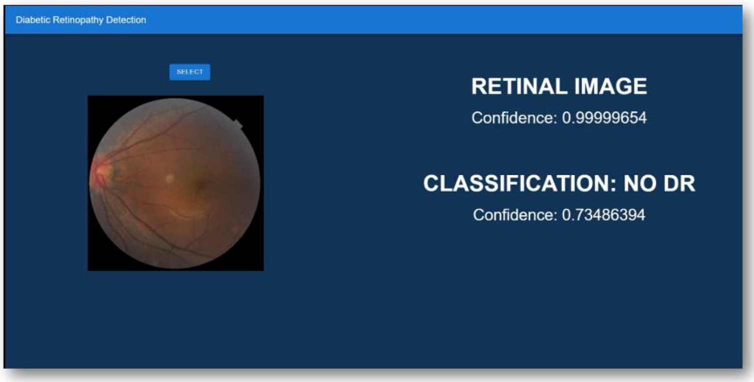


Figure 4.14: Retinal vs Non-Retinal Classification

The presence of diabetic retinopathy is determined based on the second network. This classification provides crucial information about the presence of the disease, facilitating timely medical intervention and treatment.

By leveraging the capabilities of trained models and the integration of the classification results into the app interface, users can conveniently and promptly assess the presence of diabetic retinopathy. This feature enhances the overall usability and effectiveness of the application in supporting users' health management journey.



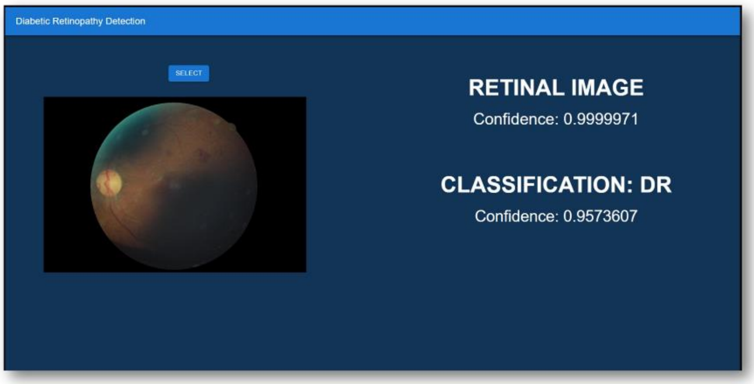


Figure 4.15: Referable vs Non-Referable Classification

4.10 Block Diagrams

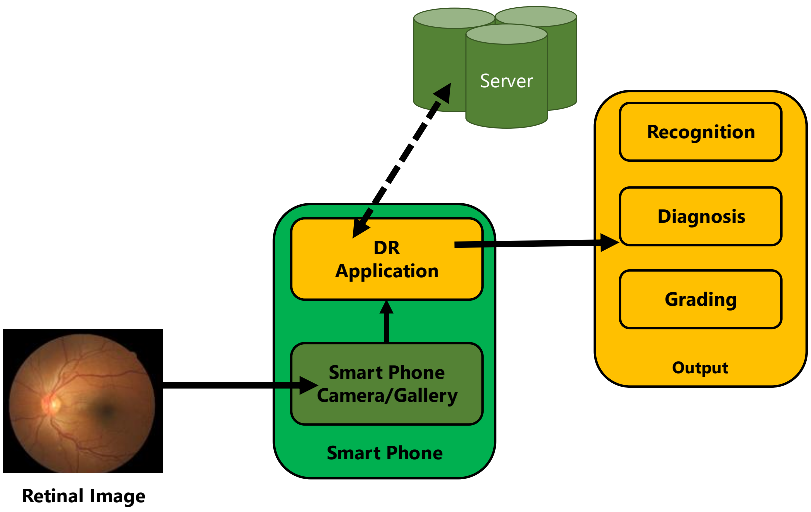


Figure 4.16: Block Diagram for Smartphone Application

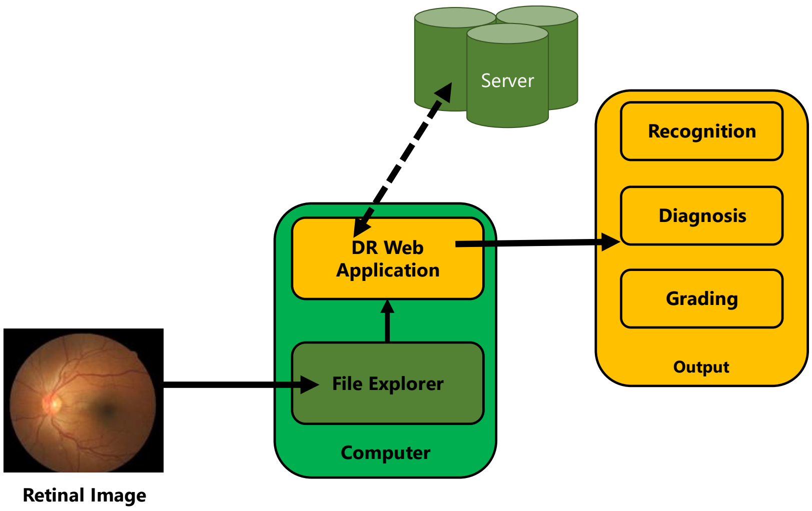


Figure 4.17: Block Diagram for Web Application

4.11 Flow Chart

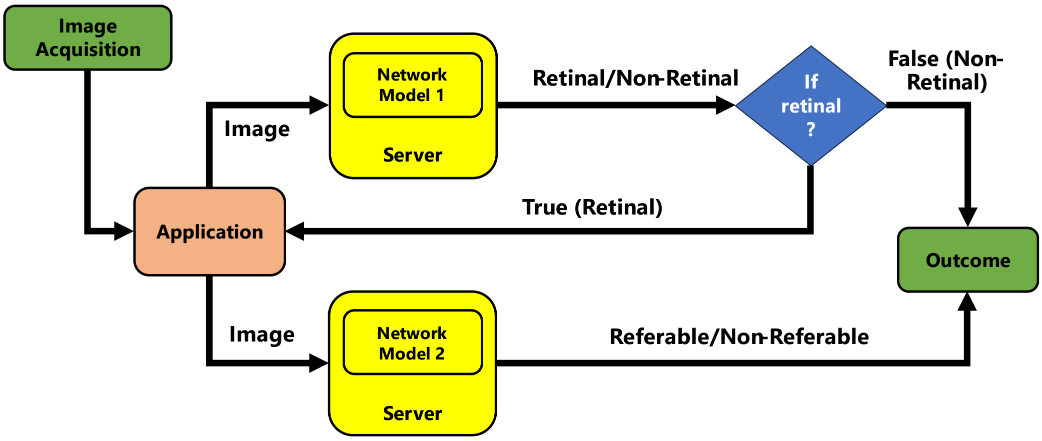


Figure 4.18: Flow Chart

Chapter 5:
5. Results and Conclusion

The dissertation aims to address the challenges faced by individuals with diabetes in accessing early diagnosis and treatment for diabetic retinopathy (DR). The installation and cost of fundus imaging devices, as well as the scarcity of ophthalmologists and healthcare centers, hinder the availability of timely diagnosis and intervention. To overcome these limitations, the proposed work explores the suitability of a smartphone-based system for detecting DR during routine health screenings.

5.1 Automatic DR Detection Algorithm Using Deep Learning

The project investigates the use of deep learning frameworks to develop automatic DR detection algorithms for retinal images captured on smartphones. Given the demand for an independent diagnostic tool that can be utilized by non-experts, various classifiers were proposed and implemented. The classification process follows 2 stages comprising the identification of retinal images in the first stage and the classification of disease prevalence in the second stage.

5.2 Integration into Smartphone and Web Application

The proposed method is integrated into smartphone and web applications, which offer an automatic system for the diagnosis of DR. This integration enables individuals with diabetes to conveniently screen and detect DR using their smartphones or they can just get to the browser, without the need for expensive equipment or expert assistance. By providing access to healthcare resources remotely, the burden of DR can be alleviated, and early intervention can be facilitated.

5.3 Contribution and Implications

This dissertation work contributes to the advancement of DR screening and detection using smartphones. By utilizing deep learning techniques and image processing algorithms, the solution proposed offers a cost-effective and accessible method to identify and classify DR. This has significant implications for improving healthcare access, especially in remote areas and for individuals lacking easy access to specialized healthcare centers or ophthalmologists.

5.4 Classification Results

Performance metrics play a vital role in assessing a machine learning model's ability to identify multiple classes of diabetic retinopathy. The widely used multi-class confusion matrix presents a table showing true positive, true negative, false positive, and false negative predictions for each class. From this matrix, performance metrics such as accuracy, recall, precision, and F1 score can be derived. It is crucial to evaluate the model using a combination of these metrics to understand its strengths and limitations.

For our project, we utilized precision, recall, F1-score, and accuracy as evaluation metrics. We employed simple ANN and MobileNetv2 in the classification process. These evaluation measures provide valuable insights into the accuracy, precision, recall, and overall effectiveness of the models in diabetic retinopathy classification.

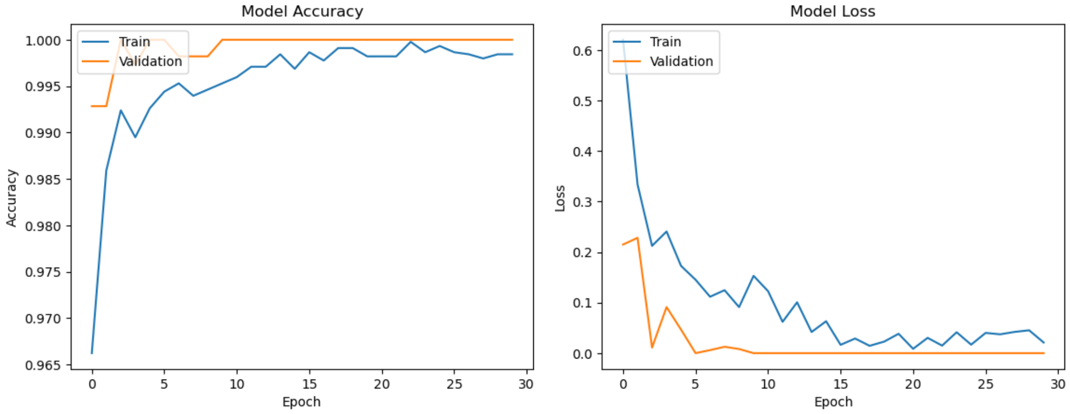


Figure 5.1: Graphs of accuracy and loss with Simple ANN

Table 5.1: Classification Report for Network 1 using ANN.

	Precision	Recall	F1-Score	Support
Non-Retinal Image	1.0	1.0	1.0	2340
Retinal image	1.0	1.0	1.0	2405
Accuracy			1.0	4745

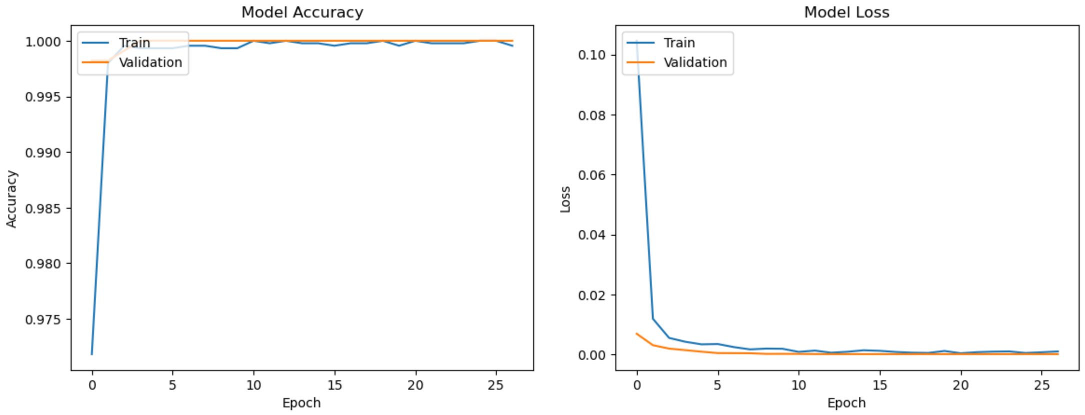


Figure 5.2: Graphs of accuracy and loss with MobileNetv2

Table 5.2: : Classification Report for Network 1 using MobileNetv2

	Precision	Recall	F1-Score	Support
Non-Retinal Image	1.0	1.0	1.0	2340
Retinal image	1.0	1.0	1.0	2405
Accuracy			1.0	4745

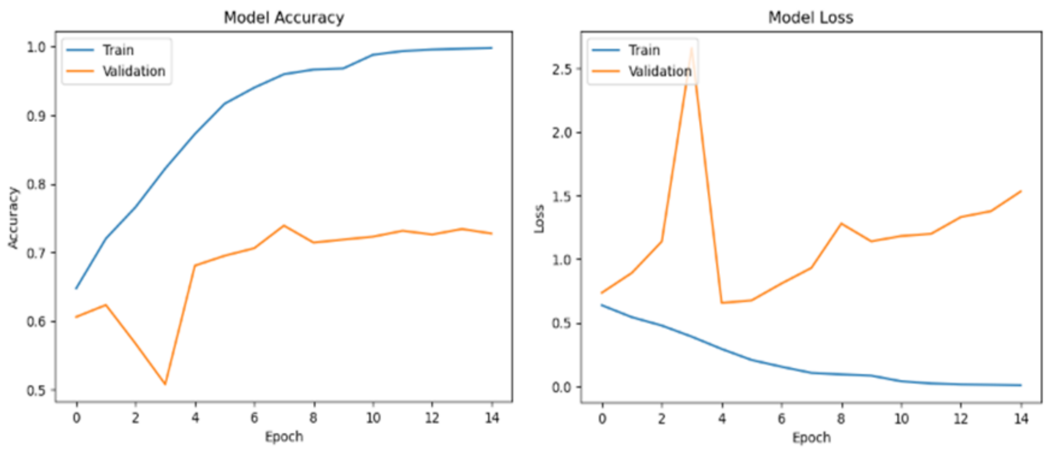


Figure 5.3: Graphs of accuracy and loss BEFORE Fine-Tuning

Table 5.3: Classification Report for Network 2 BEFORE Fine-tuning

	Precision	Recall	F1-Score	Support
Non-Referable	0.7569	0.5329	0.6254	2507
Referable	0.6393	0.8287	0.7218	2505
Accuracy			0.6807	5012

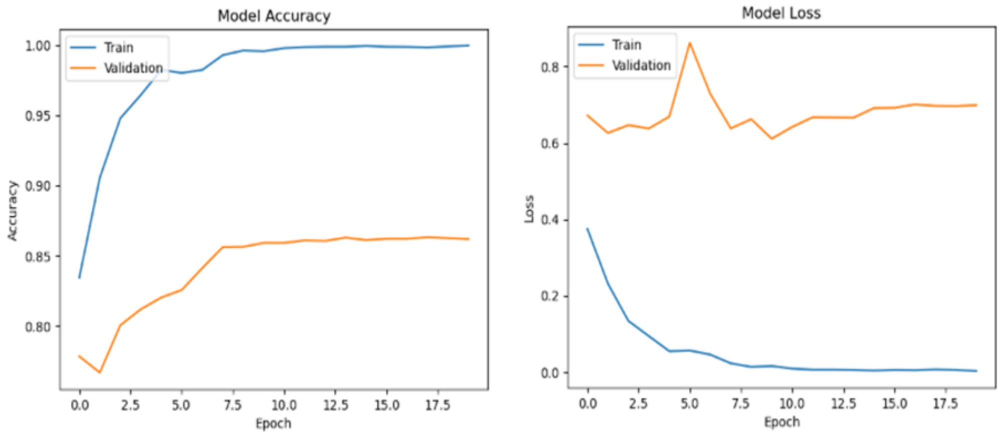


Figure 5.4: Graphs of accuracy and loss AFTER Fine-Tuning

Table 5.4: Classification Report for Network 2 AFTER Fine-tuning

	Precision	Recall	F1-Score	Support
Non-Referable	0.8304	0.9026	0.8650	2507
Referable	0.8933	0.8155	0.8526	2505
Accuracy			0.8591	5012

5.5 Future Work

The suggested solution brings notable progress in detecting DR using smartphones. However, there are opportunities for future development. Enhancing deep learning models by utilizing larger datasets and advanced network architectures can optimize their performance in classification. Incorporating more precise and severity classifications would allow for a detailed assessment of DR progression. Additionally, integrating telemedicine capabilities would enable remote consultations, improving the diagnostic and treatment process.

5.6 Limitations

Despite its benefits, the proposed solution has limitations. Factors like image quality, lighting conditions, and artifacts may impact the accuracy of DR detection. Moreover, relying on smartphone-based imaging may limit the system's ability to capture high-resolution images compared to specialized fundus imaging devices. It is important to acknowledge the limitations of the employed deep learning models, including potential biases in the training data and the need for continuous updates with new data. Additionally, the deployment of models on server may not work at times when server has a lot of load or traffic accessing the applications simultaneously.

6. References

[1] Azeem, S., Khan, U., & Liaquat, A. (2022). The increasing rate of diabetes in Pakistan: A silent killer. Annals of medicine and surgery (2012), 79, 103901. https://doi.org/10.1016/j.amsu.2022.103901

[2] Teo, Z. L., Tham, Y. C., Yu, M., Sabanayagam, C., Wong, T. Y., & Cheng, C. Y. (2021). Global prevalence of diabetic retinopathy and projection of burden through 2045: Systematic review and meta-analysis. Ophthalmology, 128(4), 716-728.https://doi.org/10.1016/j.ophtha.2021.04.027

[3] Leghari, Q., Akhtar, F., Aslam, S., Rafique, K., Majeed, S., Shanaz, S., Wajdi, M., Akram, A., & Bawani, S. (2022). Screening and prevalence study of diabetic retinopathy in Multan, district of Pakistan: A cross sectional study. Pak-Euro Journal of Medical and Life Sciences, 5(2), 429-434. https://doi.org/10.31580/pjmls.v5i2.2423

[4] Salz, D. A., & Witkin, A. J. (2015). Imaging in diabetic retinopathy. Middle East African Journal of Ophthalmology, 22(2), 145.

[5] Mohamed Akil, Yaroub Elloumi. Detection of retinal abnormalities using Smartphone-captured fundus images: A survey. Real-Time Image Processing and Deep Learning, Apr 2019, Baltimore, Maryland,, United States. ⟨10.1117/12.2519094⟩. ⟨hal-02121855⟩

[6] He, A., Li, T., Li, N., Wang, K., & Fu, H. (2021). CABNet: Category Attention Block for Imbalanced Diabetic Retinopathy Grading. IEEE transactions on medical imaging, 40(1), 143–153. https://doi.org/10.1109/TMI.2020.3023463

[7] Russo A., Mapham W., Turano R., Costagliola C., Morescalchi F., Scaroni N., and Semeraro F., “Comparison of Smartphone Ophthalmoscopy with SlitLamp Biomicroscopy for Grading Vertical Cup-to-Disc Ratio”, J Glaucoma, Volume 25, Number 9, September 2016.doi: 10.1097/IJG.0000000000000499

[8] Elloumi Y., Akil M. and Kehtarnavaz N., “A mobile computer aided system for optic nerve head detection”, Computer Methods and Programs in Biomedicine (CMPB), Vol. 162, August 2018, pages 139-148, DOI: https://doi.org/10.1016/j.cmpb.2018.05.004

[9] Fenner B.J., Raymond L. M. , Lam W.C., Tan G. S. W., Cheung G. C. M., “Advances in Retinal Imaging and Applications in Diabetic Retinopathy Screening: A Review”, Ophthalmol. Ther.. DOI: https://doi.org/10.1007/s40123-018-0153-7

[10] Rajalakshmi R., Arulmalar S., Usha M., Prathiba V., Kareemuddin K.S., Anjana R.M. and Mohan V., “Validation of Smartphone Based Retinal Photography for Diabetic Retinopathy Screening”, PLoS ONE 10(9): e0138285. doi: 10.1371/journal.pone.0138285Cham.

[11] Toy B.C., Myung D.J., He L., Pan C. K., Chang R. T., Polkinhorne A., Merrell D., Foster D., and Blumenkranz M. S., “Smartphone-Based Dilated fundus Photography And Near Visual Acuity testing as Inexpensive Screening Tools to Detect Referral Warranted

Diabetic Eye Disease", Retina,36(5):1000-8, 2016.doi:10.1097/IAE.0000000000000955.

[12] Russo A., Mapham W., Turano R., Costagliola C., Morescalchi F., Scaroni N., and Semeraro F., "Comparison of Smartphone Ophthalmoscopy With SlitLamp Biomicroscopy for Grading Vertical Cup-to-Disc Ratio", J Glaucoma, Volume 25, Number 9, September 2016.doi: 10.1097/IJG.0000000000000499.

[13] Ryan M. E., Rajalakshmi R., Prathiba V., Anjana R. M., Ranjani H., Narayan K.M.V., Olsen T.W., Mohan V., Ward L.A., Lynn M. J. and Hendrick A. M., "Comparison Among Methods of Retinopathy Assessment (CAMRA) Study", Ophthalmology, 122(10):2038-43, 2015. DOI: 10.1016/j.ophtha.2015.06.011.

[14] Rajalakshmi R., Arulmalar S., Usha M., Prathiba V., Kareemuddin K.S., Anjana R.M. and Mohan V., "Validation of Smartphone Based Retinal Photography for Diabetic Retinopathy Screening", PLoS ONE 10(9): e0138285. doi: 10.1371/journal.pone.0138285

[15] Micheletti J. M., Hendrick A. M., Khan F. N., Ziemer D. C. and Pasquel F. J., "Current and Next Generation Portable Screening Devices for Diabetic Retinopathy", Journal of Diabetes Science and Technology 2016, Vol. 10(2) 295–300. doi: 10.1177/1932296816629158.

[16] Muiesan M. L., Salvetti M., Paini A., Riviera M., Pintossi C., Bertacchini F., Colonetti E., Agabiti-Rosei C., Poli M., Semeraro F., Agabiti-Rosei E., and Russo A., "Ocular fundus photography with a smartphone device in acute hypertension", Journal of Hypertension, 35:1660–1665, Volume 35, Number 8, August 2017.doi: 10.1097/HJH.0000000000001354.

[17] Thomas S.M., Jeyaraman M.M., Hodge W.G., Hutnik C., Costella J., Malvankar-Mehta M.S., "The effectiveness of teleglaucoma versus inpatient examination for glaucoma screening: a systematic review and meta-analysis", PLoS One 2014;9(12):e113779.

[18] Wang S., Jin K., Lu H., Cheng C., Ye J. and Qian D., "Human visual system-based fundus image quality assessment of portable fundus camera photographs", IEEE Trans. on Medical Imaging, 35:1046 – 1055, April 2016. DOI: 10.1109/TMI.2015.2506902.

[19] MichelettiJ. M., Hendrick A. M., Khan F. N., Ziemer D. C. and Pasquel F. J., "Current and Next Generation Portable Screening Devices for Diabetic Retinopathy", Journal of Diabetes Science and Technology 2016, Vol. 10(2) 295–300. doi: 10.1177/1932296816629158.

[20] Koukounis D., Ttofis C., Papadopoulos A. and Theocharides T., "A high performance hardware architecture for portable, low power retinal vessel segmentation", INTEGRATION journal, Volume 47, Issue 3, June 2014, Pages 377-386, 2014. DOI: https://doi.org/10.1016/j.vlsi.2013.11.005.

[21] Ramachandran Rajalakshm, Radhakrishnan Subashini, Ranjit Mohan Anjana, Viswanathan Mohan, "Automated diabetic retinopathy detection in smartphone-based

fundus photography using artificial intelligence", Eye, 32, pages1138–1144 (2018). DOI: https://doi.org/10.1038/s41433-018-0064-9.

[22] Suriyal S., Druzgalski C., Gautam K., "Mobile Assisted Diabetic Retinopathy Detection using Deep Neural Network", 2018 Global Medical Engineering physics exchanges (GMEPE), Porto, Portugal, March, 2018. DOI: 10.1109/GMEPE PAHCE.2018.8400760.

[23] Yin F., Lee B.H., Yow A.P., Quan Y. and Wong D.W.K, "Automatic Ocular Disease Screening and Monitoring Using a Hybrid Cloud System", 2016 IEEE International Conference on Internet of Things (iThings), pp 2663-268, 2017. DOI: 10.1109/iThings-GreenCom-CPSCom-SmartData.2016.68.

[24] Jin K., Lu H., Su Z., Cheng C., Ye J. and Qian D., "Telemedicine screening of retinal diseases with a handheld portable non mydriatic fundus camera", BMC Ophthalmology,17:89, 2017. DOI: 10.1186/s12886-017-0484-5.

[25] Karakaya, M., & Hacisoftaoglu, R. E. (2020). Comparison of smartphone-based retinal imaging systems for diabetic retinopathy detection using deep learning. BMC bioinformatics, 21(Suppl 4), 259. https://doi.org/10.1186/s12859-020-03587-2

[26] Kim, T.N., Aaberg, M.T., Li, P. et al. Comparison of automated and expert human grading of diabetic retinopathy using smartphone-based retinal photography. Eye 35, 334–342 (2021). https://doi.org/10.1038/s41433-020-0849-5

[27] Welcome to Colaboratory. [online] Available at: http://colab.research.google.com/

[28] Carneiro, T., Da Nóbrega, R.V.M., Nepomuceno, T., Bian, G.B., De Albuquerque, V.H.C. and Reboucas Filho, P.P., 2018. Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. IEEE Access, 6, pp.61677-61685.

[29] Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., Erginay, A. and Charton, B., 2014. Feedback on a publicly distributed image database: the Messidor database. Image Analysis & Stereology, 33(3), pp.231-234.

[30] E-ophtha. [online] Available at: http://www.adcis.net/en/Download-Third-Party/E-Ophtha.html

[31] Kaggle. (2015). Diabetic Retinopathy Detection. [online] Available at: https://www.kaggle.com/c/diabetic-retinopathy-detection

[32] Drive dataset. [online] Available at: https://www.isi.uu.nl/Research/Databases/DRIVE/

[33] Hoover, A., Kouznetsova, V. and Goldbaum, M., 1998. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. In Proceedings of the AMIA Symposium (p. 931). American Medical Informatics Association

[34] Kälviäinen, R.V.J.P.H. and Uusitalo, H., 2007. DIARETDB1 diabetic retinopathy database and evaluation protocol. In Medical Image Understanding and Analysis (Vol. 2007, p. 61).

[35] Diaretdb1 dataset. [online] Available at: http://www.it.lut.fi/project/imageret/diaretdb1/

[36] Chase dataset. [online] Available at: http://www.chasestudy.ac.uk/

[37] Prentašić, P., Lončarić, S., Vatavuk, Z., Benčić, G., Subašić, M., Petković, T., Dujmović, L., Malenica-Ravlić, M., Budimlija, N. and Tadić, R., 2013, September. Diabetic retinopathy image database (DRiDB): a new database for diabetic retinopathy screening programs research. In 2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA) (pp. 711-716). IEEE.

[38] Zhang, Z., Yin, F.S., Liu, J., Wong, W.K., Tan, N.M., Lee, B.H., Cheng, J. and Wong, T.Y., 2010, August. Origa-light: An online retinal fundus image database for glaucoma analysis and research. In 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology (pp. 3065-3068). IEEE.

[39] Sng, C.C., Foo, L.L., Cheng, C.Y., Allen Jr, J.C., He, M., Krishnaswamy, G., Nongpiur, M.E., Friedman, D.S., Wong, T.Y. and Aung, T., 2012. Determinants of anterior chamber depth: the Singapore Chinese Eye Study. Ophthalmology, 119(6), pp.1143-1150.

[40] Nih areds dataset. [online] Available at: https://www.nih.gov/news-events/news-releases/nih-adds-first-images-major-research-database

[41] Al-Diri, B., Hunter, A., Steel, D., Habib, M., Hudaib, T. and Berry, S., 2008, August. A reference data set for retinal vessel profiles. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 2262-2265). IEEE.

[42] Eyepacs dataset. [online] Available at: http://www.eyepacs.com/eyepacssystem/

[43] Fumero, F., Alayón, S., Sanchez, J.L., Sigut, J. and Gonzalez-Hernandez, M., 2011, June. RIM-ONE: An open retinal image database for optic nerve evaluation. In 2011 24th international symposium on computer-based medical systems (CBMS) (pp. 1-6). IEEE.

[44] Sivaswamy, J., Krishnadas, S.R., Joshi, G.D., Jain, M. and Tabish, A.U.S., 2014, April. Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In 2014 IEEE 11th international symposium on biomedical imaging (ISBI) (pp. 53-56). IEEE.

[45] Aria dataset. [online] Available at: http://www.eyecharity.com/aria_online.html.

[46] drion dataset. [online] Available at: http://www.ia.uned.es/~ejcarmona/DRIONS-DB.html

[47] drion dataset. [online] Available at: http://www.ia.uned.es/~ejcarmona/DRIONS-DB.html

[48] Seed-db. [online] Available at: https://www.seri.com.sg/key-programmes/singapore- epidemiology-of-eye-diseases-seed/

[49] Wang S, Zheng J, Hu HM, Li B. Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Trans Image Process. 2013 Sep;22(9):3538-48. doi: 10.1109/TIP.2013.2261309. Epub 2013 May 2. PMID: 23661319.

[50] LeCun, Y., Boser, B.E., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W.E. and Jackel, L.D., 1990. Handwritten digit recognition with a back-propagation network. In Advances in neural information processing systems (pp. 396-404).

[51] Miotto, R., Wang, F., Wang, S., Jiang, X. and Dudley, J.T., 2017. Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics, 19(6), pp.1236-1246.

[52] Razzak, M.I., Naz, S. and Zaib, A., 2018. Deep learning for medical image processing: Overview, challenges and the future. In Classification in BioApps (pp. 323-350). Springer, Cham.

[53] Erhan, D., Bengio, Y., Courville, A., Manzagol, P.A., Vincent, P. and Bengio, S., 2010. Why does unsupervised pre-training help deep learning? Journal of Machine Learning Research, 11(Feb), pp.625-660.

[54] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. and Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis, 42, pp.60-88.

[55] Altaf, F., Islam, S., Akhtar, N. and Janjua, N.K., 2019. Going Deep in Medical Image Analysis: Concepts, Methods, Challenges and Future Directions. arXiv preprint arXiv:1902.05655.

[56] B. Graham, “Kaggle Diabetic Retinopathy Detection competition report,” pp. 1–9, 2015.