N^o d'ordre: XX/2024

THESIS SUBMITTED TO THE

UNIVERSITY OF MOHAMED BOUDIAF – MSILA

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
DEPARTMENT OF COMPUTER SCIENCE

MASTER THESIS

In partial fulfillment of the requirements for the degree of

Master in : Computer science

Specialty : Artificial Intelligence

By :

Djamel Aziz Tayeb

Obeyda Toumi

Title of the thesis

Macular Edema disease Detection and Classification

Under the supervision of

Lamri Sayad and Abdessattar Ghemougi

Composition of the jury

M.	**Said Kadri**	Lecturer A	University of Msila	President
Ms.	**Rahima Bentrcia**	Lecturer A	University of Msila	Examiner

2023/2024

Dedication

My parents, no amount of words will be enough to tell how grateful I am to you. Having parents like you is the biggest of all blessings! Thank you for all that you've done for me and all that you're still doing.

My sister. I am truly grateful for the bond we share. You have always been there for me, offering your unconditional love and support in both good times and bad. Your presence in my life brings me immense joy and comfort.

My friends, thanks for always cheering me up whenever life knocks me down. Thank you so much for being my friends!

Djamel Aziz Tayeb

إهداء

"وآخر دعواهم أن الحمد لله رب العالمين" الحمد لله على التمام، الحمد لله على الكمال.

قال تعالى: "وأما بنعمة ربك فحدث" ، سأحدثكم عن إحدى أعظم النعم، ألا وهي نعمة البصر. بفضل هذا البحث، تمكنت من إدراك هذه النعمة وفضلها. الحمد لله حمداً كثيراً طيباً مباركاً فيه.

قال تعالى: "لئن شكرتم لأزيدنكم"، وإني أشكر نعمتك التي أنعمت بها علي يا الله.

الى نفسي

قال ابن تيمية:" ما لا يكون بالله لا يكون، وما لا يكون لله لا ينفع ولا يدوم." وقال الإمام مالك: "ما كان لله دام واتصل، وما كان لغيره انقطع وانفصل."

الى عائلتي

أهدي ثمرة جهدي المتواضع إلى من وهبوني الحياة والأمل، والنشأة على شغف الاطلاع والمعرفة، ومن علموني أن أرتقي سلم الحياة بحكمة وصبر؛ برا، وإحسانا، ووفاءً لهما: والدي العزيز، ووالدتي العزيزة.

إلى من وهبني الله نعمة وجودهم في حياتي، إلى العقد المتين الذين كانوا عونا لي في رحلة بحثي: إخواني وأخواتي.

إلى من كاتفني ونحن نشق الطريق معا نحو النجاح في مسيرتنا العلمية، إلى رفيق دربي: طيب جمال عزيز.

وأخيرًا إلى كل من ساعدني، وكان له دور من قريب أو بعيد في إتمام هذه الدراسة، سائلا المولى أن يجزي الجميع خير الجزاء في الدنيا والآخرة.

ثم إلى كل طالب علم سعى بعلمه ليفيد الإسلام والمسلمين بكل ما أعطاه الله من علم ومعرفة.

عبيدة تومي

Acknowledgments

we want to express our deepest gratitude and appreciation to the following individuals and institutions that have contributed to the completion of this thesis.

First and foremost, we are immensely thankful to my thesis advisoras **Dr.Lamri Sayad Dr.Abdessattar Ghemougui**. Their guidance, expertise, and continuous support have been invaluable throughout this research journey. Their patience, constructive feedback, and unwavering commitment to my academic development have been instrumental in shaping this thesis. we are truly grateful for their mentorship. we would also like to extend our gratitude to the faculty members of the Informatics Department at Msila University, whose knowledge and expertise have greatly influenced our intellectual growth. Their dedication to teaching and research has been a constant source of inspiration.

We would also like to express our appreciation to the participants who willingly took part in this study. Their cooperation and willingness to share their experiences and insights have enriched the findings of this thesis.

Lastly, we would like to thank all those who have directly or indirectly contributed to this thesis. Your contributions, whether through discussions, feedback, or encouragement, have played a significant role in shaping the outcome of this research. To everyone mentioned above and to those whose names might have inadvertently been omitted, please accept my sincere thanks. Your support and encouragement have been crucial in making this thesis a reality.

ملخص

مرض الشبكية المرتبط بوجود حالات نضح بالقرب من البقعة هو اعتلال البقعة السكري. يمكن أن تسبب رواسب السوائل أو تراكمها داخل البقعة الوذمة البقعية،والتي يمكن أن تؤدي إلى عدم وضوح الرؤية أو تشوهها وتغير لونها. بالإضافة إلى ذلك ، فإن تلف الخلايا المخروطية في مناطق شبكية العين يضعف قدرة الشخص على رؤية التفاصيل الدقيقة ويمنعها من أداء العديد من الأنشطة القائمة على الرؤية. الغرض من هذا العمل هو تنفيذ مصنف التعلم العميق ، الشبكة العصبية التلافيفية ، لحل مشكلة الكشف عن الوذمة البقعية وتصنيفها.

الكلمات المفتاحية: اعتلال الشبكية بمرض السكري ، التعلم الآلي ، التعلم العميق، الرعاية الصحية، المعلوماتية الصحية.

Abstract

A retinal disease associated with the presence of exudates near the macula is diabetic maculopathy. Fluid deposits or accumulation within the macula can cause macular edema (ME), which can lead to blurred or distorted vision and discoloration. Additionally, damage to the cone cells in areas of the retina impairs a person's ability to see minute details and prevents them from performing many vision-based activities. The purpose of this work is to implement a deep learning (DL) classifier,convolutional neural network (CNN), to solve the problem of macular edema detection and classification.

Keywords: Diabetec retinopathy,Machine learning, Deep leaning,Healthcare, Health informatics.

Résumé

Une maladie rétinienne associée à la présence d'exsudats près de la macula est la maculopathie diabétique. Les dépôts ou l'accumulation de liquide dans la macula peuvent provoquer un œdème maculaire (EM), ce qui peut entraîner une vision floue ou déformée et une décoloration. En outre, les dommages aux cellules du cône dans les zones de la rétine altèrent la capacité d'une personne à voir les détails et l'empêchent d'effectuer de nombreuses activités basées sur la vision. Le but de ce travail est de mettre en œuvre un classificateur d'apprentissage profond (DL), un réseau de neurones convolutifs (CNN), pour résoudre le problème de la détection et de la classification de l'œdème maculaire.

Mots Clée : Rétinopathie diabétique,apprentissage automatique,apprentissage profond, soins de santé, informatique de santé.

Contents

List of Figures

Introduction

General Introduction

Diabetic Retinopathy (DR) is one of the primary causes of blindness in people who are diagnosed with diabetes, which makes it an essential public health issue globally. Given the increasing prevalence of diabetes worldwide, and hence, the high burden of DR, reliable and accurate methods for the detection and diagnosis of early-stage DR are vitally needed. Although accurate, traditional diagnostic techniques are labor-intensive and rely extensively on experienced ophthalmologists. This case necessitates support from auto-detection and classification of DR in due time.

Over the past few years, with the rapid progress in Artificial Intelligence (AI) as well as Deep Learning (DL) advancements, the landscape has witnessed a paradigm shift on novel ways to solve this challenge. Convolutional Neural Networks (CNNs), a family of Deep Learning DL models, have successively shown remarkable performance on different image classification tasks, especially in the domain of medical image analysis. CNNs are proposed as a method of learning hierarchical features from images and also have a good accuracy in detecting complex patterns underpinning DR in retinal images.

Still, building a powerful DL model for DR classification generally demands large datasets with annotations, which are not always easy to obtain in medical domain. This is where transfer learning is very handy. Above-mentioned pre-trained models are known as the transfer learning models, where we fine-tune the pre-trained model on a specific task-related data (small data set) as pre-trained models have been trained on a large real-world dataset. These high level models can be fine-tuned on retinal images to achieve high accuracy in detecting and classifying different stages of DR.

In this project work, we concentrate on developing a deep learning model on classification of Diabetic Retinopathy using Convolutional Neural Networks along with transfer learning methods. The ultimate goal is to design a reliable and automatic system that helps health professionals to detect and diagnose DR in the early stages, in an attempt to reduce vision loss among diabetic patients.

The following sections will describe in detail the methodology used in this study, including data collection and preprocessing, sample selection, training and optimization

procedures, performance evaluation If combining CNN with transfer learning, this study seeks to contribute to the growing body of knowledge medical.

Context

Diabetic Retinopathy(DR) is a serious complication of diabetes and is one of the leading causes of blindness worldwide. This complaint affects the blood vessels of the retina, causing lesions that can lead to vision loss if not detected and treated beforehand diabetes damages the small blood vessels in the retina, which can beget fluid or blood to blunder into the eye, as well as the conformation of new abnormal blood vessels.

These changes can affect vision gradationally and permanently, and in the most serious cases, lead to unrecoverable blindness.

Regular webbing for Diabetic Retinopathy(DR) is thus essential to help serious complications associated with this complaint. People with diabetes are encouraged to suffer regular eye examinations to describe any abnormalities as soon as possible.

Beforehand, discovery allows applicable treatment to begin as soon as possible, which can help decelerate the progression of the complaint and help vision loss. still, despite the significance of early discovery, numerous people with diabetes don't admit regular follow-up due to constraints in penetrating technical health services. Indeed, webbing tests for Diabetic Retinopathy(DR) frequently require technical outfits and the intervention of good health professionals, which can make access to these services delicate for numerous people.

Problem

Our society suffers from the miracle of only consulting a croaker when the complaint intensifies, indeed with age. numerous people with diabetes don't suffer regular webbing for Diabetic Retinopathy(DR), on the one hand, because of the lack of mindfulness and information about the significance of this webbing, and because of restrictions on access to technical health services.

Lack of mindfulness and information about Diabetic Retinopathy(DR) is one of the main reasons why numerous people with diabetes don't admit regular webbing for this complaint. indeed, numerous people aren't apprehensive of the threat of developing Diabetic Retinopathy(DR) and the serious complications that can affect if it isn't detected and treated beforehand. As a result, they aren't laboriously seeking regular webbing for Diabetic Retinopathy(DR), performing in delayed opinion and treatment of the complaint.

In addition, indeed for those who are apprehensive of the significance of webbing

for Diabetic Retinopathy(DR), access to technical webbing services can be delicate due to colorful constraints, similar to the lack of health centers equipped to perform comprehensive ophthalmological examinations, long delay times for an appointment with an ophthalmologist, and the high costs associated with these services.

As a result, numerous people with diabetes don't admit regular webbing for Diabetic Retinopathy(DR), performing in an increased threat of serious complications associated with this complaint is similar to vision loss and blindness. it's thus critical to find results that grease access to these webbing services and ameliorate the early discovery of Diabetic Retinopathy(DR).

Scope and Objectives

Diabetic Retinopathy(DR) is an arising exploration field in medical image processing.It can be used in real-time to help ophthalmologists diagnose retinopathy in a bettered and effective way.

This will promote the development of collaboration between engineering work and the medical field, produce a new arena, and help bridge the gap that exists between the engineering professions in exploration and assiduity in Algeria.

The purpose of this study is to develop more dependable and accurate image processing and pattern recognition styles for automated webbing Diabetic Retinopathy(DR).
The main objects of the exploration are as follows

- Design and develop Diabetic Retinopathy(DR) examination systems with state-of-the-art image processing capabilities and pattern recognition algorithms that intelligently classify retinal image abnormalities to describe and classify the threat of Diabetic Retinopathy(DR) and unforeseen vision loss. The purpose of the proposed system isn't to circumvent the croaker's part in clinical practice, but to help croakers in saving time spent on examining retinal images.

- Design fluently accessible telescreen system algorithms. a cost-effective result in saving vision in people with diabetes.

- Raise mindfulness of the content among academics similar to scholars and experimenters in the healthcare systems and technologies are demanded.

- To create a terrain for active exploration in biomedical operations in exploration and educational institutes of Algeria.

Motivation

Application of both machine literacy(ML) and Deep literacy(DL) in each area of health and medical image analysis has set out some positive trends in the identification and opinion of for presence of colorful sickness.

Thus, we are convinced that by tapping into these available technological tools, one can design a DR webbing tool that is effective, accurate, and universally available. Thus, using our proposed system to detect DR at a much earlier stage and get the appropriate care, the risk posed by this complaint could be considerably minimized and the condition improved. Living standards for various cases with diabetes.

Other than refining the healthcare issues, our result has several bottom sides. Through simplifying the first phase of case finding and treatment, one should convenience both cases and health care practitioners. In the same way, our AI-rooted webbing instrument can raise the extensiveness of DR webbing, which allows pertinent examinations in multifarious healthcare environments without technical sophistication.

This in turn can save a lot of the cost by avoiding costly medical complications that are time holding anywhere in the process. Similarly, getting an early diagnosis can help minimize the cases of requiring aggressive treatment and, probably, reparative surgeries, which would help lighten the load on both cases and the overall health management industry. Consequently, our result could not only improve the existing conditions in the health sector but also help give a better quality of life to those with diabetes cases.

Proposed System

Thus, our exactly proposed solution rests upon the concepts of Artificial Intelligence and more strictly Deep Learning in the sense of adopting Convolutional Neural Networks for the construction of our automatic Diabetic Retinopathy detector. Demospec must be perfect enough to single out findus images and find Diabetic Retinopathy signs, as the programs are trained with large amounts of data. As such, the trained model would do the feature of recognizing various aspects and also the abnormalities depicted in the fundus image including; microaneurysms, hemorrhages together with the exudates of preliminary diabetic retinopathy (DR).

The logic behind having a system, that is able to monitor the deterioration of health and consequently address it in a manner befitting what appears to be the further deterioration in health. The concerned doctors, specialists as well as other concerned individuals working in the branch of health care will be able to view the outcome of the analysis made by the system, review the conclusion made by the system for the concerned diagnostic,

and treat the patient accordingly.

Consequently, based on the AI Integration of the system developed in this paper, the designed DR screening model proposed the most accurate and efficient screening approach. This is made possible through the creation of an application that is within reach as well as easily expandable to assist in detecting the ailment early, hence improving the current patient outcomes by assisting in the identification of diabetic retinopathy and odds of vision impairment.

Dissertation Structure

The thesis dissertation is structured into the following chapters:

Chapter 1: Diabetic Retinopathy (DR).

Chapter 2 : Machine Learning (ML) and Deep Learning (DL).

Chapter 3 : Proposal.

Chapter 4 : Implementation and Results

Finally, Conclusions are drawn, and future works are presented.

Chapter 1
Diabetic Retinopathy (DR)

Chapter 1

Diabetic Retinopathy (DR)

1.1 Introduction

The modern society known for high speed, busy living habits, and frequent changes in eating habits has a plethora of challenges emerging in nature hanging over mankind. Amongst all these diseases, diabetes takes one of the most prominent positions of being one of the leading conditions that affects an individual's quality of life dramatically.

This disease is a puzzle shrouded in medical terms; The functioning of our bodies requires a mystery complex: sugar. Or the power of glucose in energizing the cardinal reagents in each cell, while the play remains dormant, and the band's symphony still concealed.

Therefore, for people suffering from this disease, it is not enough to have a 'diabetes' status; it becomes a way of life where your body and mind undergo a daily struggle to stay healthy. With the rising incidence numbers for this disease internationally, awareness becomes a need, not only for education but also for prevention, and efficient prevention and care management strategies that are obligatory must also be an inherent part of daily global life activities. The disease thus remains deserves attention not only for the details explaining it but also for the approaches of how they can be used to fight it and live a full-fledged healthful life every day.

Proceeding that, it is suggested to witness, how the disease dimensions can be unfolded and discussed. It comprehensively presents different symptoms of diabetes as well as the different types of it, the common ways through which it can be treated, and how to live a healthy life with the disease. We must be aware of how important insulin is in our body system: thus knowing that any shift of balance in this mechanism is causative of such disease.

Because knowledge is power, our journey aims to inform people about this general

health issue that affects millions of people around the world, and also, we get to know how we can fight it and live along with it.

This chapter gives a brief clinical overview of diabetic retinopathy (DR), though more detailed discussions are introduced in the next chapter. It outlines the risk factors, signs, features, and types of pathology involving the diabetic retinopathy (DR) discipline. In addition, diabetic retinopathy (DR).

1.2 Diabetes

Diabetes mellitus (DM) is a chronic disease with long-term microvascular and microvascular complications which include diabetic nephropathy, neuropathy, and retinopathy [1]. This disease causes death, disability, and blindness among persons who are 20–74 years of age [2].

Moreover, 80% of blindness in this age group is related to Diabetic Retinopathy(DR). Its global prevalence is rapidly increasing. According to the International Diabetic Federation (IDF), it was estimated that nearly 537 million adults (10.5%) all over the world suffer from Diabetes mellitus (DM) in 2021.

Almost one in two (240 million; 44.7%) adults with diabetes are unaware that they have the condition (undiagnosed diabetes) and nearly 90% of diabetic cases are type 2 DM (T2DM). This global estimate is expected to rise to 643 million (11.3%) in 2030 and to 783 million (12.2%) by 2045 [3, 4]. There are two types of Diabetes mellitus (DM) i.e., type 1 and type 2 diabetes. Figure 1.1 shows the estimation of diabetes in all regions of the world.

1.3 Types of Diabetes

1.3.1 Type 1 Diabetes

Type 1 DM results from the destruction of beta cells in the pancreas, which leads to complete insulin deficiency. There are two forms of type 1 DM. One is an immune-mediated disease with autoimmune markers such as islet cell antibodies (ICA's) and autoantibodies to glutamic acid decarboxylase (GAD65). About 85-90% of patients with fasting hyperglycemia are positive for one or more of these markers. Strong human leukocyte antigen (HLA) associations also exist.

A second form of type 1 DM, now called idiopathic diabetes, has no known cause. Only a minority of patients fall into this group, mostly of Asian origin. The onset of this type is sudden, and it can occur at any age [5].

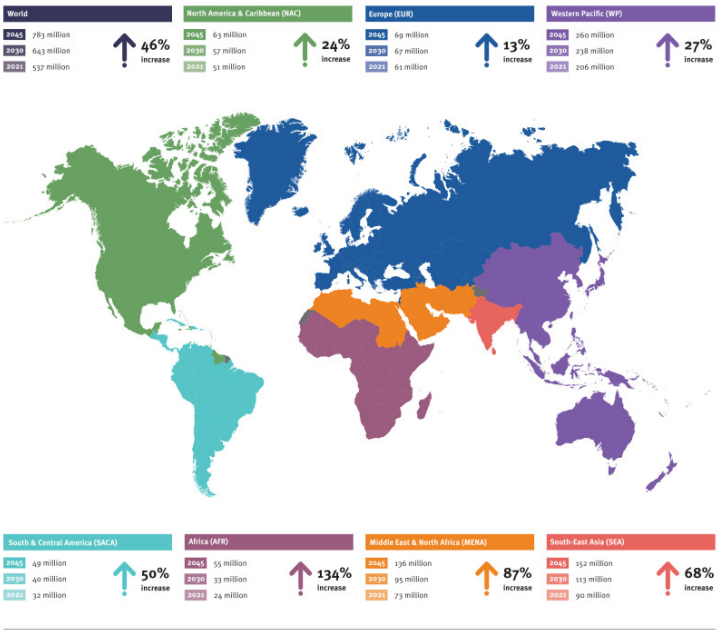


Figure 1.1: Number of people with diabetes worldwide and per IDF Region in 2021–2045 (20–79 years) [6].

1.3.2 Type 2 Diabetes

Type 2 DM is more common than type 1 DM because it is related to obesity, which means that those with a Body Mass Indicator (BMI) greater than 35 have a 37 times greater risk than those with a BMI less than 22. There is always some element of beta-cell failure whenever type 2 diabetes is diagnosed. Because the pathophysiology tends to be indolent, symptoms can sometimes be slight or so slow in onset that they are not.

Retinopathy is the commonest finding, with about 30% of all subjects newly diagnosed having detectable retinal lesions [5]. The onset is gradual, and it occurs mostly in adults. Figure: 1.2 shows the duration of DM and the presence of eye diseases during this duration.

Diabetes Type	Duration of Disease	Ocular Manifestations
Type 1	>10 years	60% have some retinopathy
	>15 years	Virtually all patients have some degree of Retinopathy ; 25% progress to proliferative diabetic retinopathy
	>20 years	50% progress to proliferative retinopathy
Type 2	At diagnosis,	20% have retinopathy
	>4 years	4% progress to proliferative retinopathy
	>15 years	60% to 80% have some retinopathy, up to 20% progress to proliferative retinopathy

Table 1.1: Duration of DM and presence of eye disease. [7]

1.4 Retina Structure

It is important to describe the structure of the retina before explaining diabetic retinopathy. Figure: 1.2 shows the anatomy of the human eye [8]. Retina has mainly two parts which are the macular (the center and main portion) and the peripheral retina. The peripheral retina gives us vision to the side, called peripheral vision. It is this part of the retina that is at work when we see something out of the corner of the eye.

In order to see no details, we must look straight ahead, using the macular, which is the central portion of the retina.

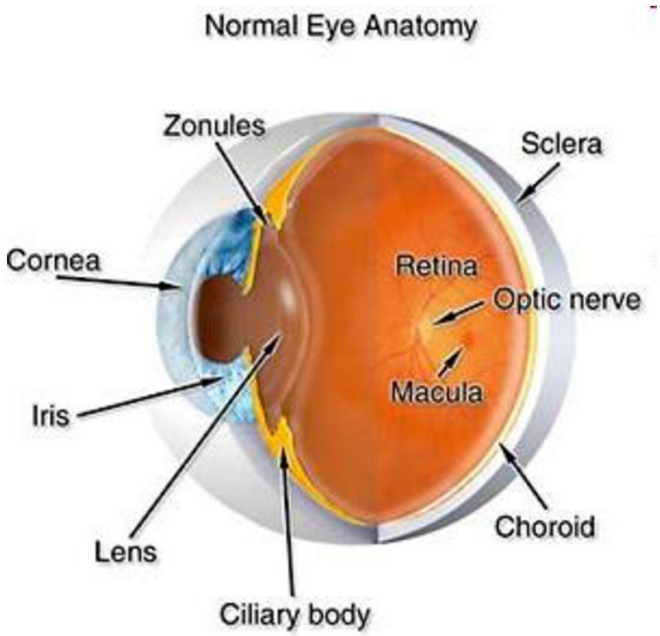


Figure 1.2: Anatomy of eye [8]

Even though the macular makes up only a small part of the retina, it is one hundred times more sensitive to detail than the peripheral retina. Figure: 1.3 shows the structure of retina and its main components. Dark curved lines are the blood vessels that bring oxygen and nutrition to the retina.

In order for the peripheral retina and macular to work properly, the blood vessels must be normal. The central part of the macular is called foveal. The brightest circular region of retina is called optic nerve, from where all blood vessels rise [9].

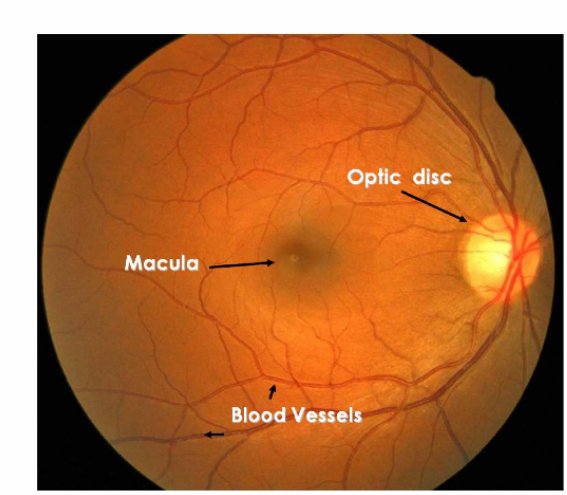


Figure 1.3: Structure of retina and its main components. [7]

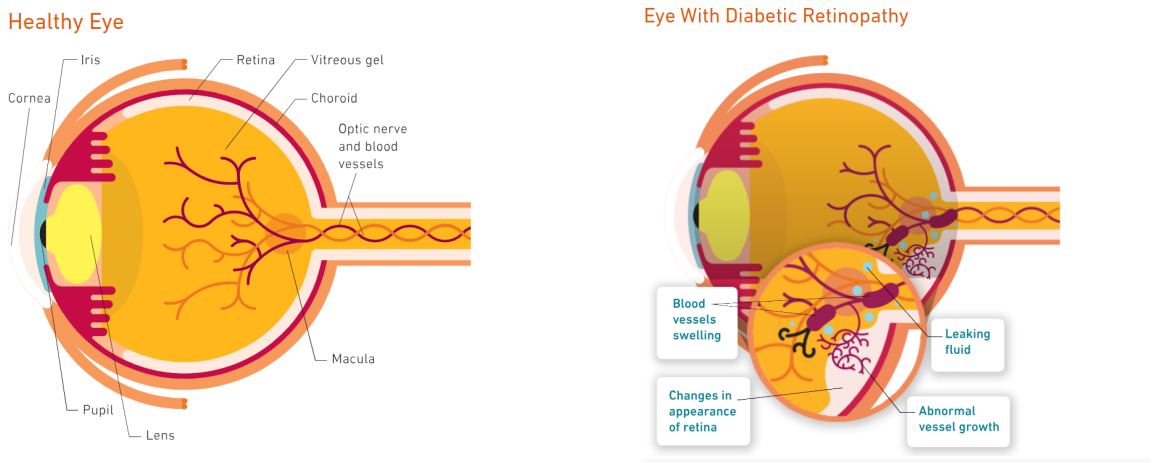
1.5 Retinal Diseases

There are many diseases affecting the retina and there are more than 20 common diseases, of which I will mention the diseases resulting from complications of diabetes which are 3: Diabetic retinopathy (DR), Diabetic macular edema (DME), Neovascular glaucoma (NG).

1.6 Diabetic Retinopathy (DR)

Diabetic Retinopathy(DR) is one of the leading causes of blindness among adults around the world, and the growth of DR is attached to Diabetes mellitus (DM) disease duration. DR is a disease that can affect and destroy blood vessels and harm the retina, which is the inside part at the end of each eye.

A person can see when the retina captures light and transforms it into signals to be decoded by the brain. There are two stages of DR [10]. The first is called non-proliferative DR and the second is called proliferative DR. The non-proliferative stage occurs near the beginning of DR, where blood vessels can be thinner and make a bulge that leads to leaking into the retina. The stage of proliferative happens when DR becomes worse, some blood vessels are blocked causing new blood vessels to expand or increase in the retina leading to dangerous vision difficulties. Figure: 1.4 shows a demonstration of how the normal eye may look in comparison to a proliferative stage where the eye with DR contains bulges, and blood vessels swelling leading to abnormal growth, leaking fluid, and some changes in the retina's appearance [11].



(a) Healthy Eye **(b)** Eye With DR

Figure 1.4: : A healthy eye vs. An eye with DR [11]

Diabetic Retinopathy(DR) is considered as epidemic disease, affecting one-third of diabetic patients. It is considering as the most frequent reason for blindness in diabetic persons. Therefore, besides personal caring programs, early exploration and treatment can reduce the risk of vision loss by 95% [12]. Figure: 1.5 (a) depicts an overview of healthy eye seeing 2 boys clearly, meanwhile Figure: 1.5 (b) shows a diabetic patient with DR seeing the same boys hardly with hazy and dark spots blocking the sharp vision.

(a) Healthy Eye **(b)** Eye With DR

Figure 1.5: scene of normal eye VS scene of eye with DR [13]

Diabetic Retinopathy (DR) directs to progressive changes in vasculature formation (including vascular tortuosity, branching angles, and calibers) and producing malformations (microaneurysms, hemorrhages, and exudates). DR is diagnosed by visually inspecting retinal fundus images for the presence of one or more retinal lesions like microaneurysms (MAs), hemorrhages (HEs), soft exudates (SEs), and hard exudates (EXs) [14] as described in Figure: 1.6.

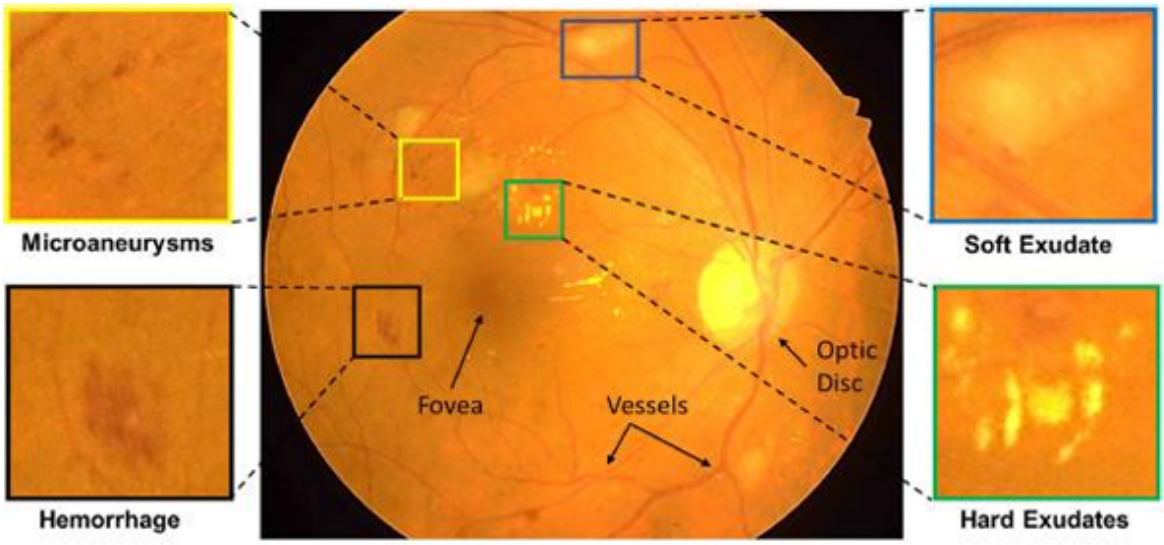


Figure 1.6: Fundus retinal image representation in the center. Pointing to normal formations (foveal, vessels, and optic disc) and deformations related to DR: The left section (MAs), and HEs and in right (SEs), and (EXs) [15]

.

1.7 Diabetic Macular Edema (DME)

The macular is the vital partition of a person's retina, allowing him to view and control the ability of detailed sight. Any leak into the macular produces macular edema. This leak is caused by the accumulation of fluid that is called diabetic macular edema (DME). It forms a sort of swelling, and in some cases, cyst formation in the macular, caused by chronic hyperglycemia [16].

Progressively, it may damage the clear vision, leading to partial or total vision loss. DME is commonly a consequence of people who previously have additional symptoms of DR and is secondary to retinal barrier rupture, where up to one-third of diabetic patients are diagnosed with DME.

The most significant molecule in the retinal barrier rupture is the vascular endothelial growth factor (VEGF). The initiation of anti-VEGF and steroid medications for treating DME has increased the understanding of pathophysiology [17]. However, the utilization of anti-VEGF drugs has shown that about one-third of patients built an immune system to intro-virtual therapy [18]. The diagnosis of macular edema is performed clinically. Traditionally, stereoscopic fundus photography was used as a standard screening system for diagnosing DME [19] but has since been replaced by Optical Coherence Tomography (OCT) images. DME can be classified into focal and diffuse. Focal macular edema is

characterized by the appearance of retinal thickening in local areas, derived from leakage of individual/clusters of microaneurysms.

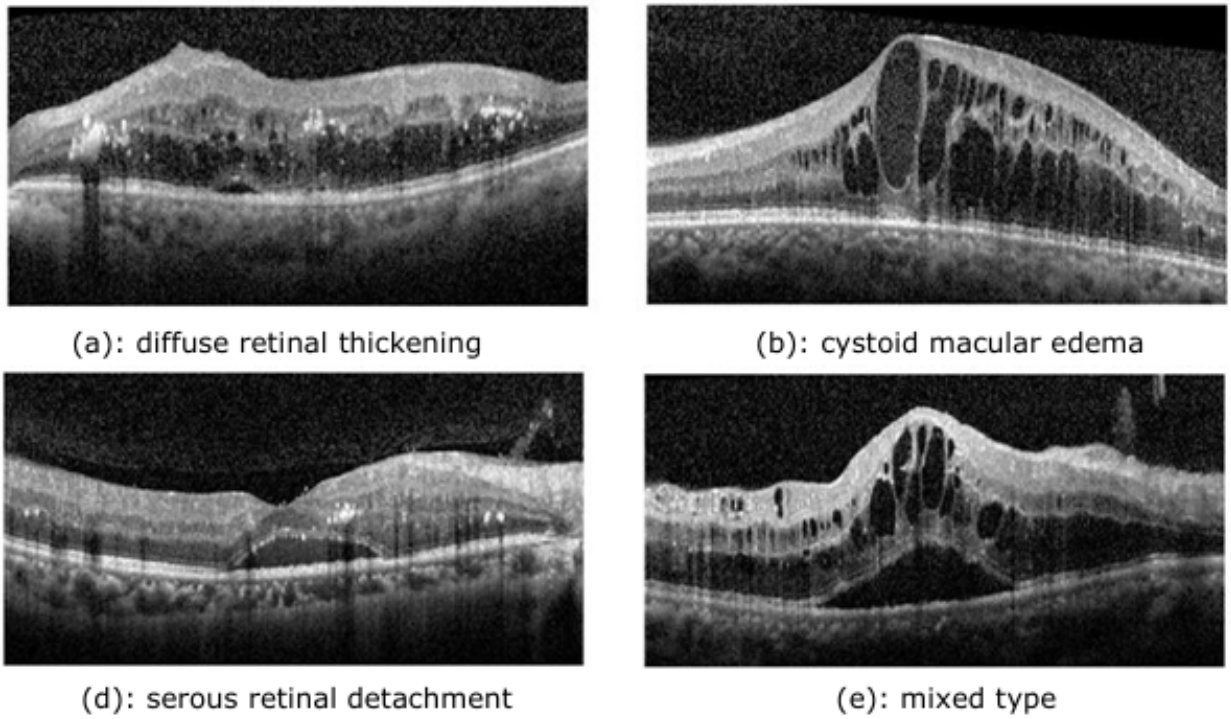


Figure 1.7: Representative OCT images of the different types of DME [20]

Diffuse macular edema is derived from damaged capillaries, microaneurysms, and arterioles. It is distinguished by a more spread thickening of the macular. Cystoid macular edema is usually linked to diffuse macular edema, it occurs from a breakdown of the retinal barrier with a swollen fluid in the outer plexiform and inner nuclear layers. The appearance or absence of cystoids does not affect the diagnosis of DME.

Figure: 1.7 presents different types of DME in Optical Coherence Tomography (OCT) scans. If the macular edema has not improved yet, the first sign of this disease will be the hazy vision in the middle of the visual scope amongst the surrounding area of one eye.

Another sign is the loss of color brightness with double vision. Unfortunately, people affected by DME will have symptoms that vary from lightly blurry vision to complete blindness eventually. A person having DR might develop into DME when there is an unnatural accumulation of fluid in the macular left untreated, caused by broken blood vessels in the retina. These blood vessels start to increase pressure in the eye and leak fluid.

Figure: 1.8 presents an eye with an abnormal vessel growth in a DR person in image (a) which is the principal cause of the swollen macular in image (b). This leakage in the

retina leads to macular edema DME [11].

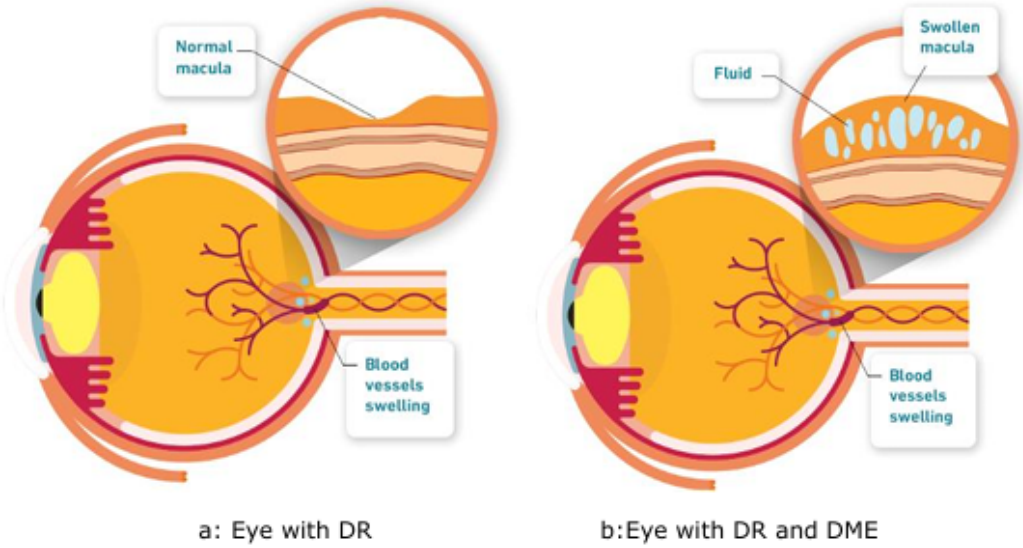


Figure 1.8: DR vs. DME eye, where swelling vessels in DR leads into leaking fluid in macular [11]

Diabetic Macular Edema(DME) can also happen after an eye's operation, which can be related to age macular deterioration, or as a result of inflammatory disorders that impact the eye [21].

Macular edema may occur within a few weeks after any eye surgery related to retinal disease. The infection probably hits the second eye, with an incidence of 50after the first eye is infected. Contrary to macular edema caused by diabetes, this inflammation is temporary and mild, which will be healed using eye-drop treatment. Diabetes as well as early unsupervised high blood glucose may lead to DR and DME.

It is necessary to control the blood sugar otherwise, problems with vision may occur.

Figure: 1.9 presents a diversity of sights, where each image indicates the progress of disease. DR in group 'A' may differ from thin to dark spots, as seen in image 'b' and 'c'. Meanwhile, group 'B' shows multiple developments of vision failure such as blurry vision like in image 'd', loss of color brightness in image 'e', or blunt one in image 'f'. Early detection of DR and DME can save lives and prevent unpleasant consequences. Thus, diabetic patients should control their blood sugar initially, then undergo a clinical examination to check the retina frequently using a powerful optical scan tool via Spectral-Domain Optical Coherence Tomography (SD-OCT) [22].

Optical Coherence Tomography (OCT) is popularly used in ophthalmology for investigating the morphology of the retina for disease discovery. This thesis applies this important optical screening system to help patients with diabetic by detecting diseases in

their eyes before any partial or total blindness may occur.

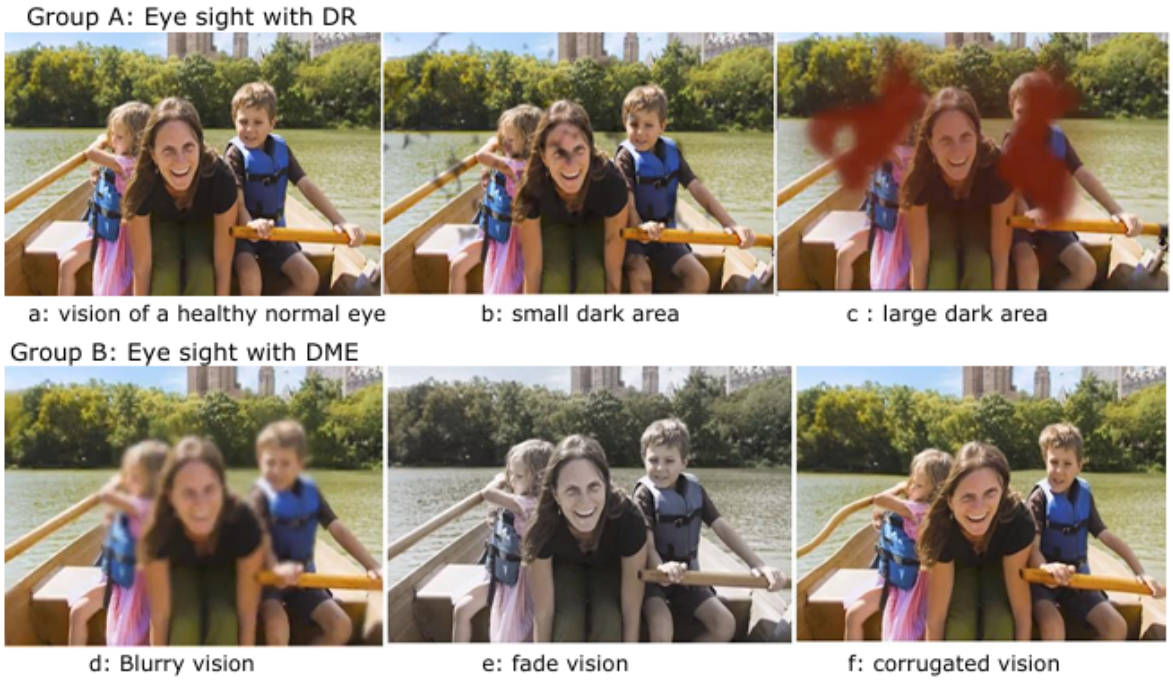


Figure 1.9: Comparison between several sights being seen by eye infected with DR (Group A) and DME (Group B) [11]

1.8 Optical Imaging Modalities

Optical imaging (especially for the retina) has experienced progress in the past century [17] to provide better knowledge of the eye in wellness and illness.

Important developments have occurred in hardware as well as in image analysis software. In this section, we will present different ocular imaging techniques that have been made to improve the visualization of ocular pathophysiology.

1.8.1 Optical Coherence Tomography (OCT)

Optical coherence tomography (OCT) is a non-invasive modality applied for cross-sectional imaging without contact. It captures the eye vessels and the subsurface structures that cannot be reached by different optical systems or operations. Using OCT imaging, ophthalmologists detect notable retinal layers that enable them to map and scale these layers that contribute to treatment [23].

It is a type of diagnostic equipment where a fast infrared laser emission is pointed at the patient's eye, with a beam reflected at a mirror for real-time visualization.

The absorbed noisy light is filtered out, leading to only consistent light being caught, producing HD images of tissue structures [24]. OCT is an important imaging modality for its fast development into clinical perspectives, given the advantages that it provides clinicians.

The light needed for OCT imaging is quite low to use in sensitive eye tissue and structures [25]. OCT is widely available in the market just after 5 years since its inception in 1991 [26]. OCT has become the recommended modality for imaging diabetic diseases like DME and DR. It can provide volumes of retinal layers that can be segmented, allowing measurement of thickness, leading to improved diagnosis, where thinning of the tissue fiber layer indicates the onset and growth of the disease. Compared to ultrasound medical mechanisms, OCT uses similar principles of waves where beams are oriented to the examined tissue.

The rapid echo waves are reflected and scrutinized using IR light delay range to unveil the depth. This delay cannot be measured immediately, therefore an interferometer is used, where a portion of the beam is pointed to the sample and another part is directed to a reference arm.

The idea of estimating low interference is the policy, where time cohesion is a property of the light source and characterizes the continuous period of waves transmitted by the source, and measured at a specific point in space. Wave chains exiting from a light source of low temporal cohesion simply maintain a stable bond phase during a very limited period of time, which corresponds to a restricted travel range, coherence length, or coherence gate.

A light source with a wide spectral frequency range consists of a set of wavelengths. The interferometer divides the emitting light from a source into two separate paths and collects the light coming from the two tracks at the output of the interferometer. The associated light strength can be measured as an electrical signal using a light detector [27]. I n the initial execution of OCT [28], the reference length was changed for several scan depths. This modification is described as time-domain OCT (TD-OCT) and the principal structure is presented in Figure: 1.10. For every sample examined, the reference wave is scanned in a depth path and the light strength is recorded on the image detector.

Thereby, an amplitude scan (A-scan) for the whole depth characterization of the reflected sample is produced. A-scan utilizes a singular one-dimensional wave to estimate the density of the aimed structure.

Diagnoses can be well performed if the aiming is accurate at the targeted formation. A-scan has the capability to identify internal swelling structure and composition. Meanwhile, to generate a brightness cross-sectional scan (B-Scan), the wave is examined alongside the sample. B-scan applies a two-dimensional array of one-dimensional A-scan waves

(produced at higher MHz frequencies) to determine the density and generate valuable images. B-scan is helpful due to its two-dimensional detailing.

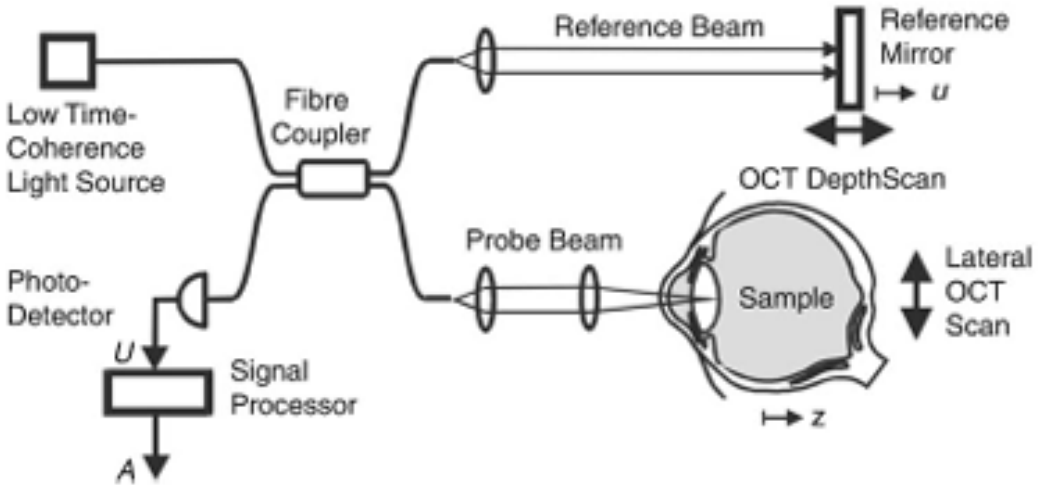


Figure 1.10: Operating method of TD-OCT: light emitting from the light source is divided into the reference wave and the middle wave. The echo light received is joined again and recorded by the detector [29]

Thus, several one-dimensional scans (A-scan) are performed at different depths to create a two-dimensional image (B-scan). Those B-scans, if obtained closely and quickly, can be translated into a volumetric image (C-scan) of a retina. There are three principal benefits of OCT over older traditional methods: non-invasive, rapid scanning, and three-dimensional figures generation. Each volumetric OCT scan can take several seconds, in comparison to about 20 minutes for techniques such as fluorescein tomography.

However, despite the speed of OCT, the patient must be held still during scanning since the body motion and eye blinking might present artifacts in measurements into imaging, which is common due to the high frequency of macular disease happening in aged patients. Typically, during OCT scanning, the blinking eye generates black lines across the image and degeneration of the signal is caused by the patient's motion [30].

Figure 1.11: Demonstration of 3D OCT-2000 SD-OCT machine and SPECTRALIS OCT device [31, 32]

Fourier domain OCT (FD-OCT or SD-OCT) is the next generation of OCT. It provides an effective implementation of the interferometer. In contradiction to TD OCT, FD-OCT uses spectral data to produce A-scans without any mechanical scanning for the optical depth length. FD-OCT was initially introduced by Fercher et al. in 1995 [33].

The principal structure is represented in Figure: 1.12. SD-OCT is related to TD-OCT, but a spectrometer changes the detector point. The spectrometer uses a diffraction component to divide the various emitting waves into a line-image captured by a fast line-scan camera.

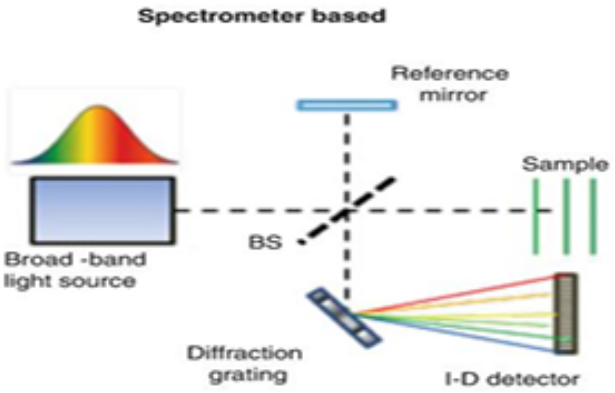


Figure 1.12: Optical implementation of spectrometer based OCT (SD-OCT) which contains a spectrometer for wave division. (Diagram is taken from Drexler et al. [34]

SD-OCT instruments operate greatly as a spectrometer performs, allowing exposed results with single exposure only. These SD-OCT systems generate image resolutions higher than 12 MP within 5 microns. Figure: 1.11 (b) illustrates an example of an SD-OCT machine named: "3D OCT-2000 Spectral Domain OCT from Topcon Medical Systems", which is a system to integrates an HD camera (12.3 MP) [31].

The appearance of SD-OCT offered the capability to surmount the restrictions of older OCT techniques. SD-OCT was able to improve image quality and capturing speed,

allowing to simultaneously imaging the entire depth information. In addition, the SPECTRALIS device has been developed by Heidelberg Engineering [32] and combines SD-OCT technology with a scanning laser fundus as seen in Figure: 1.11 (a). It was the first platform that was proposed, and it helps to find the fundus intended to be scanned with the cross-sectional OCT.

This combination allowed precise motion tracking for re-scanning at the same position. Figure: 1.13 shows a recorded sample of the eye fundus using SPECTRALIS. The left image is the optic nerve tissue. The green line indicates the selected location of the Optical Coherence Tomography OCT B-scan showed on the right image.

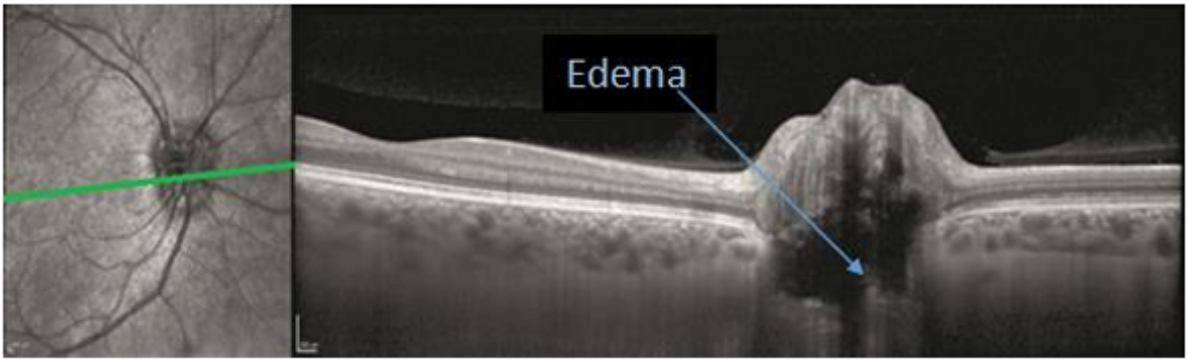


Figure 1.13: The left image is a fundus optic nerve tissue captured by SPECTRALIS machine, where the green line is represented on the right image as OCT cross-sectional B-scan. The Edema area is presented with blue arrow [35]

1.8.2 Optical coherence tomography angiography (OCTA)

Optical coherence tomography angiography (OCTA) is a non-invasive photograph modality aimed to reveal human retinal vascular networks [36].

OCTA utilizes low-coherence interferometry to cover variations in back scattered signals to distinguish blood-flow areas from static tissue areas. OCTA needs a very high density in order to obtain the required resolution of the samples to discover the thin retinal capillaries. During scanning, the variations of bulk tissue are excluded in order to control patient movement, guaranteeing that all detected variations are related to red blood flow [37]. OCTA has become widely applied clinically to detect different ophthalmological disorders, such as AMD, DR, glaucoma, and artery/vein occlusions.

Figure: 1.14 presents samples of OCTA corresponding to a young Caucasian woman's right eye. (A) Full-thickness (ILM to BM) 3 3 mm OCTA. (B) Full-thickness 6 6 mm OCTA. (C) Consequent OCT B-scan. (D) 3 3 mm retinal nerve OCTA of the retinal nerve. (E) 3 3 mm retinal GCL OCTA. (F) 3 3 mm OCTA of the internal retinal depth.

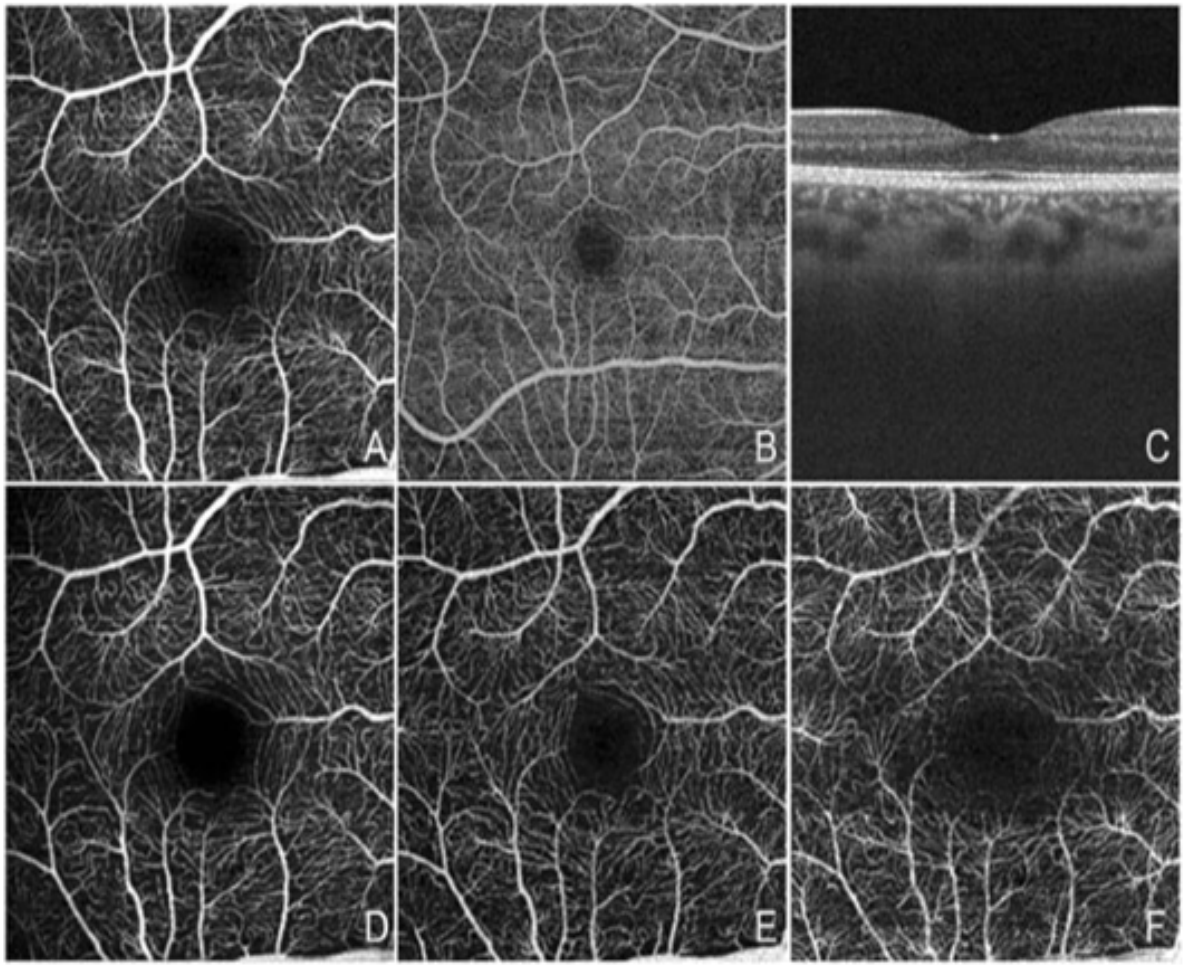


Figure 1.14: OCTA Fields of View [37]

1.8.3 Color Fundus Photography (CFP)

Compared toothed optical imaging modalities, Color Fundus Photography (CFP) which is taken by the fundus camera is cost-effective and simple [38, 39]. Despite its limitation in the field of view (FOV), CFP is frequently required for the screening purpose of eye diseases [40]. CFP is a non-invasive technique to record the fundus image so that it can be referred to in another location or time in a wide variety of ophthalmic conditions [[39].

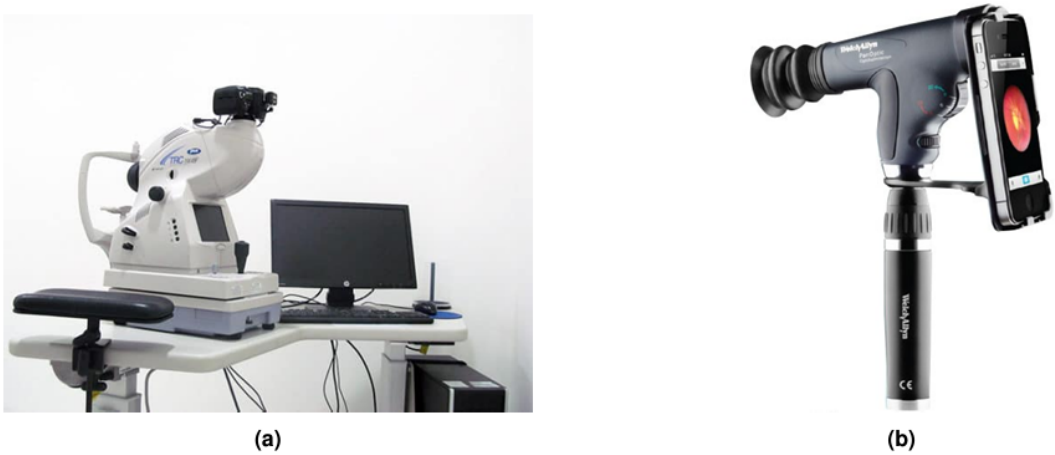


Figure 1.15: Digital fundus camera. (a) Topcon TRC-NW8Fplus [41]. (b) iExaminer [42].

The fundus camera is an indirect ophthalmoscope-based camera. It is built with an attached camera and a specialized low-power microscope. Fundus cameras are described by the field of view (FOV)- the optical angle of acceptance of the lens. The normal FOV, 30°, visualizes a retina image that is 2.5 times larger than the actual retina. Fundus camera provides capturing images between 45° and 140° [43].

The fundus image is captured when a patient sitting upright at the fundus camera while an ophthalmic photographer focuses and aligns the camera [44]. The conventional, traditional fundus cameras have a limited FOV and frequently require pupillary dilation for reliable examination of eye conditions [38]. Despite the advantage of digital retinal imaging that provides rapidly acquired, high-resolution, reproducible images [45], the fundus camera is still actively innovated to provide higher quality and lower cost [44].

Figure: 1.15 shows two different types of digital fundus cameras.

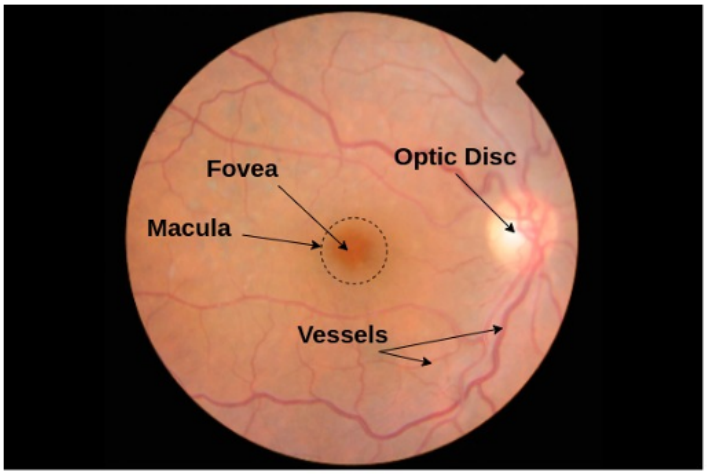


Figure 1.16: Color fundus image. [7]

The color fundus image consists of three channels (Red-Green-Blue). As seen in Figure: 1.16, the optic disc, macular, foveal, and vessels are clearly visible in a color fundus image.

The color fundus image is also helpful in interpreting the Fluorescein Angiography (FA) image since some retinal landmarks are visible in the color fundus image but not in the Fluorescein Angiography (FA) image [43]. CFP is used for clinical studies, disease documentation, telemedicine, and patient education [45].

1.9 Grading of DR

The ophthalmologists grade the retinal image into different stages of DR, depending upon the type and number of lesions on the surface of the retina. The above-mentioned proposed system would detect all abnormal regions and classify them into MAs, HM, HE, CWS, and vascular abnormalities.

Once all the lesions are detected and classified, the next step would be the grading of the patient's retina for automated screening and self-diagnostic purposes.

There are four types of grading categories that are used in the health recommendations for DR [46].

In category 1, the automated system can only separate retina into two categories, which are (i) retinal images with no or mild DR (ii) retinal images with severe levels of DR.

These categories help identification of diabetic patients with no or mild DR and those who have severe DR. The system based on category 2 can accurately identify the presence of sight-threatening DR. This category divides the patients into two groups, one having no sight-threatening DR and the other with sight-threatening DR due to the presence of DME. The second group would require immediate referral to ophthalmologists. Category 3 comprises such automated systems which can identify different levels of NPDR (mild, moderate, severe), PDR(NVE, NVD)and ME with great accuracies. The final and fourth category is actually the one that indicates the digital fundus images can be used for reliable screening and grading of DR.

This category can identify all lesions related to retina and classify retina into all stages of DR.

In proposed system, we would follow categories 3 and 4 and the grading criteria in proposed system would be like the one given in Table: 1.2.

Grade	Type	Symptom	Feature	Action
RO	No DR	None	Norma1 retina	Annual rescreen
R1	Mild non proliferative	None	Haemorrhages & Microaneurysms, Very minor IRMAs	Inform diabetes team
R2	Moderate non-proliferative	None	Extensive Microaneurysms, intra retinal haemorrhage, and hard exudates.	Refer ophthalmologist
R3	Severe non-proliferative	None	Venous abnormalities, large blot haemorrhages, cotton wool spots (small infarcts), venous beading, venous loop, venous reduplication, IRMA	Urgent refer Ophthalmologist
R4	Proli ferative retinopathy	Floaters, sudden visual loss	New vessel formation either at the disc (NVD) or elsewhere (NVE).	Urgent refer ophthalm ologist
M0			No maculopathy	Annual rescreen
M1	Diabetic Maculopathy	Blurred centra1 vision	1. exudates $<$ or $=$ 1DD of centre of fovea 2. circinate or group of exudates within macular 3. anyMicroaneurysms or haemorrhage $<$ or $=$ 1DD of center of fovea onlyis associated with a best V.A of $<$ or $=$ 6/12 retinal thickening $<$ or $=$ 1DD of center of foveal	Refer ophthalmologist

Table 1.2: Grading of DR. [7]

1.10 Conclusion

The Medical devices and Systems in the present world are in a far better position than the previous years, hence DR screening as well as the treatment is also under much better enhancement. Currently, the following solutions are available: First, diabetic patients, suspected to be Positive for DR, were observed using 45° CFP images.

Two: to provide a more accurate diagnosis and follow up with more advanced imaging that includes ultra-wide field imaging (UWF), and optical coherence tomography angiography (OCTA), in patients with retinal pathology. Group 3: The presence of DR, need for treatment, in case, by laser photocoagulation.

New developments in the diagnostic front include early detection and preliminary diagnosis, which still prove difficult in other areas such as follow-up diagnostic and accurate diagnosis. The technologies related to new medical imaging remain or are still under development of becoming popular, and the hospitals need more time and money to invest in better equipment. Furthermore, physicians require valuable time as to acquaint themselves with these enhanced imagining technologies.

Thus, the development of the potent diagnostic and follow-up system is identified as an essential need as new medical systems are being introduced and patients tend to live longer. It implies that there must be constant interaction with the patient across different specialties, constant patient education to inform and remind them of the disease, and constant screening for the disease that has very drastic effects.

By such coordinated efforts, we can thereby diminish the morbidity associated with Diabetic Retinopathy and improve the quality of needless lives of patients suffering from diabetes. Join a list of people without eye problems, and let us fight for a brighter future for our vision.

Chapter 2
Machine Learning (ML) and Deep Learning (DL)

Chapter 2

Machine Learning (ML) and Deep Learning (DL)

2.1 Introduction

Machine Learning (ML) and Deep Learning (DL) have handed us a disquisition of a whole new exploration area. As further data and better computational power become available, they've been enforced in colorful fields. The demand for Artificial Intelligence (AI) in the field of health informatics is also increasing, and we can anticipate to see the implicit benefits of Artificial Intelligence(AI) operations in healthcare. Deep Learning (DL) can help clinicians diagnose complaints, identify Diabetic Retinopathy(DR), identify medicine goods for each case, understand the relationship between genotypes and phenotypes, and prognosticate contagious complaint outbreaks with high delicacy.

In discrepancy to traditional models, its approach does not bear sphere-specific data preprocess, and it's anticipated that it'll eventually change mortal life a lot in the future. Deep Learning (DL) is popular right now because it's easy, and it works, but **what is it?** First, we must understand the base through which deep literacy surfaced

2.2 Artificial Intelligence (AI)

Artificial Intelligence (AI) is a field dedicated to creating machines that possess the ability to think and behave in a manner similar to humans.

Machine Learning (ML) and Deep Learning (DL) are both subfields of Artificial Intelligence (AI) that focus on creating algorithms and models capable of learning from data. While they share some similarities, they also have distinct differences and applications Here's a concise overview of Machine Learning (ML) and Deep Learning (DL) [47].

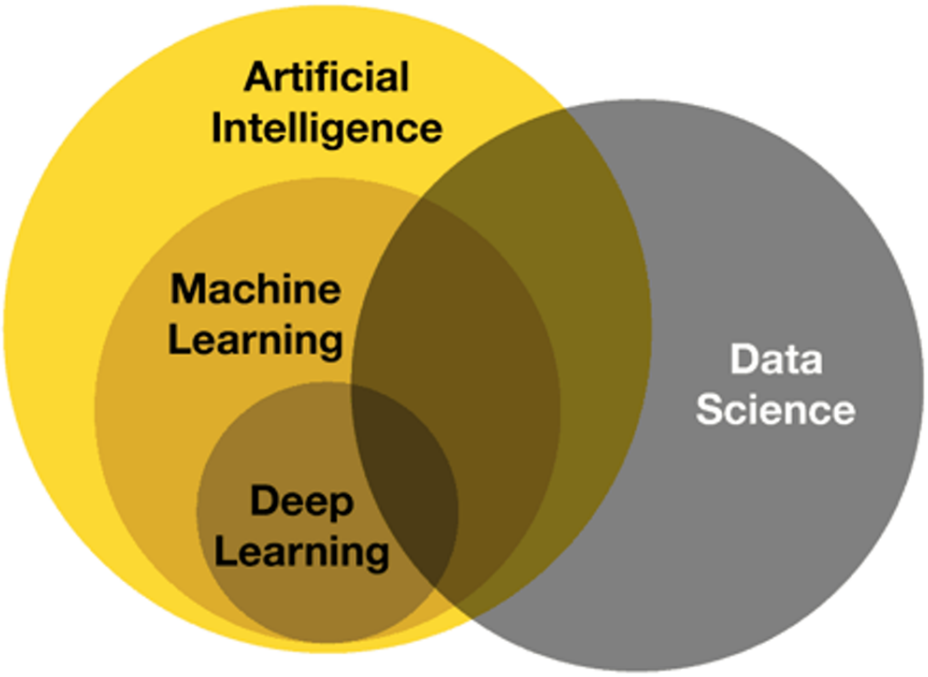


Figure 2.1: AI Vs ML Vs DS [48]

2.3 Machine Learning (ML)

Machine learning (ML) is a branch of Artificial Intelligence (AI) that systematically applies algorithms to synthesize the underlying relationships among data and information. For example, Machine learning (ML) systems can be trained on automatic speech recognition systems (such as iPhone's Siri) to convert acoustic information in a sequence of speech data into a semantic structure expressed in the form of a string of words.

Machine learning (ML) is already finding widespread uses in web search, credit scoring, stock market prediction, gene sequence analysis, behavior analysis, smart coupons, drug development, weather forecasting, big data analytics, and many more applications.

Machine learning (ML) will play a decisive role in the development of a host of user-centric innovations [49]. In 1959, Arthur Samuel described machine learning (ML) as the field of study that gives computers the ability to learn without being explicitly programmed. He concluded that programming computers to learn from experience should eventually eliminate the need for much of this detailed programming effort. According to Tom M. Mitchell's definition of Machine learning (ML): A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E [49].

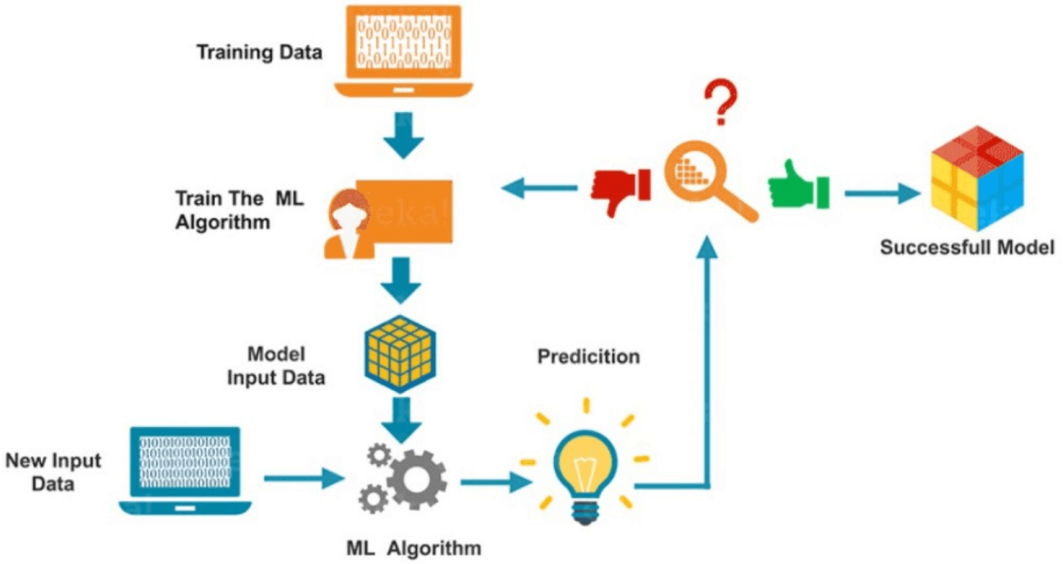


Figure 2.2: Machine learning design [50]

2.3.1 Types of Learning

Machine learning (ML) algorithms are organized into taxonomy, based on the desired outcome of the algorithm. Common algorithm types include: supervised learning, unsupervised learning, and reinforcement learning.

2.3.1.1 Supervised Learning

where the algorithm generates a function that maps inputs to desired outputs. One standard formulation of the supervised learning task is the classification problem: the learner is required to learn (to approximate the behavior of) a function that maps a vector into one of several classes by looking at several input-output examples of the function.

Supervised learning techniques: Linear Regression, Logistic Regression, CART, Naive Bayes, KNN, and Deep Learning (DL). For example, looking at the Figure 2.3, we can tell there are two classes: apples and cupcakes. And we have a bunch of training data for them. Our algorithm learns the specifications of both the classes and now can predict the class of a new coming object.

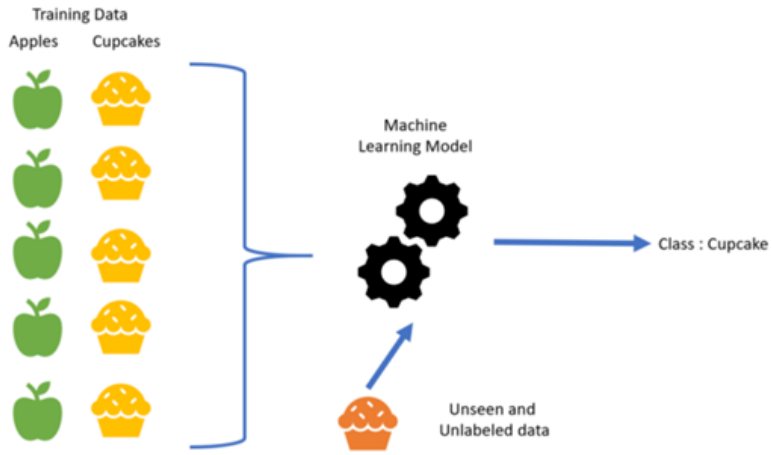


Figure 2.3: Supervised learning example [51]

2.3.1.2 Unsupervised learning

The model learns through observation and finds structures in the data. Once the model is given a dataset, it automatically finds patterns and relationships in the dataset by creating clusters in it. What it cannot do is add labels to the cluster, like it cannot say this is a group of apples or mangoes, but it will separate all the apples from mangoes [50].

Unsupervised learning techniques: Apriori, K-means, SVM and PCA. For example, looking at the Figure 2.4, we can say that if we are feeding apple, carrot and cheese as raw input data then our model will distinguish all three, but it cannot tell whether a given cluster is of apple or not as it is unlabeled, but any new data will automatically t into the clusters that are formed.

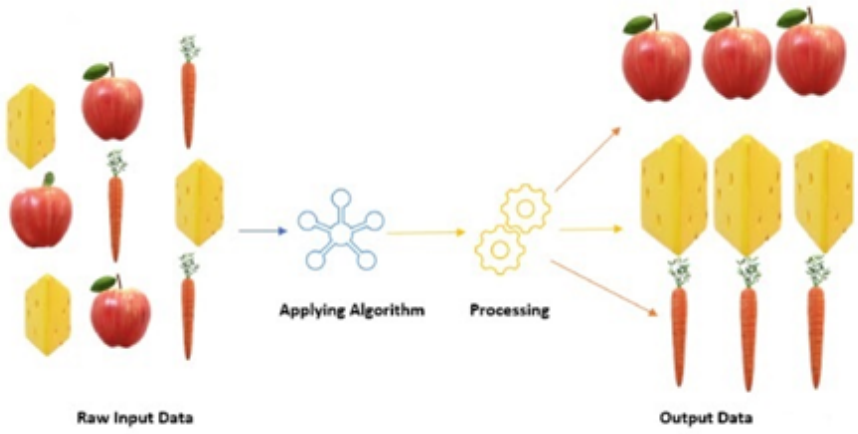


Figure 2.4: Unsupervised learning example [52]

2.3.1.3 Reinforcement learning

Reinforcement Learning is about taking suitable actions to maximize reward in a particular situation. It is employed by various software and machines to find the best possible behavior or path to take in a specific situation. Reinforcement learning differs from the supervised learning in a way that in supervised learning the training data has the answer key with it, so the model is trained with the correct answer itself whereas in reinforcement learning, there is no answer and the reinforcement agent decides what to do in order to perform the given task. In the absence of training data set, it is bound to learn from its experience [53].

The goal of reinforcement learning in this example of the Figure 2.5 is to train the dog (agent) to complete a task within an environment, which includes the surroundings of the dog as well as the trainer. First, the trainer issues a command or cue, which the dog observes (observation). The dog then responds by taking an action. If the action is close to the desired behavior, the trainer will likely provide a reward, such as a food treat or a toy; otherwise, no reward or a negative reward will be provided. At the beginning of training, the dog will likely take more random actions like rolling over when the command given is sitting, as it is trying to associate specific observations with actions and reward.

From the dogs perspective, the ideal case would be one in which it would respond correctly to every cue, so that it gets as many treats as possible. So, the whole meaning of reinforcement learning training is to tune the dogs policy so that it learns the desired behaviors that will maximize some reward. After training is complete, the dog should be able to observe the owner and take the appropriate action [54].

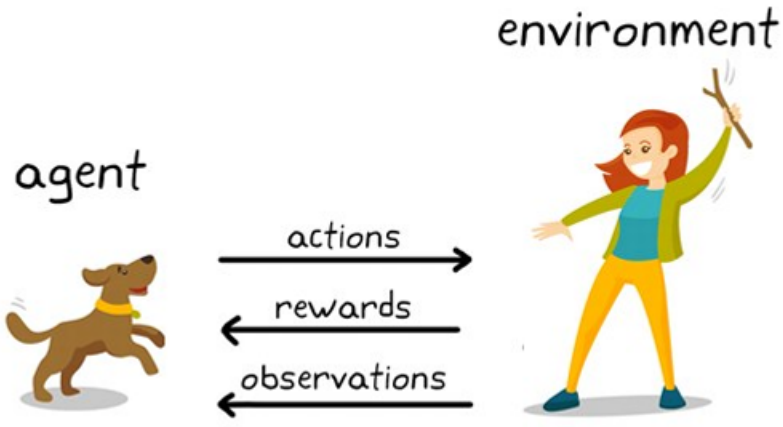


Figure 2.5: Reinforcement learning in dog training [54]

2.4 Deep Learning (DL)

Deep Learning (DL), as a new area of Machine learning (ML) research, is a process that allows the computer to learn to perform tasks which are natural for the brain like image recognition. Currently, Deep Learning (DL) methods have had a profound impact on computer vision and image analysis applications, such as image classification, segmentation, image completion and so on.

Deep learning focuses on a specific category of machine learning called Artificial Neural Networks (ANN) which is inspired by functionality of the human brain. Modern deep learning provides a very powerful framework for supervised learning.

By adding more layers and more units within a layer, a deep network can represent functions of increasing complexity. Most tasks that consist of mapping an input vector to an output vector, and that are easy for a person to do rapidly, can be accomplished via deep learning, given a sufficiently large models and sufficiently large datasets of labeled training examples [55].

2.4.1 Artificial Neural Networks (ANN)

Artificial Neural networks (ANN) have attracted considerable interest in recent years due to their ability to learn complicated maps from examples, an ability termed universal approximation. Called networks because they are typically represented by composing together many different functions called neural, because they are loosely inspired by neuroscience [55].

The learning algorithm must decide how to use those layers to produce the desired output, but the training data does not say what each individual layer should do. Although artificial neurons were inspired by the biological processes scientists were able to observe in the brain back in the 50s, but artificial and biological neurons do differ in more ways than just the materials of their containers. So, we have to talk about the basic units of a neural network.

2.4.1.1 Biological Neuron

The neuron is the basic building block of the brain and central nervous system. Neurons are specialized cells that transmit chemical and electrical signals. The brain is made up entirely of neurons and glial cells, which are non-neuronal cells that provide structure and support for the neurons. Nearly 86 billion neurons work together within the nervous system to communicate with the rest of the body. They are responsible for everything from consciousness and thought to pain and hunger [56]. The above image shows the basic

structural components of an average neuron, including the dendrite, cell body, axon, and axon terminal.

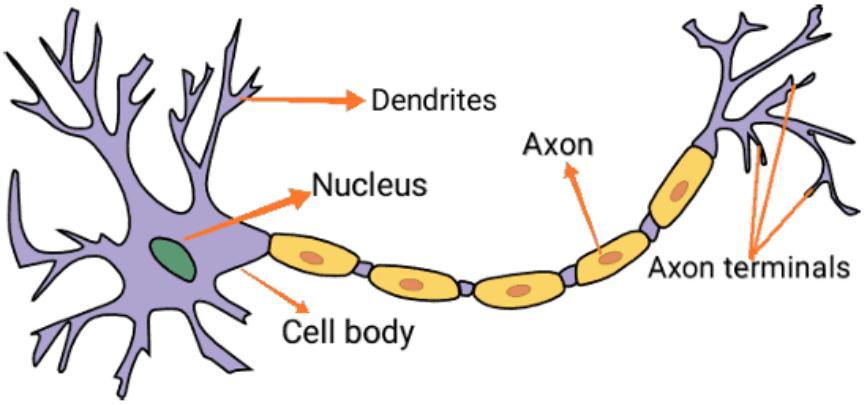


Figure 2.6: Biological neuron structure [57]

- **1. Dendrites :** are branch-like structures extending away from the cell body, and their job is to receive messages from other neurons and allow those messages to travel to the cell body.
- **2. cell body (or soma):** that contains a nucleus, smooth and rough endoplasmic reticulum, Golgi apparatus, mitochondria, and other cellular components.
- **3. An axon:** at its most basic, is a tube-like structure that carries an electrical impulse from the cell body (or from another cell's dendrites) to the structures at the opposite end of the neuron axon terminals, which can then pass the impulse to another neuron. The cell body contains a specialized structure, the axon hillock, which serves as a junction between the cell body and the axon.
- **4. The synapse:** is the chemical junction between the axon terminals of one neuron and the dendrites of the next. It is a gap where specialized chemical interactions can occur, rather than an actual structure.

Function of a Neuron: The specialized structure and organization of neurons allows them to transmit signals in the form of electric impulses from the brain to the body and back. Individually, neurons can pass a signal all the way from their own dendrites to their own axon terminals; but at a higher level neurons are organized in long chains, allowing them to pass signals very quickly from one to the other. One neuron axon will connect chemically to another neuron's dendrite at the synapse between them. Electrically charged chemicals ow from the first neuron's axon to the second neuron's dendrite, and that signal

will then ow from the second neuron's dendrite, down its axon, across a synapse, into a third neuron's dendrites, and so on [58].

2.4.1.2 Artificial Neuron

Artificial neuron is a basic building block of every artificial neural network. Its design and functionalities are derived from the observation of a biological neuron that is the basic building block of biological neural networks (systems) which includes the brain, spinal cord , and peripheral ganglia. Biological neurons and artificial neurons are Similar in design and functionalities. where the left side of Figure 2.7 represents a biological neuron with its soma, dendrites and axon and where the right side of Figure 2.6 represents an artificial neuron with its inputs, weights, transfer function, bias, and outputs [44].

In the case of a biological neuron, the information comes into the neuron via a dendrite, soma processes the information and passes it on via an axon. In the case of an artificial neuron, the information comes into the body of an artificial neuron via inputs that are weighted (each input can be individually multiplied with a weight). The body of an artificial neuron then sums the weighted inputs, bias, and processes the sum with a transfer function. At the end, an artificial neuron passes the processed information via output(s). The benefit of artificial neuron model simplicity can be seen in its mathematical description below [59]:

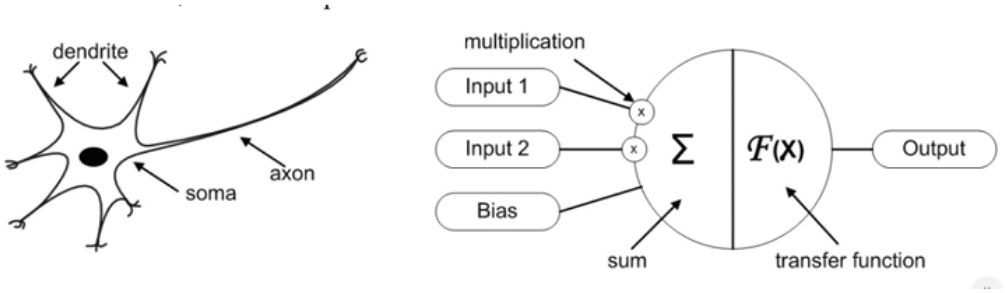


Figure 2.7: Biological and artificial neuron design [59]

$$y(k) = F(\sum_{i=0}^{m} w_i(k) \cdot x_i(k) + b) \tag{2.1}$$

Where:

-$x_i(k)$ is the input value in discrete time k where i ranges from 0 to m.

-$w_i(k)$ is the weight value in discrete time k where i ranges from 0 to m.

-b is the bias.

-f is the transfer function.

-y(k) is the output value at discrete time k.

As seen from a model of an artificial neuron and its equation 2.1 the major unknown variable of our model is its transfer function. Transfer function defines the properties of artificial neuron and can be any mathematical function. We choose it on the basis of problem that artificial neuron (artificial neural network) needs to solve and in most cases we choose it from the following set of functions: binary step function, linear function and non-linear function [59].

Binary step function:

A binary step function is a threshold-based activation function. If the input value is above or below a certain threshold, the neuron is activated and sends exactly the same signal to the next layer.

$$y = \begin{cases} 0 & \text{if } w_i x_i \geq \text{threshold} \\ 1 & \text{if } w_i x_i < \text{threshold} \end{cases} \tag{2.2}$$

The problem with a step function is that it does not allow multi-value output for example, it cannot support classifying the inputs into one of several categories [60].

Linear function:

A linear activation function takes the form: A = cx (2.3) It takes the inputs, multiplied by the weights for each neuron, and creates an output signal proportional to the input. In one sense, a linear function is better than a step function because it allows multiple outputs, not just yes and no. However, a linear activation function has two major problems:

- **Not possible to use back-propagation** (gradient descent) to train the model the derivative of the function is a constant, and has no relation to the input, X. So it's not possible to go back and understand which weights in the input neurons can provide a better prediction.

- **All layers of the neural network collapse into one** with linear activation functions, no matter how many layers in the neural network, the last layer will be a linear function of the first layer (because a linear combination of linear functions is still a linear function). So a linear activation function turns the neural network into just one layer.

A neural network with a linear activation function is simply a linear regression model. It has limited power and ability to handle complexity varying parameters of input data [60].

Non-linear function:

Modern neural network models use non-linear activation functions. They allow the model to create complex mappings between the network's inputs and outputs, which are essential

for learning and modeling complex data, such as images, video, audio, and data sets which are non-linear or have high dimensions.

Almost any process imaginable can be represented as a functional computation in a neural network provided that the activation function is non-linear.

Non-linear functions address the problems of a linear activation function:

- They allow back-propagation because they have a derivative function which is related to the inputs.
- They allow the stacking of multiple layers of neurons to create a deep neural network. Multiple hidden layers of neurons are needed to learn complex data sets with high levels of accuracy [60].

The most famous non-linear activation functions are mentioned in the table 2.1 below.

Function	Formula	Advantages	Disadvantages
Sigmoid	$S(x) = \frac{1}{1+e^{-x}}$	Smooth gradient. Output values bound. Clear predictions.	Vanishing gradient. Outputs, not zero centered. Computationally expensive.
Hyberbolic Tangent	$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}$	Zero centered.	Like the sigmoid function.
Softmax	$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j-1}^{K} e^{z_j}}$	Able to handle multiple classes Useful for output neurons.	Does not support non linearly separable data. Does not support null rejection.
Relu	y = max(0.1x, x)	Computationally efficient.	The dying Relu problem

Table 2.1: Some of non-linear activation functions

2.4.1.3 Layers

In this classic artificial neural network there are many types of layers used in the network, each type of layer is responsible for some computations.

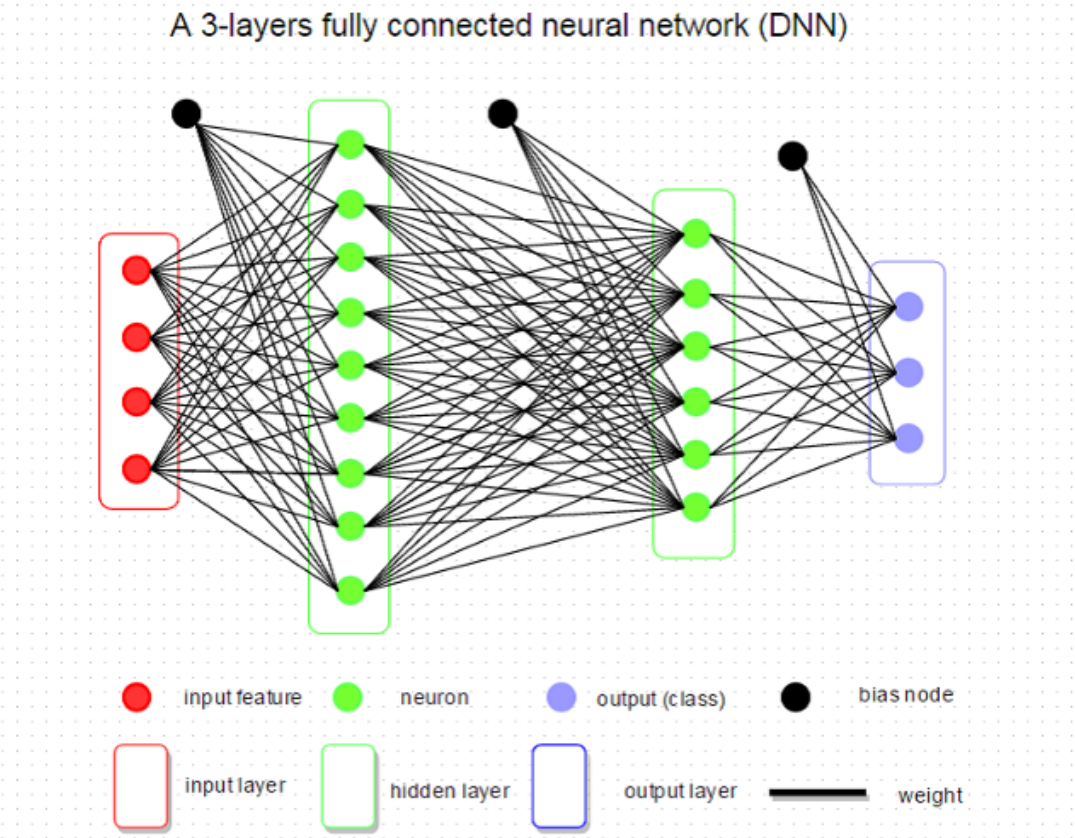


Figure 2.8: Artificial neural networks architecture [55]

- **Input layer:** Input layer is the first layer in the neural network, composed of input neurons, and brings initial data to the hidden layers for further processing.

- **Hidden layer:** The layer or group of layers between the input and output layer. Deep learning is an optimization problem looks for the optimal solution of a very complex problem, many of these computations are made in the hidden layer(s).The choice of hidden layers depends on the complexity of the data, when using a less complex data it's recommended to use few hidden layers, using many hidden layers in a simple problem can lead to overfitting, and using simple architecture in a complex the problem leads to underfitting. Each hidden layer of the network is typically vector-valued. The dimensionality of these hidden layers determine the width of the model.

- **Output layer:** The output layer is the last layer in the neural network which produces the outputs of the program, in classification tasks, the size of the output layer is equal to number of classes [55].

2.4.2 Deep Neural Network architectures

There are many different neural network architectures and with time will be more architectures, so we choose some popular and widely used neural network architectures that you should know in order to advance your knowledge about neural network architectures.

2.4.2.1 Convolutional Neural Networks (CNN)

Convolutional neural networks, or CNNs in short, are the popular choice of neural networks for different Computer Vision tasks such as image recognition. The name convolution is derived from a mathematical operation involving the convolution of different functions.
There are 4 primary layers or stages in designing a CNN:

- **Convolution:**The input signal is received at this layer.
- **Subsampling or Pooling :**Inputs received from the convolution layer are smoothened to reduce the sensitivity of the filters to noise or any other variation.
- **Activation:**This layer controls how the signal flows from one layer to the other, similar to the neurons in our brain.
- **Fully connected:**In this stage, all the layers of the network are connected with every neuron from a preceding layer to the neurons from the subsequent layer [61].

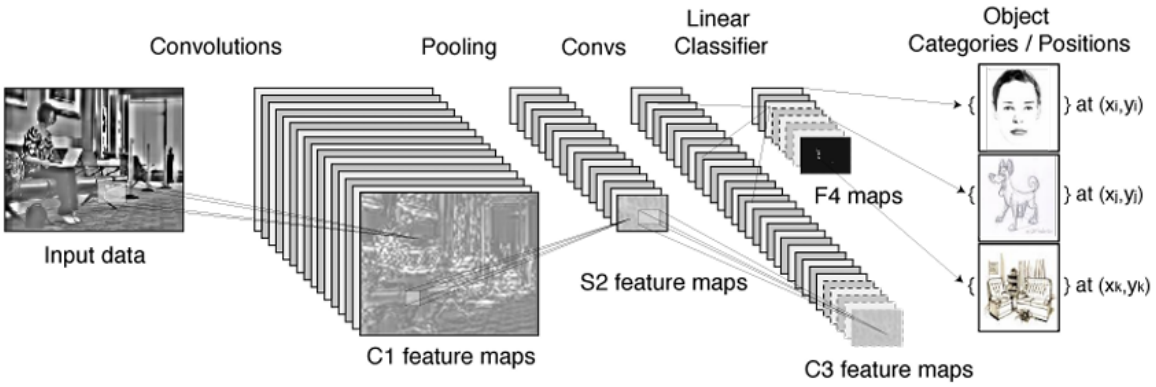


Figure 2.9: CNN architecture for visual recognition [62]

Advantages of CNN

— Very good for visual recognition.

— Once a segment within a particular sector of an image is learned, the CNN can recognize that segment present anywhere else in the image.

Disadvantages of CNN

— CNN is highly dependent on the size and quality of the training data.

— Highly susceptible to noise [63] [64].

There are various architectures of CNNs available, which have been key in building algorithms which power and shall power AI as a whole in the foreseeable future. Some of them: LeNet [65], Alex Net [66], ResNet [67], EfficientNet [68] and Inception Net [69].

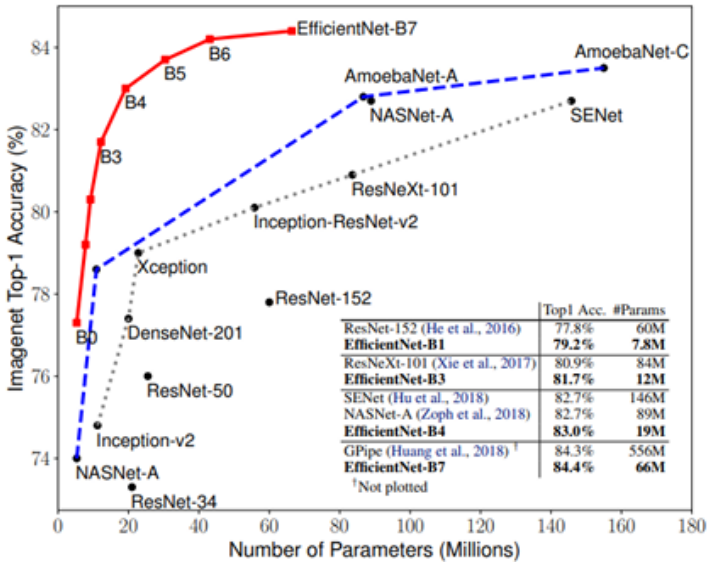


Figure 2.10: ImageNet Top-1 accuracy CNN models [68]

2.4.2.2 Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNNs) have been very popular in areas where the sequence in which the information is presented is crucial. As a result, they find a lot of applications in real-world domains such as natural language processing, speech synthesis , and machine translation. RNNs are called 'recurrent mainly because a uniform task is performed for every single element of a sequence, with the output dependent on the previous computations as well. Think of these networks as having a memory, where every calculated information is captured, store,d and utilized to calculate the final outcome [70]. Over the years, quite a few varieties of RNNs have been researched and developed:

- **Bidirectional RNN :** The output in this type of RNN depends not only on the past but also the future outcomes.
- **Deep RNN :**In this type of RNN, there are multiple layers present per step,

allowing for a greater rate of learning and more accuracy. to reduce the sensitivity of the filters to noise or any other variation.

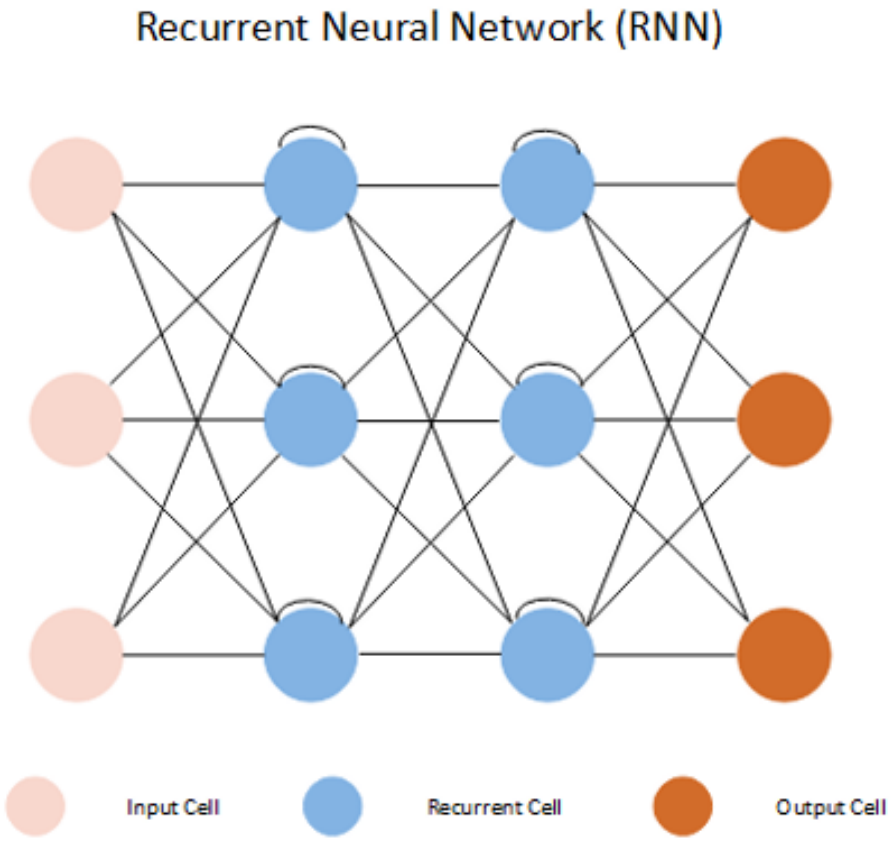


Figure 2.11: RNN architecture [71]

Advantages of RNN

— Unlike a traditional neural network, an RNN shares the same parameters across all steps. This greatly reduces the number of parameters that we need to learn.

— RNNs can be used along with CNNs to generate accurate descriptions for unlabeled images.

Disadvantages of RNN

— RNNs find it difficult to track long-term dependencies. This is especially true in case of long sentences and paragraphs having too many words in between the noun and the verb.

— RNNs cannot be stacked into very deep models. This is due to the activation function used in RNN models, making the gradient decay over multiple layers [72].

Architectures of RNNs: Identity-RNN, np-RNN, LSTM and GRU.

2.5 Conclusion

In this chapter, we introduced artificial intelligence (AI) technologies that continue to be used. These technologies are machine learning (ML) and deep learning (DL). We provided general terms and definitions and the latest knowledge of these tools for artificial intelligence (AI) in the medical field. We also cited examples of relevant work studied earlier, including Proposing K. R. Remya and his colleagues for a system based on SVM to detect and classify sugary spot edema with a resolution of more than 85%, using M. T. Esfahan and others for the ResNet34 network to classify Diabetic Retinopathy (DR) images from the Kaggle dataset with a resolution of 85%, dear development and colleagues for an automatic system to detect Diabetic Retinopathy (DR), using Cart. At 93%,

The next chapter will be devoted to our proposal.

Chapter 3
Proposal

Chapter 3

Proposal

3.1 Introduction

Now we see the problem, and we know it let's dive in this chapter and talk about how related work and how they handled this problem in last research, and we will talk about the case of our study finally we will talk about system design (datasets and CNN)

3.2 Related Works

This subsection mentions the previous works dealing with classifying retina images using machine learning techniques.

3.2.1 Computer-based detection of Diabetic Retinopathy (DR) stages using SVM

In the work by K. R. Remya, M. N. Giriprasad, and M. S. Sudhakar, a sophisticated method for the automatic detection and classification of Diabetic Macular Edema (DME) was developed using localized feature description and support vector machine (SVM) techniques. The study analyzed 400 fundus images, categorizing them into five groups: no DME, mild DME, moderate DME, severe DME, and proliferative DME. Key features such as blood vessel abnormalities, microaneurysms, hard exudates, and intranational hemorrhages were extracted from the raw images using advanced image-processing techniques, including the Time-Frequency Distribution (TFD).

The TFD was utilized to capture both temporal and frequency information from the images, enabling the extraction of more precise and localized features. These features were then fed into an SVM classifier for accurate classification. The developed system

demonstrated a sensitivity of over 85% and a specificity of 88%, showcasing its effectiveness in the early diagnosis and classification of DME [73].

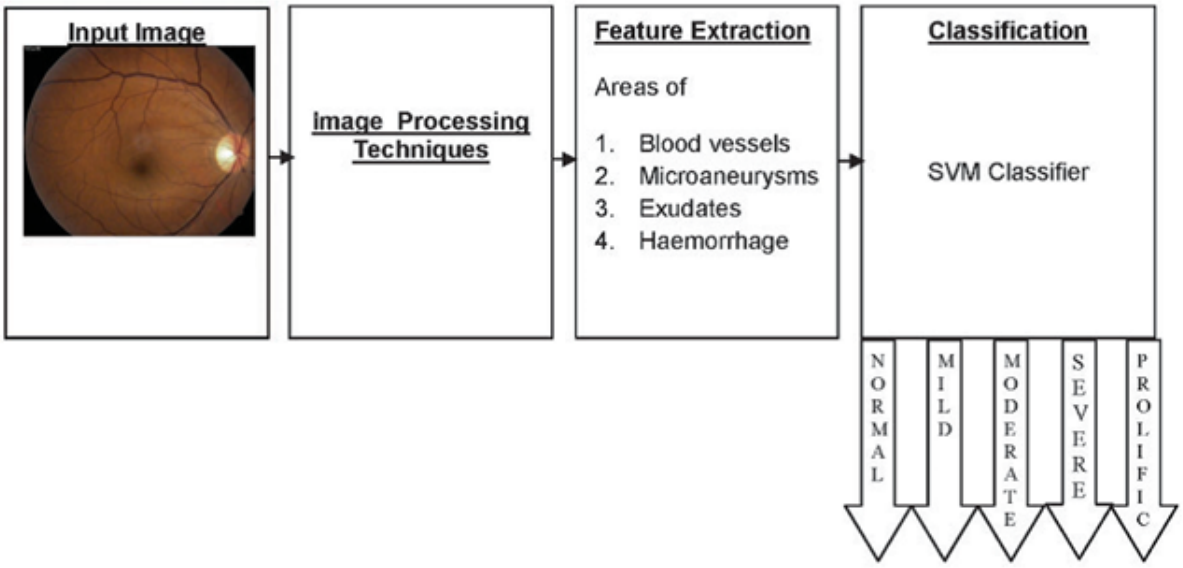


Figure 3.1: Proposed system for the detection and classification of different stages of DR [74]

3.2.2 Classification of diabetic and normal fundus images using new deep learning method

M. T. Esfahan et al. They used a known CNN, which is ResNet34 in their study to classify Diabetic Retinopathy (DR) images of the Kaggle dataset into normal or Diabetic Retinopathy (DR) image. ResNet34 is one available pretrained CNN architecture on ImageNet database. They applied a set of image preprocessing techniques to improve the quality of images. The image preprocessing included the Gaussian filter, weighted addition and image normalization. The image number was 35000 images and its size was (512,512) pixels. They reported an accuracy of 85 % and a sensitivity of 86 % [75].

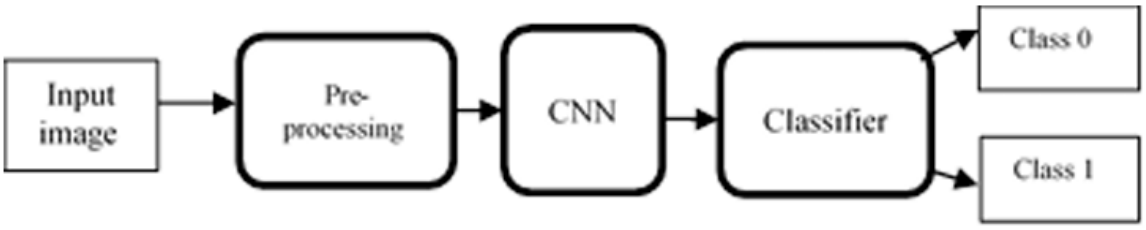


Figure 3.2: Block diagram of the proposed method (class 0= normal, class 1=DR) [75]

3.2.3 Decision tree CART algorithm for Diabetic Retinopathy (DR) classification

Aziza, Elaouaber Zineb, et al. Propose an automatic system for Diabetic Retinopathy (DR) detection from color fundus images. The proposed approach is based on the segmentation of blood vessels and extracts the geometric features, which are used in the early detection of Diabetic Retinopathy (DR). The Hessian matrix, ISODATA algorithm and active contour are used for the segmentation of the blood vessels, they have used. Finally, they have applied the decision tree CART algorithm to classify images into normal (NO-DR) or Diabetic Retinopathy (DR). The proposed system was tested on the drive and Messidor datasets and achieved an average sensitivity, specificity, and accuracy of 89%, 99% and 96%, respectively for the segmentation of retinal vessels and 91%, 100% and 93%, respectively for the classification of Diabetic Retinopathy (DR) [76].

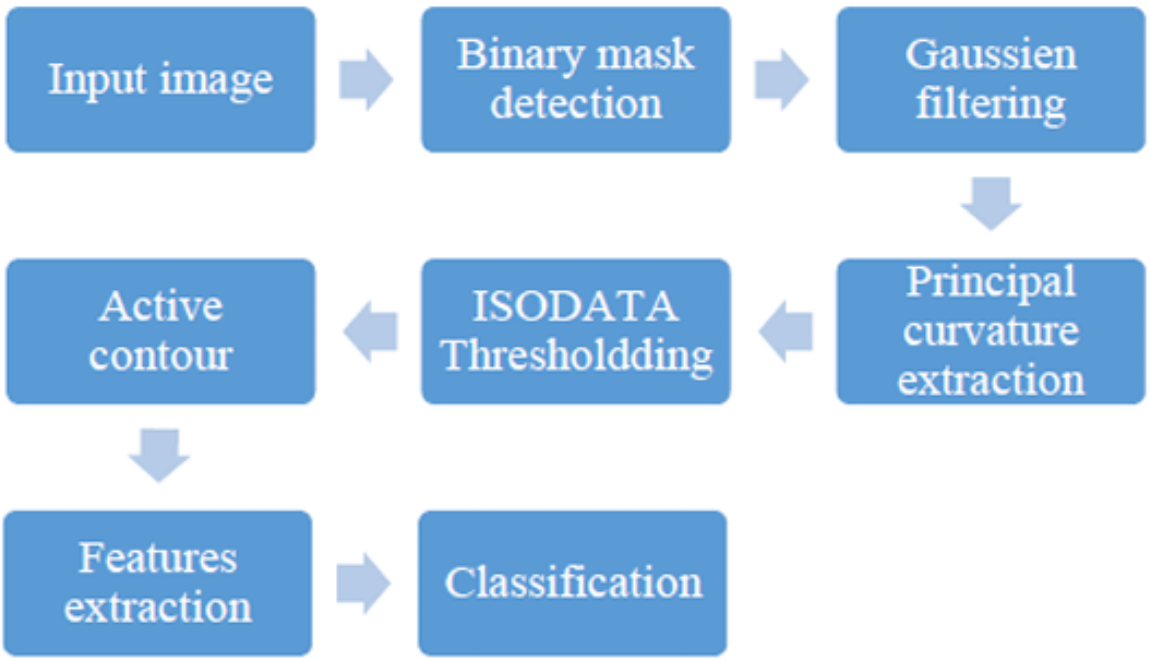


Figure 3.3: Flowchart of the proposed method for Diabetic Retinopathy (DR) detection from retinal image [76]

3.3 Case Study

The aim of this study is to create a high-resolution model using Deep Learning (DL) that can be combined with smartphone applications and to benefit from the radical improvement in the visual and sensory capabilities of phones over the past decade. Network imaging has evolved rapidly over the past 160 years and is now the mainstay of clinical care and management of patients with retinal and systemic diseases. Fundus photography is widely used to widely detect Diabetic Retinopathy (DR), age-related glaucoma and macular degeneration. With the radical improvement in the visual and sensory capabilities of smartphones over the past decade, the use of smartphones in fundus imaging has become a more powerful clinical tool, especially in low-resource areas where advanced information infrastructure and remote imaging systems are not available.

Smartphone photography and video shooting have been used at increased frequency since 2010,as first reported by Lord and colleagues. Currently, there are at least five known apparatus/software package solutions on the market: DigiSight Paxos Scope (San Francisco, USA), Peek Retina (London, UK), D-EYE (Padova, Italy), Remidio Fundus on Phone (Bangalore, India)], and Welch Allyn Panoptic with iExaminer (Skaneateles Falls, USA) [77].

(a) Handheld Remidio

(b) iExaminer

Figure 3.4: Handheld Remidio On Phone [78] and iExaminer Allyn [42]

3.4 System Design

3.4.1 DataSets used

First Dataset :

This dataset has been divided into 5 groups characterized by the severity of Diabetic Retinopathy (DR).

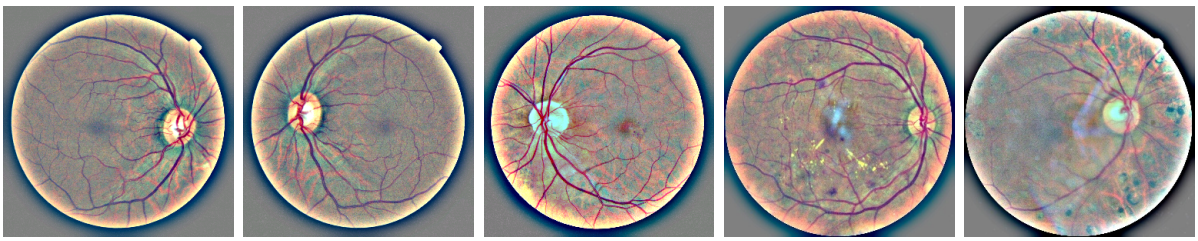
- 0: No DR
- 1: Mild Non-Proliferative DR
- 2: Moderate Non-Proliferative DR
- 3: Severe Non-Proliferative DR
- 4: Proliferative DR

The classes are balanced through under-sampling the majority class, i.e., the normal class. All classes have already been cropped, resized to (300×300) , and preprocessed by subtracting the local average color. This helps speed up the training process and avoid unnecessary computation resources for preprocessing.

Remember to use data augmentation during training because the size of the dataset is small to avoid overfitting. Try using random rotation, shifting (translating), zooming, shearing, flipping, or any other techniques [79].

Acknowledgements :

The dataset consists of multiple sources such as the eye PAC dataset and Messidor-2 dataset.

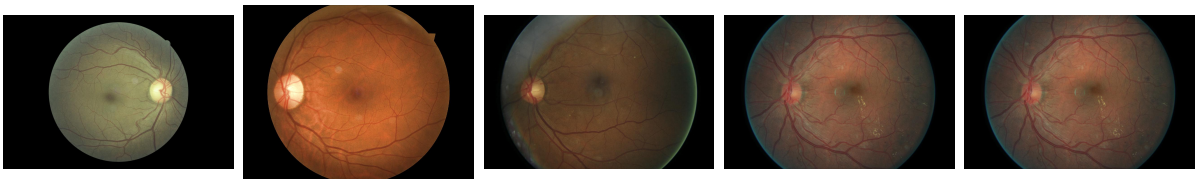


(a) NO DR **(b)** Mild **(c)** Modert **(d)** Sever **(e)** Prolifirative

Figure 3.5: images of Diabetic Retinopathy (DR) [79].

Second Dataset :

This Data contain the same classes but it's larger then the previous one and it's imbalanced



(a) NO DR **(b)** Mild **(c)** Modert **(d)** Sever **(e)** Prolifirative

Figure 3.6: images of Diabetic Retinopathy (DR) [80].

Data Preprocessing :

The preprocessing phase is a crucial phase in image classification tasks. The image must be considered.

Shape, size, noise and pixels are preliminary things, for which we use Gaussian filtering to do all of this. Our image size is (244, 244, 3). The values (224×224) refer to the width and height of the image, the third parameter in our case (3) refers to the image channels, which in this case means that the image is RGB format (Red, Green, Blue). As we found out, our second data are inbalanced and not same size, so we need to resize all images to (224×224) and we need to balance data (really hard to do) but we will try another approach that help (SMOTE) we will talk about it later.

Data Augmentation :

Data augmentation is a technique used on data that create new images by applying transforms 3 must used once (Image flips, Image Rotation and Image brightness)

3.4.2 CNN Architecture

Simple CNN Architecture

The model contains 3 blocks CONV and 2 Fully Connected layers :

first block we apply 16 filters of (3,3) kernel size, activation function Relu and Max Pooling of (2,2) that reduce size of the image

output size of first block (16,3,3) 16 :channels, 3 second bloc we use 32 filters, same activation function and pooling, last block we use 64 filters, then fully connected we applied 3 Max Pooling of (2,2) the size of image (224,224) will be (28,28) and we have 64 filters so $(64 \times 28 \times 28) = 50,176$

512 it's just hyper-parameter that we chose finally last fully connected layer (512, 5) where the 5 is a number of classes

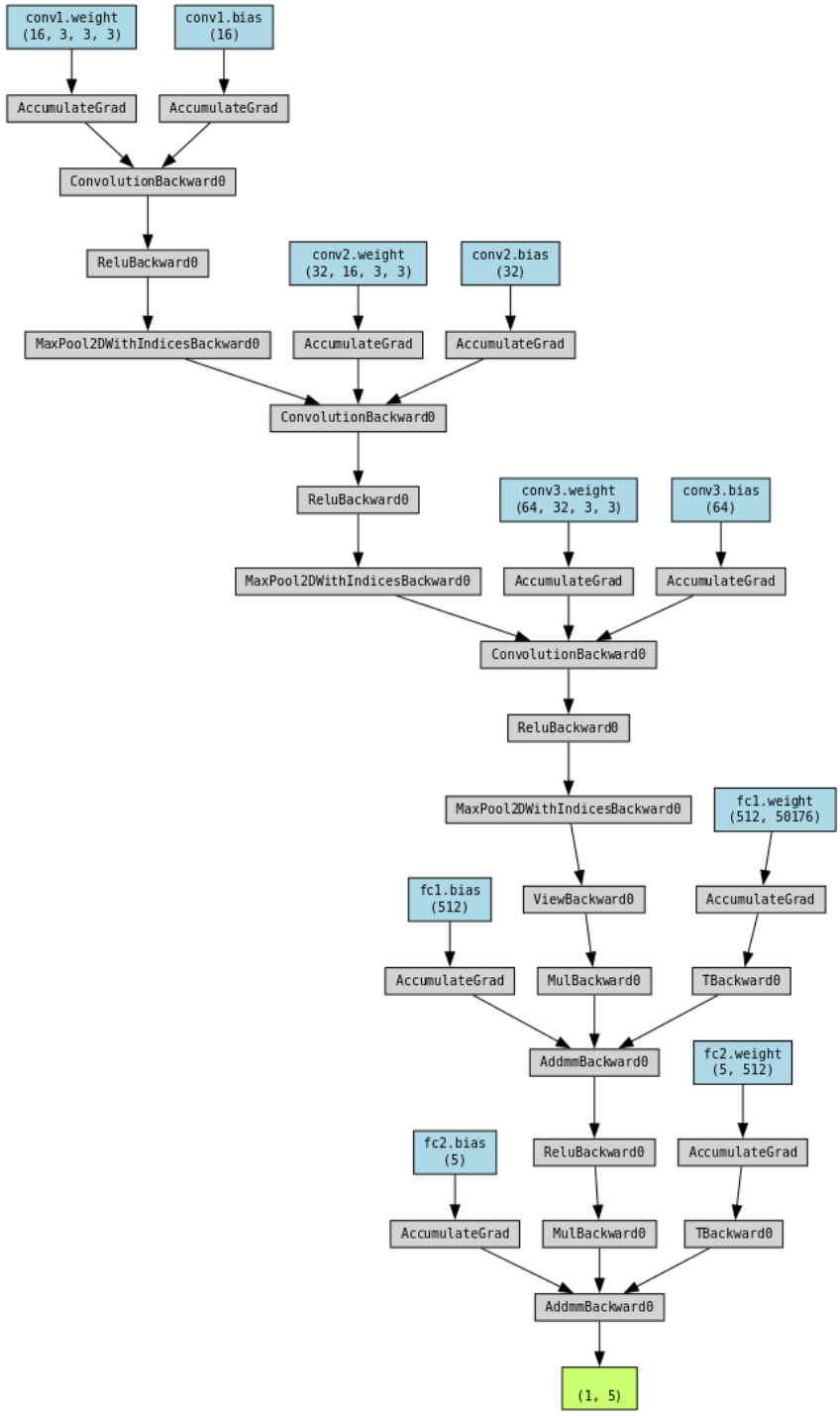


Figure 3.7: CNN Architecture

3.4.3 Transfer learning

Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem [81]. transfer learning is the use of pretrained models in image classification tasks, which has increased lately, due to the speed while training and the efficiency these models provide, this is due to the pretrained weights. in this work, we will try some pretrained models to classify our images and see if these models will give better results.

Pretrained Model Used : (3.1)

ResNet 50 :

The ResNet architecture is considered to be among the most popular Convolutional Neural Network architectures around. Introduced by Microsoft Research in 2015, Residual Networks (ResNet in short) broke several records when it was first introduced in this paper [82].

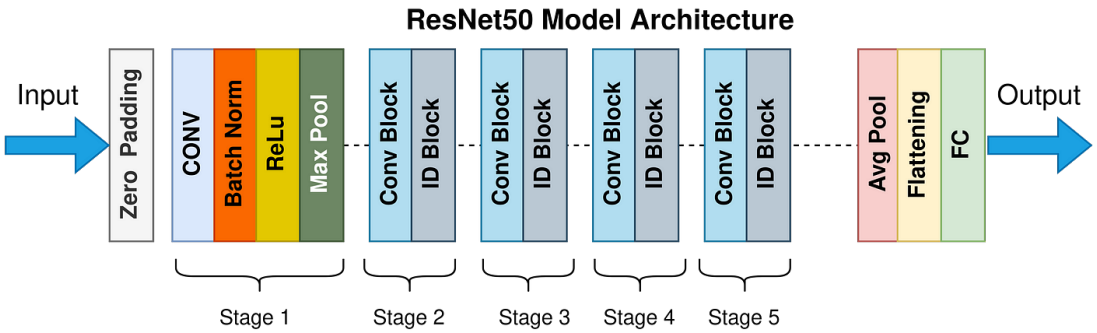


Figure 3.8: ResNet 50 [83].

Dense-Net :

Dense-Net is a powerful deep learning architecture that uses dense connectivity between layers to boost the performance of convolutional neural networks. This advanced architecture has demonstrated significant efficiency across a wide range of computer vision tasks, such as image classification, object detection, and segmentation. Its use-cases span multiple image-related applications, including face recognition, animal type identification, object detection, cancerous cell identification, among others. [84].

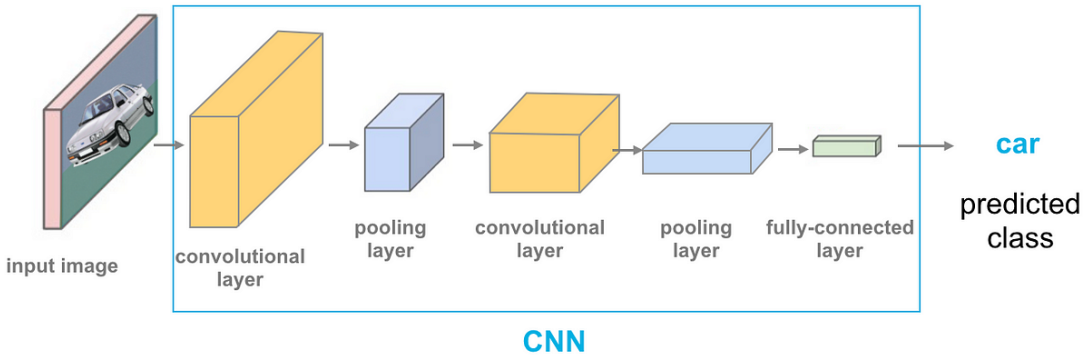


Figure 3.9: DenseNet [85].

Efficient-Net :

Efficient-Net is a convolutional neural network built upon a concept called "compound scaling." This concept addresses the longstanding trade-off between model size, accuracy, and computational efficiency. The idea behind compound scaling is to scale three essential dimensions of a neural network: width, depth, and resolution [86].

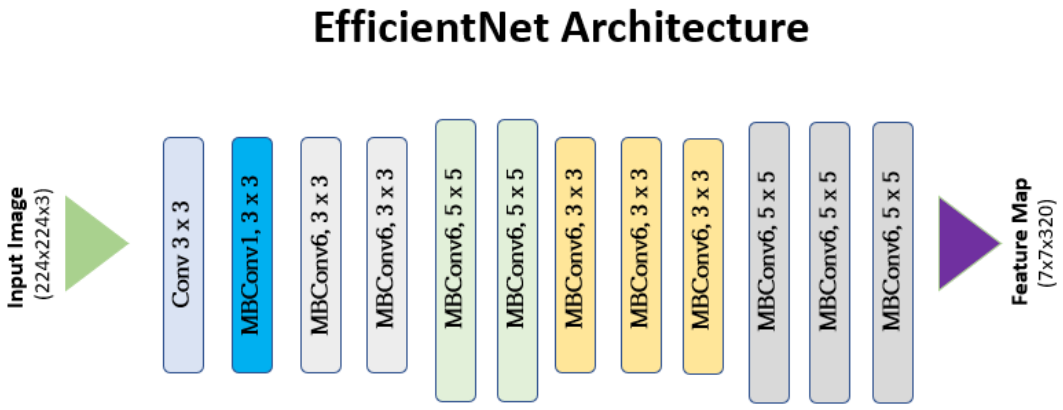


Figure 3.10: Efficient-Net [87].

3.5 Conclution

In this chapter we talked about related work and our design and the two ways to solve problem like (related work) using (ML) SVM and our system where we chose (DL) CNN and Transfer Learning

Now that we understood all this, let's dive into the implementation where we see more challenging things

Chapter 4
Implementation and Results

Chapter 4

Implementation and Results

4.1 Introduction

After detailing in the previous chapter on advanced work. In this chapter, we will address the implementation of the proposed steps to achieve the designed system. We start in the first part by selecting the dataset, showing some images of it, and then introducing the tools used in the development. After that, we will explain all the experiences we applied and the results obtained, and finally view some parts of the code.

4.2 Implementation frameworks and tools

Deep learning frameworks play a crucial role in driving the artificial intelligence revolution by providing powerful tools and pre-implemented templates for tasks such as preprocessing, classification, and evaluation. These frameworks streamline the development process, enabling scientists to focus more on the core aspects of their research rather than reinventing the wheel with each project.

This subsection examines various development frameworks and tools utilized in this thesis.

4.2.1 Python

The most used and famous programming language in data science, is a high-level programming language, and its core design philosophy is all about code readability and syntax which allows programmers to express concepts in a few lines of code. Python is developed under an OSI-approved open-source license, making it freely usable and distributable, even for commercial use. It is used successfully in thousands of real-world

business applications around the world, including many large and mission-critical systems. The python version used in this work is 3.10.12 [88].

```
[1]  import sys
     print("python version"+sys.version)

     python version3.10.12 (main, Nov 20 2023, 15:14:05) [GCC 11.4.0]
```

Figure 4.1: python version used in colab.

4.2.2 Matplotlib

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+. Matplotlib was originally written by John D. Hunter. Everything in Matplotlib is organized in a hierarchy. At the top of the hierarchy is the Matplotlib state-machine environment, which is provided by the Matplotlib.pyplot module. At this level, simple functions are used to add plot elements (lines, images, text, etc.) to the axes in the Figure. Pyplot's state-machine environment behaves similarly to MATLAB and should be most familiar to users with MATLAB experience [89].

Figure 4.2: Examples of Matplotlib works. [89]

4.2.3 Tensorflow

An open-source machine learning library for research and production developed by Google brain team, with Python and C++ programming languages as the backend, and known as the most used frameworks for deep learning tasks and with the highest activity in GitHub. At a high level, TensorFlow is a Python library that allows users to express arbitrary computation as a graph of data flows [90].

4.2.4 PyTorch

PyTorch is a software-based open-source deep learning framework used to build neural networks, combining the machine learning (ML) library of Torch with a Python-based high-level API. Its flexibility and ease of use, among other benefits, have made it the leading ML framework for academic and research communities [79].

4.2.5 Keras

Keras is an API designed for human beings, not machines. Keras follows best practices for reducing cognitive load: it offers consistent simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear and actionable feedback upon user error. Keras is also a favorite among deep learning researchers, coming in 2nd in terms of mentions in scientific papers after tensor ow [91]. Keras uses as Backend tensor ow and Theano, we have used Keras with tensor ow back end in the implementation part of this work.

4.2.6 Kaggle

Kaggle is an online community of data scientists and machine learners, owned by Google LLC. Kaggle got its start by offering machine learning competitions and now also offers a public data platform, a cloud-based workbench for data science, and short form AI education. Kaggle Kernels is a cloud computational environment that enables reproducible and collaborative analysis. Kernels supports scripts in R and Python, Jupyter Notebooks, and Markdown reports. Kaggle Kernels provide a free Cloud GPU NVIDIA K80 GPUs and 16 GB of Free RAM, the session is limited to 6 hours [92].

Our current work was inspired from a Kaggle challenge announced lately in 2019 called APTOS 2019 blindness detection [93].

4.2.7 Google Colab

Google Colab is a free cloud service with free GPU and TPU 2 with a limit of 12 hours per session and 12 GB of RAM limits. The execution of deep models with complicated architecture and huge data (as in our case) can lead to the use of extreme amounts of hardware materials, Google Colabs provides a free GPU, currently they provide GPU with specs Tesla K80, having 2496 CUDA cores, compute 3.7, 12 GB (11.439 GB Usable) GDDR5 VRAM. and a CPU of CPU: 1xsingle core hyper threaded i.e(1 core, 2 threads) Xeon Processors @2.3Ghz (No Turbo Boost), 45 MB Cache [94].

Why do we need GPU in deep learning ?

Graphics processing unit, is a processor that is good at handling specialized computations. This is in contrast to a central processing unit (CPU), which is a processor that is good at handling general computations. CPUs are the processors that power most of the typical computations on our electronic devices. A GPU can be much faster at computing than a CPU. However, this is not always the case.

The speed of a GPU relative to a CPU depends on the type of computation being performed. The type of computation most suitable for a GPU is a computation that can be done in parallel. Parallel computing is a type of computation where a particular computation is broken into independent smaller computations that can be carried out simultaneously.

The resulting computations are then recombined, or synchronized, to form the result of the original larger computation. The number of tasks that a larger task can be broken into depends on the number of cores contained in a particular piece of hardware. Cores are the units that actually do the computation within a given processor, and CPUs typically have four, eight, or sixteen cores while GPUs have potentially thousands [95].

4.2.8 Cross Validation

Cross validation is a technique used in machine learning to evaluate the performance of a model on unseen data. It involves dividing the available data into multiple folds or subsets, using one of these folds as a validation set, and training the model on the remaining folds. This process is repeated multiple times, each time using a different fold as the validation set. Finally, the results from each validation step are averaged to produce a more robust estimate of the model's performance. Cross validation is an important step in the machine learning process and helps to ensure that the model selected for deployment is robust and generalizes well to new data [96].

4.2.9 Model evaluation metrics

This subsection discuses how we can evaluate our model results, what makes a model better than another and how we can predict that our model will give good results in the testing .
Many evaluation metrics to evaluate classification output quality.

4.2.9.1 Accuracy

Accuracy is the number of correctly predicted data points out of all the data points. More formally, it is defined as the number of true positives and true negatives divided by the number of true positives, true negatives, false positives and false negatives, as it summarized in the equation: 4.1.

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \tag{4.1}$$

A true positive or true negative is a data point that the algorithm correctly classified as true or false. A false positive or false negative, on the other hand, is a data point that the algorithm incorrectly classified.

4.2.9.2 Recall

Named also Sensitivity, Recall is the ratio of correctly predicted positive observations to the all observations in actual class, in other way, recall measures the proportion of actual positives that are correctly identified.

the formula:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.2}$$

4.2.9.3 Precision

Out of all the classes, how much we predicted correctly. It should be high as possible.

the formula:

$$\text{Recall} = \frac{TP}{TP + FP} \tag{4.3}$$

4.2.9.4 Implementation

As we have mentioned before, this data set is preprocessed so we don't have to do the preprocessing step

Any model trained should pass through these three steps:

1: Data loading and Preprocessing

2: Define Model Architecture

3: Start training and validation

1: Data Loading and Preprocessing :

```
]:
class DiabetMessidorEyeDataSet(Dataset):
    def __init__(self, data_dir, transform=None):
        self.data = ImageFolder(data_dir, transform=transform)

    def __len__(self):
        return len(self.data)

    def __getitem__(self, idx):
        return self.data[idx]

    @property
    def classes(self):
        return self.data.classes
```

Figure 4.3: Data Loader class

2: Define Model Architecture (code)

```
]:
class DiabetModel(nn.Module):
    def __init__(self, num_classes=5):
        super(DiabetModel, self).__init__()
        # Define the layers of the model
        self.conv1 = nn.Conv2d(3, 16, kernel_size=3, padding=1)
        self.conv2 = nn.Conv2d(16, 32, kernel_size=3,padding=1)
        self.conv3 = nn.Conv2d(32, 64, kernel_size=3,padding=1)
        self.pool = nn.MaxPool2d(2, 2)
        self.fc1 = nn.Linear(64 * 28 * 28, 512)
        self.fc2 = nn.Linear(512, num_classes)
        self.dropout = nn.Dropout(0.5)  # Increased dropout

    def forward(self, x):
        # Apply convolutions, followed by batch normalization, activation, and pooling
        x = self.pool(F.relu(self.conv1(x)))
        x = self.pool(F.relu(self.conv2(x)))
        x = self.pool(F.relu(self.conv3(x)))
        # Flatten the tensor for the fully connected layer
        x = x.view(-1, 64 * 28 * 28)
        x = self.dropout(x)
        x = F.relu(self.fc1(x))
        x = self.dropout(x)
        x = self.fc2(x)
        return x
```

Figure 4.4: Simple CNN Architecture

3 : Training and Validation

```
model1.to(device)
num_epochs =3
scaler = GradScaler()
train_losses , val_losses = [] , []
for epoch in range(num_epochs):
    model1.train()
    running_loss = 0.0
    correct_train = 0
    total_train = 0
    for images, labels in tqdm(train_loader, "Training Loop"):
        images, labels = images.to(device), labels.to(device)
        optimizer1.zero_grad()
        with autocast():
            output = model1(images)
            loss = criterion1(output, labels)
            _, predicted = torch.max(output, 1)
            correct_train += (predicted == labels).sum().item()
            total_train += labels.size(0)
        scaler.scale(loss).backward()
        scaler.step(optimizer1)
        scaler.update()
        running_loss += loss.item() * labels.size(0)  # Use loss.item() to get the scalar value
    train_loss = running_loss / len(train_loader.dataset)
    train_accuracy = correct_train / total_train
    train_losses.append(train_loss)
```

Figure 4.5: Training Loop

```
  model1.eval()
  running_loss = 0.0
  correct_val = 0
  total_val = 0
  true_labels = []
  predicted_labels = []
  with torch.no_grad():
      for images, labels in tqdm(val_loader, 'Validation Loop'):
          images, labels = images.to(device), labels.to(device)
          with autocast():
              output = model1(images)
              loss = criterion1(output, labels)
              _, predicted = torch.max(output, 1)
              correct_val += (predicted == labels).sum().item()
              total_val += labels.size(0)
              true_labels.extend(labels.cpu().numpy())
              predicted_labels.extend(predicted.cpu().numpy())
          running_loss += loss.item() * labels.size(0)  # Use loss.item() to get the scalar value
  val_loss = running_loss / len(val_loader.dataset)
  val_accuracy = correct_val / total_val
  val_losses.append(val_loss)
  print(f"""
  Epoch {epoch+1}/{num_epochs} :
  Train loss: {train_loss}
  Train accuracy: {train_accuracy}
  Validation loss: {val_loss}
  Validation accuracy: {val_accuracy}""")
```

Figure 4.6: Validation Loop

4 : Result Of Simple CNN

```
    Epoch 1/3 :
    Train loss: 1.4416285196940104
    Train accuracy: 0.3472222222222222
    Validation loss: 1.1210801866319444
    Validation accuracy: 0.45555555555555555
Loading widget...
Loading widget...

    Epoch 2/3 :
    Train loss: 1.150192419687907
    Train accuracy: 0.43194444444444446
    Validation loss: 1.1229390462239583
    Validation accuracy: 0.48333333333333334
Loading widget...
Loading widget...

    Epoch 3/3 :
    Train loss: 1.1215344746907552
    Train accuracy: 0.4361111111111111
    Validation loss: 1.0646538628472222
    Validation accuracy: 0.4638888888888889
```

Figure 4.7: Result

As we see no validation loss improvement, so more epochs it's just an over fitting to the data.

What about transfer learning??

5: Transfer Learning Approach

```
class DiabetModel(nn.Module):
    def __init__(self , num_classes = 5):
        super(DiabetModel,self).__init__()
        self.base_model = timm.create_model('efficientnet_b4', pretrained=True)
        self.features = nn.Sequential(*list(self.base_model.children())[:-1])

        out_size = 1792
        self.classifier = nn.Sequential(
            nn.Flatten(),nn.Dropout(0.3),nn.Linear(out_size , num_classes)
        )


    def forward(self , x) :
        x= self.features(x)
        output = self.classifier(x)
        return output
```

Figure 4.8: Model Architecture

6 : Result of Transfer Learning

```
Epoch 1/3 :
Train loss: 0.5870658801661597
Train accuracy: 0.7673611111111112
Validation loss: 1.0642919699350992
Validation accuracy: 0.5722222222222222
ading widget...
ading widget...

Epoch 2/3 :
Train loss: 0.2623870308200518
Train accuracy: 0.9111111111111111
Validation loss: 1.0390099419487848
Validation accuracy: 0.6944444444444444
ading widget...
ading widget...

Epoch 3/3 :
Train loss: 0.16450613116224608
Train accuracy: 0.9423611111111111
Validation loss: 1.4902902709113226
Validation accuracy: 0.65
```

Figure 4.9: Transfer Learning Approach

Ok it's a little bit of batter but not what we expect, let's try cross-validation

7 : Cross Validation

```
Epoch 5: val_loss did not improve from 0.25131
21/21 ━━━━━━━━━━━━━━━━━━━━ 6s 286ms/step - accuracy: 0.7805 - loss: 0.6461 - val_accuracy: 0.930
6 - val_loss: 0.2543 - learning_rate: 1.0000e-08
Epoch 5: early stopping
Restoring model weights from the end of the best epoch: 1.
Average training loss:  0.7301204562187195
Average training accuracy:  0.7387607395648956
Average validation loss:  0.7935424953699112
Average validation accuracy:  0.6694444477558136
Maximum training accuracy:  0.939393937587738
Maximum validation accuracy:  0.9583333134651184
```

Figure 4.10: Transfer Learning Approach

The result is amazing 95% validation accuracy

sorry to tell you it's an overfitting, as we know cross-validation changes samples, so the model memorize all samples of our dataset when we tested it was worse than the previous model

So what is the problem and how we can solve it ???

the problem is our dataset it's too small we can't find more data just in Google because it's a medical problem so we searched in Kaggle for more data, and finally, we found another one.

4.2.10 Second dataset

4.2.10.1 About Data

This is augmented data that contains the same classes with different images; it's 10 times bigger than the previous one found in Kaggle.

Link to Data : [97]

4.2.10.2 Implementation

Now that we know how it's work, let's dive in directly

1 : Data Preprocessing

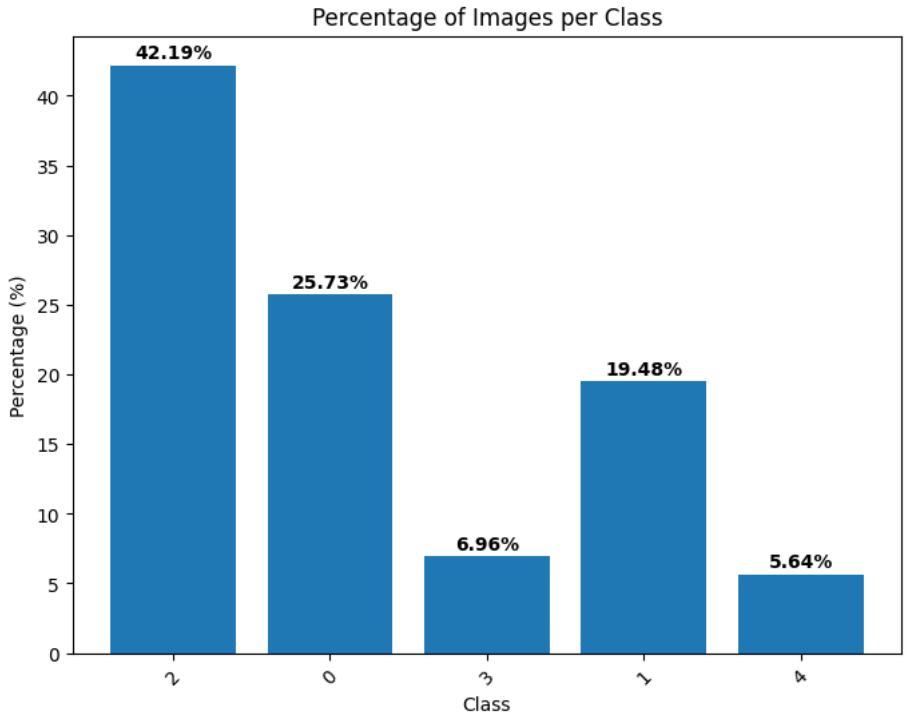


Figure 4.11: Data Visualisation

The data are imbalanced. This is a big problem for our model. Let's imagine it like this:

suppose that you have a child and you try to teach him some math. If you give him 100 multiplication equations and only 10 divisions, he will be good at multiplication and bad at division. The same is true when we talk about models and imbalanced datasets. So what can we do ????

2 : Solution Proposed :

1 : SMOTE SMOTE stands for Synthetic Minority Over-sampling Technique. It is a pre-processing technique used to handle class imbalance. It balances the data by generating synthetic samples for the minority class [98].

We tried to apply this solution, but it consumes a lot of memory space, and we were unable to run it in our notebook it's stop

2 : Change classes weight in CrossEntropyLoss :

This will simply pay attention to minority classes more than others, we can modify it as we like

```
classes_weights = torch.tensor([25.441205631568508,
                                19.551854055125915,
                                42.692841562561966,
                                6.702359706523894,
                                5.611739044219711])

criterion = torch.nn.CrossEntropyLoss(weight=classes_weights).to(device)
```

Figure 4.12: Changing Weights

3 : Result of Simple CNN (Used In Previous Data) :

Figure 4.13: Result Of Training with Simple CNN

4 : Result of Transfer Learning :

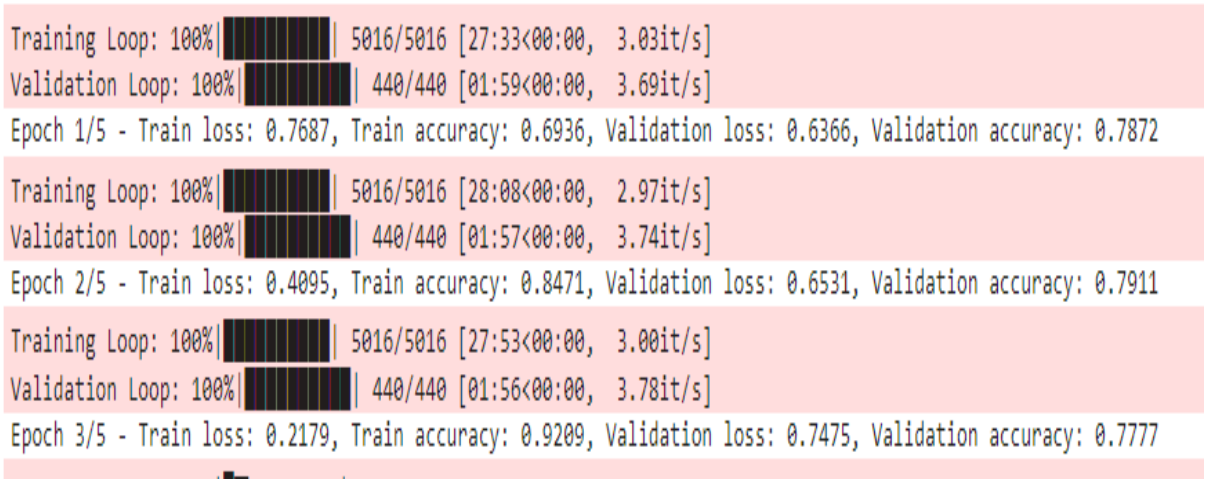


Figure 4.14: Result Transfer Learning

Left and Right Eyes Models : (4.4)

Explanation :

Actually, when we looked at the data and the names of the images, we noticed that the eyes are split too (left and right). Ok, so what's the problem? The problem is that images are only matrices, and feature extraction works with numbers, so the features of the left eye aren't like those of the right eye, even if the two eyes are the same class, for example, 'No DR'. Same class, but different features.

That will cause bad predictions even if we get good accuracy, so what we can do is build two models and combine them in one model.

The output is an array of 5 values that contain the probability of each class, so he will loop item by item and get the probability of each class.

Example :

let tell we pass an image of (left eye), class : (4: Proliferative DR)

The output of the two models will be like this:

left_model : [0.1, 0.3, 0.3, 0.5, 0.8]

right_model : [0.2, 0.4, 0.4, 0.5, 0.4]

left model gets 0.8 (80%) for class 4 and the right model only 0.4, but the average is the max value

The value be always the max one, that's why it's better.

Implementation

We use for left_model and right_model The same architecture used for the previous model

Class Combined Model

```
class CombinedModel(nn.Module):
    def __init__(self, model1, model2):
        super(CombinedModel, self).__init__()
        self.model1 = model1
        self.model2 = model2

    def forward(self, x):
        output1 = self.model1(x)
        output2 = self.model2(x)

        combined_output = (output1 + output2) / 2  # Average the predictions

        return combined_output

full_model =CombinedModel(left_model , right_model)
```

Figure 4.15: Combine Model

Results : (4.5)

Left Model :

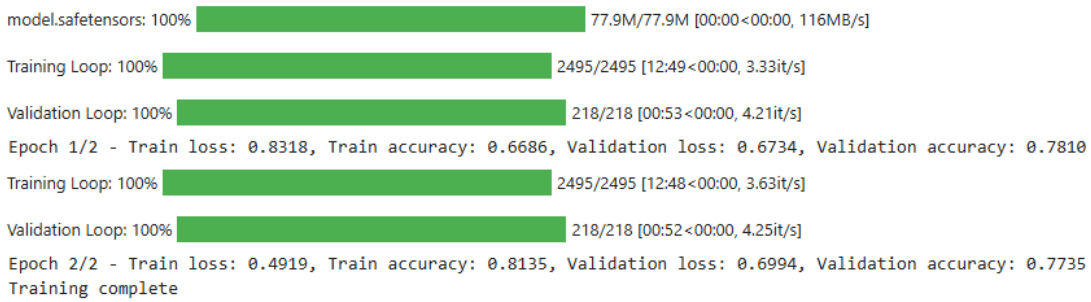


Figure 4.16: Left Model

Right Model :

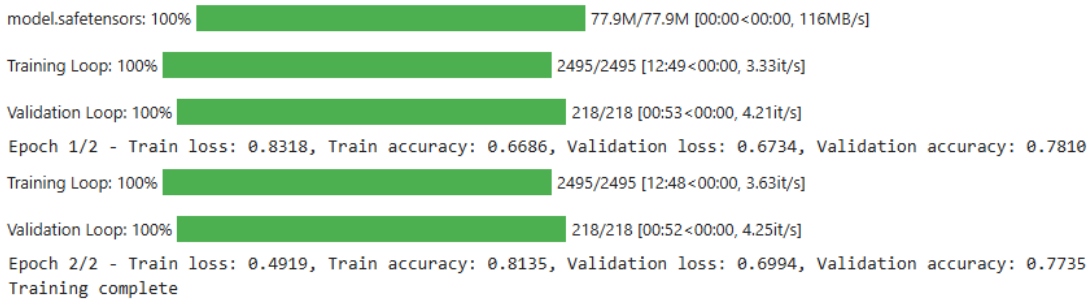


Figure 4.17: Right Model

Combined Model :

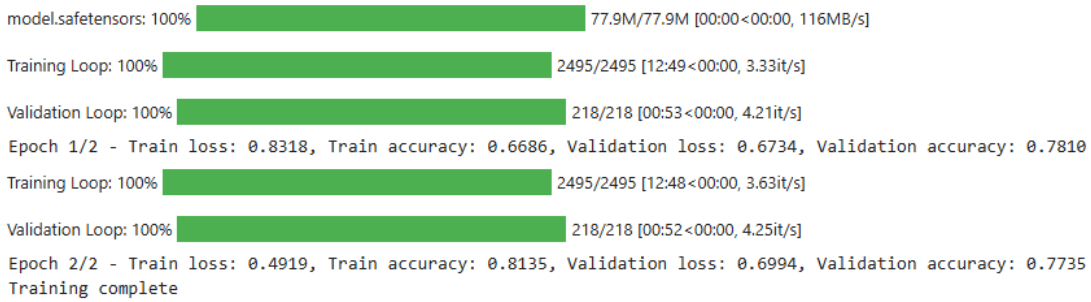


Figure 4.18: Combined Model

Actually, these results are good, but let's see the confusion metrics to understand why these are bad results not good

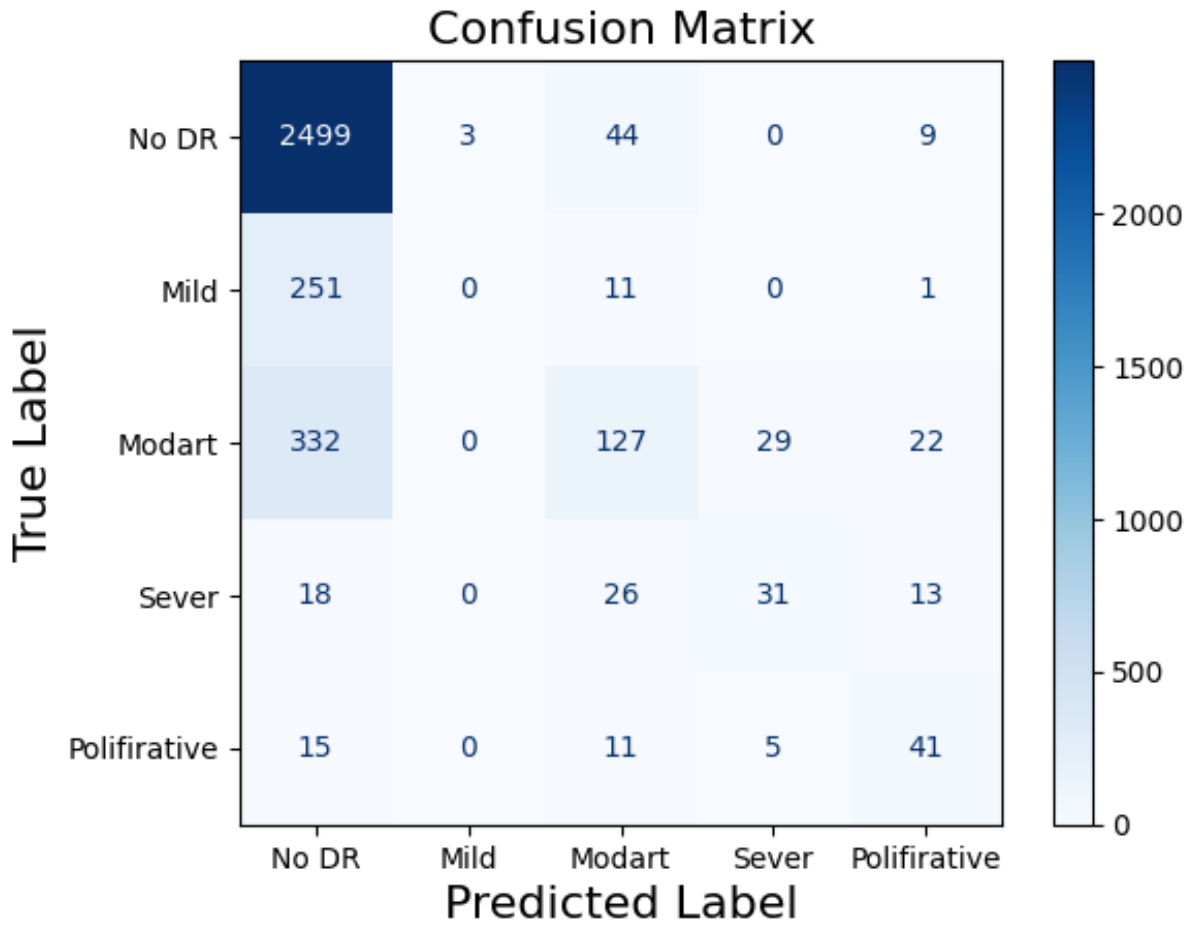


Figure 4.19: Confusion Matrix

Explanation :

to understand how the results are bad, see the predicted values are 2499 images in class (No DR) and are the true values to, so it's a good prediction

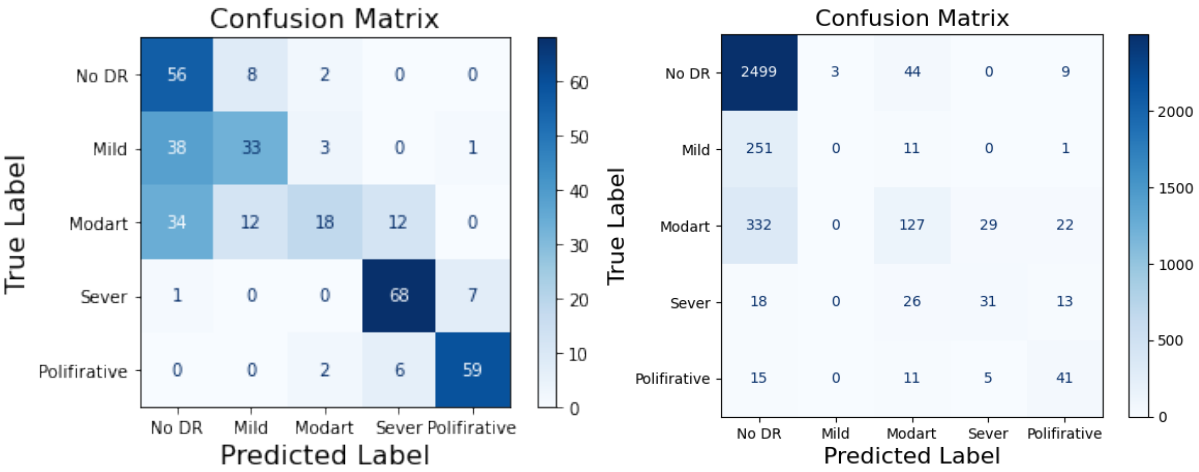
but let's see (Mild) and (Moderate) (215 and 332) are predicted as (No DR)

so how we get this accuracy is simple because data are imbalanced it's classified the other classes in (No DR) class

the real (No DR) class contains a lot of images, so it gives 77% accuracy

Comparison : (4.6)

let's compare two final models that we build using first data and second data



(a) First Model **(b)** Second Model

as we see first model is more stable and predicts more (Serve) and (Proliferative) values so it detects the disease

4.3 Application

4.3.1 Tools Used

4.3.1.1 Rect-Native :

A mobile development framework is one of the most vital components for creating feature-rich and high-performance mobile applications. Selecting the right framework has a direct impact on the quality of development and user experience. React Native is among the top mobile development frameworks used by developers across the world today. It offers robust functionality and a significant number of powerful features for both iOS and Android applications [99].

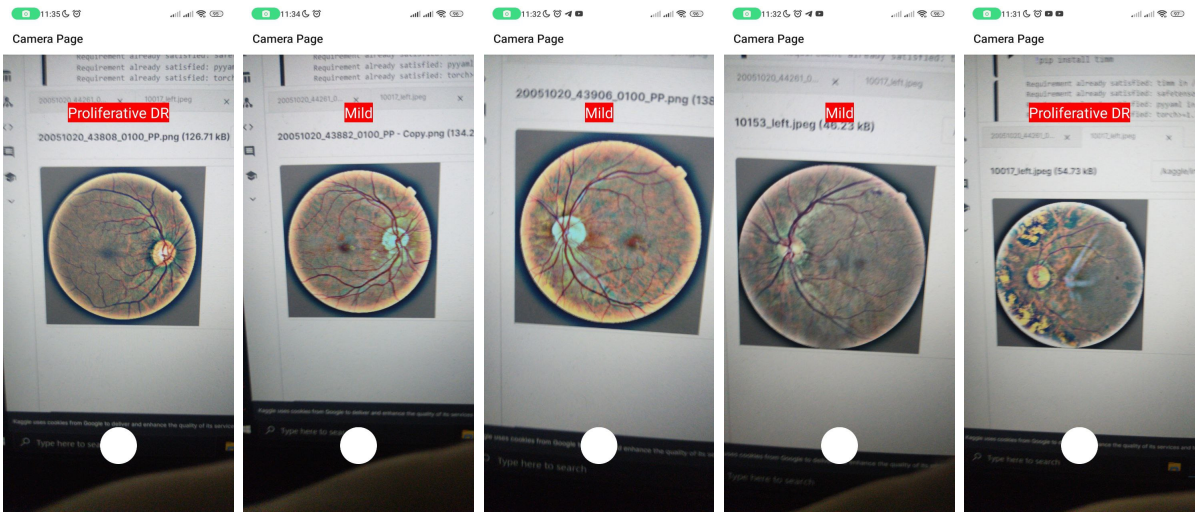
4.3.1.2 Expo :

Expo is a framework that makes developing Android and iOS apps easier. Our framework provides file-based routing, a standard library of native modules, and much more. Expo is open source with an active community on GitHub and Discord [100].

4.3.1.3 TensorFlow Lite :

TensorFlow Lite is a set of tools that enables on-device machine learning by helping developers run their models on mobile, embedded, and edge devices [101].

4.3.2 Detection of (DR) Mobile App



(a) NO DR **(b)** Mild **(c)** Modert **(d)** Sever **(e)** Prolifirative

Figure 4.21: images of Diabetic Retinopathy (DR) [79].

as we see here it's predict (Mild) and (Proliferative DR) more than other classes, it's just 5 random images that we test.

4.4 Conclusion

In this chapter, we learned that the main power of deep learning is the dataset. We have implemented a lot of things (SMOTE and visualized the data processing, transfer learning, etc.), and we tried a lot of ways to augment the accuracy and create a better model, but in this condition, when you have imbalanced data or small data, it's really difficult.

Conclusion

General Conclusion

In this final paper, we described a system for diagnosing eye diseases, particularly those leading to blindness that require early and accurate diagnosis such as diabetic retinopathy (DR) and diabetic macular edema (DME) using a deep learning (DL) approach.

First, we conducted a comprehensive study on diabetes in general, followed by an exploration of the anatomy and specific eye diseases resulting from diabetes complications, including symptoms and diagnosis. Then, we identified relevant concepts in machine learning (ML) and deep learning (DL) that pertain to our problem and proposed potential solutions. Additionally, we reviewed existing work in this context.

As our contribution, we proposed an approach based on Convolutional Neural Networks (CNN). This approach involves a classifier that predicts one of five categories (0, 1, 2, 3, and 4) for diagnosis. We first prepared the training data, focusing our study on two main themes: The first theme consists of training the proposed CNN model (training phase), and the second theme involves a forecasting procedure using the proposed model (forecast phase).

The proposed system comprises two modules:

Module 1: transfer learning model with small data

Module 2: transfer learning model with imbalanced data

This work has been carried out despite several challenges, including a small data volume, an unbalanced database, and poor image accuracy. The most significant difficulties were encountered during data preparation and the testing phase, as well as the lack of robust hardware needed for training, which led to considerable time expenditure.

For future work, we aim to address many related problems and create another models on different datasets to diagnose a broader range of retinal diseases. Additionally, we plan to integrate our model directly into fundus scanning devices to assist ophthalmologists in making decisions. We also envision embedding our model into smartphones equipped with fundus cameras, such as the Remidio fundus on phone camera, to enable patients to

diagnose their retinas conveniently.

Perspectives :

The optimal goal is to develop a large and diverse database to enhance the robustness and accuracy of our model. Future perspectives include:

Data Augmentation and Acquisition: Expanding the dataset with more diverse and higher-quality images to improve model training and validation.

Model Optimization: Refining the CNN architecture to enhance performance and reduce computational requirements, making it feasible for real-time applications.

Integration with Telemedicine: Leveraging the system within telemedicine frameworks to provide remote diagnostic services, especially in underserved regions.

User-Friendly Applications: Developing user-friendly mobile and desktop applications that allow both patients and healthcare providers to easily utilize the diagnostic tool.

Cross-Validation: Conducting extensive cross-validation studies with multiple datasets from different demographics to ensure the models generalize ability and reliability.

By pursuing these perspectives, we aim to significantly advance the field of automated retinal disease diagnosis, ultimately contributing to better healthcare outcomes and the prevention of blindness.

Bibliography

[1] Cleveland Clinic. Diabetes mellitus. https://2u.pw/klWjExKZ, 2024. Accessed: 2024-05-31.

[2] J. J. Staal, M. D. Abramo, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.

[3] World Health Organization. Diabetes 2018. https://www.who.int/news-room/fact-sheets/detail/diabetes. Accessed: 2024-05-02.

[4] International Diabetes Federation. IDF Diabetes Atlas 10th edition 2022. https://diabetesatlas.org/. Accessed: 2024-05-02.

[5] J. J. Staal, M. D. Abramo, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004. Accessed: 2024-05-31.

[6] Diabetic retinopathy: A vision health initiative. https://www.ncbi.nlm.nih.gov/books/NBK581930/, 2021. Accessed: 22-05-2024.

[7] Muhammad Usman Akram. *Retinal Image Analysis & Hybrid Classifier for Screening of Diabetic Retinopathy*. Phd thesis, National University of Sciences & Technology, Department of Computer Engineering, College of Electrical & Mechanical Engineering, 2012. Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Engineering.

[8] A. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imag.*, 19(3):203–210, 2000.

[9] A.M. Mendonca and A. Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Trans. Med. Imag.*, 25(9):1200–1213, 2006.

[10] National institute of diabetes and digestive and kidney diseases. https://bit.ly/3h3kDH1. Accessed: 2024-05-14.

[11] Lucentis ranibizumab injection. https://bit.ly/3h2a1Is. Accessed: 2024-05-14.

[12] Centers for Disease Control and Prevention (CDC). National diabetes statistics report, 2020. https://bit.ly/330murb. Accessed: 2024-05-14.

[13] Diabetic retinopathy (the basics). https://bit.ly/3jKNMsc. Accessed: 2024-05-14.

[14] T Wong, C Cheung, M Larsen, et al. Diabetic retinopathy. *Nature Rev. Disease Primers*, 2016.

[15] P Porwal, S Pachade, M Kokare, et al. Indian diabetic retinopathy image dataset (idrid): Diabetic retinopathy – segmentation and grading challenge. *Medical Image Analysis*, 59, 2020.

[16] YR Sharma, Karthikeya R, et al. Recent advances in management of diabetic macular edema. *Current Diabetes Reviews*, 11(2):79–97, 2015.

[17] The Royal College of Ophthalmologists. Diabetic retinopathy guidelines, December 2012. https://bit.ly/2Dyhi57. Accessed: 2024-05-16.

[18] J. Wells, A. Glassman, A. Ayala, et al. Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema. *The New England Journal of Medicine*, 372:193–1203, 2015.

[19] F. Bandelloa, M. B. Parodia, et al. Diabetic macular edema. In *Dev Ophthalmol*, volume 47, pages 73–110. Basel, Karger, 2010.

[20] Y Chung, YH Kim, S Ha, et al. Role of inflammation in classification of diabetic macular edema by optical coherence tomography. *Journal of Diabetes Research*, 2019:Article ID 8164250, 2019.

[21] National eye institute. https://bit.ly/3buIpKZ. Accessed: 2024-05-16.

[22] T. Wong, C. Cheung, M. Larsen, et al. Diabetic retinopathy. *Nature Reviews Disease Primers*, 2016.

[23] P. Sajda. Machine learning for detection and diagnosis of disease. *Annual Review of Biomedical Engineering*, 8:537–565, 2006.

[24] B. Povazay, K. Bizheva, A. H. Unterhuber, W. Drexler, A. Apolonski, and et al. Submicrometer axial resolution optical coherence tomography. *Optics Letters*, 27(20):1800, 2002.

[25] E. A. Swanson, J. A. Izatt, M. R. Hee, and J. G. Fujimoto. In-vivo retinal imaging by optical coherence tomography. *Optics Letters*, 18(22):1864–1866, 1993.

[26] J. Fujimoto and E. Swanson. The development, commercialization and impact of optical coherence tomography. *Investigative Ophthalmology and Visual Science*, 57(9), 2016.

[27] S. Aumann, S. Donner, J. Fischer, and et al. Optical coherence tomography (oct): Principle and technical realization. In *High Resolution Imaging in Microscopy and Ophthalmology*. Springer, Cham, 2019.

[28] D. Huang and et al. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.

[29] A. F. Fercher and et al. Optical coherence tomography - principles and applications. *Reports on Progress in Physics*, 66(2):239–303, 2003.

[30] Explaining oct scans. https://bit.ly/3hGwj2X. Accessed: 2024-05-20.

[31] Ophthalmology web. https://bit.ly/3lW6FdB. Accessed: 2024-05-20.

[32] Heidelberg engineering. https://bit.ly/2ZdIwFY. Accessed: 2024-05-20.

[33] A. F. Fercher and et al. Measurement of intraocular distances by backscattering spectral interferometry. *Optics Communications*, 117:43–48, 1995.

[34] W. Drexler and et al. Optical coherence tomography today: speed, contrast, and multimodality. *Journal of Biomedical Optics*, 19(7):071412, 2014.

[35] S. Aumann, S. Donner, J. Fischer, and F. Müller. Optical coherence tomography (oct): Principle and technical realization. In *High Resolution Imaging in Microscopy and Ophthalmology*. Springer, Cham, 2019.

[36] D. Gildea. The diagnostic value of optical coherence tomography angiography in diabetic retinopathy: a systematic review. *International Ophthalmology*, 39(10):2413–2433, 2019.

[37] E. Talisa, K. Nadia, and et al. A review of optical coherence tomography angiography (octa). *International Journal of Retina and Vitreous*, 1(1):5, 2015.

[38] X. Yao, T. Son, and J. Ma. Developing portable widefield fundus camera for teleophthalmology: Technical challenges and potential solutions. *Experimental Biology and Medicine*, 247(4):289–299, 2022.

[39] R. Besenczi, J. Tóth, and A. Hajdu. A review on automatic analysis techniques for color fundus photographs. *Computational and Structural Biotechnology Journal*, 14:371–384, 2016.

[40] B. Xiao, Q. Liao, Y. Li, F. Weng, L. Jin, Y. Wang, W. Huang, J. Yi, M. J. Burton, and J. L. Yip. Validation of handheld fundus camera with mydriasis for retinal imaging of diabetic retinopathy screening in china: a prospective comparison study. *BMJ open*, 10(10):e040196, 2020.

[41] Topcon. Topcon TRC-NW8F-Plus. https://www.ophthalmicmart.com/product/topcon-trc-nw8f-plus/. Accessed: 2024-05-21.

[42] W. Allyn. iExaminer^{TM}. https://www.welchallyn.com/en/microsites/iexaminer.html. Accessed: 2024-05-21.

[43] P. J. Saine and M. E. Tyler. *Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging*, volume 132. Butterworth Heinemann Boston, 2002.

[44] N. Panwar, P. Huang, J. Lee, P. A. Keane, T. S. Chuan, A. Richhariya, S. Teoh, T. H. Lim, and R. Agrawal. Fundus photography in the 21st century—a review of recent technological advances and their implications for worldwide healthcare. *Telemedicine and e-Health*, 22(3):198–208, 2016.

[45] C. R. Baumal. Imaging in diabetic retinopathy. In C. R. Baumal and J. S. Duker, editors, *Current Management of Diabetic Retinopathy*, chapter 4, pages 25–36. Elsevier, 2018.

[46] S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Academic, Burlington, MA, 1 edition, 1999.

[47] Flatiron School. Deep learning vs. machine learning, May 2021. Accessed on May 18, 2023.

[48] Persado. Artificial intelligence vs. machine learning vs. data science, 2021. Accessed on May 18, 2023.

[49] Mariette Awad and Rahul Khanna. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Apress, New York, NY, 2015.

[50] Edureka. What is machine learning?: Types of machine learning. https://www.edureka.co/blog/what-is-machine-learning/, May 2019. Accessed: 2024-05-23.

[51] Canburak Tumer. Machine learning basics with examples- part 2 supervised learning. https://2u.pw/909RiIt6, Aug 2018. Accessed: 2024-05-23.

[52] Analytics Steps. Introduction to machine learning: Supervised and unsupervised learning. https://2u.pw/HPBmNUhh, 2020. Accessed: 2024-05-23.

[53] Richard S Sutton and Andrew G Barto. *Introduction to Reinforcement Learning*, volume 135. MIT Press, Cambridge, MA, 1998.

[54] MathWorks. Three things to know about reinforcement learning. https://www.kdnuggets.com/2019/10/mathworks-reinforcement-learning.html, Oct 2019. Accessed: 2024-05-24.

[55] Jeff Heaton. Ian goodfellow, yoshua bengio, and aaron courville: Deep learning. *Genetic Programming and Evolvable Machines*, 19:305–307, 2018.

[56] Harvey Lodish. Overview of neuron structure and function. https://www.ncbi.nlm.nih.gov/books/NBK21535/, Jan 1970. Accessed: 2024-05-24.

[57] Kite. Introduction to artificial neural networks in python. https://roioverload.com/biology-inspires-the-artificial-neural-network/, Aug 2019. Accessed: 2024-05-24.

[58] Boundless. Boundless psychology. https://courses.lumenlearning.com/boundless-psychology/chapter/neurons/. Accessed: 2024-05-24.

[59] Andrej Krenker, Janez Bester, and Andrej Kos. Introduction to the artificial neural networks. In *Artificial Neural Networks: Methodological Advances and Biomedical Applications*, pages 1–18. InTech, 2011.

[60] MissingLink. 7 types of activation functions in neural networks: How to choose? https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/. Accessed: 2024-05-24.

[61] Amey Varangaonkar et al. Top deep learning architectures. https://hub.packtpub.com/top-5-deep-learning-architectures/, Sep 2018. Accessed: 2024-05-24.

[62] Athanasios Voulodimos et al. Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, 2018:1–13, 2018.

[63] Manjunath Jogin et al. Feature extraction using convolution neural networks (cnn) and deep learning. In *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*, pages 2319–2323. IEEE, 2018.

[64] Ping-Rong Chen et al. Dsnet: An efficient cnn for road scene segmentation. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 424–432. IEEE, 2019.

[65] Yann LeCun et al. Lenet-5, convolutional neural networks. http://yann.lecun.com/exdb/lenet, 2015. Accessed: 2024-05-26.

[66] Deep convolutional neural networks (alexnet). https://d2l.ai/chapter_convolutional-modern/alexnet.html#architecture. Colab [mxnet] Open the notebook in Colab, Colab [pytorch] Open the notebook in Colab, Colab [tensorflow] Open the notebook in Colab. Accessed: 2024-05-26.

[67] Kaiming He et al. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[68] Mingxing Tan and Quoc V Le. Efficientnet: Improving accuracy and efficiency through automl and model scaling. *arXiv preprint arXiv:1905.11946*, 2019.

[69] Christian Szegedy et al. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826, 2016.

[70] Danilo Mandic and Jonathon Chambers. *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. Wiley, Chichester, England, 2001.

[71] Neural network examples. https://www.edrawsoft.com/neural-network-examples.html. Accessed: 2024-05-26.

[72] Hongyang Xu et al. Sentiment analysis of chinese version using svm & rnn. In *Proceedings of the 6th International Conference on Information Engineering*, pages 1–5, 2017.

[73] K. R. Remya, M. N. Giriprasad, and M. S. Sudhakar. A localized feature description means assisting diabetic macular edema detection and classification. https://2u.pw/tOkRwLJG, 2023. Accessed: 2024-05-29.

[74] Udyavara R Acharya et al. Computer-based detection of diabetes retinopathy stages using digital fundus images. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 223(5):545–553, 2009.

[75] Mehdi Torabian Esfahani, Mahsa Ghaderi, and Raheleh KafiYeh. Classification of diabetic and normal fundus images using new deep learning method. *Leonardo Electronic Journal of Practices and Technologies*, 17(32):233–248, 2018.

[76] Elaouaber Zineb Aziza et al. Decision tree cart algorithm for diabetic retinopathy classification. In *2019 6th International Conference on Image and Signal Processing and their Applications (ISPA)*, pages 1–5. IEEE, 2019.

[77] Michael Abramo and Christine N. Kay. Chapter 6- image processing. In Stephen J. Ryan et al., editors, *Retina (Fifth Edition)*, pages 151–176. W.B. Saunders, London, 5th edition, 2013.

[78] Remidio fundus , we blend technology and user centric design to craft beautiful digital solutions for your ophthalmic screening and diagnosis. https://www.remidio.com/. Accessed: 2024-05-26.

[79] IBM. Pytorch: An open source machine learning library for python, 2024. Accessed: 26-05-2024.

[80] sam991. Diabeticretinopathy_agumented_data. https://www.kaggle.com/datasets/sam991/diabeticretinopathy-agumented-data, 2021. Accessed: 30-03-2024.

[81] Jeremy West, Dan Ventura, and Sean Warnick. Spring research presentation: A theoretical foundation for inductive transfer. In *Brigham Young University, College of Physical and Mathematical Sciences*, volume 1.08, 2007.

[82] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[83] Suvaditya Mukherjee. The annotated resnet-50. https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758. Accessed : 07/06/2024.

[84] https://medium.com/nocoding-ai/densenet121 760df192f12d. Densenet. https://medium.com/nocoding-ai/densenet121-760df192f12d.

[85] NOCODING AI. Densenet. https://medium.com/nocoding-ai/densenet121-760df192f12d.

[86] Petru Potrimba. What is efficientnet? https://blog.roboflow.com/what-is-efficientnet/. Accessed: 07-06-2024.

[87] wisdomml. EfficientNet and its Performance Comparison with Other Transfer Learning Networks. https://wisdomml.in/efficientnet-and-its-performance-comparison-with-other-transfer-learning-network

[88] Python core team. Python. https://www.python.org/. Accessed: 2024-05-31.

[89] matplotlib documentation. matplotlib documentation. https://matplotlib.org/faq/usage_faq.html#coding-styles. Accessed: 2024-05-31.

[90] TensorFlow. Tensorflow. https://www.tensorflow.org/. Accessed: 2024-05-31.

[91] Keras Team. Keras. https://keras.io/. Accessed: 2024-05-31.

[92] Your Machine Learning and Data Science Community. Kaggle. https://www.kaggle.com/. Accessed: 2024-05-31.

[93] APTOS 2019 Blindness Detection competition. APTOS 2019 Blindness Detection competition.

[94] Google Colab. Google colab. https://colab.research.google.com/. Accessed: 2024-05-31.

[95] CUDA Explained - Why Deep Learning uses GPUs. Cuda explained - why deep learning uses gpus.

[96] GeeksforGeeks. Cross validation in machine learning. https://www.geeksforgeeks.org/cross-validation-machine-learning/, 2021. Accessed: 30-05-2024.

[97] Sam991. Diabetic retinopathy - augmented data. https://www.kaggle.com/datasets/sam991/diabeticretinopathy-agumented-data, 2021. Accessed: 30-05-2024.

[98] GeeksforGeeks. Smote for imbalanced classification with python. https://www.geeksforgeeks.org/smote-for-imbalanced-classification-with-python/. Accessed: 30-05-2024.

[99] Back4App Blog. React native framework. https://blog.back4app.com/react-native-framework/, 2024. Accessed:30-05-2024.

[100] Expo Documentation. Get started - introduction. https://docs.expo.dev/get-started/introduction/, 2024. Accessed: 30-05-2024.

[101] TensorFlow. Tensorflow lite guide. https://www.tensorflow.org/lite/guide#:~:text=TensorFlow%20Lite%20is%20a%20set,%2C%20embedded%2C%20and%20edge%20devices., 2024. Accessed: 30-05-2024.