

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
ABU BAKR BELKAID UNIVERSITY - TLEMSEN
FACULTY OF SCIENCES
COMPUTER SCIENCE DEPARTMENT

FINAL YEAR PROJECT THESIS

For the attainment of the Master's degree in Computer Science

Specialty : Intelligent models and decision (M.I.D)

Theme

Early Detection of Diabetic Retinopathy
Through Deep Neural Networks

Realized by :

- BERGOUG Ilyes

Presented 26/06/2024 in front of the jury composed of :

Dr. BENAZZOUZ Mourtada. Supervisor.

Dr. MEZIANE-TANI Souad. President.

Dr. BRIKCI-NIGASSA Amine. Examiner.

Acknowledgements

First and foremost, I am profoundly grateful to **Allah**, the Almighty, for His blessings and guidance throughout this journey.

To my family, your unwavering love and support have been my anchor. I am deeply grateful to my parents for their constant encouragement and to my siblings for their understanding and patience.

I would like to express my deepest appreciation to my supervisor, **Mourtada BENZAOUZ**, for his invaluable guidance, support, and encouragement throughout the course of my research.

I extend my heartfelt thanks to my professors, **Souad MEZIANE-TANI**, and **Amine BRIKCI-NIGASSA** for their constructive feedback and assistance.

Table of Content

List of Figures.....	V
List of Tables	VI
Acronyms	VII
Introduction.....	1
Context	1
Problem Statement.....	1
Contribution	1
Chapter 1	2
Diabetic Retinopathy	2
1.1 Introduction	3
1.2 Diabetes Mellitus (DM).....	3
1.3 Types of Diabetes.....	3
• 1.3.1 Type 1 Diabetes.....	3
• 1.3.2 Type 2 Diabetes.....	4
• 1.3.3 Gestational Diabetes.....	4
1.4 What is Diabetic Retinopathy ?	4
1.5 Epidemiology of Diabetic Retinopathy.....	5
1.6 Pathophysiology of Diabetic Retinopathy	5
• 1.6.1 Hyperglycemia and Retinal Microvasculopathy.....	5
• 1.6.2 Inflammation.....	6
• 1.6.3 Retinal Neurodegeneration	7
1.7 Classification of Diabetic Retinopathy	8
• 1.7.1 Non Proliferative Diabetic Retinopathy (NPDR)	8
• 1.7.2 Proliferative Diabetic Retinopathy (PDR).....	8
1.8 Stages of Diabetic Retinopathy	9
• 1.8.1 Mild Nonproliferative Retinopathy.....	9
• 1.8.2 Moderate Nonproliferative Retinopathy	9
• 1.8.3 Severe Nonproliferative Retinopathy	9
• 1.8.4 Proliferative Retinopathy.....	9
1.9 Symptoms and Diagnosis Techniques	10
• 1.9.1 Symptoms	10
• 1.9.2 Diagnosis Techniques.....	10
1.10 Diabetic Retinopathy Complications.....	11

1.11 Treatment.....	11
• 1.11.1 Early Diabetic Retinopathy.....	11
• 1.11.2 Advanced Diabetic Retinopathy.....	12
1.12 Conclusion.....	13
Chapter 2	14
Deep Learning	14
2.1 Artificial Intelligence (AI)	15
2.2 Machine Learning (ML)	15
2.3 Deep Learning (DL)	15
2.4 Why Deep Learning in Today’s Research and Applications?	16
2.5 Understanding various forms of data.....	16
2.6 Convolutional Neural Networks	16
• 2.6.1 Convolutional Layers.....	17
• 2.6.2 Activation Layers.....	18
• 2.6.3 Pooling Layers	19
• 2.6.4 Flattening Layers.....	19
• 2.6.5 Dense Layers	20
• 2.6.6 Dropout Layer	20
• 2.6.7 Regularization Layers	20
• 2.6.8 Batch Normalization Layers	20
2.7 Transfer Learning	20
2.8 Data Augmentation	21
2.9 Fine-tuning.....	22
2.10 Architectures of Convolutional Neural Networks.....	23
• 2.10.1 VGG Network Architecture	23
• 2.10.2 ResNet Network Architecture.....	23
• 2.10.3 GoogLeNet Network Architecture	24
• 2.10.4 AlexNet Network Architecture.....	24
• 2.10.5 DenseNet Network Architecture	25
• 2.10.6 Xception Network Architecture	25
2.11 Conclusion.....	25
Chapter 3	26
Implementation and evaluation	26
3.1 Introduction	27
3.2 Tools and Configuration.....	27

• 3.2.1 Collaboratory	27
• 3.2.2 Software and Libraries	28
3.3 Datasets	28
• 3.3.1 First Dataset (MuReD).....	28
• 3.3.2 Second Dataset (RFMiD 2.0)	30
• 3.3.3 Third Dataset.....	30
3.4 Approaches and techniques.....	31
• VGG16 Model.....	31
• ResNet50 Model	35
• Custom Model.....	39
3.5 Deployment	43
3.6 Conclusion and interpretation	44
Conclusion and future perspectives	45
Bibliography	46
Abstract.....	48
Résumé.....	48
المُلخَص.....	48

List of Figures

Figure 1. The prevalence of Diabetes worldwide documented in 2021, along with prognoses for 2045.3

Figure 2. difference between normal retina and diabetic retinopathy.5

Figure 3. A visual representation of the pathophysiology of DR.8

Figure 4. Stages of Diabetic Retinopathy.9

Figure 5. An illustration of the relation between AI,ML and DL.15

Figure 6. An example of a Convolutional Neural Network including multiple convolution and pooling layers.17

Figure 7. Structure of a convolutional Neural Network (CNN).17

Figure 8. An example of a convolutional operation with a kernel size of 3×3 , no padding, and a stride of 1.18

Figure 9. Plot of different activation functions : (a) ReLU activation function; (b) Sigmoid activation function; and (c) Tanh activation function.19

Figure 10. a: An example of max pooling operation with a filter size of 2×2 , no padding, and a stride of 2. b: Examples of the max pooling operation on images. Note that images in the upper row are downsampled by a factor of 2, from 26×26 to 13×1319

Figure 11. Flattening 2D feature maps to 1D vector.20

Figure 12. Transfer Learning Process.21

Figure 13. Convolutional Neural Network Architecture.21

Figure 14. Image data augmentation techniques.22

Figure 15. VGG16 Architecture.23

Figure 16. Inception blocks. (a) InceptionV1. (b) InceptionV3.24

Figure 17. AlexNet Architecture.25

Figure 18. Class Distribution of DR in MuReD Dataset.29

Figure 19. Class distribution of DR in RFMiD 2.0 Dataset.30

Figure 20. Performance of VGG16 on the MuReD dataset.31

Figure 21. Classification Report of the Test Set in MuReD Dataset.32

Figure 22. Confusion Matrix of the Test Set in MuReD Dataset.32

Figure 23. Performance of VGG16 on the MuReD dataset after applying different techniques.33

Figure 24. Classification Report of the Test Set after applying different techniques.33

Figure 25. Confusion Matrix of the Test Set after applying different techniques.34

Figure 26. Performance of VGG16 on the RFMiD 2.0 dataset.34

Figure 27. Classification Report of the test set in RFMiD 2.0 Dataset.35

Figure 28. Confusion Matrix of the test set in RFMiD 2.0 Dataset.35

Figure 29. Performance of ResNet50 on the MuReD dataset before fine-tuning.36

Figure 30. Performance of ResNet50 on the MuReD dataset after fine-tuning.36

Figure 31. Classification Report of the test set in MuReD Dataset/ResNet50.37

Figure 32. Confusion Matrix of the test set in MuReD Dataset/ResNet50.37

Figure 33. Performance of ResNet50 on the third dataset before fine-tuning.38

Figure 34. Performance of ResNet50 on the third dataset after fine-tuning.38

Figure 35. Classification Report of the test set in the third Dataset/ResNet50.39

Figure 36. Confusion Matrix of the test set in the third Dataset/Resnet50.39

Figure 37. Architecture of the custom model.40

Figure 38. performance of the custom model on the third dataset.....	40
Figure 39. Classification Report of the training set in the third Dataset using Custom model.	41
Figure 40. Confusion Matrix of the training set in the third Dataset using Custom model.....	41
Figure 41. Classification Report of the validation set in the third Dataset using Custom model.....	41
Figure 42. Confusion Matrix of the validation set in the third Dataset using Custom model.	41
Figure 43. Classification Report of the test set in the third Dataset using Custom model..	42
Figure 44. Confusion Matrix of the test set in the third Dataset using Custom model.....	42
Figure 45. Illustration of our web application.	43
Figure 46. Model prediction of the uploaded retinal image.....	44

List of Tables

Table 1. configuration of the virtual machine provided by Google Colab.	27
Table 2. Software and libraries used during the research.	28
Table 3. A detailed breakdown of each column in the MuReD dataset.	29
Table 4. Accuracy of different models trained on various datasets.	42
Table 5. Diabetic Retinopathy class precision for each model	43

Acronyms

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
ReLU	Rectified linear unit
CPU	Central Processing Unit
GPU	Graphics Processing Unit
TPU	Tensor Processing Unit
RAM	Random-access memory
FC	Fully Connected
TL	Transfer Learning
VGGNet	Visual Geometry Group Network
ResNet	Residual Network
MuReD	Multi Label Retinal Disease
RFMiD	Retinal Fundus Multi-disease
DR	Diabetic Retinopathy
DM	Diabetes Mellitus
PDR	Proliferative Diabetic Retinopathy
NPDR	Non Proliferative Diabetic Retinopathy

Introduction

Context

Diabetic retinopathy represents a major complication of diabetes, affecting the blood vessels of the retina and potentially leading to vision loss if not diagnosed early. This condition arises when prolonged periods of high blood sugar cause damage to the tiny blood vessels within the retina, leading to leakage, swelling, and ultimately, the growth of new, abnormal blood vessels. The progression of diabetic retinopathy can be silent and gradual, often going unnoticed until significant damage has occurred. In the current medical context, where the prevalence of diabetes is rapidly increasing worldwide, the early detection of this condition is essential. Early diagnosis and timely treatment can prevent severe visual impairment and improve patients' quality of life. Regular screening and innovative diagnostic tools are critical components in managing the impact of diabetic retinopathy, highlighting the need for advancements in medical technology and practices.

Problem Statement

The central issue of our study lies in the design and implementation of a decision support system based on deep learning for the early detection of diabetic retinopathy. This challenge involves not only the collection and processing of massive amounts of medical data but also the development of sophisticated algorithms capable of efficiently analyzing and interpreting this data to provide rapid and accurate diagnoses. Traditional methods of diagnosing diabetic retinopathy, which rely heavily on manual examination of retinal images by trained specialists, are time-consuming and subject to human error. With the increasing availability of digital retinal imaging and advancements in artificial intelligence, there is a significant opportunity to improve diagnostic accuracy and efficiency. Our problem statement addresses the need for an automated, reliable system that can assist healthcare professionals in identifying early signs of diabetic retinopathy, thus enabling prompt intervention and better patient outcomes.

Contribution

We propose an approach based on the integration of medical imaging and artificial intelligence for the early detection of diabetic retinopathy. Our solution involves developing a deep learning model specifically trained to analyze retinal images and detect the early signs of diabetic retinopathy with high precision and sensitivity. This model will leverage large datasets of retinal images, annotated by medical experts, to learn the subtle patterns and markers indicative of the disease. However we faced several challenges during the development process, including class imbalance. In medical datasets, especially for conditions like diabetic retinopathy, images of healthy retinas often outnumber those with signs of the disease. This imbalance can lead to a biased model that performs well on the majority class but poorly on the minority class. Techniques such as image augmentation and fine-tuning were employed to ensure the robustness of our model across diverse image qualities and conditions. By combining these approaches, our contribution aims to provide healthcare professionals with an effective and reliable tool for the early screening of diabetic retinopathy, ultimately contributing to better management and prevention of this serious diabetes complication.

Chapter 1

Diabetic Retinopathy

1.1 Introduction

In this chapter we will introduce diabetes and its types, followed by an exploration of the epidemiology, pathophysiology, classification, and stages of diabetic retinopathy. We will discuss its symptoms, diagnostic methods, potential complications, and available treatments.

1.2 Diabetes Mellitus (DM)

Diabetes Mellitus is a persistent condition characterized by high blood sugar levels and disruption of fat and protein metabolism. The increase in blood sugar occurs because it is not properly metabolized in the cells due to insufficient insulin production by the pancreas or the cells' inability to use the insulin produced effectively. There are three primary types of diabetes: Type 1, where the pancreas fails to produce insulin; Type 2, where the body's cells are resistant to insulin and insulin production decreases over time; and gestational diabetes, which occurs during pregnancy and can lead to complications. Additionally, there are two other forms of glucose intolerance - impaired fasting glucose (IFG) and impaired fasting glycemia (IGT) - which are intermediate conditions between normal and diabetic blood glucose levels. People with IFG and IGT have an increased risk of cardiovascular disease compared to those with normal blood glucose levels. [1]

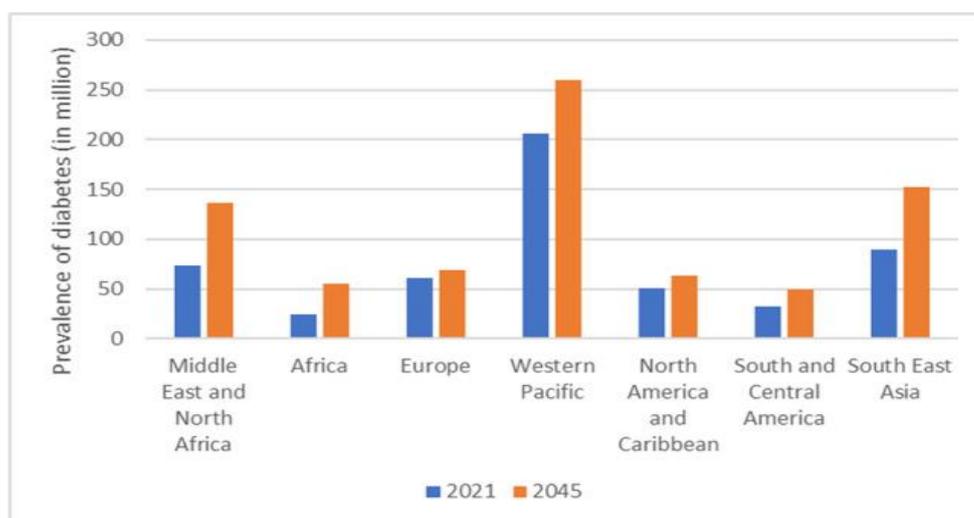


Figure 1. The prevalence of Diabetes worldwide documented in 2021, along with prognoses for 2045. [10]

1.3 Types of Diabetes

• 1.3.1 Type 1 Diabetes

Type 1 diabetes mellitus, also known as autoimmune diabetes, is a long-term condition characterized by a lack of insulin due to the loss of pancreatic β -cells, leading to high blood sugar levels. While symptoms typically appear during childhood or adolescence, they can also manifest later in life. The exact cause of the condition is not fully understood, but it is

believed to involve the destruction of β -cells by T cells. Certain autoantibodies targeting specific proteins associated with β -cells can be detected months to years before symptoms develop, helping identify individuals at risk of developing Type 1 diabetes. The progression of the disease can be divided into three stages based on the presence of hyperglycemia and associated symptoms. Although there is no cure, lifelong insulin injections are necessary for patients, and alternative insulin treatments like pumps and continuous glucose monitoring are being explored. While improved glycemic control has reduced the risk of complications, there is still a need for extensive research to achieve early diagnosis, prevent β -cell loss, and enhance treatment options for better quality of life and prognosis for those with Type 1 diabetes. [2]

- **1.3.2 Type 2 Diabetes**

Type 2 diabetes mellitus (T2DM) is a significant global health issue associated with the rise in obesity. People with T2DM face a high risk of both small blood vessel complications (like retinopathy, nephropathy, and neuropathy) and large blood vessel complications (such as heart-related conditions) due to high blood sugar and various components of the insulin resistance syndrome. Environmental and genetic factors both contribute to the disturbances that lead to impaired glucose regulation in T2DM. While insulin resistance and impaired insulin secretion are the central issues in T2DM, there are at least six other abnormalities that affect glucose metabolism. Because of the numerous disturbances in T2DM, using a combination of multiple antidiabetic medications will be necessary to maintain normal blood sugar levels. The treatment should not only be effective and safe but also improve the patient's quality of life. While there are several new medications in the works, the primary need is for treatments that improve insulin sensitivity, stop the ongoing failure of pancreatic β -cells, and prevent or reverse small blood vessel complications. [3]

- **1.3.3 Gestational Diabetes**

During pregnancy, gestational diabetes mellitus (GDM) is identified as varying levels of glucose intolerance. It is typically identified through screening pregnant women for risk factors and, for those at risk, testing for mild and often symptomless abnormal glucose tolerance. GDM seems to stem from similar physiological and genetic irregularities as diabetes outside of pregnancy. Women with GDM are at a high risk of developing diabetes when not pregnant, making it an important opportunity to research the early stages of diabetes and develop preventive measures. [4]

1.4 What is Diabetic Retinopathy ?

Diabetic retinopathy is a disease that can result in vision loss and blindness for people with diabetes. It affects the blood vessels in the retina, which is the light-sensitive tissue located at the back of the eye. [5]

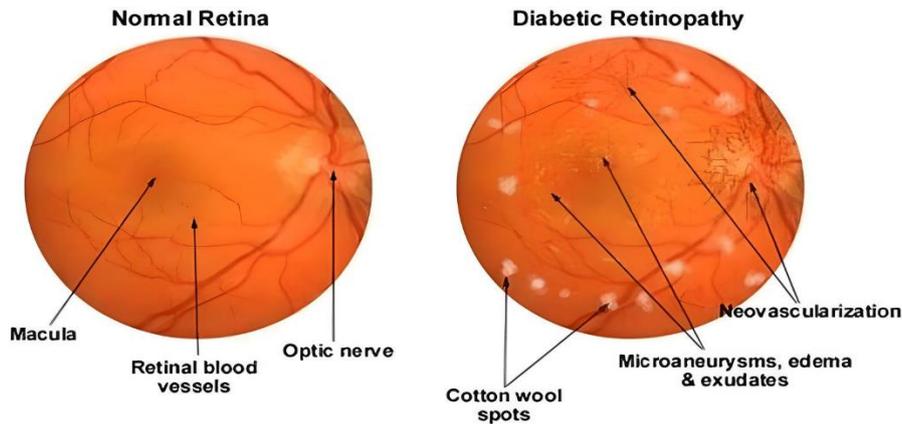


Figure 2. difference between normal retina and diabetic retinopathy. [11]

1.5 Epidemiology of Diabetic Retinopathy

Diabetes mellitus encompasses a group of metabolic conditions and is mainly categorized into type 1 and type 2 diabetes based on its mechanisms. It is linked with a wide range of macrovascular and microvascular complications, including retinopathy, neuropathy, and nephropathy. A meta-analysis in 2010 revealed that diabetic retinopathy (DR) caused blindness and visual impairment in a significant number of people worldwide. The prevalence of DR increases with the severity and duration of diabetes. Studies have shown that the prevalence of DR is higher in individuals with a shorter duration of diabetes and in those with higher levels of hemoglobin A1c, prolonged diabetes, greater systolic blood pressure, and frequent use of insulin. Additionally, the prevalence of DR is higher in diabetic men than in women. The most severe complications of DR include diabetic macular edema (DME) and proliferative retinopathy (PR). The prevalence of DR and its complications is higher among insulin-dependent diabetes mellitus (IDDM) patients compared to non-insulin-dependent diabetes mellitus (NIDDM) patients. Furthermore, the prevalence of DR varies significantly across different geographical regions, with a higher prevalence in developing countries and urban areas. The increase in prevalence is attributed to changes in lifestyle, including unhealthy diets, reduced physical activity, obesity, hypertension, and high cholesterol. The prevalence of DR is significantly higher in urban regions compared to rural areas in countries such as India and China. These findings highlight the global concern of diabetic retinopathy and emphasize the need for improved diagnostic, screening, and preventive strategies. [8]

1.6 Pathophysiology of Diabetic Retinopathy

Diabetic retinopathy develops due to a chronic imbalance in blood sugar levels. Here are the main pathophysiological mechanisms involved:

- **1.6.1 Hyperglycemia and Retinal Microvasculopathy**

DR has been known as a microvascular disease for a long time. High blood sugar levels are thought to be a major factor in causing damage to the small blood vessels in the retina.

Several metabolic pathways, including the polyol pathway, accumulation of advanced glycation end products (AGEs), the protein kinase C (PKC) pathway, and the hexosamine pathway, have been implicated in the damage caused by high blood sugar levels. When exposed to high blood sugar levels, the earliest responses of the retinal blood vessels include widening of the blood vessels and changes in blood flow. These changes are considered to be a form of self-regulation to increase retinal metabolism in individuals with diabetes. The loss of pericytes is another significant early event in the development of DR. Studies have shown evidence of pericyte cell death triggered by high glucose levels, both in laboratory experiments and in living organisms. Since pericytes provide structural support for capillaries, their loss leads to localized bulging of capillary walls, which is associated with the formation of microaneurysms, the earliest clinical sign of DR. In addition to pericyte loss, the apoptosis of endothelial cells and thickening of the basement membrane are also observed during the development of DR, collectively contributing to the impairment of the blood-retinal barrier (BRB). Furthermore, significant loss of pericytes and endothelial cells leads to the blockage of capillaries and reduced blood supply. Activation of hypoxia-inducible factor 1 (HIF-1) due to retinal ischemia or oxygen deprivation leads to an elevated synthesis of vascular endothelial growth factor (VEGF). Other evidence suggests that elevated levels of phospholipase A2 (PLA2) under diabetic conditions also trigger an increase in VEGF levels. VEGF, a crucial factor in the progression of proliferative diabetic retinopathy (PDR) and diabetic macular edema (DME), is believed to increase vascular permeability by causing phosphorylation of tight junction proteins such as occludin and zonula occludens-1 (ZO-1). In addition, as an angiogenic factor, VEGF promotes the proliferation of endothelial cells through the activation of mitogen-activated protein (MAP). Increased expression of VEGF has been observed in the retina of diabetic mice, as well as in the vitreous of patients with DME and PDR. Angiopoietins (Ang-1, Ang-2), along with other factors, are involved in the regulation of vascular permeability through their interaction with the endothelial receptor tyrosine kinase Tie2. Ang-2, an antagonist of Tie2, has been shown to increase vascular leakage in the retinas of diabetic rats. It is speculated that angiogenic factors other than VEGF may also be involved in the changes to the microvasculature during DR, potentially offering new targets for therapeutic intervention. [9]

- **1.6.2 Inflammation**

The role of inflammation is crucial in the development of diabetic retinopathy (DR). Chronic low-grade inflammation has been observed in various stages of DR in both diabetic animal models and patients. Leukostasis, the adherence of white blood cells to the blood vessel walls, has been identified as a significant process in the early stages of DR. In a study from 1991, Schröder et al. initially reported the blockage of retinal microvasculature by monocytes and granulocytes in diabetic rats induced with streptozotocin (STZ). Increased adherence of leukocytes was noted in the retinal vasculature as early as three days after the onset of diabetes in rats. The researchers also found that the increased leukostasis was spatially associated with damage to the endothelium and impairment of the blood-retinal barrier (BRB) in diabetic rats. Subsequent studies revealed that leukostasis contributed to the loss of endothelial cells and the breakdown of the BRB through the Fas (CD95)/Fas-

ligand pathway. The adhesion of leukocytes to the endothelium, mediated by adhesion molecules, has been implicated in leukostasis in diabetes. Increased leukocyte adhesion and upregulated expression of leukocyte b2-integrins CD11a, CD11b, and CD18 were reported in diabetic rats and patients. Additionally, endothelial cell adhesion molecules such as intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule (VCAM)-1, and selectins (E-selectin) were found to be increased in diabetic animals and patients. The expression of VCAM-1 and E-selectin in the plasma of patients is correlated with the severity of DR. Genetic deficiency of CD18 or ICAM-1 resulted in significantly reduced adherent leukocytes. Inhibition of CD18 or ICAM-1 with anti-CD18 F(ab9)2 fragments or antibody decreased retinal leukostasis and vascular lesions in diabetic rats. Chemokines, which regulate the attraction and activation of leukocytes, have also been shown to be involved in the pathogenesis of DR. Chemokines such as monocyte chemoattractant protein-1 (MCP-1), macrophage inflammatory protein-1 α (MIP-1 α), and MIP-1 β have been reported to be elevated in diabetic patients. MCP-1 deficiency leads to reduced retinal vascular leakage in diabetic mice. Furthermore, inflammatory cytokines such as tumor necrosis factor alpha (TNF- α), interleukin 6 (IL-6), IL-8, and IL-1 β were significantly upregulated in diabetic patients, and their expression level was correlated with the severity of DR. Dysfunction of retinal glial cells is also presumed to be involved in the initiation and amplification of retinal inflammation in DR. Glial cells in the retina, including astrocytes, Müller cells, and microglia, are responsible for providing structural support and maintaining homeostasis in the retina. Under hyperglycemic stress, microglia is activated, followed by increased secretion of TNF- α , IL-6, MCP-1, and VEGF. Later involvement of Müller cells and astrocytes is associated with the amplification of inflammation responses by producing proinflammatory cytokines. [9]

- **1.6.3 Retinal Neurodegeneration**

Retinal neurodegeneration is one of the early events in the progression of DR. In diabetic rats, the apoptosis of retinal neurons can be observed as early as one month after the induction of diabetes. Upregulation of pro-apoptotic molecules such as cleaved caspase-3, Bax, and Fas has been detected in retinal neurons of diabetic animals and individuals. Mitochondrial dysfunction has been associated with retinal degeneration in DR. In the eyes of diabetic individuals, increased retinal expression of pro-apoptotic mitochondrial proteins such as cytochrome c and apoptosis-inducing factor (AIF) has been found. Studies conducted in vitro have shown that exposure to high glucose is linked to increased mitochondrial fragmentation and cell apoptosis. Besides mitochondrial damage, the involvement of oxidative stress in diabetes-induced retinal degeneration has also been widely investigated. In the diabetic mouse retina, there is a significant increase in the generation of reactive oxygen species (ROS). The suppression of ROS generation effectively inhibited visual impairment and caspase-3-mediated retinal neuronal apoptosis. There is mounting evidence that retinal neurodegeneration may be an independent pathophysiology of DR. In a mouse model of diabetes, loss of ganglion cells and reduction in retinal thickness were observed before the presence of microvascular alterations. In diabetic patients, inner retinal thinning was detected even in the absence of DR or with minimal DR (microaneurysms). Therefore, further investigation of the molecular

mechanisms underlying retinal neurodegeneration may provide potential therapeutic targets for early intervention in DR. [9]

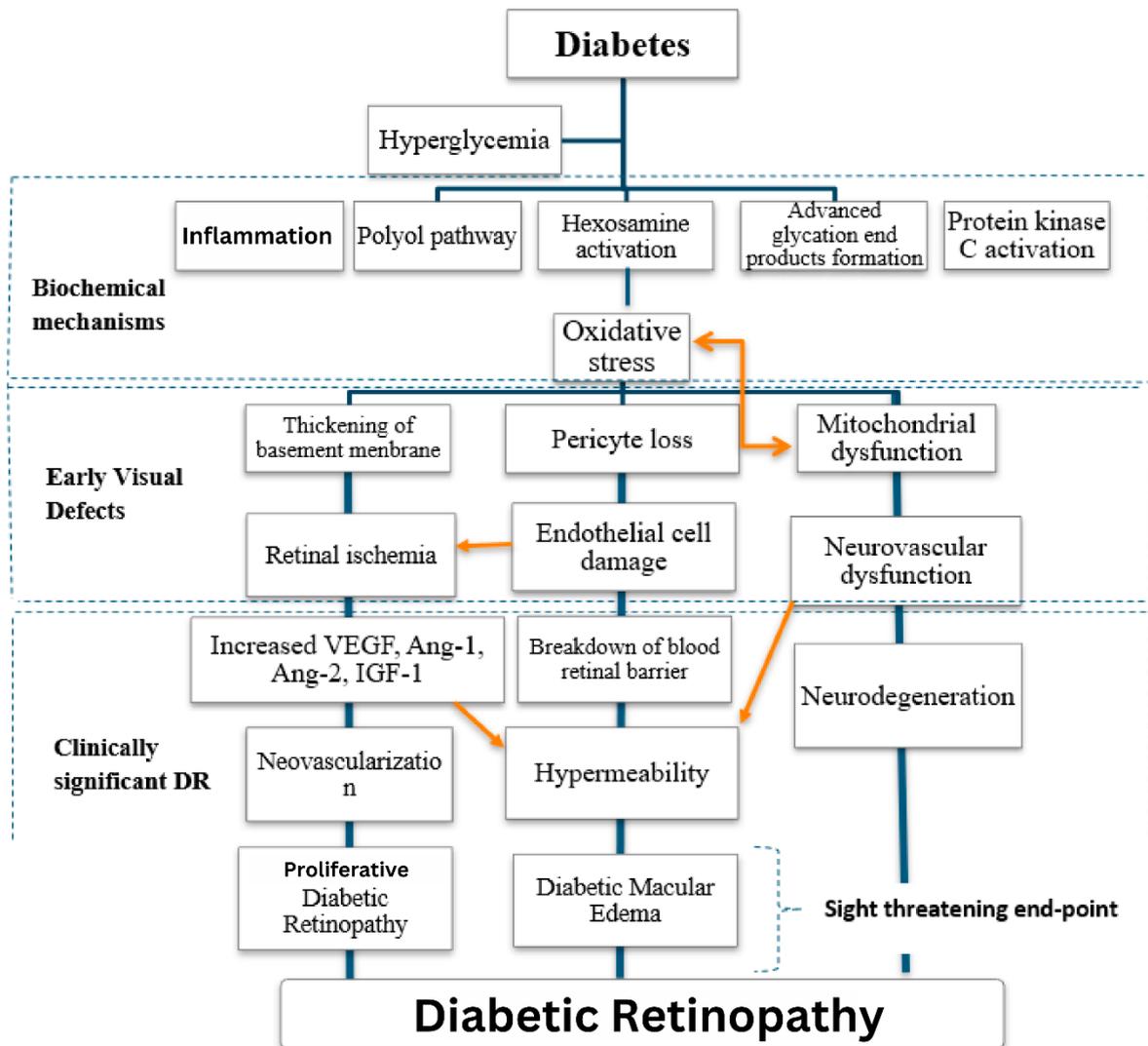


Figure 3. A visual representation of the pathophysiology of DR. [8]

1.7 Classification of Diabetic Retinopathy

• 1.7.1 Non Proliferative Diabetic Retinopathy (NPDR)

In the early stages of the disease known as NPDR, symptoms are usually mild or may not be present at all. NPDR is characterized by weakened blood vessels in the retina, which can result in the formation of small bulges called microaneurysms. These microaneurysms have the potential to leak fluid into the retina, leading to macular swelling. [7]

• 1.7.2 Proliferative Diabetic Retinopathy (PDR)

At the advanced stage of the disease known as PDR, poor circulation leads to a lack of oxygen in the retina. This can result in the growth of delicate new blood vessels in the retina and vitreous, the gel-like fluid at the back of the eye. These new blood vessels may leak

blood into the vitreous, leading to vision impairment. PDR can also lead to complications such as retinal detachment due to scar tissue formation and the development of glaucoma, a progressive eye disease that causes damage to the optic nerve. In cases of proliferative diabetic retinopathy, the nerve damage is caused by very high pressure in the eye. Without proper treatment, proliferative diabetic retinopathy can lead to significant vision loss and potential blindness. [7]

1.8 Stages of Diabetic Retinopathy

- **1.8.1 Mild Nonproliferative Retinopathy**

In the initial phase, microaneurysms develop, which are tiny swollen areas resembling balloons in the small blood vessels of the retina. [7]

- **1.8.2 Moderate Nonproliferative Retinopathy**

As the illness advances, certain blood vessels that supply the retina become obstructed.[7]

- **1.8.3 Severe Nonproliferative Retinopathy**

Numerous blood vessels are obstructed, leading to a lack of blood supply in several parts of the retina. These areas then signal the body to generate new blood vessels for nourishment.[7]

- **1.8.4 Proliferative Retinopathy**

In the advanced stage of this condition, the retina sends signals for nourishment that lead to the development of new blood vessels, a condition known as proliferative retinopathy. These new blood vessels are abnormal and delicate, growing along the retina and the surface of the clear, vitreous gel within the eye. While they don't cause symptoms or vision loss on their own, their fragile walls can lead to blood leakage, resulting in severe vision loss and potential blindness.[7]

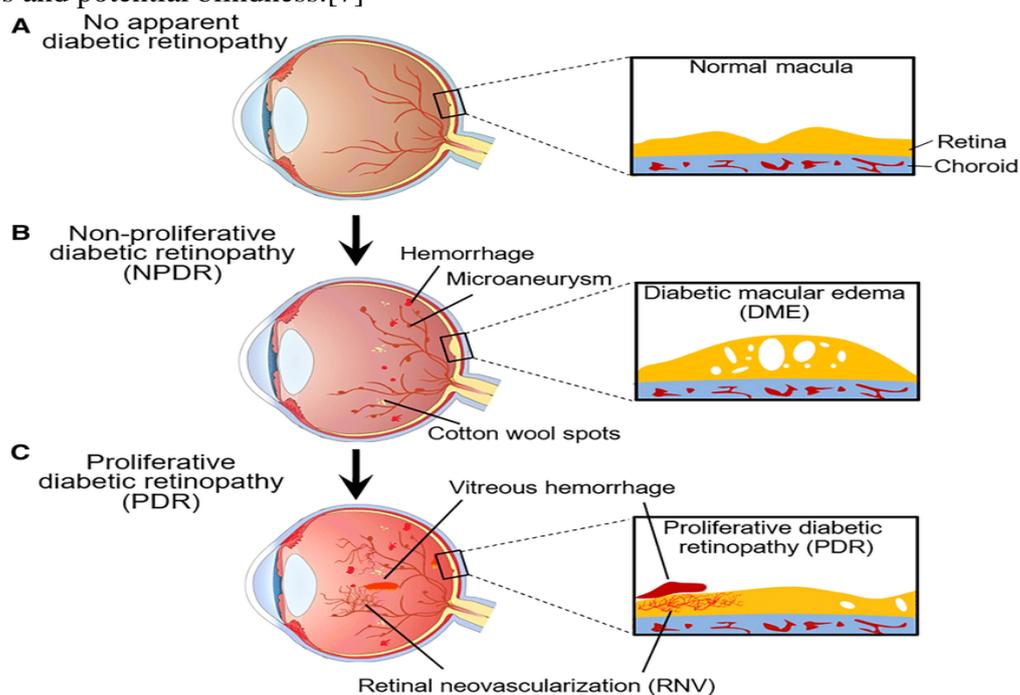


Figure 4. Stages of Diabetic Retinopathy. [23]

1.9 Symptoms and Diagnosis Techniques

- **1.9.1 Symptoms**

As diabetic retinopathy advances, symptoms may involve floaters or dark strings in your vision, blurred vision, fluctuating vision, dark or empty spots in your vision, vision loss, and difficulty perceiving colors. This condition typically impacts both eyes. [7]

Diabetic retinopathy can be categorized as either early or advanced, based on the signs and symptoms you experience.[7]

- **Early diabetic Retinopathy**

This particular form of diabetic retinopathy is known as nonproliferative diabetic retinopathy (NPDR). It's named as such because, at this stage, new blood vessels are not growing. NPDR can be categorized as mild, moderate, or severe. In NPDR, the blood vessel walls in the retina become weakened. Small bulges, called microaneurysms, protrude from the vessel walls and may sometimes leak or release fluid and blood into the retina. As the condition advances, the smaller vessels may close, and the larger retinal vessels may start to expand and become irregular in size. The retina's nerve fibers may experience swelling as well. At times, the central part of the retina (macula) may also start to swell, a condition known as macular edema.[7]

- **Advanced Diabetic Retinopathy**

The most advanced stage of diabetic retinopathy is known as proliferative diabetic retinopathy (PDR). It is termed "proliferative" because new abnormal blood vessels start growing in the retina. These unusual blood vessels have the potential to either grow or leak into the vitreous humor, which is the jelly-like substance found in the center of the eye. Over time, the growth of new blood vessels can lead to scar tissue formation, potentially causing the retina to detach from the back of the eye. Additionally, if the new blood vessels disrupt the normal fluid flow out of the eye, it can increase pressure within the eyeball, leading to glaucoma. This increased pressure can also damage the optic nerve. The risk of developing diabetic retinopathy is heightened by various factors in individuals with diabetes. These factors include the duration of the disease, suboptimal blood sugar control, hypertension, elevated cholesterol levels, pregnancy, and tobacco usage. [7]

- **1.9.2 Diagnosis Techniques**

Diabetic individuals, as well as their family members, friends, and medical professionals, should be educated about the importance of regular eye screenings to detect diabetic retinopathy (DR) at an early stage . Various diagnostic techniques, such as direct and indirect ophthalmoscopy, stereoscopic digital and fundus photography, ultrawide-field fundus fluorescein angiography (UWFA), optic coherence tomography (OCT), and optic coherence tomography-angiography (OCT-A), are used to diagnose and assess DR. While direct ophthalmoscopy is the most common method for diagnosing diabetic retinopathy, it has lower sensitivity when compared to stereoscopic seven-field color photography, especially when performed by non-eye care practitioners. The Early Treatment Diabetic Retinopathy Study (ETDRS) has certified the grading of stereoscopic color fundus

photographs as a recognized standard for the diagnosis of DR. Single-field fundus photography has also been used as an early diagnostic technique for diabetic retinopathy, showing promising results in detecting individuals with retinopathy and recommending them for ophthalmic care . Furthermore, ultrawide-field fundus fluorescein angiography (UWFA) and OCT have been increasingly utilized for the diagnosis of DR. UWFA can capture a larger area of the retina compared to conventional methods, while OCT provides high-resolution cross-sectional images of the retina, the retinal nerve fiber layer, and the optic nerve head. Additionally, optical coherence tomography angiography (OCT-A) has emerged as a non-invasive diagnostic tool for detailed imaging of the retinal vasculature, potentially allowing for the early detection of microvascular changes in diabetic eyes. While fluorescein angiography is an invasive and time-consuming procedure, it remains valuable for detecting vascular changes in diabetic retinopathy. [8]

1.10 Diabetic Retinopathy Complications

- **Vitreous hemorrhage**

Fresh blood vessels have the potential to seep into the transparent, jelly-like substance within the eye. If the bleeding is minimal, only a few dark spots or floaters may be noticed. In more serious instances, the blood can fill the vitreous cavity and obstruct vision entirely. Typically, vitreous hemorrhage does not lead to permanent vision loss. The blood often dissipates from the eye within a few weeks or months. If the retina is undamaged, vision may recover its previous clarity.[7]

- **Retinal Detachment**

Diabetic retinopathy causes the formation of abnormal blood vessels, which in turn stimulates the growth of scar tissue. This scar tissue can tug on the retina, causing it to detach from the back of the eye. As a result, floating spots may appear in vision, flashes of light may occur, or even significant vision loss can happen. [7]

- **Glaucoma**

The front part of the eye may develop new blood vessels, which can disrupt the regular drainage of fluid from the eye. This disruption can lead to increased pressure in the eye, known as glaucoma. The elevated pressure can harm the optic nerve, which transmits visual information from the eye to the brain. [7]

- **Blindness**

In the end, diabetic retinopathy, glaucoma, or a combination of both can result in total vision impairment. [7]

1.11 Treatment

The treatment for diabetic retinopathy is determined by the specific type and severity of the condition, as well as how it has responded to previous treatments.

- **1.11.1 Early Diabetic Retinopathy**

If nonproliferative diabetic retinopathy is present, immediate treatment may not be necessary. However, an eye doctor will monitor the eyes closely to assess the need for any

treatment. It might also be beneficial to collaborate with a diabetes doctor (endocrinologist) to explore additional measures to enhance diabetes management. The positive aspect is that when diabetic retinopathy is at a mild or moderate stage, effective blood sugar control can typically help slow down its progression. [7]

- **1.11.2 Advanced Diabetic Retinopathy**

Prompt surgical treatment is necessary for individuals with proliferative diabetic retinopathy. In some cases, surgery may also be advised for severe nonproliferative diabetic retinopathy [7]. Treatment options may vary based on the specific retinal issues and may include:

- **Focal laser treatment**

The laser treatment, also called photocoagulation, is used to halt or reduce the leakage of blood and fluid in the eye. This procedure is typically carried out in a doctor's office or an eye clinic. During the treatment, abnormal blood vessels are targeted with laser burns to address leaks. Focal laser treatment is usually completed in a single session. After the procedure, vision may be blurry for approximately a day. Occasionally, small spots may be noticed in the visual field that are associated with the laser treatment, but these typically vanish within a few weeks. If blurred vision was experienced due to swelling of the central macula before the surgery, vision may not fully return to normal. However, there are cases where vision does improve. [7]

- **Scatter laser treatment**

During this laser treatment, also called panretinal photocoagulation, abnormal blood vessels can be reduced in size. It is typically performed in a doctor's office or an eye clinic. Throughout the procedure, laser burns are applied to areas of the retina away from the macula. These burns cause abnormal new blood vessels to shrink and form scar tissue. Scatter laser treatment usually takes place over two or more sessions. Vision may be blurry for about a day after the procedure, and there is a possibility of experiencing some loss of peripheral or night vision afterwards. [7]

- **Vitrectomy**

The procedure is performed to eliminate blood from the vitreous (middle of the eye) and to remove any scar tissue that may be pulling on the retina. It is carried out in a surgery center or hospital under local or general anesthesia. During the procedure, a small incision is made in the eye, and delicate instruments are used to remove scar tissue and blood. These are replaced with a salt solution to maintain the eye's normal shape. In some cases, a gas bubble may be placed in the eye cavity to assist in reattaching the retina. If a gas bubble is used, maintaining a facedown position until it dissipates, which can take several days, may be necessary. An eye patch and medicated eye drops will be required for a few days or weeks. Vitrectomy may be followed by laser treatment. While surgery often slows or halts the progression of diabetic retinopathy, it is not a cure. Since diabetes is a lifelong condition, there is a possibility of future retinal damage and vision loss. Regular eye exams will be necessary even after treatment for diabetic retinopathy, and additional treatment may be recommended at some point. New treatments for diabetic retinopathy are being researched, including medications that

may prevent abnormal blood vessels from forming in the eye. Some of these medications are injected directly into the eye to address existing swelling or abnormal blood vessels. Although these treatments show promise, they have not yet been studied in long-term trials. [7]

1.12 Conclusion

We can conclude that diabetic retinopathy is a serious complication of diabetes that requires early detection and effective management. Comprehensive knowledge of its progression, symptoms, diagnosis, complications, and treatment options is essential for preventing vision loss and improving patient outcomes. The next chapter will talk about Deep Learning.

Chapter 2

Deep Learning

2.1 Artificial Intelligence (AI)

The field of Artificial Intelligence (AI) in computer science is dedicated to the creation of systems that can execute tasks normally associated with human intelligence. These tasks include learning from experience, reasoning, problem-solving, understanding natural language, perceiving the environment, and even exhibiting traits such as creativity and planning. AI aims to build machines that can simulate human cognitive functions and enhance human capabilities in various domains. [12]

2.2 Machine Learning (ML)

Within the domain of artificial intelligence (AI), machine learning is a distinct discipline that centers around the construction of algorithms and statistical models. These enable computers to carry out specific tasks without the necessity of explicit instructions. Instead, these systems learn from data, identifying patterns, making decisions, and improving over time based on experience. Machine learning algorithms are designed to build a model based on sample data, known as training data, to make predictions or decisions without being explicitly programmed to perform the task. [13]

2.3 Deep Learning (DL)

Deep learning is a subset of machine learning. It focuses on algorithms inspired by the structure and function of the brain, called artificial neural networks. Deep learning models are capable of learning from large amounts of data and are characterized by their use of multiple layers of neurons that work together to progressively extract higher-level features from the raw input.

Deep learning involves the use of neural networks with many layers, often referred to as deep neural networks (DNNs). These networks are designed to simulate the way the human brain processes information. The term "deep" refers to the number of layers through which the data is transformed. Each layer in the network processes the input data, transforms it into a slightly more abstract and composite representation, and passes it to the next layer. This hierarchical learning process allows the model to understand complex patterns and structures within the data. [14]

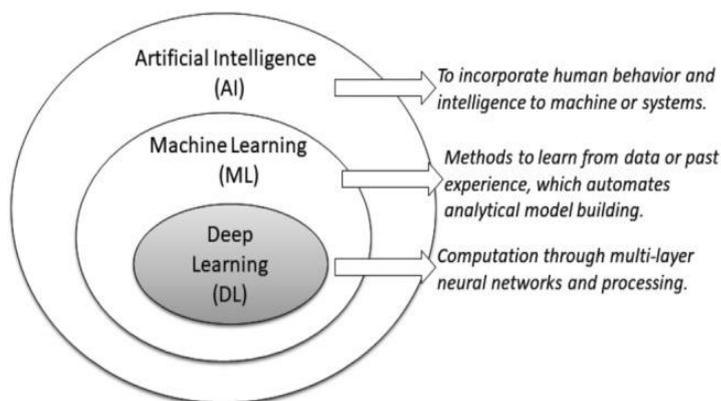


Figure 5. An illustration of the relation between AI, ML and DL. [14]

2.4 Why Deep Learning in Today's Research and Applications?

The main focus of today's Fourth Industrial Revolution (Industry 4.0) is typically technology-driven automation, smart and intelligent systems, in various application areas including smart healthcare, business intelligence, smart cities, cybersecurity intelligence, and many more. Deep learning approaches have grown dramatically in terms of performance in a wide range of applications considering security technologies, particularly, as an excellent solution for uncovering complex architecture in high-dimensional data. Thus, DL techniques can play a key role in building intelligent data-driven systems according to today's needs, because of their excellent learning capabilities from historical data. Consequently, DL can change the world as well as humans' everyday life through its automation power and learning from experience. DL technology is therefore relevant to artificial intelligence, machine learning and data science with advanced analytics that are well-known areas in computer science, particularly, today's intelligent computing. [14]

2.5 Understanding various forms of data

As DL models learn from data, an in-depth understanding and representation of data are important to build a data-driven intelligent system in a particular application area. In the real world, data can be in various forms, which typically can be represented as below for deep learning modeling:

Sequential data: Sequential data is any kind of data where the order matters. It needs to explicitly account for the sequential nature of input data while building the model. Text streams, audio fragments, video clips, time-series data, are some examples of sequential data. [14]

Image or 2D data: A digital image is made up of a matrix, which is a rectangular array of numbers, symbols, or expressions arranged in rows and columns in a 2D array of numbers. Matrix, pixels, voxels, and bit depth are the four essential characteristics or fundamental parameters of a digital image. [14]

Tabular data: A tabular dataset consists primarily of rows and columns. Thus tabular datasets contain data in a columnar format as in a database table. Each column (field) must have a name and each column may only contain data of the defined type. Overall, it is a logical and systematic arrangement of data in the form of rows and columns that are based on data properties or features. Deep learning models can learn efficiently on tabular data and allow us to build data-driven intelligent systems. [14]

2.6 Convolutional Neural Networks

Convolutional Neural Networks (CNNs or ConvNet) are a class of deep neural networks that are particularly effective for analyzing visual data. They have been widely used in various applications such as image and video recognition, image classification, medical image analysis, and natural language processing. [16]

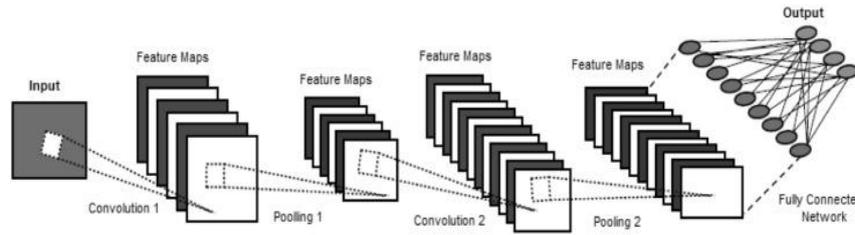


Figure 6. An example of a Convolutional Neural Network including multiple convolution and pooling layers. [14]

In CNNs, there are two types of layers: primary layers and secondary layers. Primary layers are essential and include convolution layers, pooling layers, activation layers, dense layers, and flatten layers. On the other hand, secondary layers are optional and can be added to enhance the CNN's resilience to overfitting and improve its generalization. These optional layers consist of dropout layers, regularization layers, and batch normalization layers. [15]. Figure 7 shows the structure of a convolutional neural network (CNN).

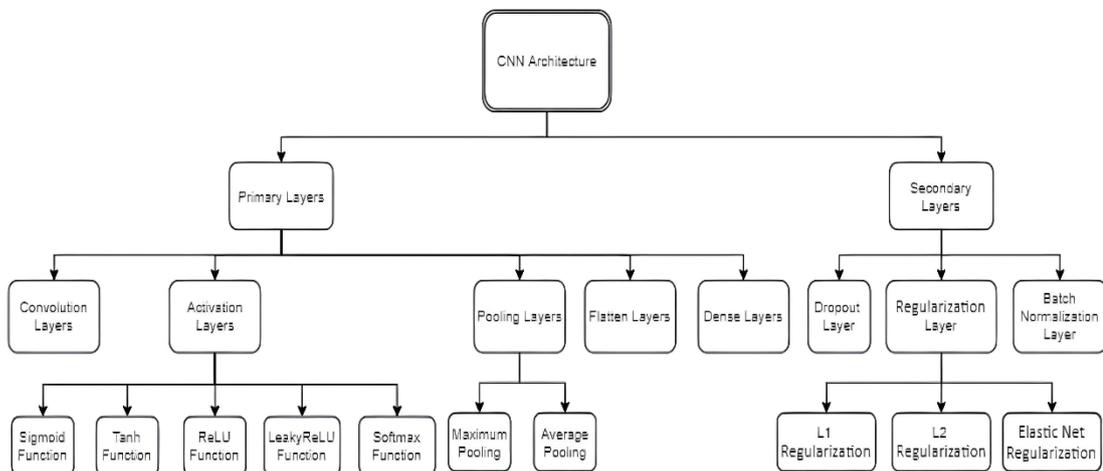


Figure 7. Structure of a convolutional Neural Network (CNN). [15]

• 2.6.1 Convolutional Layers

The primary and crucial layer in a Convolutional Neural Network (CNN) is the convolutional layer, which automatically extracts image features without the need for manual definition. In this layer, a set of filters (also called kernels) is applied to the input data to produce a set of feature maps. Each filter scans across the input data, performing element-wise multiplications and summing the results to detect specific patterns, such as edges or textures. The output of these operations is stored in feature maps, which highlight the presence of these patterns at different locations in the input [15]. Two important parameters that influence the behavior of the convolutional layer are stride and padding:

- **Stride:** This defines how far the filter moves with each step. A larger stride results in a smaller output feature map.
- **Padding:** This adds extra pixels (usually zeros) around the input data to control the spatial size of the output feature map, allowing for the preservation of the original input dimensions.

The numbers within the filters are initially set randomly and are optimized during the training process. The result of the convolutional operations is a set of feature maps that serve as inputs for the next layers in the network, enabling the CNN to learn and recognize complex patterns in the data. [15]

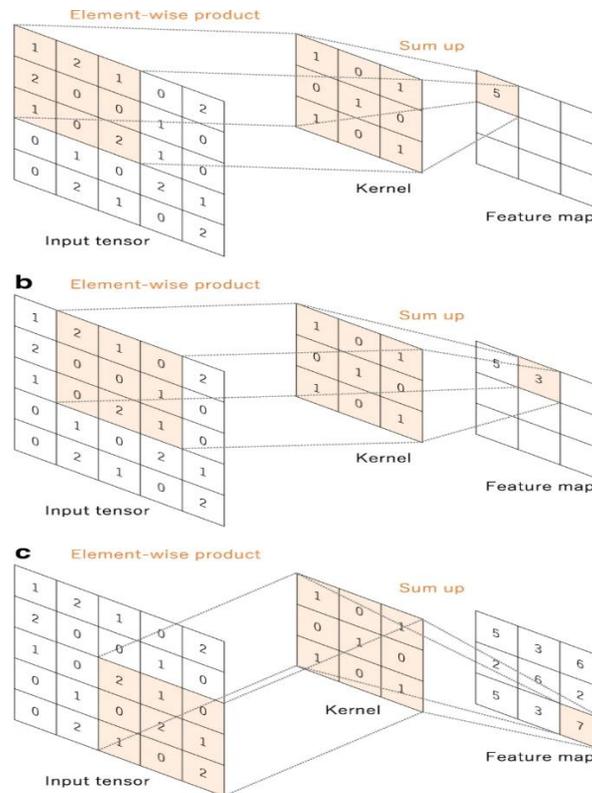


Figure 8. An example of a convolutional operation with a kernel size of 3×3 , no padding, and a stride of 1. [17]

• 2.6.2 Activation Layers

The activation layers, which typically come after the convolution layers, are crucial as they determine whether a given neuron will fire. These layers take a real number as input and apply a nonlinear function to it. Activation layers are essential as they allow the network to learn nonlinear mappings, making it more resilient to complex functions. Common activation layers used in CNNs include sigmoid, Tanh, ReLU, LeakyReLU, and softmax. These layers can be categorized as either saturated or non-saturated. If the output of the activation layer falls within finite boundaries, it is considered saturated; if it tends towards infinity, it is classified as non-saturated. Non-saturated activation functions offer several advantages over saturated ones. For example, they can significantly alleviate the exploding/vanishing gradient problem encountered during the backpropagation algorithm, which is a key training issue in CNNs [15]. Figure 9 represents different activation functions.

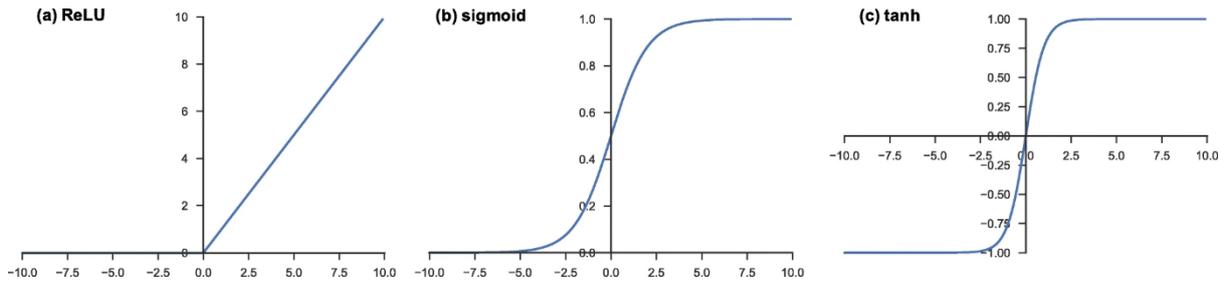


Figure 9. Plot of different activation functions : (a) ReLU activation function; (b) Sigmoid activation function; and (c) Tanh activation function. [17]

• 2.6.3 Pooling Layers

Pooling layers are typically positioned between consecutive convolution layers in order to gradually decrease the spatial size of the representation. This reduction helps in cutting down the number of parameters and computational load in a network. A pooling layer works by extracting important pixels and eliminating noise, thus reducing the output feature map of the convolution layer. One of the most widely used types of pooling operation is max pooling. This process involves extracting patches from the input feature maps, finding the maximum value in each patch, and discarding all other values. [15]

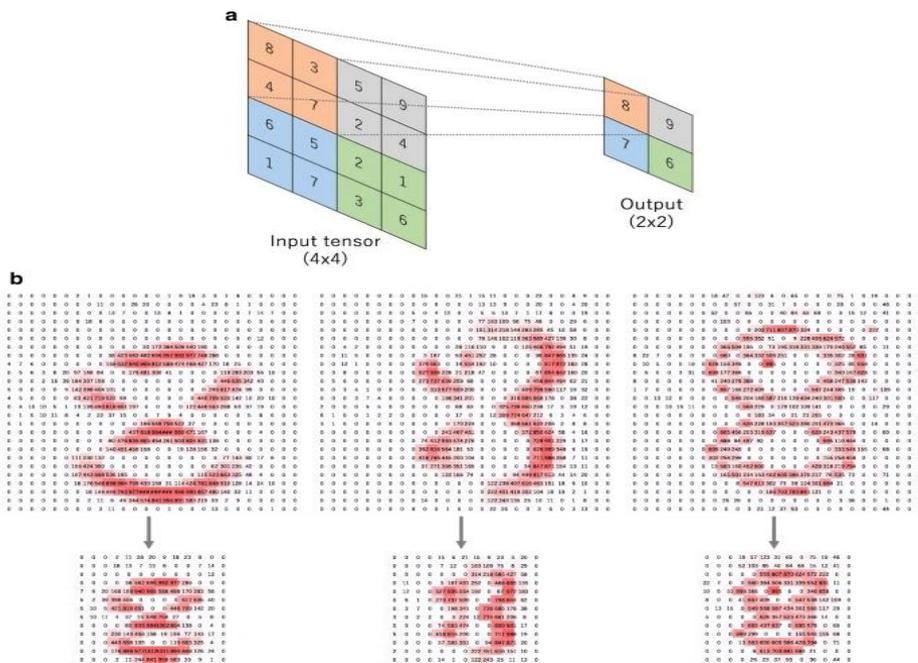


Figure 10. a: An example of max pooling operation with a filter size of 2×2 , no padding, and a stride of 2. b: Examples of the max pooling operation on images. Note that images in the upper row are downsampled by a factor of 2, from 26×26 to 13×13 . [17]

• 2.6.4 Flattening Layers

The output from the pooling layer is converted into a 1D vector of size $1D$ so that it can be fed into the subsequent dense layers [15]. This transformation is illustrated in Figure 11. The resulting vector's dimensionality is calculated as:

$$DimFlat = Dimimg * Dimimg * numcolor$$

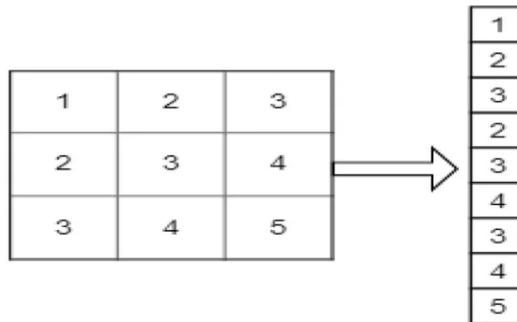


Figure 11. Flattening 2D feature maps to 1D vector. [15]

- **2.6.5 Dense Layers**

Fully connected layers, also called dense layers, are commonly positioned at the end of a network and are fed the output of the feature extraction layers as input. Their primary objective is to utilize all the features extracted from the preceding layers to classify the original image. Towards the end of the network, a softmax or sigmoid function is applied to produce the target probability. [15]

- **2.6.6 Dropout Layer**

Dropout layer can be used as a regularization technique in neural networks. This layer randomly disables some neurons during training with a specified dropout rate probability P . It's similar to the concept of bagging in neural networks. [15]

- **2.6.7 Regularization Layers**

Complex models with large weights often struggle to generalize well, as they can end up learning noise instead of the true underlying patterns. It is commonly believed that models with smaller weights tend to generalize better than those with larger weights. To address this, regularization techniques are frequently employed to prevent overfitting. Regularization functions work by adding a penalty term to the loss function, which discourages the use of large weights by the model. The primary goal of regularization is to eliminate weights that do not significantly contribute to the accuracy of the model by reducing them towards zero. Different types of regularization, such as L1, L2, and elastic nets, exist, and they differ mainly in terms of the penalties they impose. [15]

- **2.6.8 Batch Normalization Layers**

Batch normalization is a technique that can accelerate network training and enhance its resistance to overfitting. It addresses network covariance shift by normalizing the inputs of each layer it is applied to, dividing by the batch standard deviation, and subtracting the batch mean. Moreover, batch normalization introduces noise to each layer to bolster its robustness. [15]

2.7 Transfer Learning

Transfer Learning is a method that utilizes previously acquired model knowledge to tackle a new task with minimal training or fine-tuning. Unlike traditional machine learning methods, deep learning requires a large amount of training data. This poses a significant challenge, especially in domains like healthcare, where creating extensive, high-quality annotated medical

datasets is both arduous and expensive. Additionally, deep learning models require substantial computational resources, such as a GPU-enabled server, although efforts are being made to enhance this. Therefore, a transfer learning approach, could potentially offer a solution to these issues. Figure 8 illustrates the general structure of the transfer learning process, where knowledge from a pre-trained model is transferred to a new deep learning (DL) model. This approach is particularly popular in deep learning today, as it enables the training of deep neural networks with minimal data. [14]

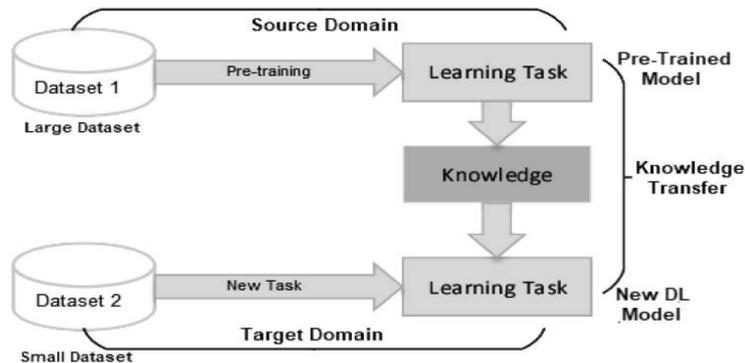


Figure 12. Transfer Learning Process. [14]

Transfer learning is particularly effective for training convolutional neural networks (CNNs) rapidly and accurately by utilizing pre-trained weights from another CNN, often trained on a larger dataset like ImageNet. Several CNN architectures trained on ImageNet have achieved high accuracy, and their weights can be used to classify entirely different datasets, avoiding the need for random weight initialization. There are four main strategies in transfer learning: using pre-trained layers as feature extractors with a new classifier, fine-tuning the entire network with a small learning rate and adding a new classifier, fine-tuning only the top layers while keeping the bottom layers frozen, and training a state-of-the-art architecture from scratch. Many researchers recommend fine-tuning only the top layers, as they detect more dataset-specific features, while the bottom layers identify generic features such as edges and circles. [15]

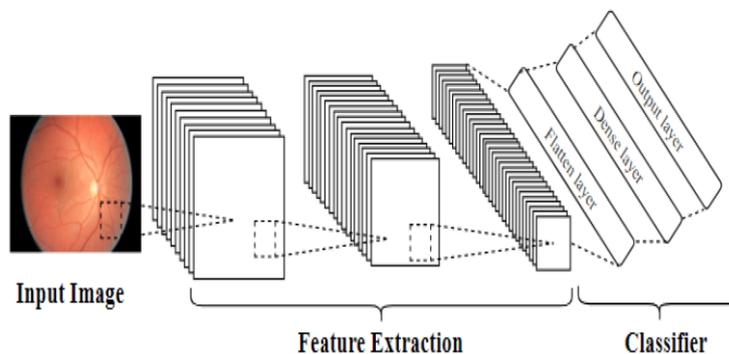


Figure 13. Convolutional Neural Network Architecture. [15]

2.8 Data Augmentation

Data Augmentation methods tackle overfitting and class imbalance at its source, the training dataset. These methods operate on the principle that by applying augmentations, additional valuable information can be derived from the original dataset. Augmentations increase the size

of the training dataset by either distorting existing data or adding synthetic samples. Distorting augmentations modify existing images while preserving their labels. This category includes transformations like geometric and color adjustments, random erasing, adversarial training, and neural style transfer. Adding synthetic samples involves creating new instances and incorporating them into the training set. This can be achieved through techniques such as image mixing, feature space adjustments, and generative adversarial networks (GANs).



Figure 14. Image data augmentation techniques. [22]

2.9 Fine-tuning

The fine-tuning approach is widely used for transfer learning in neural networks. It is defined by transferring knowledge from a generative to a discriminative model, resulting in strong generalization. The original process involve using a pre-trained network and replacing its last classifier layer with a randomly initialized one.

2.10 Architectures of Convolutional Neural Networks

• 2.10.1 VGG Network Architecture

In 2014, the Visual Geometry Group at Oxford University introduced two new architectures called VGG16 and VGG19. VGG16 achieved a top five accuracy rate of 91.90% in the 2014 ImageNet competition. The VGG16 architecture consists of 138,355,752 parameters, five convolution blocks, and three dense layers. Each block contains several convolutional layers followed by a max pool layer to downsize the block output and eliminate noise. The first two blocks have two convolutional layers each, while the last three blocks have three convolutional layers each. Throughout the network, a kernel size with a stride of 1 is utilized. Following the five blocks, a flatten layer was incorporated to convert the 3D block vector into a 1D vector for input into the fully connected layers. The initial two fully connected layers comprise 4096 neurons, and the final fully connected layer contains 1000 neurons. A softmax layer is added after the fully connected layers to ensure that the probability summation of the output is one. The primary distinction between VGG16 and VGG19 is that VGG19 includes 19 convolution layers rather than 16, resulting in an increase in the number of parameters from 138,357,544 to 143,667,240. The authors argued that these additional layers enhance the architecture's robustness and its ability to learn more complex architectures. One of the network's key advantages is its sequence of blocks, where sequential convolutional layers are stacked to reduce the amount of spatial information required. However, a notable drawback is that the authors allocate more weights to the classifier portion than to the feature extraction portion, significantly increasing the number of parameters. The ImageNet weights for the network are accessible in the Keras package. [15]

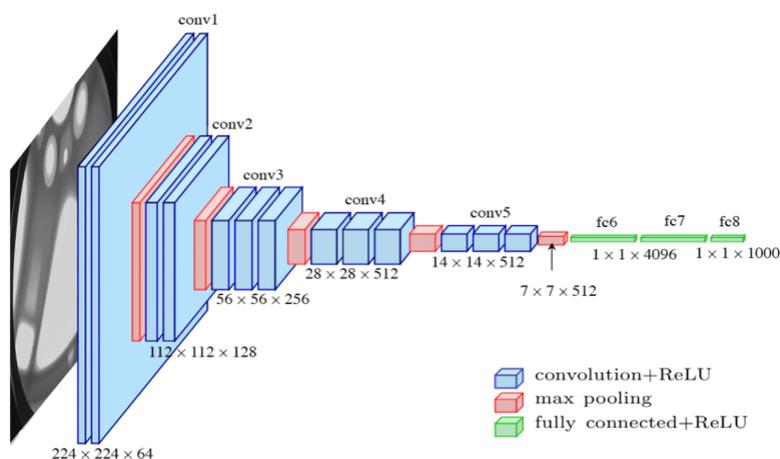


Figure 15. VGG16 Architecture. [18]

• 2.10.2 ResNet Network Architecture

ResNet, short for residual network, was developed by He et al. in 2015 and won first place in the 2015 ImageNet competition with a top five accuracy rate of 94.29%. It comprises a total of 25,000,000 parameters. Unlike other architectures, ResNet is a very deep network that can extend up to 152 layers and incorporates a unique connection known as the residual connection. This connection is applied between the convolutional layers and then passed to the ReLU

activation layer. The residual connection ensures that the weights learned from the previous layers do not vanish during backpropagation. Three versions of this network have been introduced, namely ResNet50, ResNet101, and ResNet152, which differ in the number of layers. The primary advantage of this network lies in its use of residual connections, which enables the incorporation of a large number of layers. Furthermore, increasing the network's depth (as opposed to widening it) results in fewer additional parameters. However, one of the main drawbacks of this network is the need for summation in each residual block, which keeps the filter size constant. Moreover, this network necessitates large datasets for proper training, leading to a computationally intensive training phase. The ImageNet weights for the network are accessible in the Keras package. [15]

- **2.10.3 GoogLeNet Network Architecture**

In 2014, a new network architecture known as the GoogLeNet network (also called InceptionV1 architecture) was introduced by Google researchers. They achieved a top 5 accuracy rate of 92.2% in the ImageNet competition. Following the success of InceptionV1, the authors developed other versions such as InceptionV2 and InceptionV3. The main concept of the GoogLeNet architecture is to utilize multiple convolution layers in the same block to create a network that is not only deeper but also wider, enabling it to capture various features of images. These blocks are known as Inception blocks. The most renowned GoogLeNet architectures are InceptionV1 and InceptionV3. In InceptionV1, six convolution layers are used in the inception blocks, whereas in InceptionV3, seven convolution layers are used. The InceptionV1 architecture is commonly referred to as the GoogLeNet architecture. One of the primary advantages of this network is the inclusion of an inception module, which enables the network to capture different aspect ratios of the same image by using convolution layers in parallel. However, a major drawback of this network is the significant computational effort required to train it due to its depth and width. The ImageNet weights for InceptionV3 are accessible in the Keras package. [15]

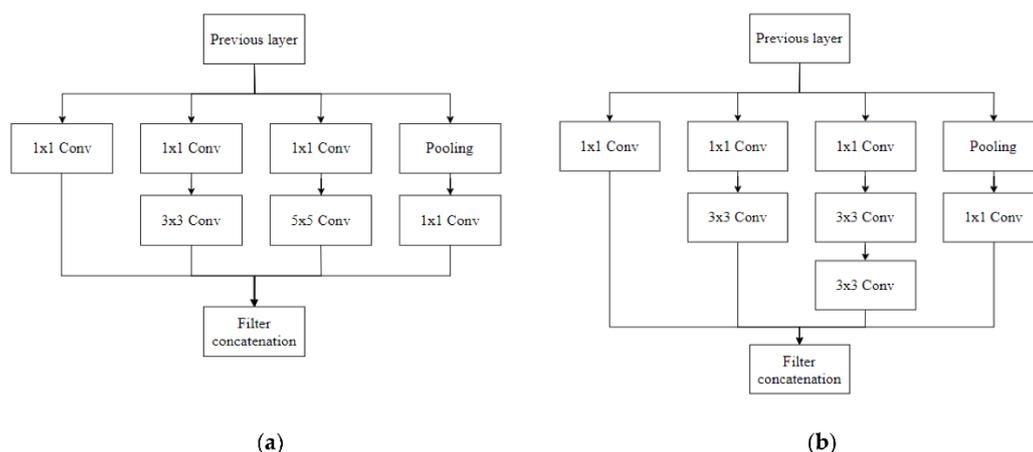


Figure 16. Inception blocks. (a) InceptionV1. (b) InceptionV3. [15]

- **2.10.4 AlexNet Network Architecture**

The AlexNet architecture, which took part in the 2012 ImageNet challenge, marked the debut of CNNs in the competition and achieved an impressive accuracy rate of 84.60%. This outperformed the shallower algorithms previously used for image classification. Since then, CNNs have emerged as the leading algorithm for image classification. With 60,000,000

parameters, five convolution layers, and three dense layers, the AlexNet introduced two significant innovations: the use of the ReLU activation function instead of the sigmoid function, and the integration of dropout to address overfitting in deep architectures. One of its key advantages is its computational efficiency during the training process compared to other considered networks. However, its depth may limit its capability to capture complex image features. [15]

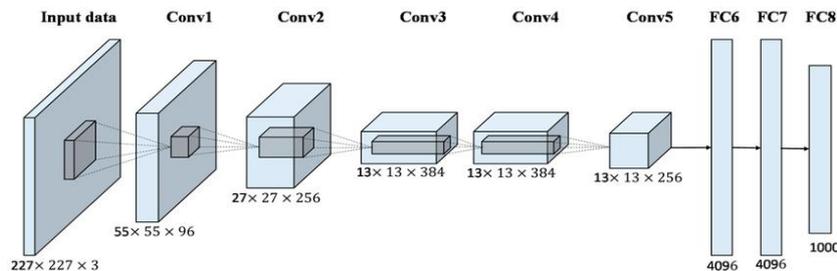


Figure 17. AlexNet Architecture. [19]

• 2.10.5 DenseNet Network Architecture

DenseNet architecture, short for densely-connected convolutional networks, draws inspiration from ResNet, but diverges from it by utilizing dense blocks instead of residual connections. A dense block comprises a series of convolution layers, relative to VGG, with each layer connected to all subsequent layers. The primary concept is for each convolution layer to receive information from all preceding layers. DenseNet boasts 8,062,504 parameters and achieved a top 5 accuracy rate of 93.34% in the ILSVCR challenge. Its key advantage lies in the presence of connections between all layers, minimizing information loss, especially in deep layers. However, its drawbacks include high computational costs during training and the necessity for large datasets to achieve optimal performance. Notably, the network's ImageNet weights are accessible in the Keras package. [15]

• 2.10.6 Xception Network Architecture

The Xception network, short for extreme inception, was created by Chollet and draws inspiration from the InceptionV3 architecture. The key concept behind the Xception design involves replacing the inception module with depthwise separable convolution, followed by a pointwise separable convolution. This network comprises 71 layers and houses 22.9 million parameters. It achieved an impressive 94.50% top 5 accuracy rate on the ILSVCR challenge. One of its primary advantages is its deep architecture with a relatively low number of parameters, making it computationally efficient compared to other deep networks. However, a notable drawback is that training this network effectively necessitates very large datasets. [15]

2.11 Conclusion

In this chapter, we explored the impact of deep learning on diabetic retinopathy detection, focusing on Convolutional Neural Networks (CNNs). We discussed the importance of CNN architectures, data augmentation techniques, and fine-tuning in enhancing model performance. Understanding these elements highlights the technological advancements driving automated diabetic retinopathy detection and sets the stage for interpreting and evaluating these models in clinical settings in the next chapter.

Chapter 3

Implementation and evaluation

3.1 Introduction

In the previous chapters, we discussed the medical background of diabetic retinopathy, including its types, stages, and treatments, as well as the fundamentals of deep learning, focusing on Convolutional Neural Networks (CNNs) and transfer learning. Now, in this chapter, we will present our contributions and results, detailing the models developed, the datasets used, the experimental setup, and the performance analysis of our deep learning approach to detect diabetic retinopathy.

3.2 Tools and Configuration

- **3.2.1 Collaboratory**

Google Colab, short for Google Collaboratory, is a cloud-based platform provided by Google that facilitates machine learning research and development. It offers a free Jupyter notebook environment that runs entirely in the cloud, eliminating the need for users to set up and maintain their own computing resources. This makes it particularly advantageous for researchers and developers who require substantial computational power but may not have access to high-end hardware locally. Google Colab provides access to GPUs and TPUs (Tensor Processing Units), allowing for accelerated training of deep learning models. It also supports seamless integration with Google Drive, enabling easy storage and sharing of notebooks and datasets. The platform comes pre-installed with popular libraries and frameworks such as TensorFlow, PyTorch, and scikit-learn, making it suitable for a wide range of machine learning tasks. Overall, Google Colab has become a preferred choice for many in the AI and machine learning community due to its accessibility, powerful capabilities, and collaborative features.

The entire experiment was conducted using Google Colab, supplemented with additional computational units purchased for enhanced performance. The following table outlines the configuration of the virtual machine (VM) provided by Google Colab.

CPU	Intel Xeon CPU @2.20 GHz
GPU	NVIDIA L4 GPU 24 GB GDDR6
RAM	64 GB
DISK	80 GB

Table 1. configuration of the virtual machine provided by Google Colab.

- **3.2.2 Software and Libraries**

Software and libraries	Description
TensorFlow	Google's open-source framework for machine learning, known for its scalability and support for deep learning models via computational graphs.
Keras	High-level neural networks API, designed for ease of use and rapid prototyping, built on top of TensorFlow and other frameworks.
PyTorch	Developed by Facebook's AI Research lab, PyTorch offers dynamic computational graphs and flexibility for building and training neural networks, favored for research.
NumPy	Fundamental package for numerical computing in Python, providing powerful arrays and mathematical functions essential for data manipulation and computation.
Pandas	Python library for data manipulation and analysis, offering easy-to-use data structures (like DataFrame) and tools for cleaning and processing data.
Scikit-learn	A popular Python library for machine learning tasks, offering a straightforward interface and a comprehensive range of tools for both supervised and unsupervised learning.
Seaborn	Statistical data visualization library in Python, built on matplotlib, simplifying the creation of informative and attractive statistical graphics.
OpenCV	Open Source Computer Vision Library, providing extensive tools for real-time computer vision tasks such as image and video processing.
VS code (Visual Studio Code)	Lightweight but powerful source code editor developed by Microsoft, known for its built-in support for debugging, syntax highlighting, and extension ecosystem. It is widely used for various programming languages and development tasks.
Django	High-level Python web framework that encourages rapid development and clean, pragmatic design. It is known for its "batteries-included" philosophy, facilitating the creation of robust web applications.

Table 2. Software and libraries used during the research.

3.3 Datasets

- **3.3.1 First Dataset (MuReD)**

The MuReD (MULTI-LABEL RETINAL DISEASE) dataset is curated for research in diabetic retinopathy and related conditions, offering a collection of 2451 retinal images. This dataset consists of two CSV files: 'train_data.csv' and 'test_data.csv'. 'train_data.csv' contains 1764 rows and 21 columns, with each row identified by an 'ID' column containing names of retinal images. The columns include 'DR' (Diabetic Retinopathy), 'NORMAL', 'MH', 'ODC', 'TSLN', 'ARMD', 'DN', 'MYA', 'BRVO', 'ODP', 'CRVO', 'CNV', 'RS', 'ODE', 'LS', 'CSR', 'HTR', 'ASR',

'CRS', and 'OTHER', all represented as integer values except for 'ID'. The 'DR' column serves as the target with 1368 instances labeled as 0 and 396 as 1. 'test_data.csv' comprises 444 rows with similar column names and types, including 345 instances labeled as 0 and 99 as 1 in the 'DR' column. Additionally, the dataset includes a folder containing 2451 retinal images corresponding to the 'ID' values in the CSV files. The following table provides a detailed breakdown of each column in the dataset. [20]

NORMAL	Normal retina.
MH	Media Haze.
ODC	Optic disc cupping.
TSLN	Tessellation.
ARMD	Age-related macular degeneration.
DN	Drusen.
MYA	Myopia.
BRVO	Branch retinal vein occlusion.
ODP	Optic disc pallor.
CRVO	Central retinal vein occlusion.
CNV	Choroidal neovascularization.
RS	Retinitis.
ODE	Optic Disc Edema.
LS	Laser scars.
CSR	Central serous retinopathy.
HTR	Hypertensive retinopathy.
ASR	Arteriosclerotic retinopathy.
CRS	Chorioretinitis.
OTHER	Other diseases.

Table 3. A detailed breakdown of each column in the MuReD dataset.

We can see that in both the training and test datasets, the class distribution for diabetic retinopathy (DR) is imbalanced and shows that approximately 77.6% of instances belong to class 0 (no diabetic retinopathy), while around 22.4% belong to class 1 (diabetic retinopathy). This distribution is visually represented in Figure 18, illustrating the proportions of each class in the datasets.

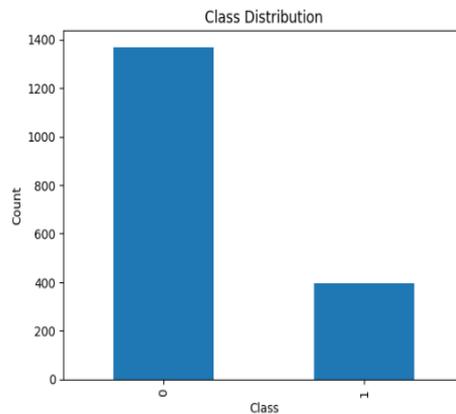


Figure 18. Class Distribution of DR in MuReD Dataset.

- **3.3.2 Second Dataset (RFMiD 2.0)**

The RFMiD2.0 (Retinal Fundus Multi-disease Image Dataset) consists of 860 images, designed to support research in diagnosing multiple retinal diseases using fundus images. The dataset is divided into three subsets: training, validation, and test. The training set contains 509 images and is accompanied by a CSV file named 'training_labels.csv', which includes 509 rows and 53 columns, with 52 integer columns and one float column. The validation set comprises 177 images with a corresponding 'validation_labels.csv' file, featuring 177 rows and 53 columns, similarly with 52 integer columns and one float column. The test set includes 174 images and a 'testing_labels.csv' file with 174 rows and 53 columns, maintaining the same structure of 52 integer columns and one float column. In all subsets, the 'DR' column represents the target variable for diabetic retinopathy. [21]

Similar to the MuReD dataset, the class distribution for 'DR' in the RFMiD2.0 dataset is imbalanced. In the training set, approximately 91.7% of instances are labeled as 0 (no diabetic retinopathy) and 8.3% are labeled as 1 (diabetic retinopathy). The validation set shows a similar distribution, with around 92.1% labeled as 0 and 7.9% labeled as 1. In the test set, about 92% of instances are labeled as 0 and 8% are labeled as 1. This distribution highlights the prevalence of each class within the dataset. The following figure shows the class distribution of DR in the training set.

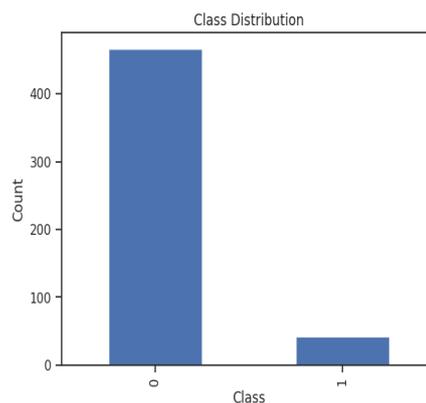


Figure 19. Class distribution of DR in RFMiD 2.0 Dataset.

- **3.3.3 Third Dataset**

The third dataset, obtained from Kaggle, is organized into three files: train, validation, and test. Each file contains two subfolders, 'DR' and 'No_DR'. In the training set, the 'DR' subfolder contains 1050 images, while the 'No_DR' subfolder contains 1026 images. The validation set similarly has a 'DR' subfolder with 245 images and a 'No_DR' subfolder with 286 images. The test set includes a 'DR' subfolder with 113 images and a 'No_DR' subfolder with 118 images. Importantly, this dataset is balanced, meaning that each subset contains an equal number of images with and without diabetic retinopathy. This balance ensures fair evaluation and comparison of classification models across different sets.

Note: 'DR' indicates diabetic retinopathy, and 'No_DR' indicates no diabetic retinopathy.

3.4 Approaches and techniques

- **VGG16 Model**

At first, we chose the VGG16 architecture, known for its depth and effectiveness in image recognition tasks. Developed by the Visual Geometry Group at the University of Oxford, VGG16 consists of 13 convolutional layers, each using a 3x3 filter size, and is interspersed with max-pooling layers for spatial downsampling. This architecture culminates with three fully connected layers, making it adept at learning intricate patterns and high-level features essential for image classification. By initializing VGG16 with pre-trained weights from the ImageNet dataset and excluding the fully connected layers, we could utilize its powerful feature extraction capabilities while adapting the top layers specifically for detecting diabetic retinopathy.

To adapt VGG16 for our purposes, we added a flattening layer (`Flatten`) to transform the 3D output of the convolutional base into a 1D feature vector. This was followed by a densely connected layer with 1024 units, activated by ReLU to introduce non-linearity and enhance the model's capacity to learn complex relationships within the data. Finally, a sigmoid-activated dense layer was appended to produce a binary classification output, indicating the likelihood of diabetic retinopathy presence in retinal images. The model was compiled with an Adam optimizer using a learning rate of $1e-4$ (0.0001), binary cross-entropy loss, and accuracy metrics for evaluation. Additionally, the model was trained for 10 epochs to optimize its performance on the training data.

We applied various data augmentation techniques using TensorFlow's Keras API ImageDataGenerator during training. These techniques included horizontal and vertical flips, random rotations, width and height shifts, zooming, shearing, and nearest-fill mode for pixel filling. Each technique introduced controlled variations into the training images, expanding our dataset and exposing the model to a wider range of inputs. This approach aimed to enhance the model's ability to generalize to unseen data, which is crucial in medical imaging where variability in image quality and patient conditions can impact diagnosis accuracy.

The VGG16 model achieved promising results during training on MuReD dataset, demonstrating an accuracy of 82.26% with a corresponding loss of 38.92%. In the test phase, the model exhibited an accuracy of 85.36% with a reduced test loss of 33.23%.

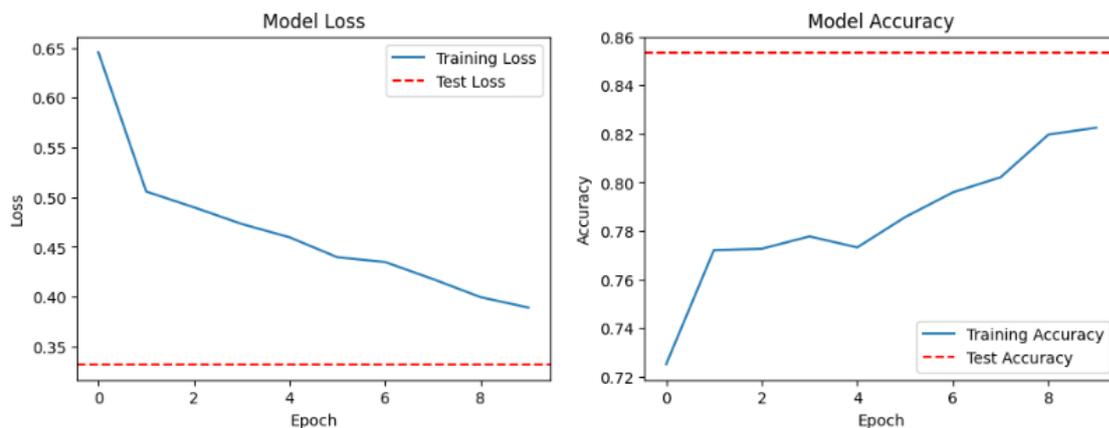


Figure 20. Performance of VGG16 on the MuReD dataset.

Despite the high overall accuracy, further analysis through the classification report revealed a significant bias due to class imbalance. The classification report on the test set highlights this imbalance, showing that the model performed well in identifying 'No_DR' cases, achieving a precision of 78% and recall of 90%. However, the model struggled with 'DR' (diabetic retinopathy) cases, achieving lower precision (28%) and recall (13%). This disparity is evident in the lower F1-score for 'DR' (0.18) compared to 'No_DR' (0.84). The macro-average F1-score was 0.51, indicating an overall imbalance in performance across classes. The weighted average F1-score, considering class imbalance, was 0.69, reflecting the model's general performance across all classes.

```

Classification Report (Test Set):
              precision    recall  f1-score   support

   No_DR       0.78        0.90        0.84        345
    DR         0.28        0.13        0.18         99

 accuracy              0.73        444
 macro avg           0.53        0.52        0.51        444
 weighted avg        0.67        0.73        0.69        444

```

Figure 21. Classification Report of the Test Set in MuReD Dataset.

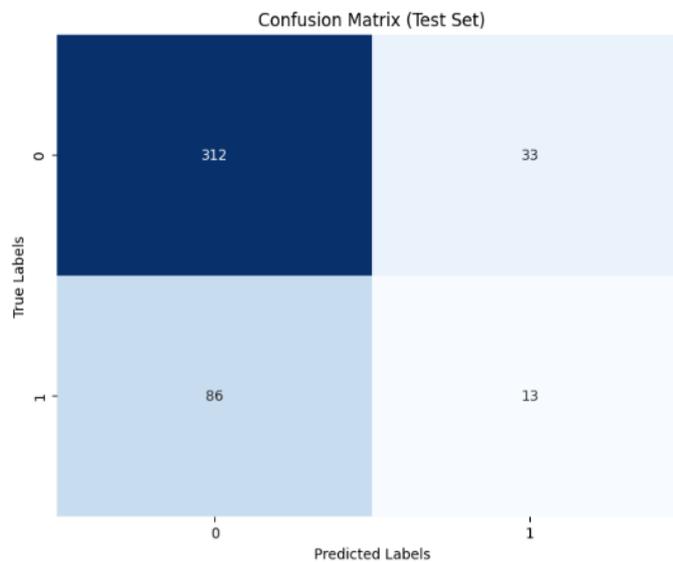


Figure 22. Confusion Matrix of the Test Set in MuReD Dataset.

These results underscore the importance of addressing class imbalance to improve the model's ability to generalize to minority classes, such as diabetic retinopathy cases, in medical image analysis tasks.

To improve the model's performance, we implemented two additional techniques: fine-tuning the last two blocks of the VGG16 model and adjusting for class imbalance using class weights.

First, we fine-tuned the last two blocks of the VGG16 model by unfreezing the layers in these blocks and allowing them to be retrained. This step enabled the model to adapt its deeper, more

complex features specifically to the nuances present in our dataset. We also used a smaller learning rate of $1e-5$ (0.00001) to ensure that the fine-tuning process was more controlled and precise.

Second, we addressed the class imbalance issue by incorporating class weights during training. The class weights were calculated to ensure that the minority class (diabetic retinopathy) received more attention during the training process. we maintained the same data augmentation techniques as before and the model was then retrained for 5 epochs with these enhancements.

After implementing these improvements, the model achieved a training accuracy of 77.27% with a training loss of 40.43%. On the test set, the accuracy was 81.53% with a test loss of 49.45%.

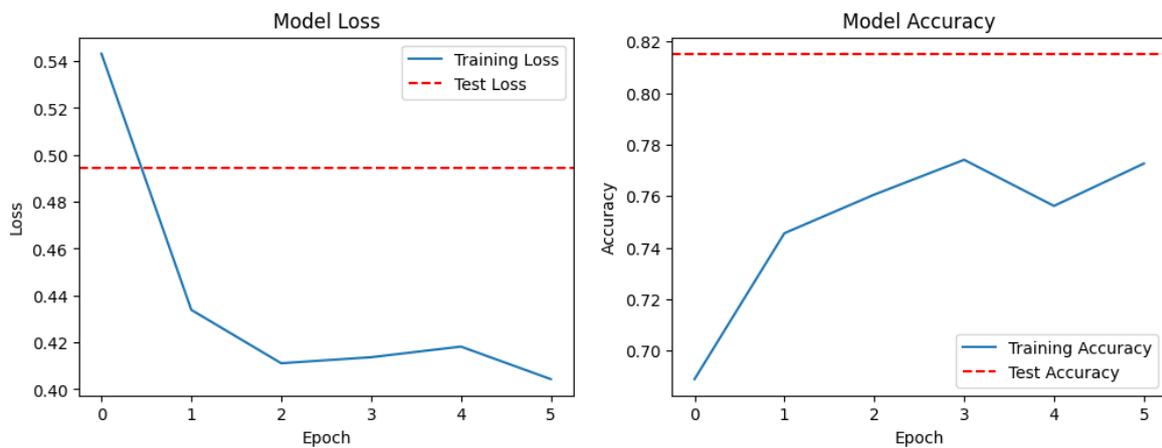


Figure 23. Performance of VGG16 on the MuReD dataset after applying different techniques.

While these results show a reduction in performance metrics compared to the initial results, the classification report provides a more nuanced view:

```

Classification Report (Test Set):
              precision    recall  f1-score   support

     0         0.75         0.61         0.67         345
     1         0.17         0.27         0.21          99

 accuracy                   0.54         444
 macro avg              0.46         0.44         0.44         444
 weighted avg           0.62         0.54         0.57         444
    
```

Figure 24. Classification Report of the Test Set after applying different techniques.

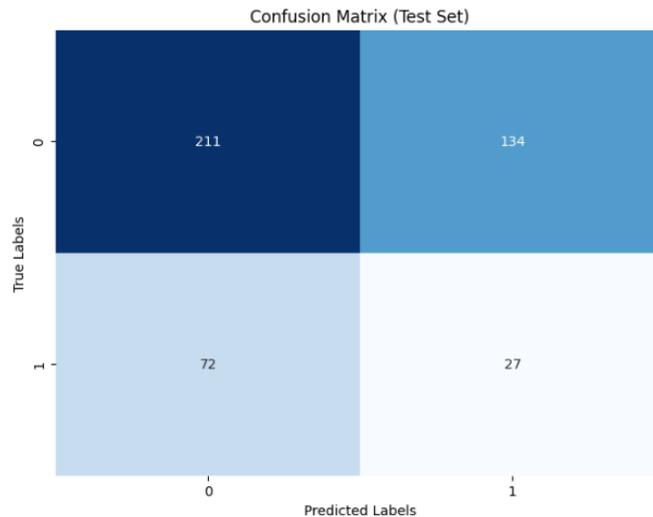


Figure 25. Confusion Matrix of the Test Set after applying different techniques.

We can see that there is a slight improvement in detecting the positive cases (diabetic retinopathy) in MuReD dataset. While the overall accuracy and loss metrics have decreased, the precision, recall, and F1-score for the positive class show some improvement, indicating that the adjustments in class weights and fine-tuning have helped the model better identify cases of diabetic retinopathy.

Using the same VGG16 model with fine-tuning of the last two blocks and identical data augmentation techniques, we trained the model on the RFMiD2.0 dataset for 10 epochs. The results were as follows:

- **Training Set:** 99.01% accuracy, 3.74% loss
- **Validation Set:** 93.22% accuracy, 15.79% loss
- **Test Set:** 95.29% accuracy, 5.98% loss

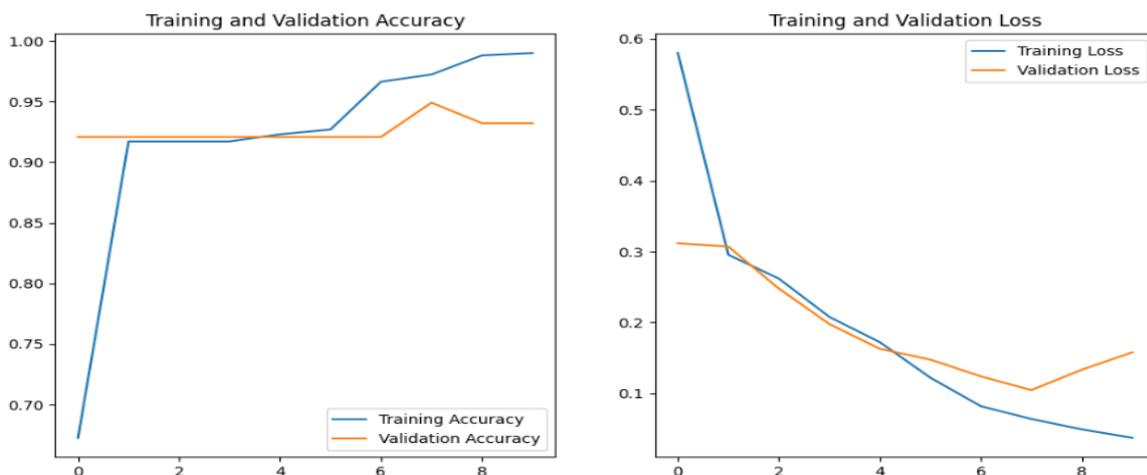


Figure 26. Performance of VGG16 on the RFMiD 2.0 dataset.

The classification report for the test set is detailed below:

Classification Report (Test Set):				
	precision	recall	f1-score	support
0	0.92	0.94	0.93	156
1	0.17	0.14	0.15	14
accuracy			0.87	170
macro avg	0.55	0.54	0.54	170
weighted avg	0.86	0.87	0.87	170

Figure 27. Classification Report of the test set in RFMiD 2.0 Dataset.

Despite the high overall accuracy and low loss metrics, the classification report reveals a significant imbalance in the model's performance. The precision, recall, and F1-score for the positive class (diabetic retinopathy) are low, indicating that while the model performs well on the majority class, it struggles to accurately identify the minority class as shown in the confusion matrix below.

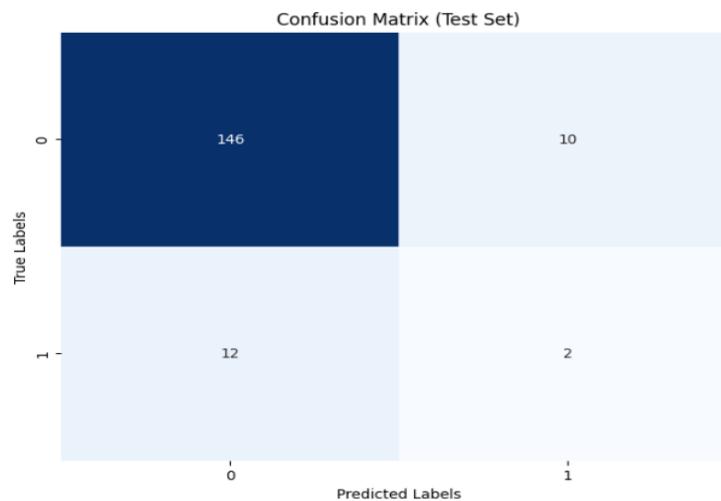


Figure 28. Confusion Matrix of the test set in RFMiD 2.0 Dataset.

- **ResNet50 Model**

For the next phase of our research, we explored a different model: ResNet50. Initially, we loaded the pre-trained ResNet50 model using the weights from ImageNet. The model was configured to exclude the top fully connected layers, focusing on feature extraction from the input images. To adapt ResNet50 for our purposes, we added custom layers on top of the base model. Specifically, we appended a Global Average Pooling layer to reduce the spatial dimensions of the output volume. This was followed by a densely connected layer with 1024 units and ReLU activation to introduce non-linearity and enhance learning. Finally, a sigmoid-activated dense layer was added to produce a binary classification output, indicating the likelihood of diabetic retinopathy presence in retinal images.

This model was trained on the MURED dataset using a combination of class weights and data augmentation techniques. Initially, the model was trained for 5 epochs with a learning rate of $1e-5$ (0.00001).

Similar to our previous approach, we applied various data augmentation techniques to enhance the model's generalization capabilities. These techniques included horizontal and vertical flips,

random rotations, width and height shifts, zooming, shearing, and nearest-fill mode for pixel filling. These augmentations introduced variability into the training data, helping the model to better handle different scenarios and reduce overfitting.

After the initial training phase, we fine-tuned the last 10 layers of the ResNet50 model. Fine-tuning involved unfreezing these layers to allow them to learn specific features from our dataset. The model was then retrained for an additional 15 epochs with a reduced learning rate of $1e-6$ (0.000001). This lower learning rate helped ensure that the fine-tuning adjustments were more precise and controlled.

For the first five epochs of training before fine-tuning, the ResNet50 model achieved a training accuracy of 56.86% and a training loss of 80.87%. This initial phase helped the model begin learning the distinguishing features for diabetic retinopathy detection.

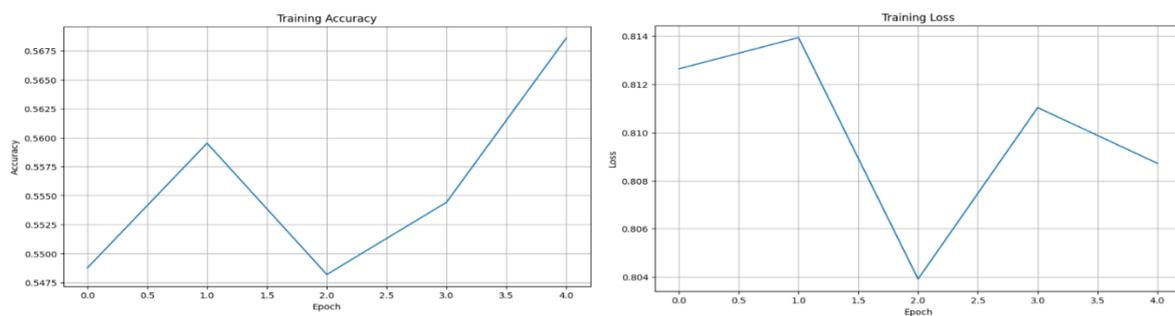


Figure 29. Performance of ResNet50 on the MuReD dataset before fine-tuning.

After fine-tuning the last 10 layers and training for an additional 15 epochs, the model's performance improved, achieving a training accuracy of 59.75% and a training loss of 68.84%. This fine-tuning step allowed the model to adapt more specifically to the MuReD dataset, refining its ability to identify relevant features.

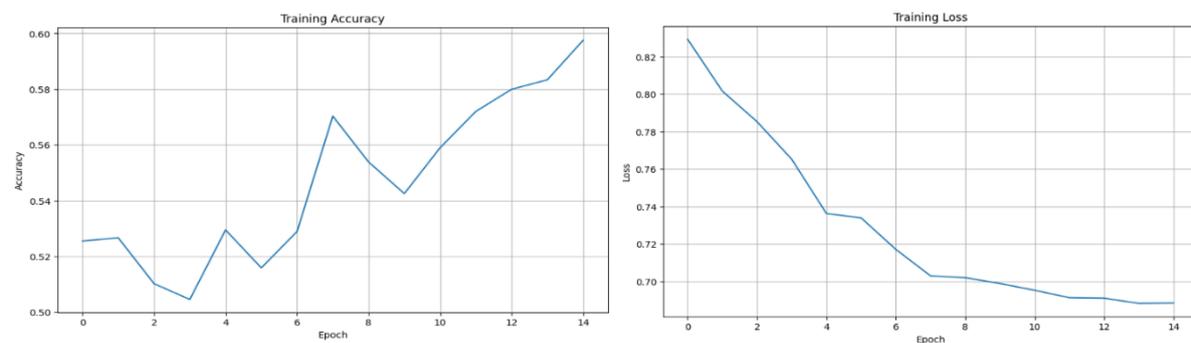


Figure 30. Performance of ResNet50 on the MuReD dataset after fine-tuning.

When evaluated on the test set, the model achieved an accuracy of 70.05% and a test loss of 63.56%. Despite the improvement in accuracy, the classification report highlighted that the model still faced challenges due to class imbalance:

Classification Report (Test Set):				
	precision	recall	f1-score	support
No_DR	0.78	0.73	0.75	345
DR	0.22	0.26	0.24	99
accuracy			0.63	444
macro avg	0.50	0.50	0.50	444
weighted avg	0.65	0.63	0.64	444

Figure 31. Classification Report of the test set in MuReD Dataset/ResNet50.

The classification report indicates that while the model performed reasonably well for the 'No_DR' class, it struggled with the 'DR' class, reflecting the impact of the imbalanced dataset on model performance. The weighted average f1-score of 0.64 shows an overall moderate performance, with significant room for improvement in detecting positive cases of diabetic retinopathy. To further analyze the model's performance, we also generated a confusion matrix for the test set, which provides a detailed breakdown of the true positive, true negative, false positive, and false negative predictions.

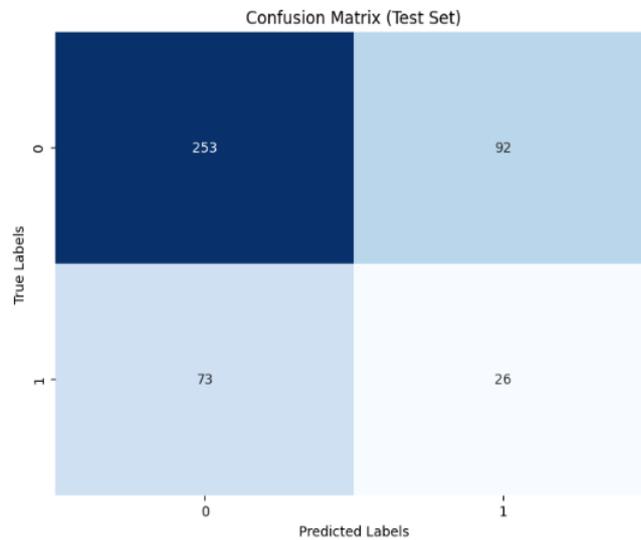


Figure 32. Confusion Matrix of the test set in MuReD Dataset/ResNet50.

For the third dataset, we applied the ResNet50 model along with comprehensive data augmentation techniques to enhance model generalization and robustness. These techniques included horizontal and vertical flips, random rotations up to 30 degrees, width and height shifts, zooming, shearing, and nearest-fill mode for pixel filling. These augmentations were crucial in expanding the dataset and exposing the model to a broader range of variations, thereby improving its ability to generalize to unseen data.

Initially, the model was trained for 25 epochs with a dynamic learning rate starting from 1e-5 (0.00001). The learning rate was scheduled to decrease by a factor of 0.5 if the validation loss did not improve for 5 consecutive epochs. This adaptive learning rate strategy aimed to optimize the model's convergence and performance during training.

After 25 epochs, the model achieved a training accuracy of 55.92% and a training loss of 68.72%. Validation results showed an accuracy of 45.66% and a validation loss of 69.03%, with the model using a learning rate of 2.5e-6 (0.0000025) at this stage.

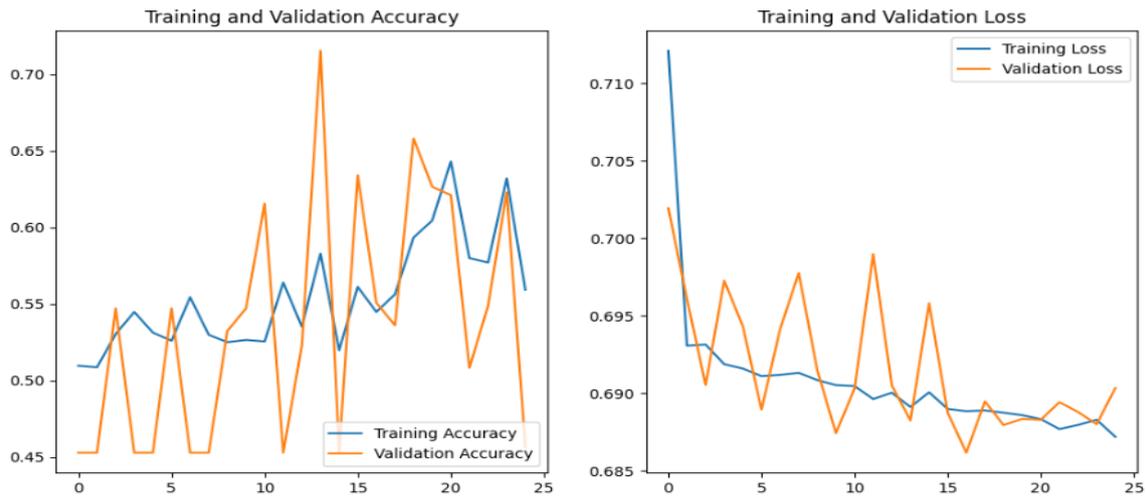


Figure 33. Performance of ResNet50 on the third dataset before fine-tuning.

Following the initial training phase, we fine-tuned the last 10 layers of the ResNet50 model to adapt its features specifically to the nuances of the third dataset. This fine-tuning step continued for an additional 25 epochs, with a dynamic learning rate starting from $1.25e-6$ (0.00000125). The adjusted learning rate further refined the model's learning process, potentially enhancing its performance on the specific characteristics of the dataset.

Upon fine-tuning the model for an additional 25 epochs, significant improvements were observed. The training accuracy increased to 73.70% with a training loss of 59.47%. Validation accuracy substantially improved to 89.65%, accompanied by a validation loss of 53.96%. During fine-tuning, the learning rate was adjusted to $6.25e-7$ (0.000000625).

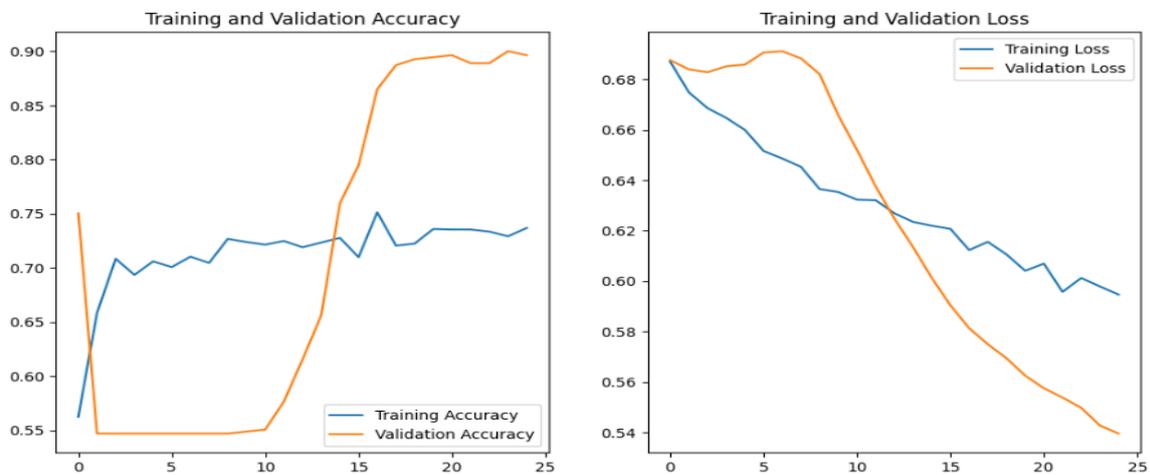


Figure 34. Performance of ResNet50 on the third dataset after fine-tuning.

Evaluation on the test set produced promising results, with an accuracy of 88.31% and a test loss of 54.16%. The classification report reflects balanced performance across precision, recall, and F1-scores for both classes, indicating robust learning and effective prediction capabilities of the model.

Classification Report (Test Set):				
	precision	recall	f1-score	support
No_DR	0.53	0.49	0.51	118
DR	0.51	0.55	0.53	113
accuracy			0.52	231
macro avg	0.52	0.52	0.52	231
weighted avg	0.52	0.52	0.52	231

Figure 35. Classification Report of the test set in the third Dataset/ResNet50.

To further analyze the model's performance, we also generated a confusion matrix for the test set, which provides a detailed breakdown of the true positive, true negative, false positive, and false negative predictions.

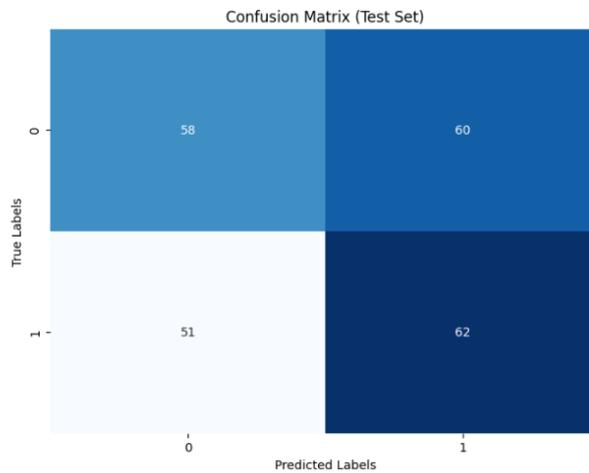


Figure 36. Confusion Matrix of the test set in the third Dataset/Resnet50.

- **Custom Model**

We also proposed another model from scratch that follows a convolutional neural network architecture tailored for image classification tasks. It comprises four convolutional layers designed to progressively extract features from input images. Each convolutional layer is followed by a rectified linear unit (ReLU) activation function to introduce non-linearity, enhancing the model's capability to learn complex patterns. Max pooling operations are applied after each convolutional layer to reduce spatial dimensions and control overfitting. The network culminates in two fully connected layers: the first aggregates features extracted by the convolutional layers, and the second produces the final output for binary classification using a sigmoid activation function. Data augmentation techniques were employed during training to enhance model generalization, and a dynamic learning rate strategy was adopted to optimize training efficiency by adjusting the learning rate based on validation performance. The following figure shows the architecture of our model.

Layer (type)	Output Shape	Param #
conv2d_5 (Conv2D)	(None, 253, 253, 8)	224
max_pooling2d_5 (MaxPooling2D)	(None, 126, 126, 8)	0
conv2d_6 (Conv2D)	(None, 124, 124, 16)	1168
max_pooling2d_6 (MaxPooling2D)	(None, 62, 62, 16)	0
conv2d_7 (Conv2D)	(None, 60, 60, 32)	4640
max_pooling2d_7 (MaxPooling2D)	(None, 30, 30, 32)	0
conv2d_8 (Conv2D)	(None, 28, 28, 64)	18496
max_pooling2d_8 (MaxPooling2D)	(None, 14, 14, 64)	0
flatten_1 (Flatten)	(None, 12544)	0
dense_2 (Dense)	(None, 100)	1254500
dropout_1 (Dropout)	(None, 100)	0
dense_3 (Dense)	(None, 2)	202

Figure 37. Architecture of the custom model.

We trained the model for 60 epochs using various data augmentation techniques such as resizing, random horizontal and vertical flips, random rotation, and normalization to enhance the training data. The learning rate was dynamically adjusted using a ReduceLRonPlateau scheduler, which reduces the learning rate by a factor of 0.5 if the validation loss does not improve for 20 epochs.

The results of our model training indicate strong performance across all datasets. On the training set, we achieved an accuracy of 94% with a training loss around 12%. The validation set showed a similar performance with an accuracy of 93% and a validation loss also around 12%, indicating that the model generalizes well to unseen data.

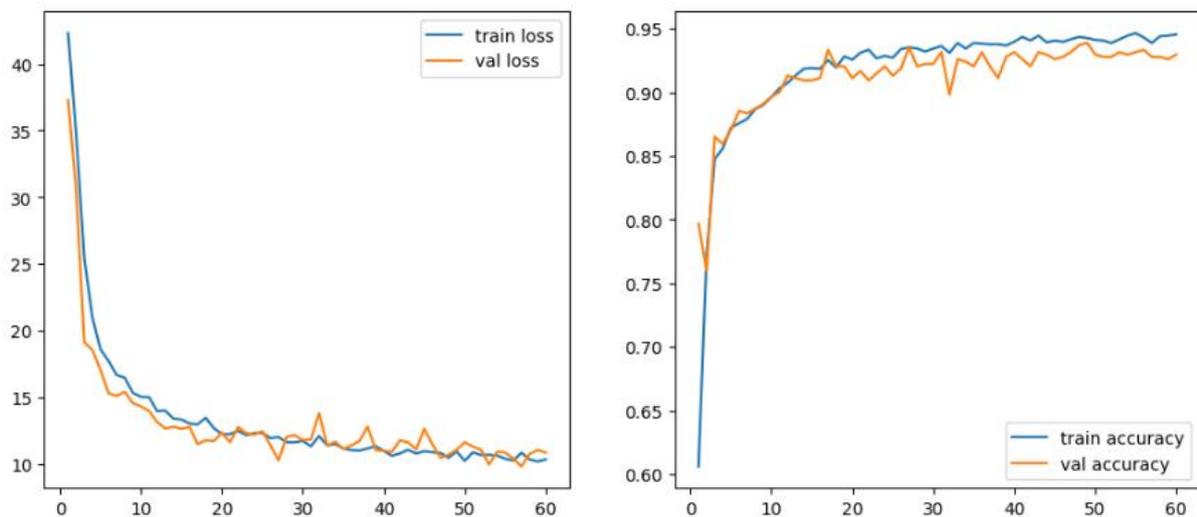


Figure 38. performance of the custom model on the third dataset.

	precision	recall	f1-score	support
0.0	0.93	0.96	0.94	1050
1.0	0.95	0.93	0.94	1026
accuracy			0.94	2076
macro avg	0.94	0.94	0.94	2076
weighted avg	0.94	0.94	0.94	2076

Figure 39. Classification Report of the training set in the third Dataset using Custom model.

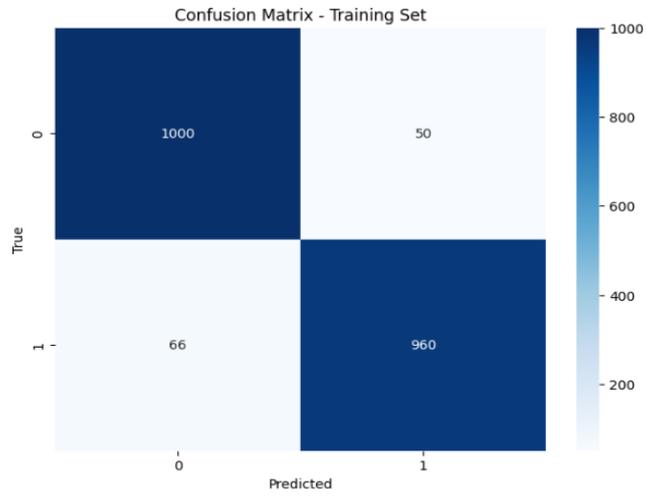


Figure 40. Confusion Matrix of the training set in the third Dataset using Custom model.

	precision	recall	f1-score	support
0.0	0.92	0.93	0.92	245
1.0	0.94	0.93	0.94	296
accuracy			0.93	541
macro avg	0.93	0.93	0.93	541
weighted avg	0.93	0.93	0.93	541

Figure 41. Classification Report of the validation set in the third Dataset using Custom model.

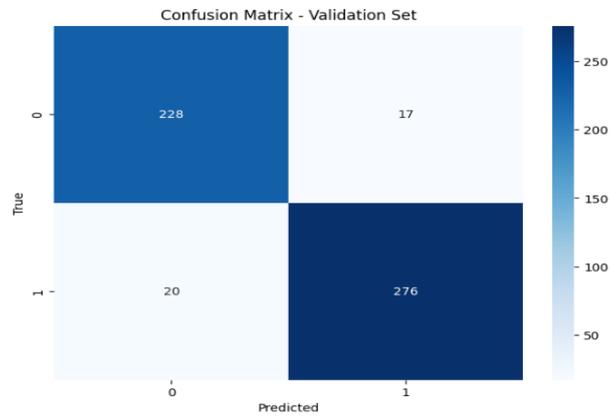


Figure 42. Confusion Matrix of the validation set in the third Dataset using Custom model.

For the test set, the model maintained a high accuracy of 94%. The classification report and confusion matrix demonstrate that the positive class was successfully detected, reflecting the model's ability to accurately distinguish between classes as shown in figure 43 and figure 44.

```

                precision    recall  f1-score   support

 0.0           0.93       0.94       0.93         113
 1.0           0.94       0.93       0.94         118

 accuracy              0.94         231
 macro avg           0.93       0.94       0.94         231
 weighted avg       0.94       0.94       0.94         231

```

Figure 43. Classification Report of the test set in the third Dataset using Custom model..

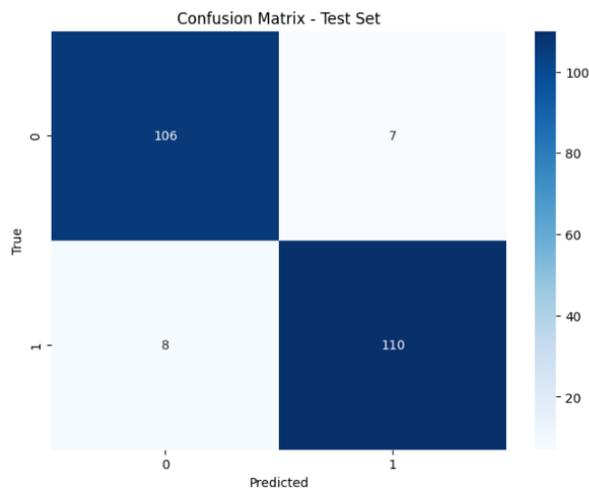


Figure 44. Confusion Matrix of the test set in the third Dataset using Custom model.

The table below shows the accuracy of different architectures on various datasets.

	RFMiD2.0	MuReD	Third dataset
VGG16 model	95.29%	85.36%	x
ResNet50 model	x	70.05%	88.31%
Custom model	x	x	94%

Table 4. Accuracy of different models trained on various datasets.

	RFMiD2.0	MuReD	Third dataset
VGG16 model	17%	28%	x
ResNet50 model	x	22%	51%
Custom model	x	x	94%

Table 5. Diabetic Retinopathy class precision for each model

In both tables, cells marked with 'x' indicate cases where computational constraints prevented the execution of experiments with certain models and datasets.

We can see from table 4 that the lowest accuracy is 70.05% which is still reasonable indicating that the models have done a great work, however table 5 shows the precision for the positive class which is our main goal in order to detect diabetic retinopathy. And it showed that the models struggled to identify the positive class which is the most important class except for the custom model that had a 94% precision of the positive class.

3.5 Deployment

To make our research more impactful we developed a web application using Django that enables users to upload images and choose between VGG16, ResNet50, or our custom CNN model for predicting diabetic retinopathy. This application enhances accessibility by allowing users to obtain predictions quickly and efficiently based on their chosen model. It serves as a practical tool for leveraging deep learning models in a real-world application, facilitating informed decision-making regarding diabetic retinopathy detection directly from user-uploaded images.

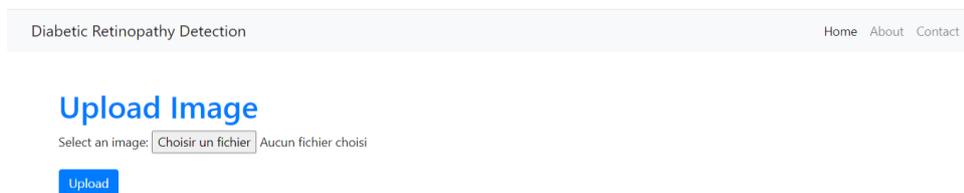


Figure 45. Illustration of our web application.

Image Uploaded Successfully



Result: Negative

Probability: 0.0%

[Upload Another Image](#)

Figure 46. Model prediction of the uploaded retinal image.

3.6 Conclusion and interpretation

In this chapter, we explored various datasets and developed different models, the VGG16 model did not perform well due to class imbalance with 17% and 28% positive class precision on both RFMiD 2.0 and MuReD datasets respectively. The ResNet50 model showed a slightly better performance especially when trained on the third dataset which is more balanced with a 51% positive class precision, the custom model showed a very good results on the third dataset with a 94% positive class precision and 94% accuracy.

Conclusion and future perspectives

Throughout this master's thesis, our focus has centered on leveraging deep neural networks for the detection of diabetic retinopathy. A central challenge we encountered was addressing class imbalance, a common issue in medical image analysis. To mitigate this challenge, we explored diverse datasets and implemented different techniques such as data augmentation and fine-tuning. These efforts were pivotal in enhancing model robustness and generalization capabilities.

We proposed and evaluated various models, noting that certain architectures struggled with class imbalance while others performed significantly better when trained on more balanced datasets. Our custom model, in particular, demonstrated strong positive class precision and overall accuracy, underscoring the importance of using appropriate datasets and tailored neural network architectures. Our exploration involved developing and evaluating various neural network architectures aimed at optimizing diagnostic accuracy in medical imaging. This process underscored the complexities and nuances inherent in applying deep learning to the critical medical domain of diabetic retinopathy detection.

Looking ahead, continual refinement and the exploration of innovative methodologies are essential to advance the efficacy and reliability of these detection systems. Future research should focus on exploring additional datasets to ensure comprehensive training and validation of models across diverse populations and conditions. Additionally, experimenting with different neural network architectures and hybrid models may uncover new pathways to improve detection accuracy and robustness.

We aim to translate these promising results into practical applications by integrating them into routine clinical practice and obtaining feedback from healthcare professionals. By doing so, we hope to ensure that our models not only achieve high performance metrics but also effectively assist in the early detection and management of diabetic retinopathy in real-world settings.

Bibliography

1. Roglic, G. (2016). WHO Global report on diabetes: A summary. *International Journal of Noncommunicable Diseases*, 1(1), 3-8.
2. Katsarou, A., Gudbjörnsdóttir, S., Rawshani, A., Dabelea, D., Bonifacio, E., Anderson, B. J., ... & Lernmark, Å. (2017). Type 1 diabetes mellitus. *Nature reviews Disease primers*, 3(1), 1-17.
3. DeFronzo, R. A., Ferrannini, E., Groop, L., Henry, R. R., Herman, W. H., Holst, J. J., ... & Weiss, R. (2015). Type 2 diabetes mellitus. *Nature reviews Disease primers*, 1(1), 1-22.
4. Buchanan, T. A., & Xiang, A. H. (2005). Gestational diabetes mellitus. *The Journal of clinical investigation*, 115(3), 485-491.
5. Hassan, T., Rashid, M., Zafar, E., & Alotaibi, S. (2022). Performance Analysis of Deep-Neural-Network-Based Automatic Diagnosis of Diabetic Retinopathy. *Sensors*, 22(1), 205.
6. Fort, P. E., Losiewicz, M. K., N. Reiter, C. E., J. Singh, R. S., Nakamura, M., Abcouwer, S. F., Barber, A. J., & Gardner, T. W. (2011). Differential Roles of Hyperglycemia and Hypoinsulinemia in Diabetes Induced Retinal Cell Death: Evidence for Retinal Insulin Resistance. *PLOS ONE*, 6(10), e26498. <https://doi.org/10.1371/journal.pone.0026498>
7. Kumar, K. S., Bhowmik, D., Harish, G., Duraivel, S., & Kumar, B. P. (2012). Diabetic retinopathy-symptoms, causes, risk factors and treatment. *The Pharma Innovation*, 1(8).
8. Ansari, P., Tabasumma, N., Snigdha, N. N., Siam, N. H., Panduru, R. V., Azam, S., ... & Abdel-Wahab, Y. H. (2022). Diabetic retinopathy: an overview on mechanisms, pathophysiology and pharmacotherapy. *Diabetology*, 3(1), 159-175.
9. Wang, W., & Lo, A. C. (2018). Diabetic retinopathy: pathophysiology and treatments. *International journal of molecular sciences*, 19(6), 1816.
10. Kropp, M., Golubnitschaja, O., Mazurakova, A., Koklesova, L., Sargheini, N., Vo, T. T. K. S., ... & Thumann, G. (2023). Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation. *Epma Journal*, 14(1), 21-42.
11. “Diabetic Retinopathy - Can Lead to Complete Blindness”. - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Difference-between-Normal-Retina-and-Diabetic-Retinopathy_fig2_282609747
12. Ertel, W. (2018). *Introduction to artificial intelligence*. Springer.
13. Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, 9(1), 381-386.
14. Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2(6), 420.

15. Kandel, I., & Castelli, M. (2020). Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review. *Applied Sciences*, 10(6), 2021.
16. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12), 6999-7019.
17. Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9, 611-629.
18. Nash, W., Drummond, T., & Birbilis, N. (2018). A review of deep learning in the study of materials degradation. *npj Materials Degradation*, 2(1), 37.
19. Han, X., Zhong, Y., Cao, L., & Zhang, L. (2017). Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. *Remote Sensing*, 9(8), 848.
20. Rodriguez Rivera, Manuel Alejandro; Al-Marzouqi, Hasan; Liatsis, Panos (2022), "Multi-Label Retinal Diseases (MuReD) Dataset", Mendeley Data, V1, doi: 10.17632/pc4mb3h8hz.1
21. Sachin Panchal, Ankita Naik, Manesh Kokare, Samiksha Pachade, Rushikesh Naigaonkar, Prerana Phadnis, Archana Bhange. (2023). Retinal Fundus Multi-Disease Image Dataset (RFMiD) 2.0. IEEE Dataport. <https://dx.doi.org/10.21227/mrd2-ap11>
22. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
23. Ji, L., Tian, H., Webster, K. A., & Li, W. (2021). Neurovascular regulation in diabetic retinopathy and emerging therapies. *Cellular and Molecular Life Sciences*, 78, 5977-5985.

Abstract

Diabetic retinopathy, a severe complication of diabetes, can lead to vision loss if not detected early. This Master's thesis develops a decision support system using deep learning to assist healthcare professionals in diagnosing diabetic retinopathy from retinal images. Our approach integrates medical imaging and artificial intelligence, employing various deep learning models trained on extensive datasets. Several techniques were used to enhance model robustness. Multiple models were tested on different datasets, achieving reasonable results. This system aims to improve diagnostic accuracy and efficiency, contributing significantly to the early detection and management of diabetic retinopathy.

Keywords : Diabetic retinopathy, deep learning, medical imaging, early detection, retinal images, healthcare diagnostics.

Résumé

La rétinopathie diabétique, une complication grave du diabète, peut entraîner une perte de vision si elle n'est pas détectée tôt. Ce mémoire développe un système d'aide à la décision utilisant le deep learning pour aider les professionnels de santé à diagnostiquer la rétinopathie diabétique à partir d'images rétiniennes. Notre approche intègre l'imagerie médicale et l'intelligence artificielle, en utilisant divers modèles d'apprentissage profond formés sur de vastes ensembles de données. Plusieurs techniques ont été utilisées pour améliorer la robustesse du modèle. Plusieurs modèles ont été testés sur différents ensembles de données, obtenant des résultats raisonnables. Ce système vise à améliorer la précision et l'efficacité du diagnostic, contribuant ainsi de manière significative à la détection précoce et à la gestion de la rétinopathie diabétique.

Mots-clé : rétinopathie diabétique, apprentissage profond, imagerie médicale, détection précoce, images rétiniennes, aide à la décision.

الملخص

يمكن أن يؤدي اعتلال الشبكية السكري، أحد المضاعفات الخطيرة لمرض السكري، إلى فقدان البصر إذا لم يتم اكتشافه مبكرًا. تعمل هذه الأطروحة على تطوير نظام دعم القرار باستخدام التعلم العميق لمساعدة الأطباء على تشخيص اعتلال الشبكية السكري بناءً على صور الشبكية. يدمج نهجنا التصوير الطبي والذكاء الاصطناعي، باستخدام نماذج التعلم العميق المختلفة المدربة على مجموعات البيانات الكبيرة. تم استخدام العديد من التقنيات لتحسين متانة النموذج. تم اختبار عدة نماذج على مجموعات بيانات مختلفة، وتم الحصول على نتائج معقولة. ويهدف هذا النظام إلى تحسين دقة التشخيص وكفاءته، وبالتالي المساهمة بشكل كبير في الكشف المبكر عن اعتلال الشبكية السكري وإدارته.

الكلمات المفتاحية : اعتلال الشبكية السكري، التعلم العميق، التصوير الطبي، الكشف المبكر، صور الشبكية، نظام دعم القرار، نماذج التعلم العميق.