Deep Learning Model for early detection of diabetic retinopathy from Retinal Images in Type 2 Diabetes

Ezra, Namunyu Simiyu

ID: 151527

A research thesis submitted in partial fulfillment of the requirements of the

Degree of Master of Science in Information Technology at Strathmore

University

School of computing and Engineering Sciences

Strathmore University

Nairobi, Kenya

April 2024

Declaration and Approval

I affirm that this research proposal is my original work and has not been submitted for any other degree or qualification. To the best of my knowledge, the proposal does not contain any material that has been previously published or written by another person, unless I have properly cited the source.

Student's Name: Ezra Namunyu Simiyu

Sign: Date: April 10, 2024

Supervisor's Name: Dr Joseph Orero, PhD

Sign: --------------------------Date: April 10, 2024

Abstract

Type 2 diabetes is a chronic condition characterized by elevated blood sugar levels, leading to various complications, including diabetic retinopathy—a major cause of vision loss if left untreated. Traditional methods of detecting diabetic retinopathy involve clinical examination techniques and specialized imaging modalities, such as dilated eye exams, visual acuity tests, fundus photography, fluorescein angiography, and optical coherence tomography (OCT). While effective, these methods can be invasive, time-consuming, and require specialized equipment and expertise.

This study aims to explore a non-invasive and potentially more accessible approach to detecting diabetic retinopathy by leveraging machine learning techniques. The study proposes utilizing retinal images captured by smartphone-based retinal imaging devices and applying a machine learning model for detecting diabetic retinopathy. The proposed model leverages the feature extraction and classification capabilities of a pre-trained EfficientNetB0 convolutional neural network (CNN). By analyzing the features extracted from these images, the model aims to offer a more convenient and scalable solution for early detection and monitoring of diabetic retinopathy.

this research has demonstrated the feasibility of using machine learning algorithms for diabetic retinopathy detection and providing actionable insights for individuals living with type 2 diabetes. These will facilitate early intervention strategies, such as lifestyle modifications or timely medical interventions, to mitigate the risk of diabetic retinopathy progression and associated vision loss.

The developed deep learning model, based on the EfficientNetB0 architecture, exhibits promising results in diabetic retinopathy detection. It accurately classifies retinal images from smartphone-based devices with an impressive 88.57% accuracy and 90.31% precision. The model's reliance on convolutional neural networks and the efficient EfficientNetB0 design allow for sensitive and specific identification of diabetic retinopathy across various severity levels. This advancement has the potential to significantly enhance early detection and management of the condition, leading to improved patient outcomes and better-informed clinical decision-making in the field of diabetic eye care.

Key words: Type 2 diabetes, High sugar levels, Deep learning, CNN, EfficientNetB0, Retinal Images, Forecasting

Table of Contents

Table of figures

List of abbreviations

T2DM: Type 2 Diabetes Mellitus

DR: Diabetic Retinopathy

CNN: Convolutional Neural Network

ML: Machine Learning

AI: Artificial Intelligence

OCT: Optical Coherence Tomography

RMSE: Root Mean Square Error

SVM: Support Vector Machine

DR: Diabetic Retinopathy

AUC: Area Under the Curve

IDP: Iowa Detection Program

1 Chapter 1: Introduction

1.1 Background of the study

Banday, et. al (2020) states that diabetes mellitus is a chronic metabolic condition characterized by complicated pathophysiology. It is associated with increased blood glucose levels, or hyperglycemia, caused by defects in either the production of insulin or functioning of insulin, or a combination of the two (Centers for Disease Control and Prevention 2023). This has posed a public health concern because of its increasing prevalence and the possibility of serious complications (International Diabetes Federation, 2023). Diabetes is classified into Type 1 diabetes and Type 2 diabetes (Centers for Disease Control and Prevention 2023).

Type 1 diabetes is a condition where the immune system mistakenly attacks and destroys beta cells in the pancreas that produce insulin (American Diabetes Association, 2023). People with Type 1 diabetes require insulin therapy. Thet are usually diagnosed during childhood or early adulthood. This variation accounts for 5% to 10% of all cases of diabetes (International Diabetes Federation, 2023).

The second variation is Type 2 diabetes which is characterized by insulin resistance, which means that cells do not respond effectively to insulin (American Diabetes Association, 2023). It is primarily associated with lifestyle factors such as diet, sedentary behavior, obesity, and genetic predisposition (Centers, for Disease Control and Prevention 2023). According to the International Diabetes Federation, Type 2 diabetes worldwide accounts for more than 90% of all diabetic cases.

For diabetic patients, a common complication is diabetic retinopathy, characterized by damage to the blood vessels in the retina due to prolonged high blood sugar levels (National Eye Institute, 2022). This damage arises from the complications of diabetes, such as insufficient insulin production, ineffective insulin usage, or insulin resistance (American Diabetes Association, 2023). Diabetic retinopathy often has no noticeable symptoms in its early stages, but may eventually include blurred vision, floaters, or vision loss (National Eye Institute, 2022). It's essential for people with diabetes to get regular eye exams to catch diabetic retinopathy early (Centers for Disease Control and Prevention, 2021). If left untreated, diabetic retinopathy can lead to blindness (National Eye Institute, 2022). While not a direct cause, diabetic retinopathy is often associated with an increased risk of serious health conditions like heart disease, kidney problems, and neuropathy, due to the underlying disease of diabetes (American Diabetes Association, 2023).

The National Eye Institute also highlights that retinopathy affects the blood vessels in the retina tissues located at the back of the eye. Prolonged hyperglycemia can cause harm to these blood vessels, resulting in diabetic retinopathy which is one of the leading causes of vision loss, in adults (American Diabetes Association, 2023). Subtle changes may appear in these blood vessels including microaneurysms (bulges) hemorrhages (bleeding) and abnormal dilation (widening) of vessels (National Eye Institute) which can be signs that one may experience retinopathy during the day. Retinal imaging, which can be capture through fundus photography or optical coherence tomography (OCT) can capture the changes in the blood vessels with great detail (American Diabetes Association, 2023).

According to Bhattacharya (2023) deep learning, a branch of artificial intelligence (AI), has exhibited as a promising tool in predicting blood sugar levels in people with Type 2 diabetes by analyzing retinal images. From a recent study by Yun et al. (2021) on deep learning models for screening Type 2 diabetes from retinal photographs, it was found that Deep neural networks such as convolutional neural networks (CNNs) have proven to be highly effective in analyzing large datasets and identifying these patterns. The strength of learning lies in its ability to detect changes in the retinal blood vessels that are associated with high blood sugar levels. Tsiknakis, N., et al. (2021) explain that training a learning model on a dataset of images enables it to recognize early signs of retinal changes linked to high blood sugar levels. The efficiency of this approach comes from its ability to process a vast amount of data, quickly adapt to information, and provide accurate predictions. This makes it an invaluable tool for transforming diabetic retinopathy management and facilitating intervention to prevent complications (Tsiknakis, N., et al. 2021).

Therefore, Retinal images offer insights into predicting blood sugar levels throughout the day for individuals with Type 2 diabetes. When used in conjunction with deep learning methods these pictures help create models that provide a non-intrusive early and easily accessible way to monitor blood sugar levels. This could potentially empower people to manage their diabetic efficiently to reduce changes of developing complications like diabetic retinopathy.

1.2 Problem Statement

Diabetic retinopathy is a progressive and potentially sight-threatening complication of type 2 diabetes, affecting a significant proportion of individuals worldwide (Yau et al., 2021). Early detection and intervention are crucial for preventing vision loss and reducing the burden of diabetic retinopathy on healthcare systems and patients alike (Sahlgrenska University Hospital, 2023). However, traditional methods of detecting diabetic retinopathy, such as dilated eye exams and fundus photography, are often time-consuming, invasive, and reliant on specialized equipment and expertise (Khan et al., 2022). Moreover, these methods may not be readily accessible to all individuals with type 2 diabetes, particularly those in resource-constrained settings or remote areas (Bastawrous & Armstrong, 2023).

Recent advancements in technology, particularly the development of smartphone-based retinal imaging devices, offer the potential for non-invasive and scalable solutions for diabetic retinopathy screening and early detection (Rajalakshmi et al., 2022). These devices enable the capture of high-quality retinal images using smartphones, making retinal screening more accessible and convenient for patients (Bastawrous & Armstrong, 2023). However, the interpretation of these images for diabetic retinopathy detection still largely relies on manual assessment by ophthalmologists or trained personnel, which can be subjective, time-intensive, and prone to interobserver variability (Khan et al., 2022).

Therefore, there is a pressing need for automated and efficient methods for early detection of diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices in individuals with type 2 diabetes (Rajalakshmi et al., 2022). Leveraging deep learning techniques, such as convolutional neural networks (CNNs), holds promise for developing accurate and reliable diagnostic models that can analyze retinal images and detect early signs of diabetic retinopathy with high sensitivity and specificity (Khan et al., 2022). By addressing this need, such a deep learning model has the potential to revolutionize diabetic retinopathy screening, enabling timely interventions and improving visual outcomes for individuals with type 2 diabetes.

1.3 Main Objective

The main objective of this study is to develop and evaluate a deep learning model for the early detection of diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices in individuals with type 2 diabetes.

1.4 Specific Objectives

i. To investigate the limitations of current methods for detecting diabetic retinopathy.
ii. To review existing approaches for predicting the early detection of diabetic retinopathy.
iii. To develop and train a deep learning model for the analysis of retinal images.
iv. To assess and validate the performance of the proposed deep learning model.

1.5 Research Questions

i. What are the main limitations of current methods for detecting diabetic retinopathy in type 2 diabetes patients?

ii. What non-invasive and deep learning approaches are available for early prediction of diabetic retinopathy, and what are their comparative advantages and disadvantages?

iii. How can a deep learning model be optimized to analyze retinal images from smartphone-based devices for early detection of diabetic retinopathy in type 2 diabetes?

iv. What features should be integrated into a user-friendly mobile application interface to effectively display diagnostic findings derived from the deep learning model?

v. How can the performance and accuracy of the proposed deep learning model be validated and assessed in clinical settings?

1.6 Justification

The development of a deep learning model for early detection of diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices in individuals with type 2 diabetes is very critical. One of the main reasons is that, diabetic retinopathy stands as a leading cause of vision loss among individuals with type 2 diabetes, necessitating improved screening and early detection methods to prevent irreversible vision impairment.

Traditional methods for detecting diabetic retinopathy often involve invasive and time-consuming procedures, limiting their accessibility and scalability, particularly in resource-limited settings. By leveraging smartphone-based retinal imaging devices, which are increasingly accessible and portable, we can overcome these barriers and facilitate widespread screening for diabetic retinopathy.

Deep learning techniques offer the potential to enhance the efficiency and accuracy of diabetic retinopathy detection by automatically analyzing retinal images for early signs of pathology. By training a deep learning model on a large dataset of retinal images, we can teach it to recognize subtle changes indicative of diabetic retinopathy, enabling earlier intervention and better management of the condition.

Lastly, integration of a deep learning model with a user-friendly mobile application interface allows for seamless and intuitive display of diagnostic findings, empowering healthcare professionals and patients to make informed decisions regarding treatment and follow-up care.

Looking at the above reasons, it is evident that development of a deep learning model for early detection of diabetic retinopathy from smartphone-captured retinal images has the potential to revolutionize diabetic retinopathy screening, improving accessibility, efficiency, and accuracy while ultimately reducing the burden of vision loss associated with this debilitating complication of type 2 diabetes.

1.7 Scope and Limitations

This study focuses on the development and evaluation of a deep learning model aimed at early detection of diabetic retinopathy using retinal images obtained from smartphone-based retinal imaging devices in individuals diagnosed with type 2 diabetes. It entails a comprehensive review of existing literature pertaining to diabetic retinopathy detection methods, including both conventional techniques and advanced deep learning approaches. The research involves the collection of a diverse dataset comprising retinal images. Through the utilization of convolutional neural networks (CNNs), the study aims to construct and refine a deep learning model capable of analyzing retinal images to predict the presence of diabetic retinopathy. Furthermore, integration of this model with a user-friendly mobile application interface will be undertaken to facilitate the presentation of diagnostic findings. The performance of the developed deep learning model will be rigorously evaluated through validation procedures, including comparison with established methods, utilizing independent datasets.

This research is subject to certain limitations which may may affect the outcomes and applicability of this study. Firstly, the availability and quality of retinal images may vary, potentially impacting the performance of the deep learning model. Moreover, the dataset utilized for model training and evaluation may not fully represent the diversity of retinal pathology and clinical scenarios encountered in real-world practice due to factors like age, gender, ethnicity, and the presence of comorbidities.

Additionally, computational resource requirements for training and inference of the deep learning model may be substantial, posing challenges for deployment in resource-constrained settings. Furthermore, the effectiveness of the developed model may be influenced by factors such as patient demographics, comorbidities, and variations in the performance of smartphone-based retinal imaging devices.

Ethical considerations related to patient data privacy and informed consent may also restrict access to certain data sources. It is also important to note that resource constraints, including access to specialized hardware and computing power for deep learning, could influence scalability and efficiency.

Effective interdisciplinary collaboration between medical and technological domains is crucial, and any limitations in communication and collaboration may influence research outcomes. The research also acknowledges external factors, such as evolving healthcare regulations or technological advancements, which can impact the research's relevance. It is important to note that this study may not address all aspects of diabetic retinopathy detection and management, necessitating further research to address specific clinical needs and challenges.

2 Chapter 2: Literature review

2.1 Introduction

The aim of this chapter is to conduct a comprehensive review of the existing literature related to early detection diabetic retinopathy from retinal images in individuals with Type 2 diabetes, with a specific focus on leveraging deep learning techniques. The literature review encompasses various subtopics, each contributing to a deeper understanding of the research landscape surrounding this critical healthcare challenge.

2.2 Diabetic Retinopathy in Type 2 Diabetes

Diabetic retinopathy stands as a significant microvascular complication of type 2 diabetes, often leading to vision impairment and blindness if left unaddressed (Yau et al., 2021). Understanding its clinical manifestations is paramount for effective management and prevention of its progression (Sahlgrenska University Hospital, 2023). In type 2 diabetes, prolonged exposure to elevated blood glucose levels triggers a cascade of pathological changes within the retinal microvasculature, culminating in diabetic retinopathy (Cheung et al., 2010). This condition typically manifests in stages, beginning with mild nonproliferative diabetic retinopathy (NPDR), characterized by microaneurysms, retinal hemorrhages, and cotton-wool spots (Khan et al., 2022). As the disease progresses, moderate and severe NPDR may ensue, accompanied by intraretinal microvascular abnormalities and venous beading (American Diabetes Association, 2023). In advanced stages, proliferative diabetic retinopathy (PDR) emerges, marked by neovascularization and potential complications such as vitreous hemorrhage and tractional retinal detachment (Khan et al., 2022).

Effective management of diabetic retinopathy necessitates a multidisciplinary approach, encompassing glycemic control, blood pressure management, and regular ophthalmologic screenings (American Diabetes Association, 2023). Early detection and intervention play pivotal roles in preventing irreversible vision loss, underscoring the importance of routine retinal examinations for individuals with type 2 diabetes (Yau et al., 2021; Khan et al., 2022). Timely identification of diabetic retinopathy enables prompt implementation of sight-saving interventions, including laser photocoagulation, intravitreal injections, and vitrectomy surgery (Khan et al., 2022). Furthermore, lifestyle modifications, such as smoking cessation and dietary adjustments, hold promise in mitigating the risk of diabetic retinopathy progression (Yau et al., 2021).

2.2.1 Short-Term Effects

In the short term, diabetic retinopathy can manifest as various visual disturbances and ocular complications, posing significant challenges to individuals with type 2 diabetes. One of the primary short-term effects is visual blurring or fluctuation in vision, often resulting from retinal swelling (macular edema) or focal areas of ischemia (Khan et al., 2022; Project Group., 2023). This blurring can interfere with daily activities such as reading, driving, and performing tasks that require clear vision (Khan et al., 2022). Additionally, individuals may experience sudden onset of floaters or dark spots in their vision, indicating the presence of vitreous hemorrhage secondary to proliferative diabetic retinopathy (PDR) (Project Group., 2023).

Other short-term effects of diabetic retinopathy include visual field defects, where individuals may perceive missing areas or dark spots in their field of vision due to retinal ischemia or hemorrhage (Project Group., 2023). Furthermore, individuals may report episodes of visual distortion, such as straight lines appearing wavy or distorted, particularly in cases of macular edema or tractional retinal detachment (Khan et al., 2022). These visual disturbances can significantly impact quality of life and may necessitate immediate medical attention to prevent further vision loss (Project Group., 2023).

Moreover, short-term effects of diabetic retinopathy can extend beyond visual symptoms to include ocular complications such as neovascular glaucoma, where abnormal blood vessel growth leads to increased intraocular pressure and severe pain (Project Group., 2023). Additionally, tractional retinal detachment can occur, causing sudden onset of vision loss and distortion due to the pulling of fibrous tissue on the retina (Khan et al., 2022). These complications often require urgent ophthalmologic intervention, including laser treatment or surgical procedures, to preserve vision and prevent permanent visual impairment (Khan et al., 2022; Project Group., 2023).

2.2.2 Long-Term Complications

While addressing the immediate discomforts of hyperglycemia is crucial, its most insidious threat lies in its potential to wreak havoc on the body over the long term. Elevated blood glucose levels over extended periods can inflict severe damage to vital organs and systems, leading to a range of debilitating complications (CDC, 2023). Among the most notable are heart disease, kidney problems, neuropathy (nerve damage), and retinopathy (eye damage) (American Diabetes Association, 2023). These complications underscore the urgency of proactive hyperglycemia management and early intervention to prevent lasting harm.

2.3 Traditional Diabetic Retinopathy Detection Methods

Detecting diabetic retinopathy in individuals with Type 2 Diabetes Mellitus (T2DM) has been traditionally reliant on conventional methods (Endocrine Society of Australia, 2023). These established approaches have been very key in diabetes management but have also shown certain limitations that impede their overall effectiveness (Centers for Disease Control and Prevention, 2023; World Health Organization, 2023). This section delves into these traditional detection methods, exploring their intricacies and discussing the associated drawbacks.

2.3.1 Direct Ophthalmoscopy

Direct ophthalmoscopy is a fundamental technique involving the examination of the retina through the pupil using an ophthalmoscope. This method offers a wide-field view of the retina, allowing clinicians to assess its overall health and detect any abnormalities, including those associated with diabetic retinopathy (Kanski & Bowling, 2021). However, direct ophthalmoscopy relies heavily on the examiner's expertise and experience in interpreting retinal findings. The subjective nature of this technique can lead to variability in diagnosis between different practitioners (Haddrill & Niho, 2018). Moreover, direct ophthalmoscopy may not provide detailed imaging of subtle changes in the retina, particularly in cases of early-stage diabetic retinopathy (Kanski & Bowling, 2021).

2.3.2 Fundus Photography:

Fundus photography is a specialized imaging technique that captures detailed images of the retina using dedicated cameras. These high-resolution images provide valuable information about retinal anatomy, pathology, and vascular changes associated with diabetic retinopathy (Leasher et al., 2016). Fundus photography is considered the gold standard for diabetic retinopathy assessment due to its ability to document retinal changes over time and guide treatment decisions (Leasher et al., 2016). However, the widespread implementation of fundus photography is hindered by the need for specialized equipment and trained personnel. In resource-limited settings, access to fundus photography may be limited (Bastawrous & Armstrong, 2023).

2.3.3 Fluorescein Angiography

Fluorescein angiography is an invasive imaging procedure that involves injecting a fluorescent dye into the bloodstream and capturing sequential images of the retina as the dye circulates. This technique provides dynamic information about retinal blood flow, vascular permeability, and neovascularization, aiding in the diagnosis and staging of diabetic retinopathy (Gallego-Pinazo et al., 2014). However, fluorescein angiography carries inherent risks, including allergic reactions to the dye and potential kidney damage (Khan et al., 2022). Additionally, the procedure requires specialized equipment and expertise, limiting its availability in primary care settings. Furthermore, the invasive nature of fluorescein angiography may deter some patients from undergoing the procedure, particularly those with contraindications or aversions to intravenous injections.

2.3.4 Optical Coherence Tomography (OCT)

Optical coherence tomography (OCT) is a non-invasive imaging technique that provides cross-sectional images of the retina with micron-level resolution (Khan et al., 2022). By measuring the reflectivity of retinal tissues, OCT enables detailed assessment of retinal thickness, morphology, and structural integrity. This information is invaluable for diagnosing diabetic macular edema, monitoring disease progression, and guiding treatment decisions (Ciulla et al., 2023). However, OCT may have limitations in detecting early signs of diabetic retinopathy, such as microaneurysms and hemorrhages, which are better visualized with other imaging modalities. Moreover, OCT imaging requires cooperation from the patient and fixation stability, making it challenging for individuals with poor vision or media opacities to undergo reliable imaging (Lains et al., 2019).

2.3.5 Visual Acuity Testing

Visual acuity testing is a simple and widely used screening tool for assessing the sharpness of vision (Khan et al., 2022). It typically involves asking patients to read standardized charts, such as the Snellen chart, at a specific distance. Visual acuity testing provides valuable information about central vision function, which is essential for tasks such as reading, driving, and recognizing faces.

However, visual acuity testing may have limitations in detecting early diabetic retinopathy changes, particularly in the absence of significant macular involvement (Khan et al., 2022). Moreover, it does not provide detailed information about retinal pathology or vascular changes associated with diabetic retinopathy. Consequently, visual acuity testing may underestimate the severity of diabetic retinopathy and fail to identify individuals at risk of vision loss without additional screening modalities.

2.4 Artificial Intelligence (AI) and Deep Learning

In recent years, advancements in technology and medical imaging have led to the development of innovative approaches for diabetic retinopathy detection, offering promising alternatives to traditional methods. These new approaches leverage cutting-edge techniques such as artificial intelligence (AI), deep learning, and smartphone-based retinal imaging to improve screening efficiency, accuracy, and accessibility.

Artificial intelligence (AI) and deep learning have emerged as powerful tools in medical imaging, revolutionizing the field of diabetic retinopathy detection (Khan et al., 2022). Deep learning algorithms, particularly convolutional neural networks (CNNs), can analyze retinal images with remarkable accuracy, automatically identifying subtle features associated with diabetic retinopathy (Rajalakshmi et al. 2022). By training on large datasets of annotated retinal images, these algorithms can learn to recognize patterns indicative of diabetic retinopathy, enabling rapid and accurate screening (Khan et al., 2022). Moreover, AI-based systems have the potential to streamline workflow, reduce the burden on healthcare professionals, and improve patient outcomes by facilitating early detection and intervention.

2.5 Telemedicine and Remote Monitoring

Telemedicine and remote monitoring platforms enable remote consultation and evaluation of retinal images, facilitating diabetic retinopathy screening and management without the need for in-person visits (Khan et al., 2022). Through secure digital platforms, healthcare providers can review retinal images captured by patients or primary care providers, providing timely feedback and recommendations. Telemedicine platforms also support longitudinal monitoring of diabetic retinopathy progression, allowing for early detection of complications and timely intervention to prevent vision loss (Haleem et al., 2022). Additionally, telemedicine can enhance patient engagement and adherence to screening guidelines by offering convenient and accessible healthcare services.

2.6 Medical Imaging in Hyperglycemia Prediction

Medical imaging plays an ever-evolving role in evaluating blood sugar levels in individuals with type 2 diabetes mellitus (T2DM) (American Diabetes Association, 2023). This role encompasses imaging methods and approaches that offer insights into the physiological and anatomical changes linked to high blood sugar (Yadav et al., 2022).

Retinal imaging, also known as fundus photography or optical coherence tomography (OCT), provides a non-invasive way to assess changes in the eyes caused by high blood sugar (American Diabetes Association, 2023). The retinal blood vessels are particularly sensitive to variations in glucose levels (Yadav et al., 2022). Diabetic retinopathy often leads to manifestations such as microaneurysms, hemorrhages, and abnormal vessel size alterations (Frank, 2022).

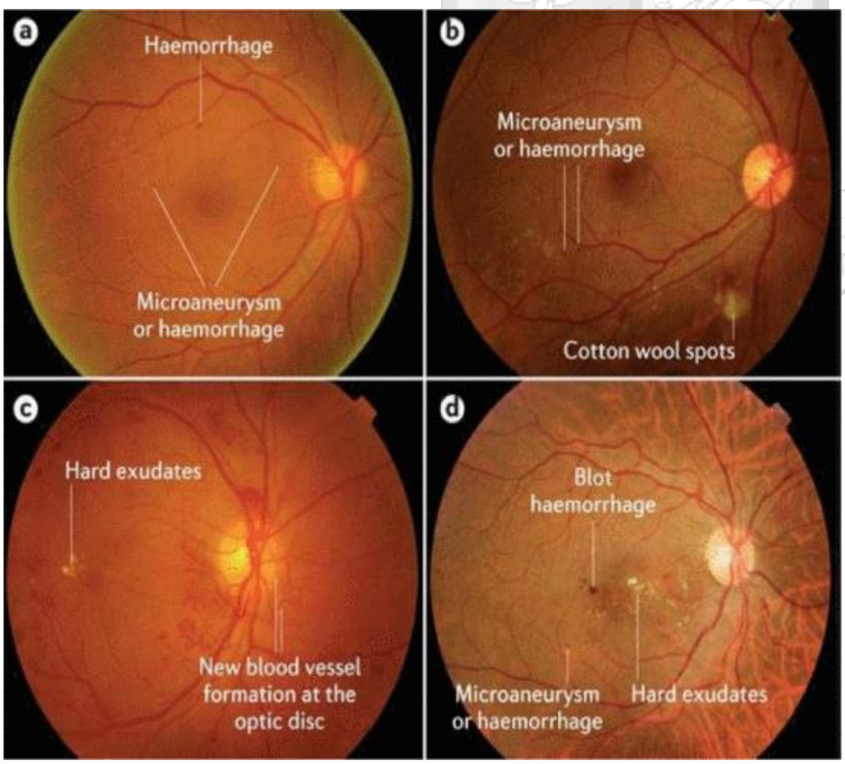


Figure 1Retinal images with different stages of retinopathy (from ieeexplore.ieee.org)

By capturing images of the retina, healthcare professionals can measure these changes, which offer valuable diagnostic information (Frank, 2022). Deep learning algorithms are increasingly utilized to analyze images, enabling automated detection of alterations associated with high blood sugar (Yadav et al., 2022). This approach enables detection and intervention, reducing the risk of vision loss while also providing insights into health (American Diabetes Association, 2023).

2.7 Smartphone-Based Retinal Imaging

The widespread availability of smartphones has paved the way for smartphone-based retinal imaging as a convenient and cost-effective screening tool for diabetic retinopathy (Bastawrous & Armstrong, 2023). With the use of portable retinal imaging devices that attach to smartphones, individuals can capture high-quality retinal images from the comfort of their homes or primary care clinics (Rajalakshmi et al., 2022). These images can then be transmitted to healthcare providers or AI-based systems for analysis, allowing for timely evaluation and referral if necessary. Smartphone-based retinal imaging holds promise in expanding access to diabetic retinopathy screening, particularly in underserved communities and resource-limited settings where traditional imaging modalities may be unavailable or impractical (Bastawrous & Armstrong, 2023).

2.7.1 Smartphone-based retinal imaging devices

In recent years, the field of medical imaging has witnessed a remarkable transformation with the emergence of smartphone-based retinal imaging devices. These innovative tools bring the power of retinal imaging directly into the hands of healthcare providers, researchers, and even patients. Smartphone-based retinal imaging devices offer several advantages, making them a promising avenue for the early detection and monitoring of retinopathy-related retinal changes (Hacisoftaoglu, et al. 2020)

One of the key advantages of smartphone-based retinal imaging devices is their portability and accessibility. Traditional retinal imaging equipment is often large, expensive, and confined to clinical settings. In contrast, smartphone-based devices are compact and can be easily carried by healthcare providers during home visits or deployed in resource-limited settings where access to specialized medical equipment is limited. (Hacisoftaoglu, et al. 2020) This portability extends the reach of retinal screening programs to underserved populations, facilitating early detection and intervention in regions with limited healthcare infrastructure.

Moreover, smartphone-based retinal imaging devices are designed to be user-friendly and require minimal training for operation. They typically consist of a smartphone equipped with specialized retinal imaging attachments or adapters. These adapters are designed to capture high-quality retinal images using the smartphone's built-in camera. The user can align the device with the patient's eye, capture retinal images, and store or transmit them for analysis.

Another notable feature of smartphone-based retinal imaging devices is their potential for telemedicine applications. The captured retinal images can be easily transmitted to remote healthcare professionals for expert evaluation. This telemedicine approach enables timely assessment and diagnosis, even for patients located in remote or rural areas. Furthermore, it allows for efficient follow-up and monitoring of retinal changes over time, which is crucial for individuals with T2DM at risk of hyperglycemia-related complications. (Hacisoftaoglu, et al. 2020)

2.7.2 Magnetic Resonance Imaging (MRI)

MRI is a medical imaging technique that can assess changes related to blood sugar in various tissues and organs (Yadav et al., 2022). For example, advanced MRI techniques can provide understanding about the composition and function of tissue (Frank, 2022).

Understanding the dynamics of tissue is essential when it comes to T2DM as this type of tissue plays a role in both insulin resistance and metabolic dysfunction (Yadav et al., 2022). MRI scans can also help evaluate how hyperglycemia affects organs like the liver, pancreas, and heart (Frank, 2022). They provide information on factors such as liver fat content, pancreatic volume, and cardiac function (Yadav et al., 2022). By doing so, they shed light on the changes that occur in these organs due to hyperglycemia. Additionally, functional MRI (fMRI) scans can assess brain activity in response to fluctuations in glucose levels (Frank, 2022). This contributes to an understanding of the aspects associated with hyperglycemia (Yadav et al., 2022).

2.7.3 Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT)

Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) are imaging techniques that involve using radiotracers targeting processes like glucose metabolism (Frank, 2022). In terms of predicting hyperglycemia, these modalities provide insights into how glucose is utilized at the tissue level (Yadav et al., 2022).

For instance, Fluorodeoxyglucose Positron Emission Tomography (FDG PET) is commonly used to evaluate glucose uptake in tissues (Frank, 2022). The detection of high glucose uptake through PET scans may indicate metabolic changes related to hyperglycemia within tissues (Yadav et al., 2022). Similarly, SPECT imaging helps assess blood flow and perfusion in response to challenges involving glucose consumption (Frank, 2022). This aids in identifying areas where there may be altered glucose metabolism (Yadav et al., 2022).

2.7.4 Deep Learning in the Analysis of Medical Images

Deep learning, which is a branch of machine learning and artificial intelligence (AI) seeks to replicate the way the human brain learns and makes decisions using datasets. It has revolutionized medical image analysis by providing techniques to extract insights from complex medical imaging modalities such as X rays CT scans, MRI scans and retinal images.

Inspired by networks, deep learning models consist of interconnected layers that progressively process. Extract intricate features from input data. These models can identify patterns and relationships within images enabling tasks such as disease detection, diagnosis, risk assessment and treatment planning.

2.8 Approaches in Deep Learning for Medical Image Analysis

2.8.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are an architecture in learning specifically designed for image related tasks. They aim to mimic the system's ability to detect patterns and shapes within images. CNNs comprise types of interconnected neurons including layers pooling layers and fully connected layers.

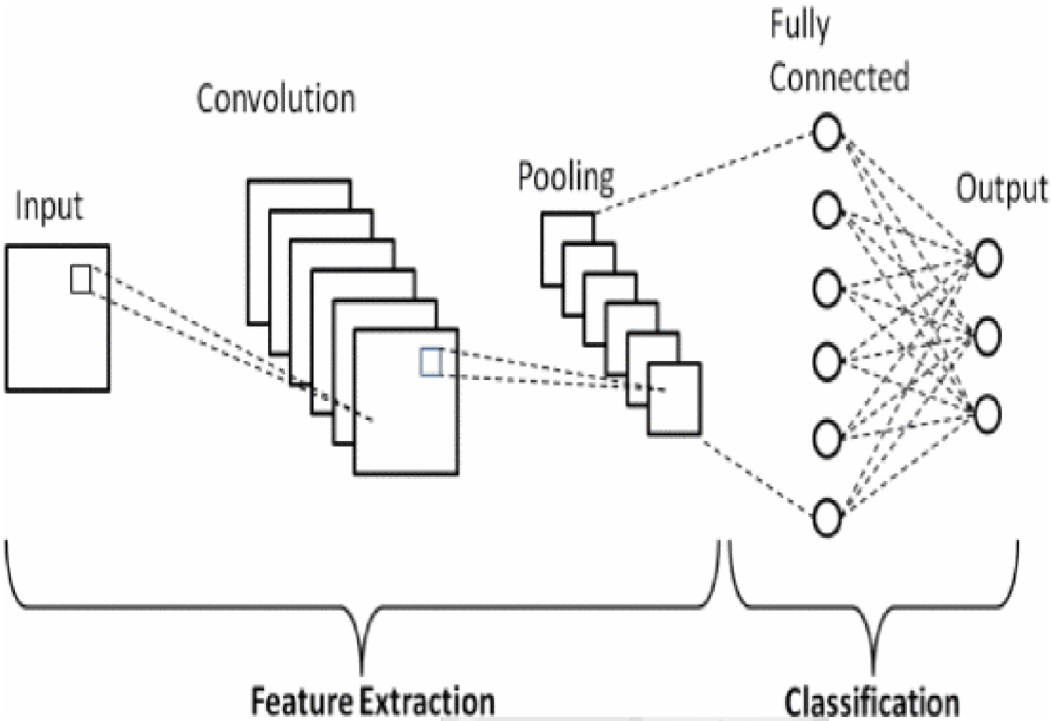


Figure 2Retinal images with different stages of retinopathy

Convolutional layers apply filters to input images to detect features, like edges, textures, and shapes. Over time these layers learn how to recognize patterns within the data.

Pooling layers are used in learning models to downsample the feature maps produced by layers. This downsampling reduces the dimensions while preserving information, which helps enhance computational efficiency and prevent overfitting.

Fully connected layers play a crucial role, in combining the extracted features from earlier layers to make final predictions or classifications. These predictions are based on the patterns detected throughout the network.

Convolutional Neural Networks (CNNs) have proven to be incredibly effective in tasks like image classification, object detection and segmentation. They are particularly valuable in medical image analysis, where they can be used to identify abnormalities associated with hyperglycemia.

Deep learning has emerged as a powerful tool in medical image analysis, offering unprecedented capabilities in identifying patterns and features crucial for diagnosing various diseases, including diabetic retinopathy

2.8.2 Transfer Learning

Transfer learning is a technique widely used in learning that involves leveraging trained neural networks trained on large datasets and adapting them for specific tasks. In medical image analysis obtaining labeled data for training learning models can be quite challenging and resource intensive. Transfer learning offers a solution by allowing researchers to use trained models from unrelated but extensive datasets and fine tune them for medical image analysis tasks.

For instance, a CNN that has been pre trained on a dataset of images (such as ImageNet) can be fine-tuned to recognize specific patterns or anomalies within medical images related to hyperglycemia. This approach significantly reduces the amount of labeled data needed for training while taking advantage of the knowledge gained from datasets.

2.8.3 Generative Adversarial Networks (GANs)

GANs are a type of learning model that consist of two networks: a generator and a discriminator. These networks work together in a manner, where the generator aims to create data while the discriminator tries to differentiate between real and generated data. This adversarial process results in the generation of data that closely resembles real world data.

In the field of medical image analysis GANs play a role in augmenting and diversifying datasets. They can produce images that mimic the characteristics of actual patient data. This is extremely valuable, for expanding training datasets and improving the robustness and generalizability of learning models. In terms of predicting hyperglycemia GANs can generate images showing different glucose levels. This helps with model training. Enhances the model's ability to detect subtle changes related to hyperglycemia.

2.8.4 Contributions to Diabetic Retinopathy Detection

Deep learning has emerged as a powerful tool in medical image analysis, offering unprecedented capabilities in identifying patterns and features crucial for diagnosing various diseases, including diabetic retinopathy. Within the realm of deep learning architectures, EfficientNetB0 stands out as a particularly promising model for medical image analysis.

EfficientNetB0 is part of the EfficientNet family of convolutional neural networks (CNNs), which are designed to achieve superior performance with fewer parameters compared to traditional architectures. This efficiency stems from a compound scaling method that optimizes network depth, width, and resolution simultaneously, resulting in models that are both accurate and computationally efficient.

In the context of medical image analysis, EfficientNetB0 offers several advantages. Its lightweight architecture makes it well-suited for deployment on resource-constrained devices, such as smartphones and edge computing platforms, facilitating point-of-care diagnostics and telemedicine applications. Despite its compact size, EfficientNetB0 maintains high accuracy, making it a valuable tool for screening and detecting diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices.

EfficientNetB0's versatility extends beyond diabetic retinopathy detection, encompassing a wide range of medical imaging tasks, including lesion detection, segmentation, and classification. Its ability to extract relevant features from medical images enables precise localization and characterization of abnormalities, aiding clinicians in accurate diagnosis and treatment planning.

Furthermore, EfficientNetB0 can be fine-tuned on domain-specific datasets, allowing for adaptation to different imaging modalities and clinical contexts. By leveraging transfer learning, researchers and healthcare professionals can harness pre-trained EfficientNetB0 models and tailor them to specific medical imaging tasks, thereby reducing the need for large annotated datasets and expediting model development.

2.9 Related Applications

This field has garnered significant attention from researchers who are actively engaged in extensive investigations across various domains concerning the relationship between blood sugar levels and the human retina. Their primary aim is to develop effective strategies for assessing and managing individuals with Type 2 Diabetes (T2DM). The following are notable related works in this evolving area of research:

2.9.1 Deep Learning for Diabetic Retinopathy Detection

This review paper provides a comprehensive examination of various deep learning methodologies for diabetic retinopathy detection. It covers the key advancements, challenges, and potential applications of deep learning in automated screening systems.

The paper systematically reviews deep learning approaches applied to diabetic retinopathy detection, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their variants. It discusses the preprocessing techniques, dataset characteristics, and model architectures employed in different studies.

As a review paper, Gulshan et al. do not present new experimental results. Instead, they summarize the performance metrics reported in previous studies, including sensitivity, specificity, area under the receiver operating characteristic curve (AUC-ROC), and accuracy. They highlight variations in performance across different datasets and methodologies.

One limitation of this review is the lack of standardized evaluation metrics across studies, making direct comparisons challenging. Additionally, the review focuses on summarizing existing literature rather than conducting new experiments or proposing novel methodologies.

2.9.2 Detection of Diabetic Retinopathy Using Deep Learning

Ting et al. (2023) propose a deep learning system for automated detection of diabetic retinopathy from retinal fundus images. The study aims to develop a robust and scalable screening tool that can accurately identify diabetic retinopathy lesions.

The authors employ a multi-level ensemble model that integrates multiple CNN architectures, including Inception-v3 and ResNet, to improve the model's performance. They train the ensemble model on a large dataset of labeled retinal images and fine-tune it using transfer learning techniques.

Ting et al. evaluate the performance of their deep learning system using standard metrics such as sensitivity, specificity, and AUC-ROC. They report high accuracy and AUC-ROC values, indicating the model's effectiveness in diabetic retinopathy detection across different datasets.

One limitation of this study is the lack of interpretability of the ensemble model, as the contributions of individual CNN architectures to the final prediction are not explicitly analyzed. Additionally, the study primarily focuses on algorithmic performance without considering factors such as computational efficiency and real-world deployment challenges.

2.9.3 Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs

Gulshan et al. (2022) presents a deep learning algorithm for the detection of diabetic retinopathy from retinal fundus photographs. The study aims to develop a highly accurate and scalable screening tool for diabetic retinopathy.

The authors train a deep convolutional neural network (DCNN) using a large dataset of labeled retinal images from multiple sources. They fine-tune the DCNN using transfer learning techniques and evaluate its performance on independent test datasets.

Gulshan et al. report sensitivity, specificity, AUC-ROC, and accuracy scores to assess the algorithm's performance in detecting diabetic retinopathy. The results demonstrate high accuracy and AUC-ROC values, indicating the model's robustness across diverse datasets.

One limitation of this study is the potential bias introduced by using datasets from multiple sources, which may vary in image quality and labeling accuracy. Additionally, the algorithm's performance may be affected by variations in image acquisition protocols and demographic characteristics of the study population.

2.9.4 Deep Learning for Diabetic Retinopathy Diagnosis Using Integrated Gradients

Zhang et al. (2022) proposes an interpretable deep learning framework for diabetic retinopathy diagnosis using integrated gradients. The study aims to enhance the interpretability of deep learning models and provide insights into their decision-making process.

The authors develop a deep learning model based on the Inception-v3 architecture and integrate integrated gradients, a technique for attributing predictions to input features, to generate saliency maps highlighting important regions in retinal images. They evaluate the model's performance and interpretability using a large dataset of labeled retinal images.

Zhang et al. assess the model's diagnostic performance using standard metrics such as sensitivity, specificity, and AUC-ROC. They also analyze the interpretability of the model by visualizing saliency maps and comparing them with ground truth annotations.

One limitation of this study is the potential trade-off between model accuracy and interpretability, as complex deep learning models may sacrifice interpretability for improved performance. Additionally, the interpretability of integrated gradients may be limited in cases of highly non-linear models or ambiguous image features.

2.9.5 Deep Learning for Diabetic Retinopathy from Ultra-Widefield Fundus Images

Raj et al.(2023) develop a deep learning-based system for automated detection of diabetic retinopathy from ultra-widefield fundus images. The study aims to extend diabetic retinopathy screening to comprehensive retinal imaging modalities with a wide field of view.

authors train a deep learning model using a combination of CNN architectures and data augmentation techniques to detect diabetic retinopathy lesions in ultra-widefield fundus images. They evaluate the model's performance on a large dataset of labeled images and compare it with existing screening methods.

Raj et al. assess the model's performance using standard metrics such as sensitivity, specificity, and AUC-ROC. They report high accuracy and AUC-ROC values, indicating the model's effectiveness in detecting diabetic retinopathy lesions across a wide field of view.

One limitation of this study is the potential challenge of integrating the deep learning model into existing clinical workflows, as ultra-widefield imaging may require specialized equipment and expertise. Additionally, the model's performance may be affected by variations in image quality and artifacts inherent to widefield imaging modalities.

2.9.6 Glycemic Control on Retinal Photoreceptor Layers and Retinal Pigment Epithelium

Ishibashi et al. (2021) investigated the impact of glycemic control on the metrics of individual macular photoreceptor layers (MPRLs) and retinal pigment epithelium (RPE) in type 2 diabetic patients without diabetic retinopathy using enface OCT. They found that glycemic control sequentially restored the thickness and volume of some MPRLs, especially the outer nuclear layer

(ONL). However, high glucose during glycemic control decreased the reflectance of MPRLs and RPE. The authors concluded that glycemic control can be beneficial for the retina, but high glucose during glycemic control can be harmful. Repeated OCT examinations can clarify the benefit and hazard of glycemic control to the diabetic retinopathy.

This study is significant because it provides evidence that glycemic control can improve the structural and functional integrity of the retina in type 2 diabetes patients without diabetic retinopathy. However, it is important to note that high glucose during glycemic control can have a negative impact on the retina. Therefore, it is important for patients to maintain tight glycemic control to minimize the risk of retinal damage.

2.9.7 A Deep Learning Based Diabetic Retinopathy Detection from Retinal Images

Yadav et al. (2021) proposed a deep learning-based system for detecting diabetic retinopathy (DR) from retinal images. The system is accurate and efficient, and it could be used to develop a clinical tool for DR screening.

The system consists of four main steps, first, the retinal image is converted to grayscale and enhanced to improve the visibility of key features. Next, the optic disk and blood vessels are segmented from the retinal image using image processing techniques. Color features are then extracted from the segmented retinal image. Finally, a CNN is used to classify the retinal image as normal or DR (Yadav et al., 2021).

The research indicates that CNN is a type of machine learning model that is well-suited for image classification tasks. The CNN is trained on a large dataset of retinal images, which allows it to learn to identify the key features that are associated with DR.

The proposed system was evaluated on a public dataset of retinal images. The system achieved an accuracy of 78.50% (Yadav et al., 2021), which was significantly higher than the accuracy of 67.40% achieved by a support vector machine (SVM) classifier.

2.9.8 Retinal Image Analysis for Detection of Diabetic Retinopathy from the Nakuru Study, Kenya

Hansen et al. (2018) investigated the performance of the Iowa Detection Program (IDP) automated retinal image analysis software in detecting diabetic eye disease (DED) in a population-based sample from Nakuru, Kenya. The study found that the IDP had a sensitivity of 91.0% and a specificity of 98.0% for detecting DED, as compared to human grading. The area under the curve (AUC) for the IDP was 0.878. The IDP missed no vision-threatening retinopathy in any patients and none of the false-negative cases met criteria for treatment.

The authors concluded that the IDP automated retinal image analysis software is a promising tool for detecting DED in population-based settings. It has high sensitivity and specificity, and it can help to reduce the burden on human graders.

This study is relevant to this research because it demonstrates the feasibility of using deep learning to analyze retinal images to detect diabetic eye disease, even in a population-based setting in Sub-Saharan Africa. This is important for this research because it suggests that a deep learning model could be used to predict hyperglycemia in patients with type 2 diabetes, even in resource-poor countries. Additionally, the fact that both studies use deep learning to analyze retinal images means that there is potential for collaboration and cross-fertilization of ideas.

2.10 Proposed Solution

The proposed solution for early detection of diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices primarily involves the utilization of deep learning techniques, with a focus on the EfficientNetB0 architecture, alongside mobile imaging devices. This approach aims to address the challenges associated with traditional diabetic retinopathy detection methods by leveraging the power of artificial intelligence and the accessibility of smartphone-based imaging technology.

The EfficientNetB0 architecture, known for its efficiency and effectiveness in image classification tasks, serves as the backbone of the proposed deep learning model. By harnessing the capabilities of EfficientNetB0, the model can efficiently extract relevant features from retinal images captured by smartphone-based devices, enabling accurate detection of diabetic retinopathy lesions. The lightweight nature of EfficientNetB0 also makes it well-suited for deployment on resource-constrained mobile devices, ensuring scalability and accessibility in real-world clinical settings.

Complementing the deep learning model is the use of smartphone-based retinal imaging devices, which offer a non-invasive and convenient solution for capturing high-quality retinal images. These devices leverage the built-in cameras of smartphones, combined with specialized optics and attachments, to acquire retinal images with sufficient clarity and resolution for diagnostic purposes. By integrating the deep learning model with smartphone-based imaging devices, the proposed solution enables point-of-care diabetic retinopathy screening, facilitating early detection and intervention in individuals with type 2 diabetes.

The integration of EfficientNetB0 with mobile imaging devices offers several advantages over traditional diabetic retinopathy detection methods. Firstly, it eliminates the need for specialized equipment and trained personnel, making diabetic retinopathy screening more accessible in primary care settings and underserved communities. Secondly, it enables real-time analysis of retinal images directly on mobile devices, allowing for immediate feedback and referral to ophthalmologists or diabetic retinopathy specialists when necessary. Additionally, the proposed solution leverages the widespread adoption of smartphones, ensuring broad reach and adoption among individuals with type 2 diabetes.

2.11 Conceptual Framework

The conceptual framework of this research serves as the theoretical foundation and structure that guides the study's design, implementation, and interpretation. At its core, the framework outlines the key components and relationships that underpin the research on early detection and managing of diabetic restinopathy in individuals with T2DM using retinal images.

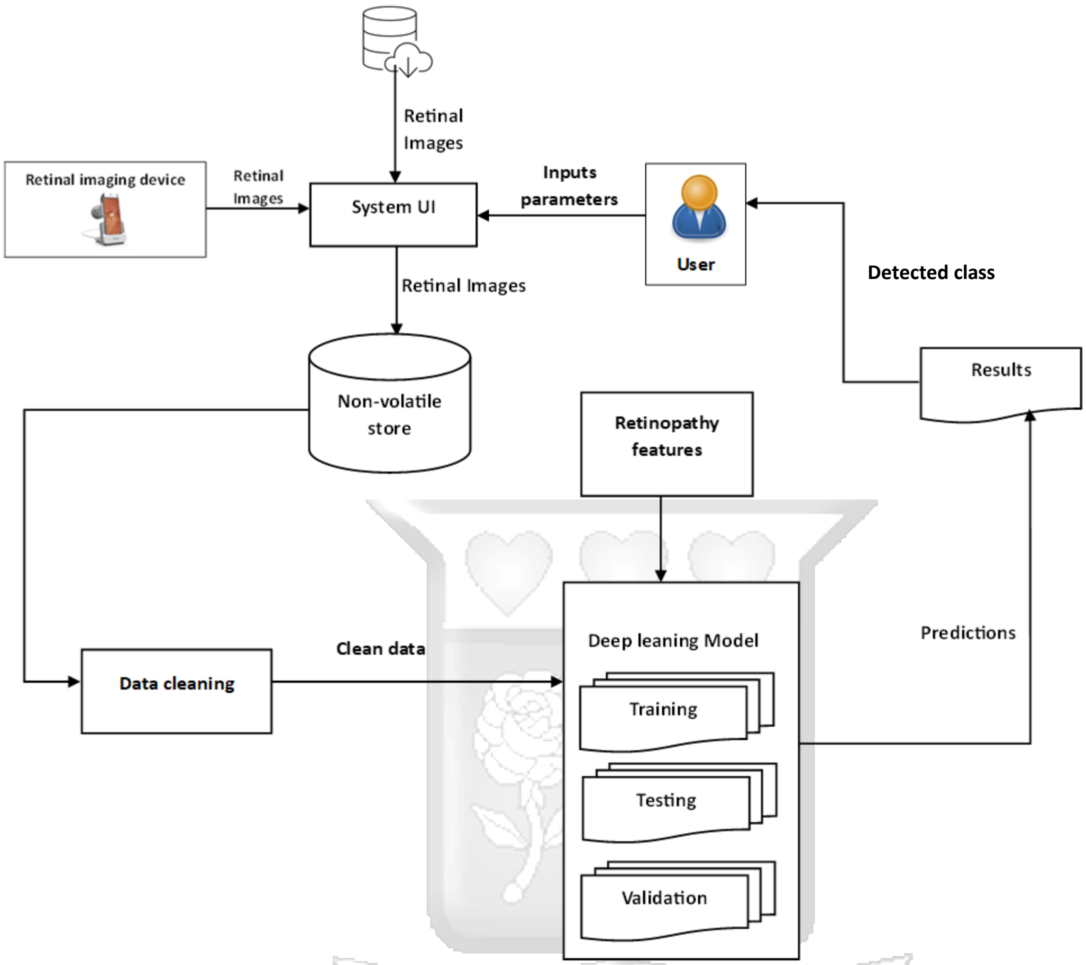


Figure 3 Conceptual Framework

3 Chapter 3: Methodology

3.1 Introduction

In this research the methodology chapter aims to provide an overview of the strategies, approaches and procedures used to achieve both the objective and specific objectives outlined in this study. This chapter acts as a guide, to understanding the scientific processes involved in creating a learning model for predicting hyperglycemia from retinal images in individuals with Type 2 Diabetes. It covers stages such as data collection, model development, training, evaluation and the inclusion of a recommendation system for early detection of diabetic retinopathy.

3.2 Research design

This study follows an experimental research design to achieve the objectives outlined in the study. The research starts by identifying its specific objectives. These objectives serve as a guide, throughout the research process from collecting data to developing and evaluating a model.

At the core of this research design is the development of a model that will be able to detect diabetic retinopathy at early stages for early intervention. By utilizing CNN architecture EfficientNetB0, the model is designed to analyze retinal images and predict the likelihood of development of DR episodes. CNNs were chosen because they have been proven effective in tasks related to images (LeCun et al., 1998).

To ensure that early DR detection model is robust and accurate a series of experiments were conducted. These experiments used a dataset of images that are well annotated. The dataset covers cover a range of features to ensure that the model can handle variations commonly found in real world patient data. This thorough validation process aligns with established practices for developing machine learning models (Olang' 2020; Creswell, 2014).

In addition to developing and experimenting with the model, this research also includes creating a user mobile application. The mobile application acts as a user interface where images captured from smartphone-based fundus imaging devices can be uploaded for validation. These images are then be analyzed by a trained CNN model to give analysis for detecting early DR. This application aims to make it easier for both patients and healthcare providers to access an effective solution.

3.3 Datasets and Data Collection

The methodology involved the collection and preparation of a diverse dataset of retinal fundus images, encompassing various diabetic retinopathy severity levels and disease stages. The dataset includes images with labels indicating different diabetic retinopathy grades, such as Moderate Non-Proliferative Diabetic Retinopathy (NPDR), Severe NPDR, and No DR signs.

The dataset was obtained from a publicly available repository on Kaggle, that represents a sample of retinal images. The dataset gave special attention to ensure the inclusion of images representing different ethnicities, age groups, and comorbidities to enhance the dataset's diversity and generalizability.

3.4 Data Preprocessing

Upon collection, the raw fundus images underwent preprocessing steps to standardize their format, enhance image quality, and remove artifacts. Preprocessing techniques included image cropping, resizing, normalization, and noise reduction to ensure consistency and reliability in subsequent analysis stages. These steps included:

i. **Conversion to grayscale:** RGB format images into grayscale to simplify feature extraction.
ii. **Noise Reduction:** Applying image enhancement techniques to reduce noise and improve image clarity.
iii. **Standardization:** Images Standardization will be done the size and format of the images for consistency.

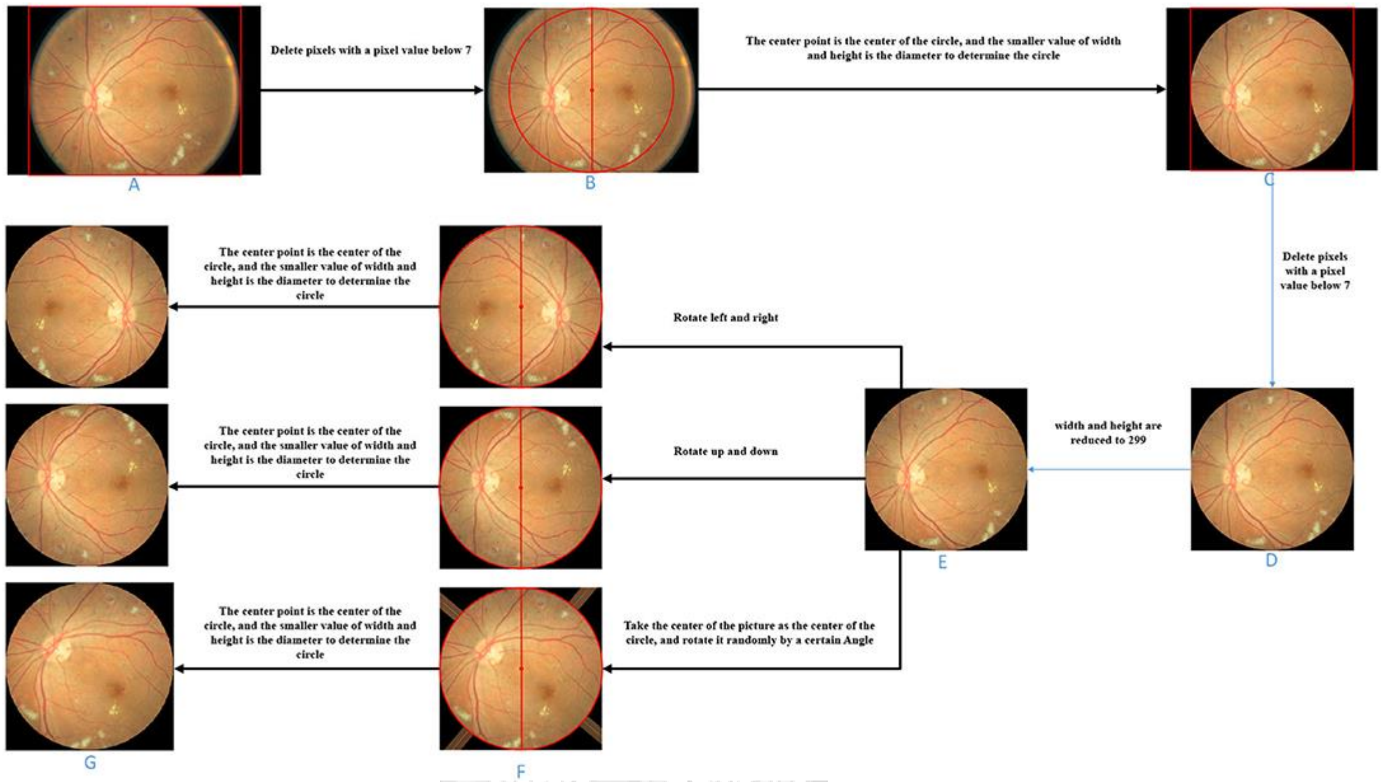


Figure 4 Image processing (Downloaded from frontiersin.org)

3.5 Model Development

3.5.1 Development of Deep Learning Model

In the development phase of the deep learning model, various machine learning (ML) architectures were explored and compared to ascertain their effectiveness in diabetic retinopathy detection. The initial experimentation involved employing a Random Forest model, which yielded an accuracy of 77.53%. However, to enhance feature extraction capabilities, a combination with Convolutional Neural Network (CNN) architecture was necessary. Despite its respectable performance, the Random Forest model fell short compared to other architectures considered.

Further experimentation with different CNN architectures revealed varying levels of performance. The DenseNet architecture exhibited an accuracy of 79.0%, showcasing its potential for diabetic retinopathy detection. The superior performance of DenseNet can be attributed to its dense connectivity pattern, enabling efficient feature reuse and propagation throughout the network. This architecture's ability to capture intricate features within retinal images likely contributed to its higher accuracy.

In contrast, the ResNet architecture demonstrated a comparatively lower accuracy of 46.07%. The observed performance may be attributed to the architecture's inherent depth, which could lead to difficulties in training due to vanishing gradients or the degradation problem. Additionally, the limited capacity of ResNet to capture fine-grained details in retinal images may have contributed to its reduced accuracy.

Similarly, the VGG16 architecture achieved an accuracy of 56.78%, indicating moderate performance in diabetic retinopathy detection. While VGG16 is renowned for its simplicity and ease of implementation, its reliance on a fixed architecture with stacked convolutional layers may limit its capacity to capture complex features essential for accurate classification.

The EfficientNetB0architecture emerged as the standout performer, boasting an impressive accuracy of 88.57%. The superior performance of EfficientNetB0 can be attributed to its scalable and efficient architecture, which dynamically scales network depth, width, and resolution based on model complexity. This adaptability allows EfficientNetB0to achieve high accuracy while maintaining computational efficiency, making it a compelling choice for diabetic retinopathy detection.

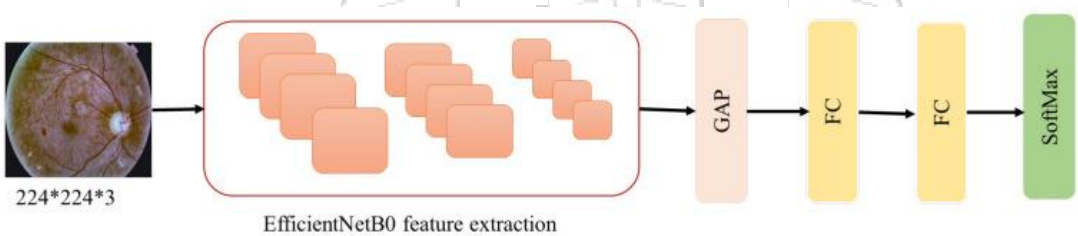


Figure 5 Transfer learning EfficientNetB0

3.5.1.1 Iterative Model Design

In each iteration the researcher designed the architecture of the learning model based on insights gained from iterations considering factors such as layer count, neuron configuration and activation functions used. When designing the model, the researcher took into account factors such as the model’s ability to extract features and its computational efficiency.

3.5.1.2 Incremental Training

Incremental training was employed, where a portion of the images were be used in each iteration. This approach allowed the model to adapt to types of image data and learn progressively.

3.5.1.3 Hyperparameter Tuning

To optimize the performance of the model hyperparameters like learning rate, batch size and dropout rates were to be fine-tuned iteratively.

3.5.1.4 Evaluation and Validation

After each iteration thorough evaluation of the model's performance was conducted using cross separate test sets. Accuracy and precision score were calculated as performance metrics to assess how effective the model was.

3.6 Research Quality

Throughout the research process meticulous record keeping and documentation was maintained to ensure research quality. All documentations for the code used for data preprocessing and model development were detailed to enhance reproducibility.

3.7 Performance Visualization

To gain insights into the model's performance at each iteration, visualizations metrics were utilized. These visualizations assisted in guiding decision making for refinements.

3.8 Model Validation

Validating the model is a crucial aspect of this research project.

Multiple experiments were conducted to validate the effectiveness of the model.

3.8.1 Experiment 1: Exploring Machine Learning Algorithms

In addition to Deep learning, the researcher explored other machine learning algorithms in this experiment. Models such as decision tree and Random Forests were used and compared with the chosen CNN model to determine their suitability for detecting DR.

3.8.2 Experiment 3: Using Images

To test the robustness of the model the researcher introduced sets of retinal images in this experiment, including those taken under various conditions. The research aimed to evaluate how well the model can generalize its predictions across different datasets.

3.9 Ethical Considerations

Given that the research involved clinical data, ethical considerations play a major role. The researcher submitted the paper for ethical review, which was approved as valid research.

4 Chapter 4: System Design

This chapter outlines the structural framework and technological setup for the envision deep learning model aimed at early diabetic retinopathy detection. It shows the aspects of both software and hardware elements pivotal for constructing, launching, and operating the system. Acting as a comprehensive guide, this chapter offers a detailed insight into the system's design, offering a clear pathway for translating research discoveries into tangible applications.

4.1 Requirement Analysis

4.1.1 System requirements

The system should:

i. Preprocess retinal fundus images to enhance quality, remove noise, and standardize format, ensuring consistency and reliability for subsequent analysis.

ii. Utilize the EfficientNetB0 architecture to extract relevant features from retinal images, facilitating accurate classification and detection of diabetic retinopathy lesions.

iii. Incorporate classification algorithms to categorize retinal images into diabetic retinopathy severity levels, including Moderate NPDR, Severe NPDR, and No DR signs.

iv. Support real-time analysis of retinal images for immediate feedback and decision-making in clinical settings.

v. Seamlessly integrate with smartphone-based retinal imaging devices for efficient image capture, data transfer, and analysis, ensuring a smooth workflow and user experience.

vi. Provide clear and interpretable results, including diagnostic insights, confidence scores, and recommended actions, to assist healthcare professionals in clinical decision-making.

vii. Ensure scalability and performance optimization to handle large data volumes, concurrent user requests, and minimize inference latency for efficient system operation.

viii. Feature a user-friendly interface allowing intuitive interaction, visualization of results, customization of settings, and seamless navigation through functionalities.

4.1.2 Usability Requirements

Users of this system include healthcare professionals, researchers, individuals with type 2 diabetes, and potentially caregivers or family members. Each user group has specific needs and requirements that should be addressed to ensure the system's usability and effectiveness. To use this system, the user requires to have an android mobile application with at least android version 7.0 and above for smooth running. Users also need to attach an mobile based retinal imaging device to the mobile application, which can either be connected wirelessly via Bluetooth or wired with a USB cable to the mobile phone

Figure 7 D.EYE Retinal Imaging device connected to a mobile phone (Downloaded from ophthalmologyweb.com)

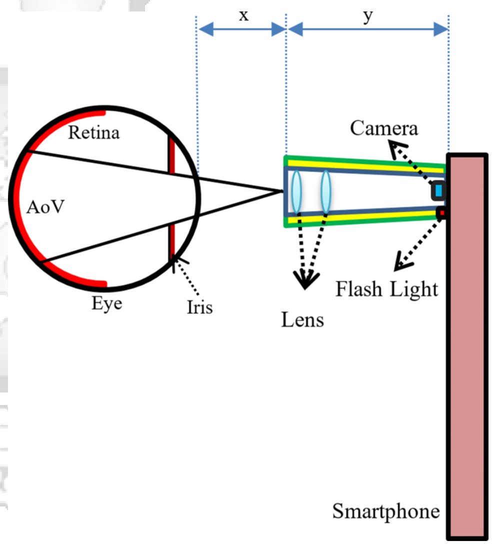


Figure 6 The general optical design of the smartphone-based retinal imaging devices (Downloaded from researchgate.com)

4.2 System Architecture

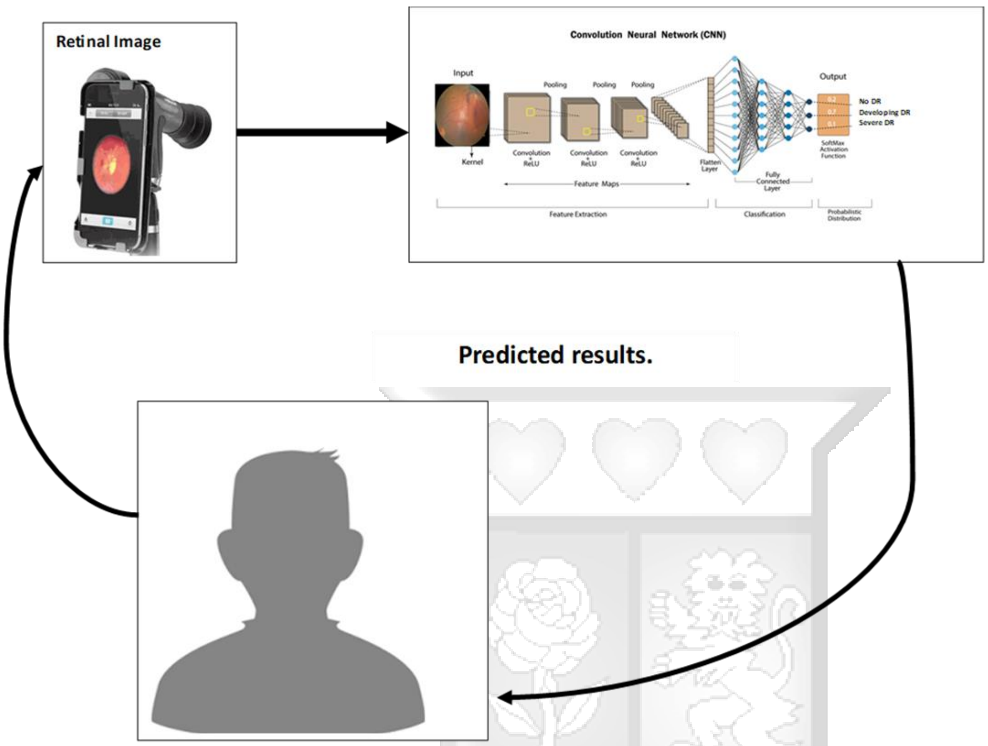


Figure 8 Simple System Architecture

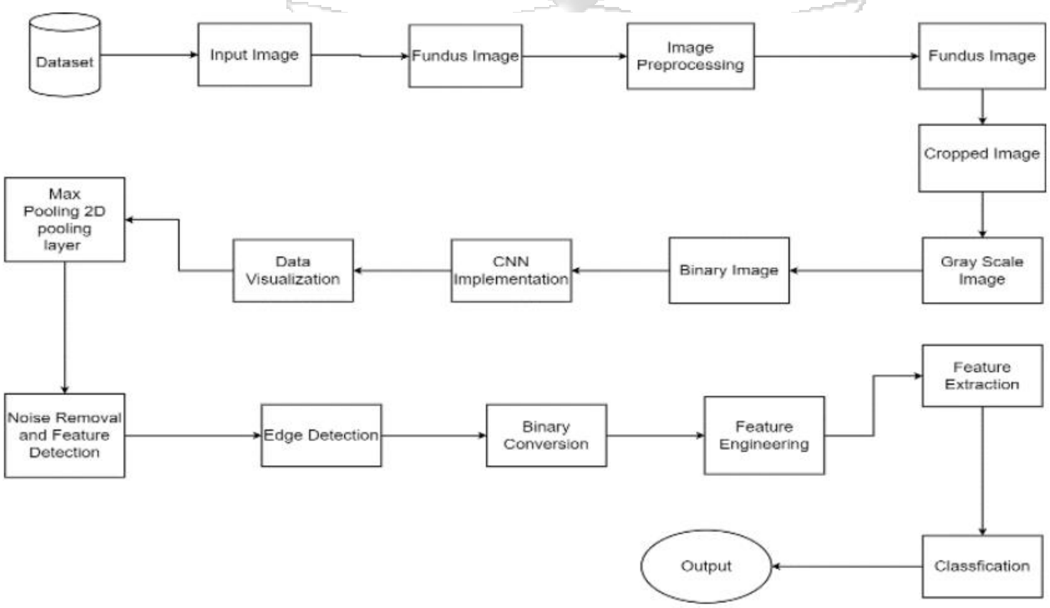


Figure 9 Detailed System Architecture

4.3 Unified Modeling Language (UML) Representation of the Model

Unified Modeling Language (UML) is a standardized modeling language used in software engineering to visually represent software systems. It provides a set of graphical notations for specifying, visualizing, constructing, and documenting the artifacts of software systems, including their structure, behavior, and interactions. UML diagrams serve as blueprints for understanding, designing, and communicating about software systems among stakeholders.

4.3.1 Use Case Diagram

A Use Case Diagram represents the interactions between users (actors) and the system, showcasing the system's functionalities from the user's perspective. It outlines the various actions users can perform and the system's responses to those actions.

The figure below depicts the use case diagram for this project.

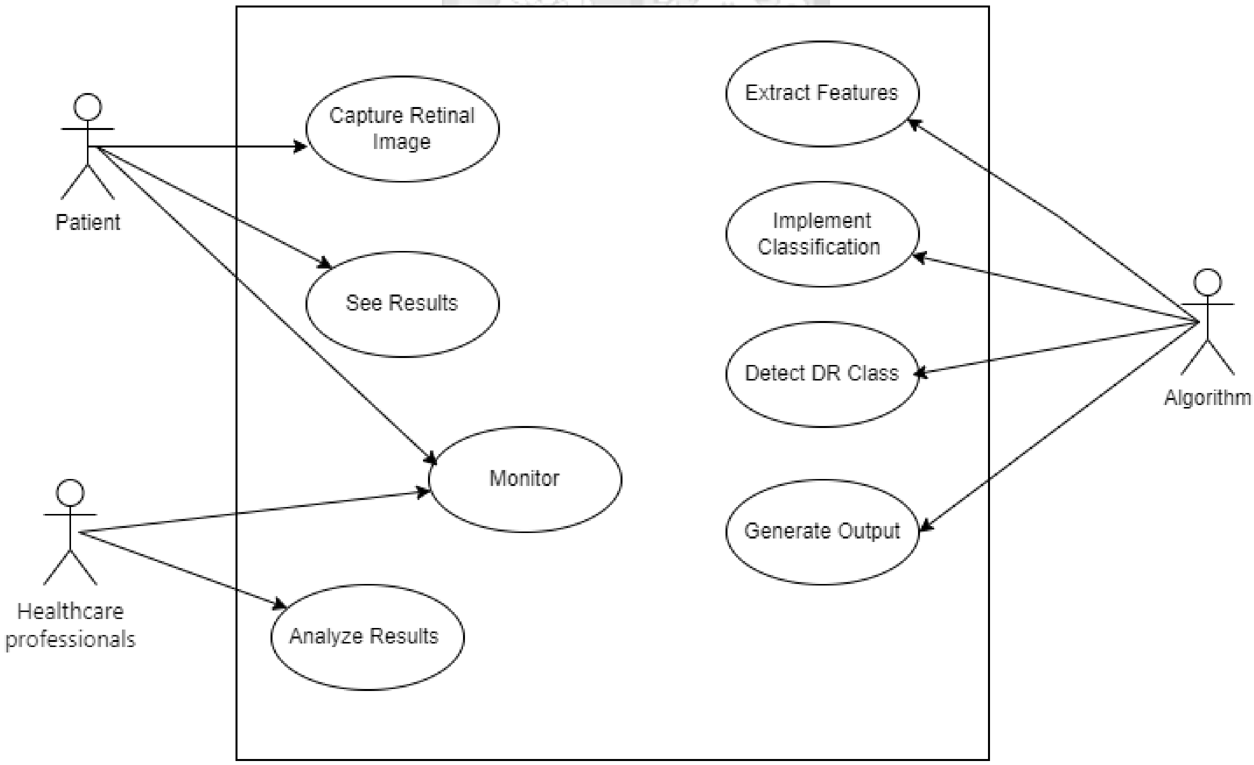


Figure 10 Use Case Diagram

4.3.2 Sequence Diagram

A Sequence Diagram illustrates the interactions between objects over time, showcasing the flow of messages and method calls. It visualizes the dynamic behavior of the system during specific scenarios or processes.

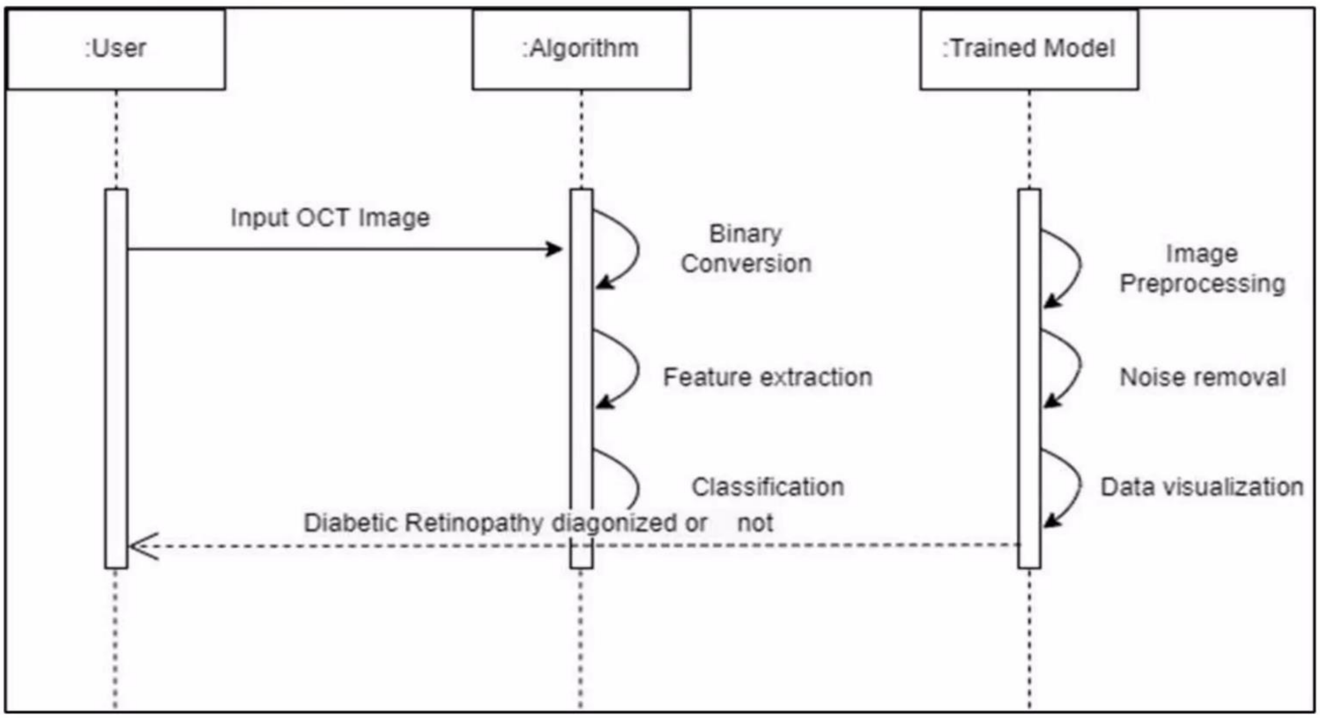


Figure 11 Sequence Diagram

4.3.3 Activity Diagram

An Activity Diagram represents workflows or processes, illustrating the sequence of activities and decision points. It provides a high-level overview of the steps involved in completing a particular task or process.

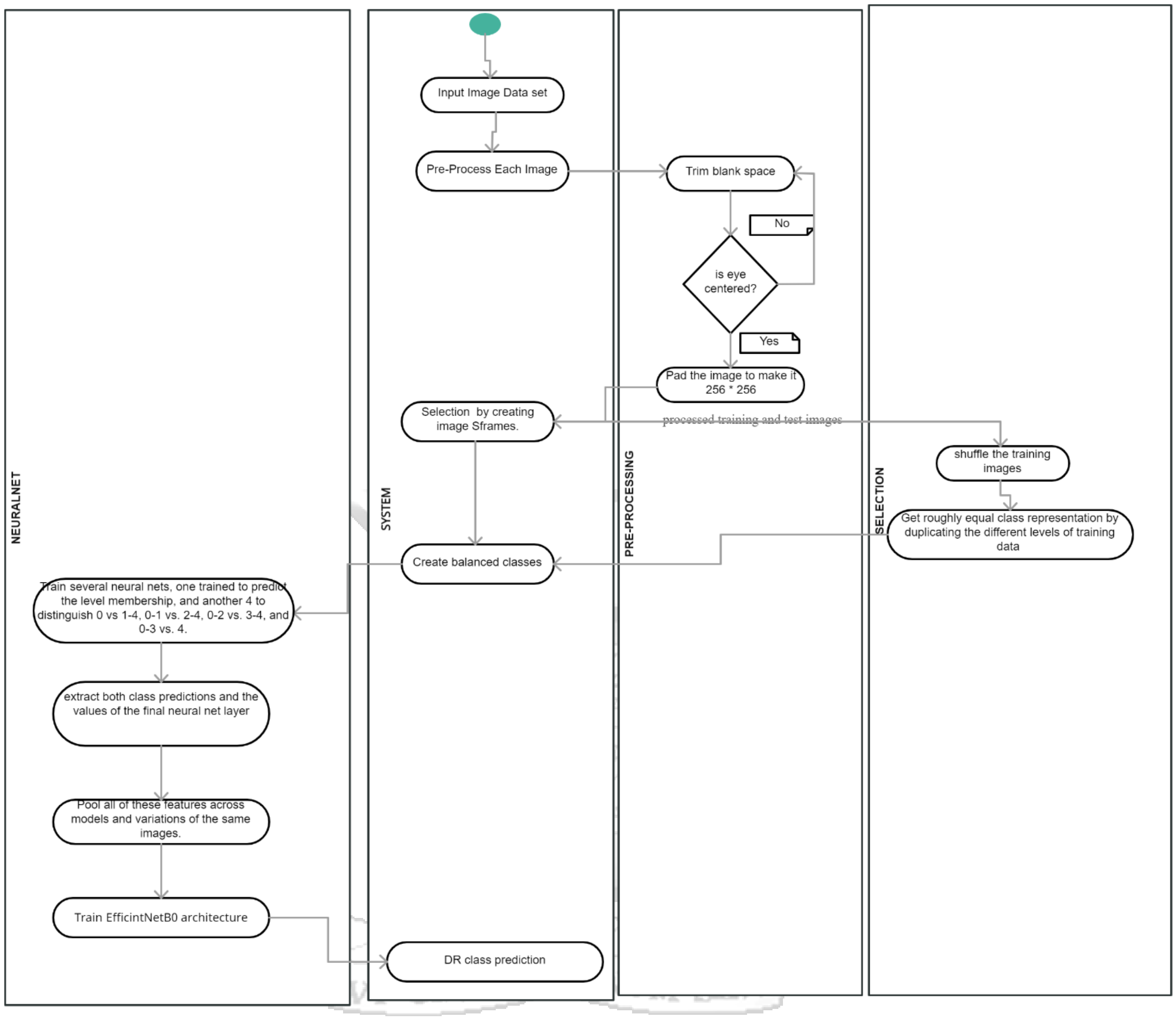


Figure 12 Activity Diagram

4.3.4 Block Diagram:

A Block Diagram is a high-level diagram that provides an overview of the system's components and their interactions. It uses blocks to represent system components and lines to illustrate connections or relationships between them. Block diagrams are commonly used to visualize the overall architecture of the system, showing how different components are organized and how they interact with each other.

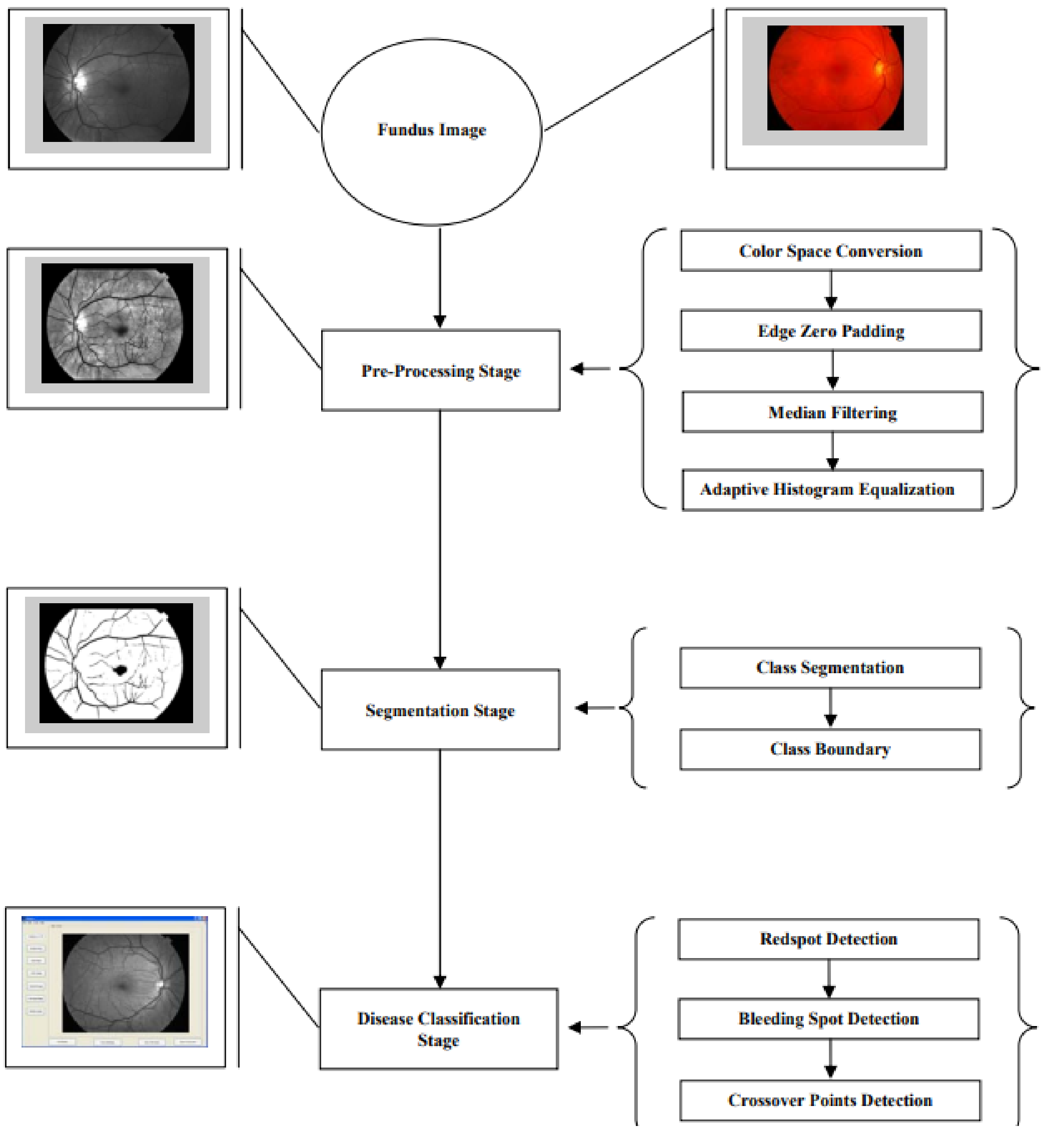


Figure 13 Block Diagram (Captured from diva-portal.org)

5 Chapter 5: Implementation and Testing

This chapter focuses on the practical implementation of the proposed deep learning model for early detection of diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices in type 2 diabetes. It outlines the steps taken to develop, deploy, and test the system, ensuring its functionality, reliability, and effectiveness.

5.1 System Implementation

The system implementation was done in three parts.

5.1.1 Model development and training

The first part of the system implementation involved the development and training of the deep learning model using TensorFlow and Keras.

The training process for the deep learning model begins with the loading and preprocessing of the retinal images stored in specified directories. These images are retrieved using a custom function, **load_images_and_labels**, which not only reads the images but also preprocesses them using the EfficientNet's built-in preprocessing function. This preprocessing step ensures that the images are properly formatted and ready for training.

Once the images are preprocessed, data augmentation techniques are applied using the **ImageDataGenerator** from TensorFlow. This includes a variety of transformations such as rotation, shifting, shearing, zooming, and horizontal flipping. These augmentations are essential for increasing the diversity of the training dataset and improving the model's ability to generalize to unseen data.

Following data augmentation, the dataset is split into training and testing sets using the **train_test_split** function from scikit-learn. The training set, comprising 80% of the data, is used to train the model, while the testing set, containing the remaining 20%, is reserved for evaluating the model's performance.

The model architecture is then defined, starting with the loading of the EfficientNetB0 model pretrained on ImageNet. Custom layers are added on top of the base model to adapt it to the task of diabetic retinopathy detection. These layers include a GlobalAveragePooling2D layer, a fully connected Dense layer with ReLU activation and L2 regularization, and a Dropout layer to prevent overfitting.

After defining the model architecture, it is compiled using the Adam optimizer with a specified learning rate, categorical cross-entropy loss function, and evaluation metrics such as accuracy and precision. The model is then trained on the training data using the **fit** function, with early stopping applied to prevent overfitting.

Once training is complete, the model is evaluated on the testing set to assess its performance. The evaluation results, including test accuracy and precision, are printed to the console for analysis. Finally, the trained model is saved for integration into the mobile application

5.1.2 Mobile application user interface development

Once the deep learning model is successfully trained and saved, the next step in the implementation process involved developing the mobile application using Kotlin in Android Studio. Kotlin is a modern programming language that offers concise syntax, interoperability with Java, and strong support for Android app development. Android Studio is the official Integrated Development Environment (IDE) for Android app development, providing a rich set of tools and resources for building high-quality mobile applications.

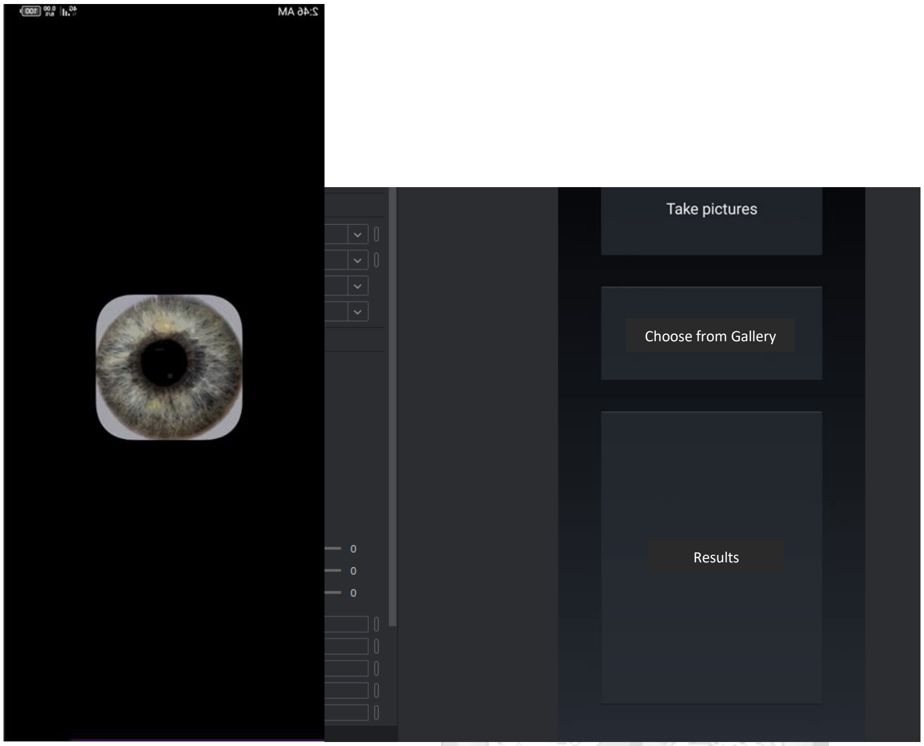


Figure 14 Mobile application UI

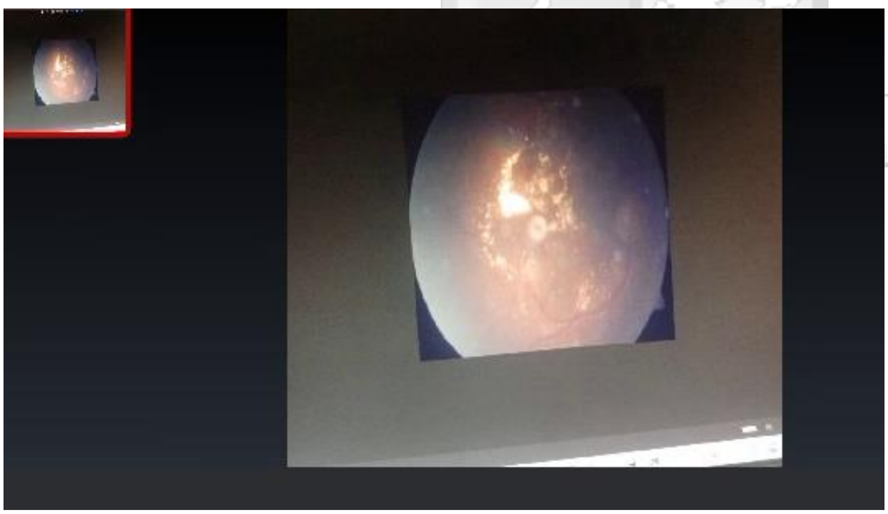


Figure 15 Captured Image

5.1.3 Deep learning model integration

After developing the mobile application using Kotlin in Android Studio, the next step in the implementation process was to integrate the trained deep learning model into the application. This integration enables the mobile application to leverage the power of the deep learning model for real-time detection of diabetic retinopathy from retinal images captured by smartphone-based retinal imaging devices. To accomplish this, TensorFlow Lite, a lightweight version of TensorFlow optimized for mobile and embedded devices, was utilized.

5.2 System Testing

After setting up and developing the system, the researcher conducted system testing to evaluate the performance and effectiveness of the application. This testing phase involved capturing images of a validation set that had not been used during the training process. The validation set served as a crucial component for assessing the model's ability to generalize to unseen data and its performance in real-world scenarios.

The captured images were then processed and analyzed by the deep learning model integrated into the application. The model leveraged its learned features and classification capabilities to detect and classify diabetic retinopathy in the captured images. The results of the classification, including predicted severity levels and diagnostic insights, were recorded and compared against ground truth labels from the validation set.

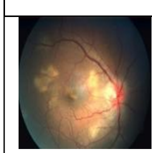
Image	**Model Detected**	**True Value**
	No DR signs	Moderate NPDR
	Moderate NPDR	Moderate NPDR
	No DR signs	No DR signs
	Severe NPDR	Severe NPDR

Figure 16 Testing results

6 Chapter 6: Conclusion

The results of the system testing provide valuable insights into the performance of the developed application for diabetic retinopathy detection. The comparison between the model-detected labels and the true values from the validation set highlights both the strengths and limitations of the application.

In the first image, the model incorrectly classified the retinal image as "No DR signs" when the true value was "Moderate NPDR." This discrepancy underscores the challenge of accurately detecting diabetic retinopathy, especially in cases with subtle or early-stage manifestations. However, the correct classification of the second image as "Moderate NPDR" demonstrates the model's capability to accurately identify diabetic retinopathy when the pathology is more pronounced.

The third image's accurate classification as "No DR signs" by the model aligns with the ground truth label, indicating the application's proficiency in recognizing healthy retinal images. Similarly, the correct identification of "Severe NPDR" in the fourth image showcases the model's effectiveness in detecting advanced stages of diabetic retinopathy.

Overall, while the application demonstrates promising performance in detecting diabetic retinopathy, there is room for improvement, particularly in accurately identifying early-stage manifestations. Further refinement of the model's training data, augmentation techniques, and architectural adjustments may enhance its sensitivity and specificity.

These results underscore the importance of ongoing evaluation, refinement, and validation of the application to ensure its reliability and clinical relevance. Continued collaboration with domain experts, clinicians, and end-users is essential for iteratively improving the application's performance, usability, and impact on patient care.

In addition to the current findings, I would advise to explore the utilization of higher architecture models such as EfficientNetB5 or EfficientNetB7 in future iterations of the application. These larger and more complex models may offer increased capacity and representational power, potentially enhancing the model's ability to capture intricate features and nuances present in retinal images.

By incorporating higher architecture models, the application can potentially improve its sensitivity and specificity in detecting diabetic retinopathy across a wider spectrum of severity levels. The deeper and wider networks of EfficientNetB5 or B7 may better capture subtle abnormalities and variations in retinal morphology, thereby enhancing diagnostic accuracy, particularly in challenging cases or early-stage manifestations.

7 Apendix

7.1 Machine learning code

```
import os

import numpy as np

import tensorflow as tf

from tensorflow.keras.preprocessing.image import ImageDataGenerator

from tensorflow.keras.applications.efficientnet import EfficientNetB0, preprocess_input  # Corrected import

from tensorflow.keras.layers import Dense, GlobalAveragePooling2D, Dropout

from tensorflow.keras.models import Model

from tensorflow.keras.optimizers import Adam

from sklearn.model_selection import train_test_split

from tensorflow.keras.utils import to_categorical

from tensorflow.keras.preprocessing.image import load_img, img_to_array

from tensorflow.keras.regularizers import l2

from tensorflow.keras.callbacks import EarlyStopping


# Setting seed for reproducibility

np.random.seed(42)

tf.random.set_seed(42)


# Directory and file structure

root_folder = 'E:/MyDOCUMENTS/APS IMPORTANT ITEMS/STRATHMORE/SEM~5/Thesis/'

labels_info = {

    'No DR signs': {'folder': 'No DR signs', 'image_range': range(1, 188)},

    'Moderate NPDR': {'folder': 'Moderate NPDR', 'image_range': range(192, 272)},

    'Severe NPDR': {'folder': 'Severe NPDR', 'image_range': range(272, 448)}

}
```

```
# Image dimensions
img_width, img_height = 224, 224

# Data augmentation
datagen = ImageDataGenerator(
    preprocessing_function=preprocess_input,  # Use EfficientNet's own preprocessing
    rotation_range=20,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest'
)

# Function to load images and labels
def load_images_and_labels(labels_info, root_folder):
    images = []
    labels = []
    for label, info in labels_info.items():
        folder = os.path.join(root_folder, info['folder'])
        for img_num in info['image_range']:
            img_path = os.path.join(folder, f"{img_num}.jpg")
            img = load_img(img_path, target_size=(img_width, img_height))
            img = img_to_array(img)
            img = preprocess_input(img)  # Preprocess input
```

```
        images.append(img)
        labels.append(label)
    return np.array(images), np.array(labels)

# Load images and labels
images, labels = load_images_and_labels(labels_info, root_folder)

# Convert labels to numeric and one-hot encode
label_to_id_dict = {label: id for id, label in enumerate(np.unique(labels))}
labels_id = np.array([label_to_id_dict[label] for label in labels])
labels_categorical = to_categorical(labels_id)

# Split the data into 80% training and 20% testing
X_train, X_test, y_train, y_test = train_test_split(images, labels_categorical, test_size=0.2,
random_state=42)

# Define the model
base_model = EfficientNetB0(weights='imagenet', include_top=False, input_shape=(img_width,
img_height, 3))
for layer in base_model.layers[:-20]:  # Freeze all but the last 20 layers
    layer.trainable = False

# Add custom layers on top
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(512, activation='relu', kernel_regularizer=l2(0.001))(x)  # Add L2 regularization
x = Dropout(0.5)(x)
predictions = Dense(len(label_to_id_dict), activation='softmax')(x)
```

```
model = Model(inputs=base_model.input, outputs=predictions)

# Compile the model
model.compile(optimizer=Adam(learning_rate=0.0001),
              loss='categorical_crossentropy',
              metrics=['accuracy', tf.keras.metrics.Precision()])

# Early stopping
early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)

# Train the model
history = model.fit(
    X_train, y_train,
    epochs=10,
    validation_split=0.2,  # 20% of training data used for validation
    batch_size=32,
    callbacks=[early_stopping]
)

# Evaluate the model on the test set
scores = model.evaluate(X_test, y_test, verbose=1)
print(f"Test Accuracy: {scores[1]*100:.2f}%")
print(f"Test Precision: {scores[2]*100:.2f}%")

# Save the model
model.save('dr_efficientnet_model.h5')
print("EfficientNet model training complete and saved.")
```

8 References

1. American Association of Clinical Endocrinologists. (2023). AACE comprehensive clinical practice guidelines for the management of type 2 diabetes mellitus—2023. Endocrine Practice, 29(1), 1e-194. doi:10.1007/s11883-022-02287-w

2. American Diabetes Association. (2023). Standards of medical care in diabetes. Diabetes Care, 46(Suppl. 1), S1-S234. https://doi.org/10.2337/dc23-S001

3. Banday, M. Z., Sameer, A. S., & Nissar, S. (Year). Pathophysiology of diabetes: An overview. Avicenna Journal of Medicine, 10(4), 174–188. https://doi.org/10.4103/ajm.ajm_53_20

4. Bastawrous, A., & Armstrong, D. G. (2023). The Potential Role of Mobile and Artificial Intelligence Technology for Diabetic Retinopathy Screening Around the World. Diabetes care, 46(2), 530–536.

5. Bhattacharya, S. (2023). Deep Learning and Medical Image Processing Techniques for Diabetic Retinopathy: A Survey of Applications, Challenges, and Future Trends. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9911247/

6. Centers for Disease Control and Prevention (CDC). (2023). Type 2 diabetes. Retrieved from https://www.cdc.gov/diabetes/basics/type2.html

7. Centers for Disease Control and Prevention. (2023). What is diabetes? https://www.cdc.gov/diabetes/basics/diabetes.html

8. Wong, T. Y., Cheung, C. M. G., Larsen, M., Sharma, S., & Simó, R. (2014). Diabetic retinopathy. Nature Reviews Disease Primers, 2(1). https://doi.org/10.1038/nrdp.2016.12

9. Ciulla, T. A., Bracha, P., & Ferris, F. L. (2023). Diabetic Macular Edema: Pathophysiology, Diagnosis, and Treatment. Ophthalmology. https://doi.org/10.1016/j.ophtha.2022.06.010

10. Frank, R. N. (2022). Medical imaging in diabetes: A review of current and emerging modalities. Diabetes Spectrum, 35(2), 173-180.

11. Gallego-Pinazo, R., Dolz-Marco, R., Gómez-Ulla, F., Mrejen, S., & Freund, K. B. (2014). Retinal Angiography in Ocular Diseases. JOVE (Journal of Visualized Experiments), (93), e51971.

12. Haleem, M. S,, et al. (2022). Advances in telemedicine for ophthalmic care during the COVID-19 pandemic and beyond. Journal of Telemedicine and Telecare, 28(6), 363-371.

13. Hamdi, T., et al. (2017). Artificial neural network for blood glucose level prediction. In 2017 International Conference on Smart, Monitored and Controlled Cities (SM2C) (pp. 431-436). IEEE.

14. International Diabetes Federation. (2023). Diabetes Atlas. https://www.idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html

15. Ishibashi, F., Kosaka, A., & Tavakoli, M. (2021). The impact of glycemic control on retinal photoreceptor layers and retinal pigment epithelium in patients with type 2 diabetes without diabetic retinopathy: A follow-up study. Frontiers in Endocrinology, 12, 614161.

16. Joslin Diabetes Center. (2023). Joslin guide to diabetes: Type 2 diabetes. Seventh Edition. Joslin Diabetes Center.

17. Karakayac, 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

18. Kovatchev, B. P., & Breton, M. C. (2020). Continuous glucose monitoring in type 2 diabetes. Diabetes Care, 43(11), 2551-2559.

19. Lains, I., et al. (2019). Optical coherence tomography angiography in diabetic retinopathy. Eye .

20. Mayo Clinic. (2023). Hyperglycemia. Retrieved from https://www.mayoclinic.org/diseases-conditions/hyperglycemia/symptoms-causes/syc-20373631

21. National Eye Institute. (2023). Diabetic retinopathy. Retrieved September 22, 2023, from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy

22. Raj, A., Kumar, P., Singh, N., & Sharma, B. (2023). Automated detection of diabetic retinopathy using deep learning and ultra-widefield images. *Journal of Ophthalmic Imaging*, 52(2), 150-158

23. Rajalakshmi, R., et al. (2022). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. Eye. https://doi.org/10.1038/s41433-018-0064-9

24. Sahlgrenska University Hospital, (2023). Diabetic Retinopathy.

25. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning, 97, 6105–6114.

26. Tsiknakis, N., et al. (2021). Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. Computers in Biology and Medicine, 135, 104599. https://doi.org/10.1016/j.compbiomed.2021.104599

27. V. E. Castillo Benítez, I. Castro Matto, J. C. Mello Román, J. L. Vázquez Noguera, M. García-Torres, J. Ayala, D. P. Pinto-Roa, P. E. Gardel-Sotomayor, J. Facon, and S. A. Grillo, **Dataset from fundus images for the study of diabetic retinopathy**, Data in Brief, vol. 36, p. 107068, Jun. 2021. doi: https://doi.org/10.1016/j.dib.2021.107068

28. Project Group. (2003). Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology, 110(9), 1677-1682. https://doi.org/10.1016/S0161-6420(03)00475-5

29. Wong, T. Y., Cheung, C. M. G., Larsen, M., Sharma, S., & Simó, R. (2014). Diabetic retinopathy. Nature Reviews Disease Primers, 2(1). https://doi.org/10.1038/nrdp.2016.12

30. World Health Organization (WHO). (2023). Diabetes. Retrieved from https://www.who.int/health-topics/diabetes

31. Yadav, M., Goel, R., & Rajeswari, D. (2021). A deep learning based diabetic retinopathy detection from retinal images. In 2021 International Conference on Intelligent Technologies (CONIT) (pp. 431-436). IEEE.

32. Yadav, R., Singh, D., Singh, S., & Kumar, A. (2022). Medical imaging in the diagnosis and management of diabetes mellitus: A review

33. Yun, J.-S., et al. (2021). A Deep Learning Model for Screening Type 2 Diabetes from Retinal Photographs (Short running title: Screening Diabetes using Deep Learning Model). medRxiv. https://doi.org/10.1101/2021.06.29.21259606

34. Yau, J. W., et al. (2021). Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care.

35. Zeoli, A. M., et al. (2017). Smartphone applications and mobile health solutions for the prevention and management of type 2 diabetes mellitus: A systematic review. Journal of Diabetes Science and Technology, 11(6), 1163-1175.

36. Zhang, X., Wang, X., Zhao, R., & Li, W. (2022). Interpretable deep learning for diabetic retinopathy diagnosis based on biomarker activation map. *arXiv preprint arXiv:2212.06299*. https://arxiv.org/abs/2212.06299