Democratic and Popular Republic of Algeria

Ministry of Higher Education and Scientific Research

UNIVERSITY M'HAMED BOUGARA-BOUMERDES

Faculty of Technology

Department of Electrical Systems Engineering

Master's thesis

Presented by

BIDA Manel and BOUARAKI Wissam

Field of Study: Telecommunications

Major: Network and Telecommunications

DIABETIC RETINOPATHY DETECTION USING CONVOLUTIONAL NEURAL NETWORKS

Defended on 02/07/2023 in front of the jury composed of:

ACHELI	Dalila	Pr	UMBB	Chair
MESSAOUDI	Noureddine	MCA	UMBB	Examinator
GUERBAI	Yasmine	MCA	UMBB	Supervisor

Academic Year: 2022/2023

ACKNOLDGMENT

First and foremost, we would like to express our profound gratitude to God for His blessings and guidance throughout this project. It is through His grace that we have accomplished this significant milestone, and we are humbled by His unwavering support.

We extend our heartfelt thanks to our esteemed supervisor, Dr. GUERBAI. Her exceptional dedication, encouragement, and unwavering belief in our abilities have been instrumental in our success. We are deeply grateful for her invaluable guidance, insightful feedback, and continuous motivation that have shaped this project into what it is today.

We would like to extend our sincere appreciation to the esteemed members of the jury, Pr. ACHELI and Mr. MESSAOUDI for accepting to evaluate our work. Their expertise, dedication, and valuable feedbacks refines our project and elevates its quality.

We also wish to extend our gratitude to the University of Boumerdes, Faculty of Technology and the Incubator of Boumerdes' team. The supportive environment and resources provided by our institution have been instrumental in our growth and learning. We are grateful for the opportunities it has offered us in our field.

A special word of thanks goes to Mr. MESSAOUDI, the head of the department, for his exceptional leadership and countless efforts in fostering an environment of innovation and excellence. His guidance and support have been invaluable throughout our academic journey.

We cannot overlook the invaluable contributions of our colleagues, BOUKANDOURA OMAR, RABIA ALLALOU and SAIBI ASMA. Their collaboration, unwavering support, and camaraderie have added immeasurable value to our journey. Their dedication and assistance have not only enhanced the outcome of our project but have also made this experience truly remarkable. We are deeply grateful for their invaluable presence and unwavering support throughout this endeavor.

In the face of diabetic retinopathy, we extend our deepest empathy and wishes for courage. May you find strength and be surrounded by the support and resources needed to overcome this disease.

Lastly, we would like to express our heartfelt gratitude to everyone who has believed in us, we appreciate all those who have extended a helping hand or offered a kind word, even the smallest gestures as each contribution has made a significant difference in our success.

MANEL & WISSAM.

DEDICATION

A special word of thanks goes to my parents, who have supported me throughout my journey and life. I consider myself fortunate to be their daughter, as they have always been by my side, providing unwavering support.

I would like to express my appreciation to my brother, WASSIM and my sisters, IKRAM and NIHAL, for their constant support. I am grateful for their encouragement and guidance. I would also, to extend my sincere thanks to all my extended family, uncles, and aunts, especially my aunts Samya , DJAMILA, NACIRA and my grandmother, for their support and encouragement.

I would like to extend a special dedication to my partner, MANEL. Her support, constant motivation, and unwavering commitment have been a continuous source of inspiration throughout the project. Together, we have overcome obstacles, solved problems, and create something we can be proud of. Her trust, dedication, and collaborative spirit have significantly contributed to the success of this work. Thank you, for your invaluable support and for being an exceptional partner in this journey.

Lastly, I want to express my heartfelt gratitude to my friends and every person who has supported me with a kind word or gesture. Your contribution has made a significant impact, and I am truly grateful.

Thank you all from the bottom of my heart.

WISSAM.

إهداء

قال رسول الله صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: (مَنْ صَنَعَ إِلَيْكُمْ مَعْرُوفًا فَكَافِئُوهُ، فَإِنْ لَمْ تَجِدُوا مَا تُكَافِئُونَهُ فَادْعُوا لَهُ حَتَّى تَرَوْا أَنَّكُمْ قَدْ كَافَأْتُمُوهُ). رواه أبو داود

إلى من ترك في قلبي غصة لن ترممها السنين، إلى من يرتعش قلبي لذكراه، إلى من أودعني الله، الراحل الباقي في قلبي، جدي -رحمه الله -وأسكنه فسيح جناته.

إلى من كلله الله بالهيبة والوقار، إلى من علمني العطاء دون انتظار، إلى من أحمل اسمه بكل افتخار، أبي الغالي، حفظه الله ورعاه.

إلى من تهوّن علي المشقة، إلى اليسر الوحيد بين كل العسر، دعواتها سند، ابتسامتها حياة و رضاها غاية، أمي، فاللهم إني استودعتك عافيتها إنك خير المستودعين.

إلى من بهم أفتخر وعليهم أعتمد، سندي وكياني إخوتي نسيمة وزينب حفظكم الله ورعاكم.
إلى ظلي، سكينتي، ضلعي ونصف قلبي، أخي، فاللهم أخي أينما حلت خطاه.

إلى التي لم تتوانى في مد يد العون لي، تحلت بالإخاء، الوفاء والعطاء، بين طيات التعب حفتني بالطمأنينة، عاملتني بالحسنى وشدت أزري نحو القمة، ويسام فاللهم ابعث الطمأنينة في قلبها وابعد عنها كل ما يؤذيها، ارح قلبها، نور طريقها واجعل ايامها جميلة شبيهة بيها.

إلى القريبين من القلب، أقرباء وأصدقاء، الداعمين، المساندين في السراء والضراء وفقكم الله وسدد خطاكم.

-الحمد لله في كل بدءٍ ومختتم-

منال.

ملخص

أحد التحديات التي يواجهها الاطباء في مجال طب العيون هي صعوبة الكشف عن اعتلال الشبكية السكري، الذي يعتبر حالة صحية خطيرة تؤثر على العين، والتي تستلزم التشخيص المبكر لتجنب فقدان البصر. الهدف الرئيسي لمشروعنا هو الكشف المبكر على اعتلال الشبكية السكري باستخدام الذكاء الاصطناعي من خلال تقنيات التعلم العميق وذلك بتطويرنا لنظامٍ دقيق لتحليل صور الشبكية باستعمال نماذج الشبكة العصبية الترابطية، لهذه النماذج قدرة استثنائية على تحديد العلامات المبكرة لإعتلال الشبكية السكري، مما يسمح بتجنب فقدان البصر. نهدف إلى تزويد أخصائيي طب العيون بأداة موثوقة وداعمة لتسهيل الكشف. يُحدث هذا المشروع تقدماً كبيراً في مجال طب العيون، ويتيح آفاقاً جديدة لتحسين تشخيصات اعتلال الشبكية السكري.

كلمات مفتاحية: اعتلال الشبكية السكري، الذكاء الاصطناعي، التعلم العميق، الشبكة العصبية الترابطية

Abstract

The project aimed to improve the early detection of diabetic retinopathy, a serious eye condition, by utilizing artificial intelligence (AI) and deep learning techniques. Through the development of a precise retinal image analysis system using CNN models, the project provided ophthalmologists with a reliable tool for detecting early warning signs of diabetic retinopathy. This advancement in ophthalmic care enables timely intervention and prevention of vision loss. By integrating AI into medical practice, the project offers new perspectives for improving outcomes for diabetic retinopathy patients.

Keywords: Diabetic Retinopathy, AI, Deep-learning, CNN.

Résumé

Les progrès technologiques permettent d'améliorer la détection précoce de la rétinopathie diabétique grâce à l'intelligence artificielle (IA) et l'apprentissage profond. Notre projet développe un système d'analyse précis des images rétiniennes en utilisant des modèles de réseaux de neurones convolutionnels (CNN). Ce système reconnaît les signes précurseurs de la rétinopathie diabétique, favorisant une intervention précoce et la prévention de la perte de vision. Il offre aux ophtalmologistes un outil fiable de soutien et ouvre de nouvelles perspectives pour les soins Ophtalmiques en intégrant l'IA dans la pratique médicale.

Mots Clés: La rétinopathie diabétique, IA, l'apprentissage profond, les réseaux de neurones convolutionnels (CNN).

Table of content

Table of figures

Table List

Acronyms

ADAM	Adaptive Moment Estimation.
AI	Artificial Intelligence.
ANN	Artificial Neural Networks.
API	Application Programming Interface.
CNN	Convolutional Neural Network.
CSV	Comma Separated Values.
CSS	Cascading Style Sheets.
CPU	Central Processing Unit.
DL	Deep Learning.
EFFICIENTNET	Efficient Neural Network.
Ep	Epochs.
ER	Error Rate.
FA	Fluorescein Angiography.
FN	False Negative.
FP	False Positive.
GPU	Graphics Processing Unit.
GRU	Gated Recurrent Unit.
HTML	Hypertext Markup Language.
LSTM	Long Short-Term Memory.

ML	Machine Learning.
MLP	Multilayer Perceptron.
MOBILENET	Mobile Network.
NPDR	Nonproliferative Diabetic Retinopathy.
OCT	Optical Coherence Tomography.
P	Precision.
PDR	Proliferative Diabetic Retinopathy.
RAM	Random Access Memory.
RELU	Rectified Linear Unit.
RGB	Red Green Blue.
RNN	Recurrent Neural Network.
S	Sigma.
SGD	Stochastic Gradient Descent.
TN	True Negative.
TP	True Positive.
VGG	Visual Geometry Group.

General Introduction

Diabetic retinopathy is a common and potentially sight-threatening complication of diabetes mellitus. Detecting the condition early and intervening in a timely manner are crucial for preventing vision loss and improving patient outcomes. In recent years, artificial intelligence (AI) has emerged as a powerful tool in healthcare, revolutionizing medical imaging and diagnostics. By utilizing advanced algorithms and machine or deep learning techniques, AI has shown promising results in the detection and diagnosis of diabetic retinopathy.

Traditional methods of screening for diabetic retinopathy involve manual examination of retinal images by ophthalmologists. However, this approach is time-consuming, subjective, and susceptible to human error. AI provides a more efficient and objective alternative by automating the screening process and delivering accurate and consistent results.

The application of AI in the detection of diabetic retinopathy entails developing deep learning models trained on extensive datasets of retinal images. These models can analyze intricate features and patterns within the images, enabling them to identify early indicators of retinopathy, such as microaneurysm and, hemorrhages. AI not only enhances the accuracy and speed of detection but also reduces the workload on healthcare professionals, allowing them to focus on more critical tasks.

This study aims to delve into the evolution of AI in diabetic retinopathy detection and its transformative potential in ophthalmic care. We will explore the techniques and algorithms employed, examine the prospective benefits for patients and healthcare professionals, and outline future prospects for the application of AI in the field of ocular health. That is why we will deepen our research to find methods for detecting diabetic retinopathy. So, the question that arises is how can we proceed to detect diabetic retinopathy using a deep learning approach?

The objectives behind our study are to predict and detect diabetic retinopathy in order to propose an approach based on deep learning models. Additionally, we aim to offer an aid to the ophthalmologists and promote telemedicine in our country.

In order to effectively carry out this project, we have structured this thesis as follows:

Chapter 1: "Generalities of Diabetec Retinopathy" primarily defines diabetes and its types, diabetic retinopathy, classification of diabetic retinopathy, and detection methods.

Chapter 2: "Methods Used for Retinopathy Detection" presents the definition of AI, the concepts of machine and deep learning, CNN (Convolutional Neural Network), and the models and architectures used in CNN.

Chapter 3: "Conception and Implementation" presents the solution to the posed problem, evaluate the results of the tests conducted in relation to the set evaluation criteria

Chapter.1

Generalities about diabetic retinopathy

1.1 Introduction

Diabetic Retinopathy (DR) is a problem with diabetes that reasons blood vessels of the retina to swell and to leak fluids and blood. Currently it is commonly known for causing blindness worldwide. This chapter presents, the definition of diabetes and its types, the general anatomy of the eye, the definition of DR, classification, causes, symptoms and methods of detection and treatments.

1.2 Diabetes definition

Diabetes is a medical condition in which the body is unable to effectively regulate the amount of sugar (glucose) in the blood. The human body breaks down most of the food into sugar and releases it into the bloodstream. When the blood sugar goes up, it signals the pancreas to release insulin [1], which is a hormone that regulates blood glucose [2], which acts like a key to let the blood sugar into the body's cells for use as energy.

1.3 Types of diabetes

1.3.1 Type one diabetes

The pancreas doesn't make or makes very little insulin which helps blood sugar to enter the cells in the body for use as energy. Without insulin, blood sugar can't get into cells and builds up in the bloodstream [3].

1.3.2 Type two diabetes

The cells don't respond normally to insulin, this is called insulin resistance. The pancreas makes more insulin to try to get cells to respond. Eventually the pancreas can't keep up, and the blood sugar rises [3].

Both types of diabetes can lead to serious health complications over time and causes many of the symptoms and complications of diabetes, it damages the eyes specially, may cause vision loss, and even blindness.

1.4 Eye definition

The eye is the primary organ of vision and sight that allows us to see. It takes in the light and send visual information to the brain. The eye can see about 200 degrees in all directions, Parts of the eye work together to allow us to see images, movement, depth and millions of colors in varying shades [4].

1.4.1 Eye Anatomy

The parts of the eye include:

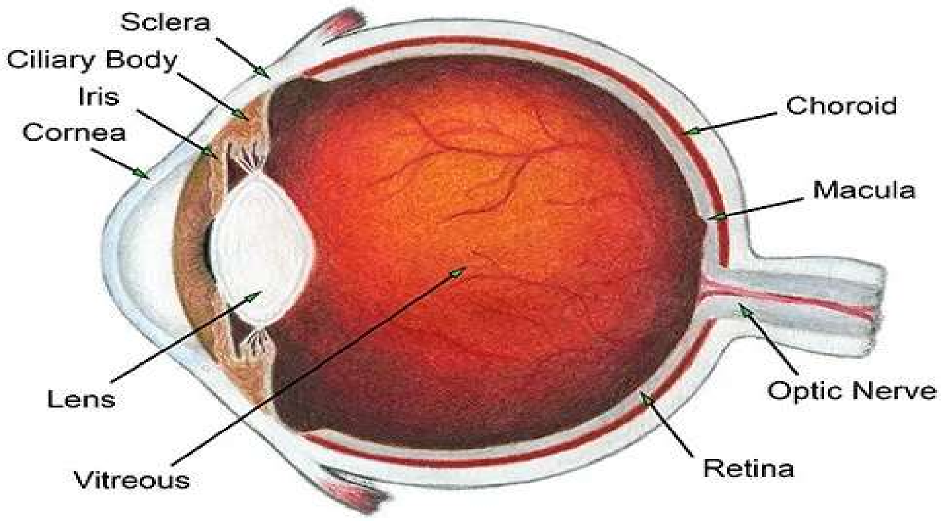


Fig.1.1: The eye anatomy [15].

Iris: the colored area of the eye, depending on the eye color.

Cornea: a clear layer that extends over the iris. Water and collagen make up the cornea. The tears protect the cornea and keep it lubricated.

The sclera: the white part of the eye that surrounds the iris.

Conjunctiva: a clear, thin tissue that covers the sclera and lines the inside of the eyelids.

Lens: which sits behind the pupil. It focuses the light that comes into the eye and sends light to the back of the eye.

Retina :a collection of cells that line the inside of the back of the eye. Part of the nervous system, the retinas send light and convert it into electrical impulses or neural signals. The retina has rods (cells that help to see in low light) and cones (cells that detect color).

Macula: a small area that is a part of the retina. It’s responsible for central vision and helping to see fine details and color.

Optic nerve: which is behind the retina. It carries signals from the retina to the brain which then interprets that visual information to tell us what we are seeing.

Muscles: which control the eye’s position and movement, how much light gets into the eye and the eyes’ ability to focus.

Vitreous: a transparent gel that fills the entire eye. It protects and maintains the shape of the eye [4].

1.5 diabetes and vision loss

When the blood glucose is too high over time, it can damage the tiny blood vessels in the back of the eye. This damage can begin during prediabetes, when blood glucose is higher than normal, but not high enough for being diagnosed with diabetes. Damaged blood vessels mayleak fluid and cause swelling. New, weak blood vessels may also begin to grow. These blood vessels can bleed into the middle part of the eye and cause dangerously high pressure inside the eye.

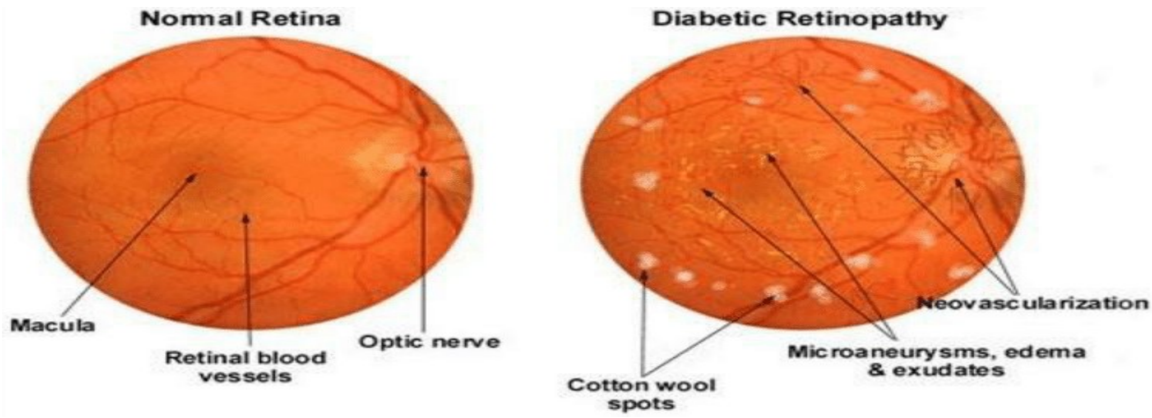


Fig.1.2: Normal retina and diabetic retinopathy retina [16].

Many conditions, several diseases and disorders whether they are eye-related or not, can affect how the eye work, may cause low vision and even vision loss that can lead to blindness which is also known as, Diabetic Retinopathy.

1.5.1 Diabetic Retinopathy definition

Diabetic retinopathy is a complication of diabetes and a leading cause of blindness. It occurs when diabetes damages the tiny blood vessels inside the retina, the light-sensitive tissue at the back of the eye. A healthy retina is necessary for good vision. Diabetic retinopathy can get worse and cause vision loss and even blindness, usually affects both eyes [5].

Developing diabetic retinopathy is related to the duration of diabetes. Type 2 diabetes has an insidious onset and can go unnoticed for years. As a result, patients may already have DR at the time of diagnosis. Type 1 diabetics, on the other hand, are diagnosed early in the course of their disease, and they typically do not develop retinopathy until years after the diagnosis is made. The risk of developing retinopathy increases after puberty. Twenty years after the diagnosis of diabetes, 80% of type 2 diabetics and nearly all type 1 diabetics show some signs of retinopathy [6].

1.5.2 Pathophysiology of DR

The pathophysiology of diabetic retinopathy (DR) is complex and involves various cellular and molecular mechanisms. Hyperglycémie, a hallmark of DR, occurs when blood sugar levels are high for a long period, affecting the small blood vessels of the retina. These vessels undergo structural modifications, such as membrane epaississement, microaneurrisms, and increased permeability. This leads to reduced blood flow to the retina, causing ischemia and a decrease in nutrients. This leads to the formation of new blood vessels in the retina and other parts of the eye.

1.5.3 Classification of DR

Diabetic retinopathy falls into two main classes:

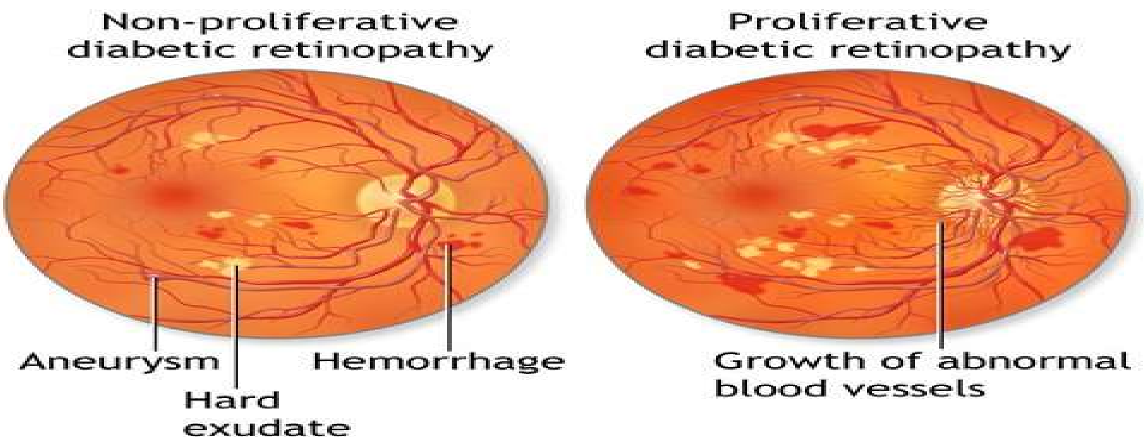


Fig.1.3: Non and proliferative diabetic retinopathy retina [17].

1.5.3.1 Non-proliferative diabetic retinopathy NPDR

Non-proliferative diabetic retinopathy (NPDR) is the most common form of diabetic retinopathy. Early stages consist of edema, lipid that has leaked from abnormal blood vessels, in the central retina, resulting in blurred central vision.

NPDR is further classified into mild, moderate, and severe stages based on the severity of the damage to the blood vessels.

- **Mild NPDR:** This is the earliest stage of diabetic retinopathy, characterized by tiny areas of swelling in the blood vessels of the retina.

- **Severe NPDR:** A larger section of blood vessels in the retina become blocked, causing a significant decrease in blood flow to this area. At this point, the body receives signals to start growing new blood vessels in the retina.
- **Moderate NPDR**: Increased swelling of tiny blood vessels starts to interfere with blood flow to the retina, preventing proper nourishment. This causes an accumulation of blood and other fluids in the macula.

1.5.3.2 proliferative diabetic retinopathy PDR

This is an advanced stage of the disease, in which new blood vessels form in the retina. Since these blood vessels are often fragile, there's a higher risk of fluid leakage. This triggers different vision problems such as blurriness, reduced field of vision, and even blindness [13].

Which is further classified into two stages:

- **Early PDR:** new blood vessels growing in the retina.
- **High-risk PDR:** In addition to the new blood vessels, scar tissues will also be formed on the retina, which will cause the retina to dissolve from the eyes, resulting in serious loss of vision or blindness.

1.5.4 Causes of DR

1.5.4.1 Early diabetic retinopathy

New blood vessels aren't growing (proliferating), the walls of the blood vessels in the retina weaken. Tiny bulges protrude from the walls of the smaller vessels, sometimes leaking fluid and blood into the retina. Larger retinal vessels can begin to dilate and become irregular in diameter as well. NPDR can progress from mild to severe as more blood vessels become blocked. Sometimes retinal blood vessel damage leads to a buildup of fluid (edema) in the center portion (macula) of the retina. If macular edema decreases vision, treatment is required to prevent permanent vision loss [7].

1.5.4.2 Advanced diabetic retinopathy

Diabetic retinopathy can progress to this more severe type, known as proliferative diabetic retinopathy. In this type, damaged blood vessels close off, causing the growth of new, abnormal blood vessels in the retina.

Eventually, scar tissue from the growth of new blood vessels can cause the retina to detach from the back of eye. If the new blood vessels interfere with the normal flow of fluid out of the eye, pressure can build in the eyeball. This buildup can damage the nerve that carries images from eye to the brain (optic nerve), resulting in glaucoma [7].

1.5.5 Symptoms of Diabetic Retinopathy

These symptoms could signify advanced stage diabetic retinopathy or a vision- threatening emergency:

- Mild blurriness or blurred vision.
- Periodic vision changes from blurry to clear.
- Floaters, including spots, dots, and cobweb-like strings in your field of vision.
- Blank or dark areas in your field of vision.
- Poor or declining night vision.
- Color changes, such as colors appearing washed out, different, or abnormal.
- Sudden or any loss of vision.
- Any changes in vision [8].

1.5.6 Method of detection

1.5.6.1 Using Medicine

The most common tests include the following:

- **Color Fundus Photography**

  Fundus photography is a reproducible technique for detecting diabetic retinopathy and has been used in large clinical research.

Fig.1.4: Example of Color Fundus Photography [85].

- **Optical Coherence Tomography OCT**

  Optical coherence tomography provides high-resolution imaging of the vitreoretinal interface, neurosensory retina, and subretinal space. OCT can be used to quantify retinal thickness, identify vitreomacular traction

Fig.1.5: Example of OCT [86]

- **Fluorescein Angiography FA**

  Fluorescein angiography may detect areas of untreated retinal capillary nonperfusion that could explain persistent retinal or disc neovascularization after previous scatter laser surgery

- **Ultrasonography**

  Ultrasonography is an extremely valuable diagnostic tool that enables assessment of the status of the retina in the presence of a vitreous hemorrhage or other media opacity [9].

How is diabetic retinopathy treated?

Treatment for diabetic retinopathy will depend on your symptoms, age and general health. It will also depend on how severe the condition is.

People with advanced retinopathy have a good chance of keeping their vision if they are treated before the retina becomes severely damaged. Treatment for diabetic retinopathy may include one or a combination of the following:

- **Laser surgery**: This is often used to treat proliferative retinopathy and sometimes macular edema. It involves shrinking the abnormal blood vessels, or sealing the leaking ones.
- **Vitrectomy**: Vitrectomy is a procedure that involves removing the jelly-like substance (vitreous) that fills the center of the eye. The vitreous is replaced with a balanced saline solution.
- **Injections**: Certain medications can be injected into the eye to slow the growth of the abnormal vessels of the retina and to treat macular edema [84].

An eye examination plays a vital role in identifying and treating eye conditions. It can also serve as a valuable tool to alert healthcare providers if diabetes management needs improvement. However, to enhance early diagnosis and promote telemedicine, the integration of artificial intelligence (AI) is necessary. AI can enable quicker identification of eye-related issues and facilitate remote medical consultations.

1.5.6.2 Using artificial intelligence

The main reason behind choosing this topic for this study is that we live in a developing country where there is always shortage of resources to overcome any problem. Our country has a population of 42 million and around 2.8 million people [10], have diabetes, from where approximately 47% of them have diabetic retinopathy. In developing countries, people usually treat themselves with a whole-body checkup for diabetes but they do not treat their eyes because they have a little knowledge about the fact that diabetes can affect their eyes. Therefore, we want to raise an awareness that diabetic retinopathy can be very dangerous if it is not treated properly in its early stages. With the help of image processing, we want to help our medical facilities so that the detection process becomes easier and none of the patient goes blind because of diabetic retinopathy. Therefore, using deep learning we are going to train our dataset to give the best possible outcome to the ophthalmologist [11]. Our AI extracted the number of DR hemorrhages as the bleed count and the total number of pixels for DR hemorrhages as the bleed area. These data were used to compare the accuracy of the classification method with mild-or-worse non-proliferative DR (NPDR) and moderate-or-worse NPDR [12]

So, the ophthalmologists they can worry less about the detection of diabetic retinopathy and focus more on the proper treatment of the patients. Since everything around us is getting digitalized, we wanted to incorporate the medical sector with computer science and this will bring better changes in people's life [11].

1.6 Conclusion

This chapter presents the definition of diabetes and its types, the definition of a human eye and its anatomy, the affection of diabetes on a human's eye (diabetic retinopathy). Explained what diabetic retinopathy is, its symptom, treatment, classification and last detection which are divided into lastly two methods using medicines and artificial intelligence. Hence, we propose in the next chapter some models of deep learning for detecting diabetic retinopathy.

Chapter.2

Methods used for retinopathy detection

2.1 Introduction

In this chapter, we will define artificial intelligence and the concept of machine learning, before moving on to deep learning, with the presentation of neural network architectures and the various models used to detect diabetic retinopathy, such as VGG16, MobileNetV1, and EfficientNetB0.

2.2 Artificial Intelligence (AI)

Artificial intelligence (AI) is the simulation of human intelligence processes by machines that are able of reasoning, learning, planning and creativity [18].
AI systems work by ingesting large amounts of labeled training data, analyzing the data for correlations and patterns, and using these patterns to make predictions about future states. In this way, a chatbot that is fed examples of text can learn to generate lifelike exchanges with people, or an image recognition tool can learn to identify and describe objects in images by reviewing millions of examples [19].

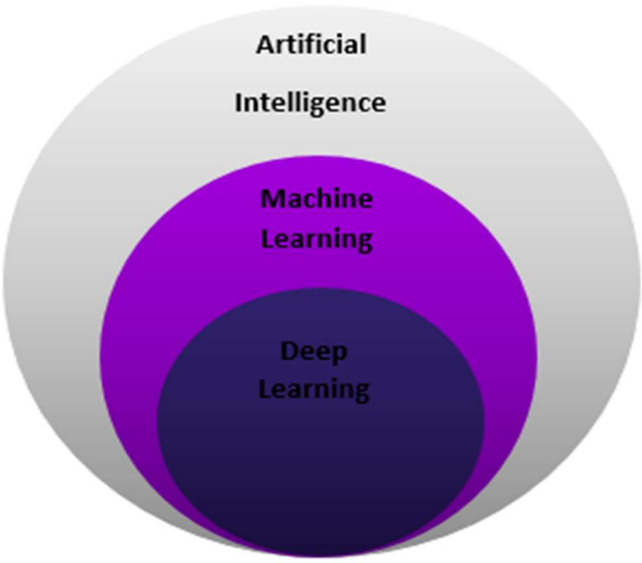


Fig. 2.1: The concept of artificial intelligence.

2.3 Machine Learning (ML)

2.3.1 Definition

Machine learning is a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior. It helps the artificial intelligence systems to perform complex tasks in a way that is similar to how humans solve problems. Machine learning gathers data like numbers, photos, or text and prepares it to be used as training by supplying the data, it lets the computer model train itself to find patterns or make predictions the more data, the better the program [20].

2.3.2 Subcategories of machine learning

There are three subcategories of machine learning:

Fig 2.2: The subcategories of machine learning.

2.3.2.1 Supervised learning

The model is supervised while it's learning, and trained with labeled data sets, which allow the models to learn and grow more accurate over time. Supervised machine learning is the most common type used today, it is effective for a variety of business purposes, including sales forecasting, inventory optimization, and fraud detection.

Some examples:

- Predicting real estate prices
- Classifying whether bank transactions are fraudulent or not
- Finding disease risk factors [20] [21].

2.3.2.2 Unsupervised learning

Unsupervised machine learning looks for less obvious patterns in unlabeled data. It can find patterns or trends that people aren't explicitly looking for. This machine learning type is very helpful when you need to identify patterns and use data to make decisions, which is widely used to create common applications.

Some examples:

- Creating customer groups based on purchase behavior
- Grouping inventory according to sales and/or manufacturing metrics [20] [21].

2.3.2.3 Reinforcement machine learning

Reinforcement learning is the closest machine learning type to how humans learn, which is through trial and error that allows it to take the best action by establishing a reward system, and by telling the machine when it made the right decisions, which helps it learn over time what actions it should take.

Some examples:

- Teaching cars to park themselves and drive autonomously
- Dynamically controlling traffic lights to reduce traffic jams [20] [21].

2.3.3 The challenges faced during the machine learning process

Difficulties can arise during the training of a machine learning system, either due to the dataset used for training or the learning system itself.

The potential difficulties may include:

- **Lack of Training Data**

Lack of training data is the most common yet fixable machine learning issue, in general, machine learning models need a huge amount of data information and examples representing exactly what we want them to do in order to achieve good results [22].

- **Poor Quality of Data**

The number one problem facing Machine Learning is the lack of good data. It is essential to have a good dataset for the learning system to be trained successfully, fully functional, and effective [23]. Data is of poor quality when the dataset contains errors, missing values, and noise. That's why data quality tools are designed to remove formatting errors, typos, redundancies, missing entries, and other issues that can reduce the quality of data [24].

- **Under fitting of Training Data**

Under fitting in machine learning model refers to a model that can neither model the training data nor generalize the new data, it's a model that fails to sufficiently learn the problem and performs poorly on a training dataset and does not perform well on a holdout sample. However, under fitting provides a useful contrast to the problem of over fitting [25].

- **Over fitting of Training Data**

  Over fitting in machine learning refers to a model that may perform very well on the training data but will not perform as well on new data, because it has essentially memorized the training data instead of learning the underlying patterns, it is a common problem in machine learning that can lead to poor model performance [25].

2.4 Deep Learning (DL)

2.4.1 Definition

Deep learning, also known as deep neural learning, is a fundamental technology within the field of machine learning. The term "deep" refers to its heavy reliance on artificial neural networks. It encompasses a range of algorithms that have the ability to mimic the functioning of the human brain by utilizing artificial neural networks. These networks consist of numerous interconnected layers, each comprising sets of information processing units, which represent neurons. Each layer receives and analyzes data from the preceding layer, enabling the progressive interpretation of information [26].

2.5 Structure of neural network

2.5.1 Architecture of neural network

Neural Networks are complex structures made of artificial neurons that can take in multiple inputs to produce a single output. This is the primary job of a Neural Network to transform input into a meaningful output. Usually, a Neural Network consists of an input and output layer with one or multiple hidden layers within. It is also known as Artificial Neural Network or ANN. ANN architecture in Neural Network functions just like a human brain.

In a Neural Network, all the neurons influence each other, in hence, they are all connected. The network can acknowledge and observe every aspect of the dataset at hand and how the different parts of data may or may not relate to each other. This is how Neural Networks are capable of finding extremely complex patterns in vast volumes of data [27].

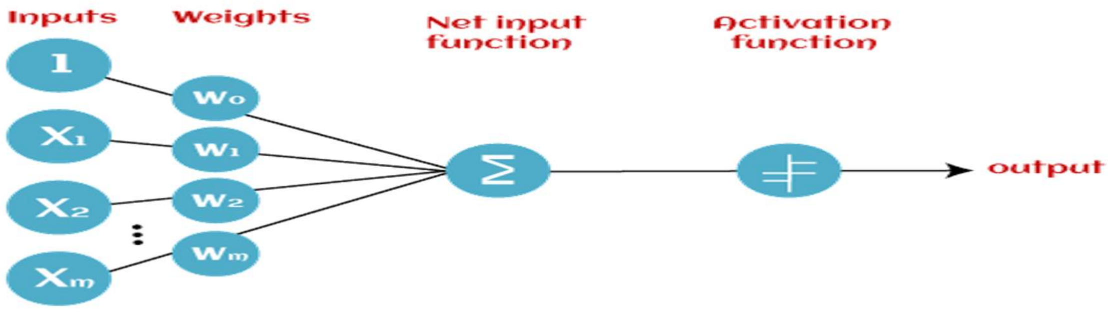


Fig.2.3: Example of architecture of neural network [28].

2.5.2 Types of neural networks

Neural networks can be classified into different types, which are used for different purposes. The most common types of neural networks are:

- **Feed forward neural networks, or Multi-Layer Perceptron (MLPs):**

They are comprised of an input layer, a hidden layer or layers, and an output layer. Data usually is fed into these models to train them, and they are the foundation for computer vision, natural language processing, and other neural networks.

- **Recurrent Neural Networks (RNNs):**

These learning algorithms are primarily leveraged when using time-series data to make predictions about future outcomes, such as stock market predictions or sales forecasting.

- **Convolutional Neural Networks (CNNs):**

Are similar to feed forward networks, but they're usually utilized for image recognition, pattern recognition, and/or computer vision. These networks harness principles from linear algebra, particularly matrix multiplication, to identify patterns within an image.

2.5.3 Functions used in neural network

- **Activation function**

Activation function is a function used in artificial neural network which outputs a small value for small inputs, and a larger value if its inputs exceed a threshold. If the inputs are large enough, the activation function "fires", otherwise it does nothing. In other words, an activation function is like a gate that checks that an incoming value is greater than a critical number**.** There are several types of activation functions: Sigmoïde, ReLu and Softmax [29].

- **loss function**

Loss functions are used to determine the error "the loss" between the output of our algorithms and the given target value. Loss functions are used in optimization problems with the goal of minimizing the loss [30].

- **Gradian decent**

Gradient Descent is known as one of the most commonly used optimization algorithms which is widely used in linear regression and classification to train machine learning models by means of minimizing errors between actual and expected results. Further, gradient descent is also used to train Neural Networks [31].

- **ADAM**

Adam optimizer is one of the most popular and famous gradient decent optimization algorithms. The algorithm is a cross between SGD and momentum. It inherits the benefits of both approaches, such as changing the learning rate using estimates of the gradient's first and second moments [32][33].

- **Back propagation**

Backpropagation algorithm trains neural networks by calculating the gradient of the loss function with respect to the network weights. The weights of the network are then updated using this gradient to reduce the loss function. By propagating the mistake backwards from the output layer to the input layer and changing the weights at each layer along the way, the backpropagation method operates. Until the network converges to a set of weights that yields an acceptable degree of performance on the training data, this process is repeated iteratively.

2.6 Neural Networks vs Deep Learning

Deep Learning and neural networks tend to be used interchangeably, the “deep” in deep learning is referring to the depth of layers in a neural network. A neural network that consists of more than three layers, which would be inclusive of the inputs and the outputs, can be considered a deep learning algorithm. A neural network that only has two or three layers is just a basic neural network.

2.7 Convolutional neural networks (CNN)

A CNN is a multilayer neural network that was biologically inspired by the animal visual cortex. The architecture is particularly useful in image-processing applications. As a deep network, early layers recognize features (such as edges), and later layers recombine these features into higher-level attributes of the input.

The CNN architecture is made up of several layers that implement feature extraction and then classification (figure 2.4). The image is divided into receptive fields that feed into a convolutional layer, which then extracts features from the input image. The next step is pooling, which reduces the dimensionality of the extracted features (through down-sampling) while retaining the most important information (typically, through max pooling). Another convolution and pooling step are then performed that feeds into a fully connected multilayer perceptron. The final output layer of this network is a set of nodes that identify features of the image (in this case, a node per identified number). To train the network we use back-propagation [34].

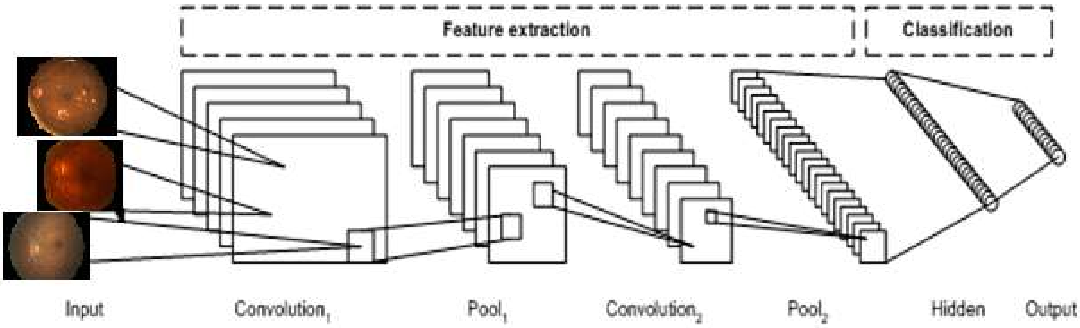


Fig.2.4: Example architecture of CNN.

The use of deep layers of processing, convolutions, pooling, and a fully connected classification layer opened the door to various new applications of deep learning neural networks. In addition to image processing, the CNN has been successfully applied to video recognition and various tasks within natural language processing.

Example of use:

- Image recognition
- Video analysis
- Natural language processing.

2.7.1 Models of CNN

2.7.1.1 VGG16

The Visual Geometry Group has proposed the CNN architecture known as VGG16 by A. Zisserman and K. Simonyan in 2014[28]. The VGG16 architecture was successfully applied to the ImageNet dataset of 14 million pictures and 1000 classes, with a top-5 test accuracy of 92.7%. The structure of the VGG16 model is depicted in the Figure (2.5) below, which contains an input layer, a sequence of convolutional and pooling layers, and an output layer [35].

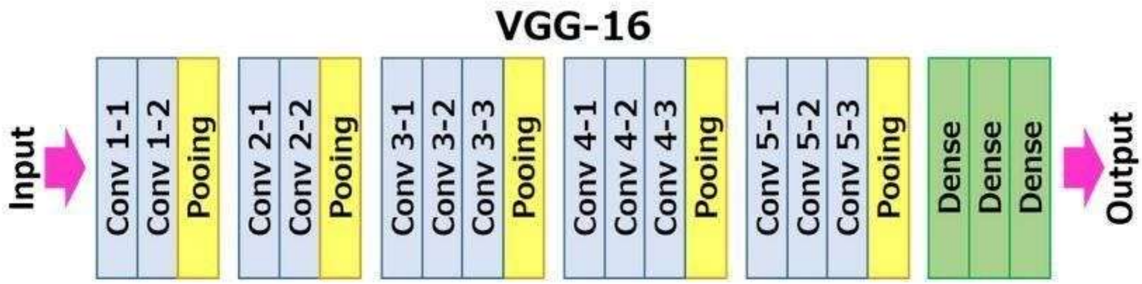


Fig.2.5: The architecture of VGG16[36].

The input of VGG is set to an RGB image of 224x244 size having dimensions set as (224, 224, 3). The average RGB value is calculated for all images on the training set image, and then the image is input as an input to the VGG convolution network [37]. It then passes through two convolutional layers where 3X3 filters are used. It followed by pooling layer that uses max pooling with stride of 2 × 2. It again passed through a set of two convolutional layers followed by one pooling layer. It then passes through the sequence of convolutional layers and pooling layers. At the end, there are three fully connected layers available that follow a stack of convolutional layers. These layers are having different architecture than its predecessor layers. The first two dense layers consists 4096 channels each. The last dense layer is a softmax layer having 1000 channels as it predicts from 1000 classes available in the ImageNet dataset [38].

- **Input**: VGGNet receives a 224×224 image input. In the ImageNet competition, the model's creators kept the image input size constant by cropping a 224×224 section from the center of each image.
- **Convolutional layers**: The convolutional filters of VGG use the smallest possible receptive field of 3×3. VGG also uses a 1×1 convolution filter as the input's linear transformation.
- **Hidden layers**: All the VGG network's hidden layers use ReLu instead of local response normalization. The latter increases training time and memory consumption with little improvement to overall accuracy.
- **Pooling layers**: A pooling layer follows several convolutional layers this helps reduce the dimensionality and the number of parameters of the feature maps created by each convolution step. Pooling is crucial given the rapid growth of the number of available filters from 64 to 128, 256, and eventually 512 in the final layers.

- **Fully connected layers**: VGGNet includes three fully connected layers. The first two layers each have 4096 channels, and the third layer has 1000 channels, one for every class [39].

2.7.1.2 MobileNetV1

MobileNet network is known as MobileNetV1. The main goal in creating this network is to keep costs and complexity low, making it easy to use for detection and classification on mobile devices or other devices with constrained resources like memory and energy. Low energy consumption models can benefit medical devices and resource-constrained developing nations. It demonstrates potency in extracting features, segmentation, and object detection [40].

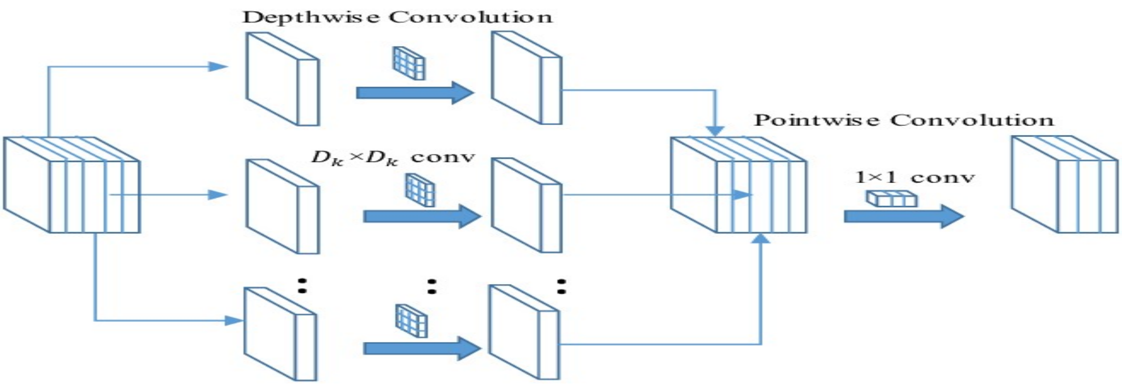


Fig.2.6: The architecture of Depth-wise Separable Convolution [41].

In MobileNet, a single filter is applied to each input channel. The outputs are combined by applying a 1×1 convolution using the pointwise convolution. A conventional executes filters and combines inputs into a new set of outputs in a single step. Two layers are employed for this, which are separated by depthwise separable convolution. As a result, there is a distinct layer available for filtering and a separate layer available for combining. This has had a significant influence on size and computation reduction. MobileNet includes 28 layers, including different levels for depthwise and pointwise convolutions [42].

- **Depth Wise Separable Convolutions**

It has two major components. The first one is Depth-wise convolution and the second one is point-wise convolution. We will look into each of them one by one:

- **Depth-wise convolution:**

In this type, convolution is applied to a single channel at a time and not like the simple convolutions in which it is done for all the channels together.

- **Point-wise convolution:**

In point-wise operation, a 1×1 convolution operation is applied on the M channels [43].

2.7.1.3 Training of MobileNetV1 and VGG16

To prevent overlearning and boost model performance, sophisticated models like MobileNetV1 and VGG16 require a lot of data when being trained. Building such a dataset, however, can be challenging in some supervised tasks, which can restrict the model's performance. The MobileNetV1 and VGG16 models have continued to employ this method:

- **Fine tunnig**

The pre-trained MobileNet and VGG16 model's final few layers are unfrozen for fine-tuning and taught at a slower learning rate. The pre-learned weights are kept by freezing the previous layers, while the last layers are trained on the fresh dataset. This approach enables the pre-trained model to enhance its overall performance by learning more precise features relevant to the current dataset. Unfrozen and trained with a lower learning rate are the model's final layers. The original pre-trained weights are still present in the model's previous layers, which are still

frozen. This enables the model to maintain the information it acquired from the first dataset while fine-tuning and improving its performance on the current dataset [44][45].

2.7.1.4 EfficientNetB0

EfficientNetB0 is a convolutional neural network trained on over one million photos from the ImageNet collection. Also, delivers a model network trained on the ImageNet dataset [46].

It is intended to be more accurate while requiring fewer parameters and more computationally efficient than other cutting-edge models. EfficientNetB0 optimizes network architecture by combining strategies such as effective scaling of network depth, width, and resolution, as well as the use of a novel compound scaling method. The model is built on the MobileNetV2 architecture's backbone and uses a combination of depth-wise-separable and standard convolutions to improve accuracy while reducing computational cost [47].

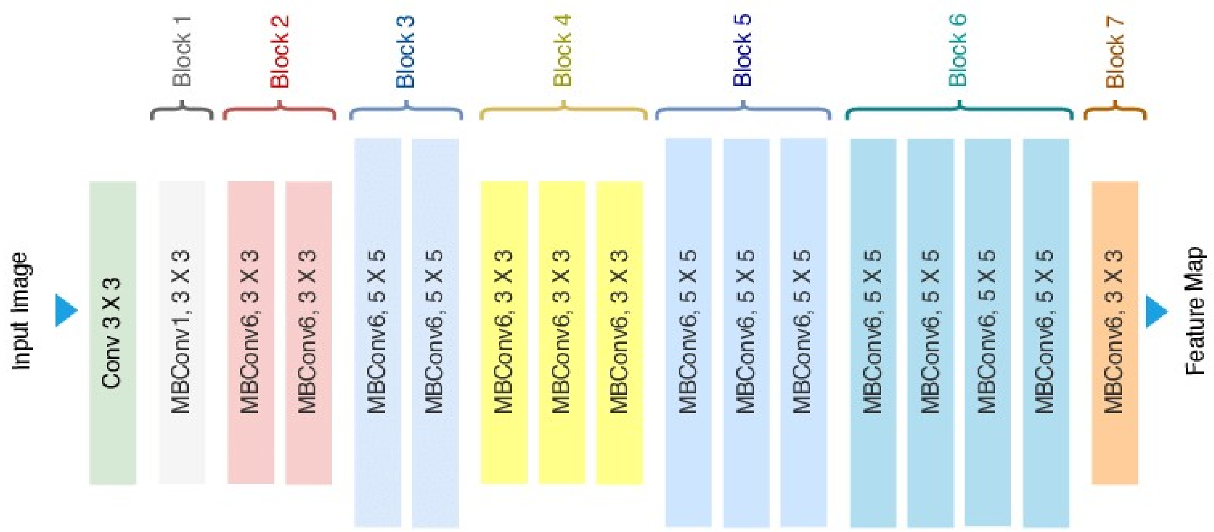


Fig.2.7: Architecture of EfficientNetb0 [48].

- **MBConv:**

In order to make the information flows in deep neural networks, we usually use residual blocks. Residual blocks connect the beginning and the ending of a convolutional block with skip connections. The channels start being wide in the beginning of the convolutional block, then they get narrower along the depth of the block till the end where they become wide again because of the added information.

(a) residual block :

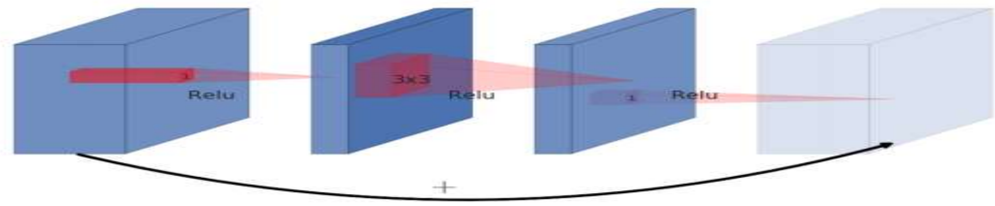


Fig 2.8: Structure of a normal residual block [51].

(b) inverted residual block:

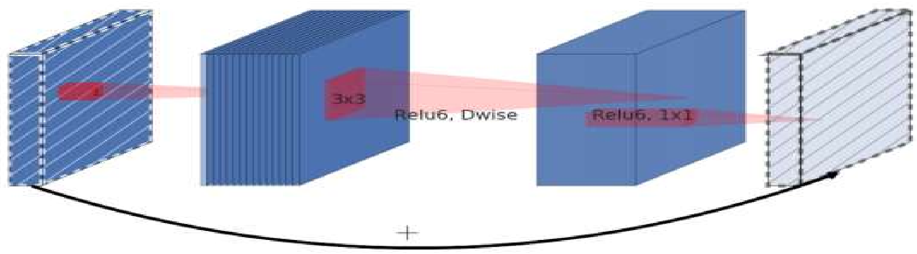


Fig.2.9: Structure of an inverted residual block (MBconv) [51].

- **Squeeze and Excitation (SE) Block:**

The Squeeze-and-Excitation Block is an architectural unit designed to improve the representational power of a network by enabling it to perform dynamic channel-wise feature recalibration. The process is:

- The block has a convolutional block as an input.
- Each channel is "squeezed" into a single numeric value using average pooling.
- A dense layer followed by a ReLU adds non-linearity and output channel complexity is reduced by a ratio.
- Another dense layer followed by a sigmoid gives each channel a smooth gating function.
- Finally, we weight each feature map of the convolutional block based on the side network; the "excitation" [50].

2.7.1.5 Transfer Learning with EfficientNetB0

Recently, transfer learning has been successfully used in various field applications, such as manufacturing, medical and baggage screening. This removes the requirement of having large dataset and also reduces the long training period, as is required by the deep learning algorithm when developed from scratch [49].

To train the model on our dataset, replace the last classification layer with a new layer that matches the number of classes in our dataset and freeze the weights of all levels except the last classification layer.

Overall, transfer learning using EfficientNetB0 can assist us in achieving cutting-edge performance on a variety of computer vision tasks while using fewer data and processing resources.

2.8 Conclusion

In this chapter, we initiated our study by introducing the concept of artificial intelligence, machine learning and its subcategories that exist alongside it, and the challenges faced during the learning process.

Then, in the second part, we presented deep learning and its neural network models with their architectures and models (VGG16, MobileNetV1 and EfficientNetB0). In the next chapter, we will present our study on diabetic retinopathy detection using various deep learning models, along with a comparison of the results obtained from different architectures for each model.

Chapter.3

Conception and implementation

3.1 Introduction

We will present in this chapter the solution that we propose to address the issue of early detection of diabetic retinopathy, where we initiate with the theoretical part and then move on to the technical part, where we will introduce the CNN models used, the system's conception, the environment, the libraries, experimentations and results evaluations. We will finalize with the presentation of the application of the detection system.

3.2 Diabetic Retinopathy Detection and Classification Using VGG-16, MobileNetV1 and EfficientNetB0 Pre-trained Networks

3.2.1 Definition

We conducted a research study aimed at analyzing a publicly available retinal fundus images dataset on Kaggle by utilizing pre-trained Convolutional Neural Network (CNN) models which are VGG16, MobileNetV1 and EfficientNetB0. Our process included applying both pre-processing techniques to improve detection accuracy features. We collected a sample size of 35,126 retinal funds images with five class labels for experimentation purposes using transfer learning and fine-tuning existing networks including pre-trained VGG16, MobileNetV1 and EfficientNetB0.

3.2.2 Problem and Constraints

Our study revolves around a medical diagnostic aid system meant for detecting symptoms linked with diabetic retinopathy pathology. With our project, we aspire to accommodate all stages of this condition from Normal (NO DR) to Terminal stage (Proliferate DR). Each of these stages correspond with unique classes within the pathology.

- **Reliability:** The system should be dependable and consistently deliver accurate results.
- **Cost-effective high performance:** It should achieve efficient and speedy execution while keeping costs low.
- **Lifelong learning capability:** Allowing the system to continuously acquire new knowledge and improve over time.

The goal here is to have a system that achieves the best performance rates according to the specified constraints (limited computing capacity and storage).

3.3 Conception of the system

We plan that our system will consist of a unit that will test the input provided by the user using our pre-trained deep learning model. This model is trained on a dataset that utilizes layers to detect various visual patterns and structures, starting from low-level features like edges and textures and progressing to higher-level features for learning purposes.

When the user selects a test input, the system loads the chosen model and transforms the input into to preprocessed retinal image. Finally, the model tries to predict the results based on the learned features. We compare the input with the entire dataset to find the most similar images to the users. The figure below provides an overview of this architecture.

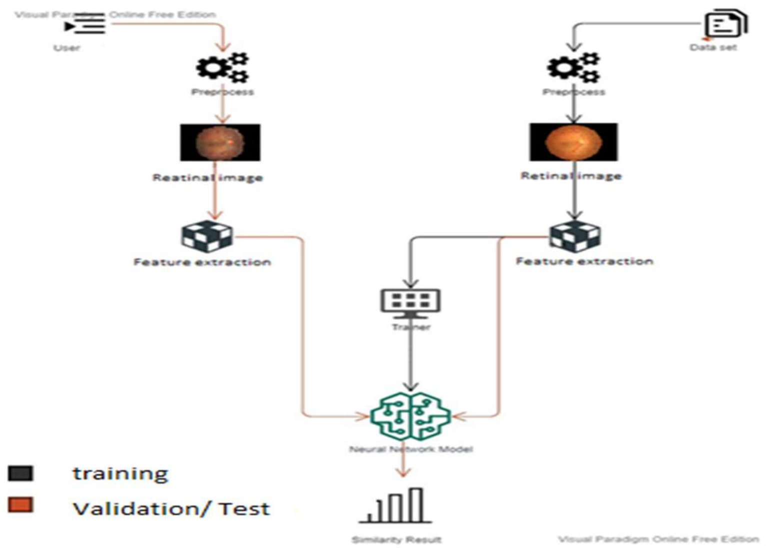


Fig.3.1: The architecture of our unit.

3.4 Implementation of the system

Moving on from system design (conception) to its implementation phase, we commence with dataset preprocessing followed by executing multiple deep learning models. Finally, we showcase our developed application, along with citing all the relevant tools utilized in this project.

3.4.1 Description and Architecture of the datasets

1st Dataset

The dataset used for this project was provided by Kaggle, named Diabetic retinopathy 2015 data colored resized, the dataset consists of 35,126 (left and right) retina scan images of medium quality.

The original dataset is available at Diabetic Retinopathy Detection [58] on Kaggle**, t**hese images are resized into 224x224 pixels, so that they can be readily used with many pre-trained deep learning models, it is accompanied by a CSV file containing details about the data, such as the stage of the pathology and eye information.

All of the images are saved into their respective folders according to the stage of diabetic retinopathy, the fundus images in the dataset are classified into 5 categories (classes), which are presented in fig.3.1 and fig.3.2 bellow, where each class represents a stage of the disease.

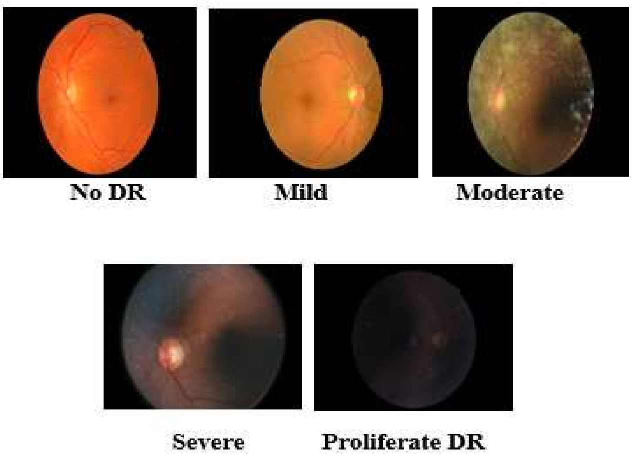


Fig.3.2: Retinal images for diabetic retinopathy with disease grades [58].

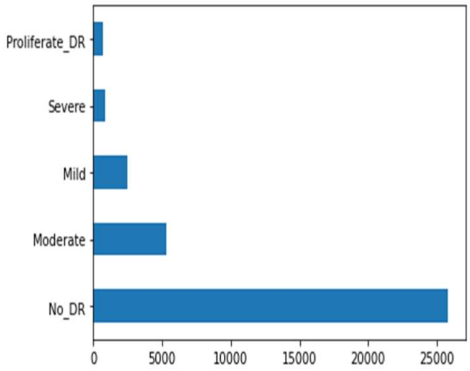


Fig.3.3: The repartition of 1st dataset images.

2nd Dataset

The dataset consists of Gaussian filtered retina scan images that are used to detect diabetic retinopathy. The original dataset is available at_APTOS 2019 Blindness Detection [57] on Kaggle, the dataset consists of 3662 Gaussian filtered retina scan images of medium quality. To ensure compatibility with various pre-trained deep learning models, the images have been resized to 224x224 pixels.

The dataset is organized into different folders in fig.3.4 and fig.3.5 bellow, each corresponding to the severity or stage of diabetic retinopathy. The organization of the folders is based on the information provided in the train.csv file.

The Gaussian filtering applied to the retina scan images helps enhance specific features and patterns related to diabetic retinopathy.

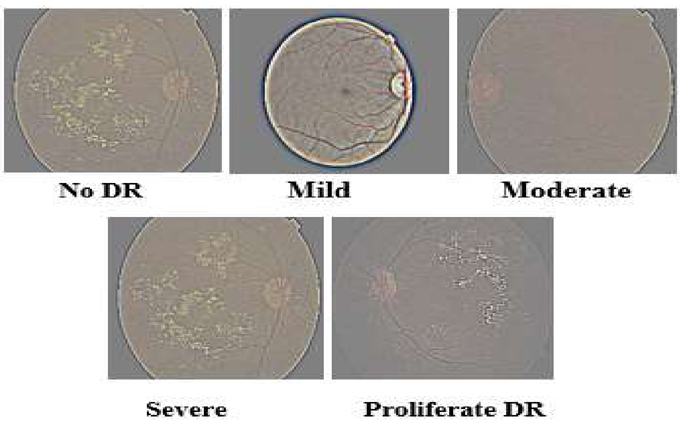


Fig.3.4: Retinal gaussian filtered images for diabetic with disease grades [57].

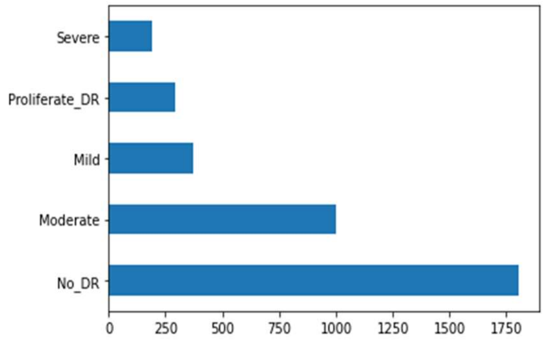


Fig.3.5: The repartition of the 2nd dataset.

3.4.2 Preprocessing on Datasets

1st Dataset

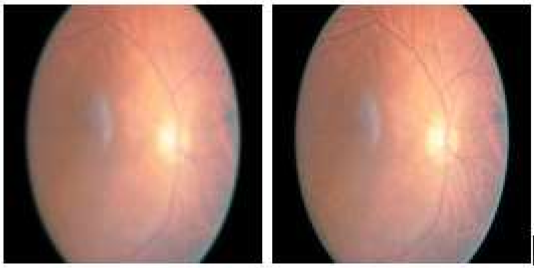
For our first-hand experience we tried to enhance the detection and understanding of diabetic retinopathy, we followed a meticulous plan. We curated a diverse dataset with meticulous annotations representing various stages of the disease. Initially exploring gaussian filtering, we later decided to work with the original RGB format of the images to preserve their visual information. We then normalized the pixel values to ensure consistency and split the dataset into training, validation and testing sets. Through feature normalization and pre-trained models, we extracted relevant features for accurate detection and classification of diabetic retinopathy.

- **Data Acquisition**

After a successful acquisition process, we have obtained a dataset that meets our requirements. It includes a diverse collection of retinal images with meticulous annotations for diabetic retinopathy. Our careful selection process ensured the inclusion of various disease stages. This enables our models to effectively learn and identify patterns and indicators of diabetic retinopathy, leading to improved detection and understanding of the disease.

- **Gaussian Filtering**

Initially, we apply gaussian filtering which is a common image processing technique that involves convolving the image with a gaussian kernel, to the retinal images, exploring different sigma values that determine the amount of smoothing applied to the image, with higher sigma values resulting in more blurring, and kernel sizes that refer to the dimensions of the gaussian kernel matrix, which determines the neighborhood of pixels considered during the convolution operation [79]. The objective of this particular step was to enhance the quality of the data by smoothing the images and reducing noise. However, after evaluating the results, we determined that this approach did not yield satisfactory outcomes for our specific task, that's why we decided to remove the gaussian filtering step from our preprocessing. Instead, we found that working with the retinal images in their original RGB format produced better outcomes and preserved important visual information.



Filtetered image **RGB image**

Fig.3.6: Gaussian filter image and a RGB image.

- **RGB Image Format**

Upon discarding the Gaussian filtering step, we delve into the retinal images in their RGB format. This format encapsulates the essence of an image through the three fundamental colors: Red (R), Green (G), and Blue (B). Each pixel in the image embodies three distinct color channels, one for red intensity, one for green intensity, and one for blue intensity. The interplay of these channels enables an expansive range of colors to be depicted, thereby retaining the unprocessed visual data embedded in the images [80].

- **Normalization**

After confirming the use of RGB images, we proceed with normalizing the pixel values. Normalization is a crucial preprocessing step that involves scaling the pixel values of the retinal images to a standardized range [81]**.** It ensures that the pixel intensities are transformed to a common scale, in our case we scaled it to [0, 1]. The purpose of normalization is to bring the pixel values of different images into a comparable range, regardless of variations in their original intensities**,** this step aids in achieving better convergence during model training and facilitates consistent data processing, achieving several benefits during model training and evaluation which lead to improved performance and accurate diabetic retinopathy detection and classification.

- **Train-Test-validation Split**

Split the preprocessed dataset into training, validation and testing sets. The training comprising 70% of the data set, which is used to train the VGG16 and MobileNetV1 models. While the validation and testing sets accounting for 15% of the data, used to evaluate the models' performance on unseen data and have the ability to make predictions.

- **Feature Extraction**

Use the pre-trained VGG16 and MobileNetV1 models to extract features from the preprocessed retinal images by using the output of the last convolutional layer as the extracted features.

For our second-hand experience, we decided to explore the application of binarization on the same RGB dataset. Binarization involves converting the dataset into a binary format by assigning binary labels or values to the data instances, this process aimed to categorize retinal images into two classes, typically representing the presence or absence of the disease (DR) and (No DR) the goal was to simplify the classification task and to simply categorize retinal images as either having (DR) not having the disease (No DR). The outcome of the binarization experiment produced mixed results. On one hand, it provided a simplified and straightforward way to categorize the images.

2nd dataset

In our experimental analysis, we utilized the APTOS 2019 Blindness Detection dataset, which is pre-trained. The dataset has undergone pre-processing, including the application of a Gaussian filter. Importantly, we have preserved the dataset in its original, unaltered form, ensuring its immediate availability and suitability for our study.

3.4.3 Hyperparameter optimization

Hyperparameter optimization is the process of searching for the best set of hyperparameter values for our model. Where hyperparameters are the configuration settings that influence the training process, such as learning rate, batch size, number of hidden layers.

The goal of hyperparameter optimization is to find the optimal combination of these settings that yields the best performance on our dataset. This process typically Contributes the model can achieve improved accuracy, convergence speed, and generalization ability [74].

- **Learning rate**

Learning rate is a hyperparameter that controls the speed of a model's convergence towards optimal weights during learning, it determines how fast or slow we will move towards the optimal weight, a lower rate results in smaller updates, potentially improving accuracy but increasing learning time. Selecting the right learning rate is crucial to avoid convergence issues or slow learning, which was done through experimentation [75].

- **Number of neurons and number of layers in hidden Layer**

The number of hidden layers in a neural network refers to the intermediate layers between the input and output layers. The number of hidden neurons is typically chosen to be between the size of the input layer and the size of the output layer. Adding more hidden layers can increase the network's ability to learn intricate patterns but may lead to overfitting. Determining the appropriate number of hidden layers depends on the problem complexity and available data, which was done through experimentation and model validation [76].

- **Batch size**

Batch size refers to the number of training examples processed together in a single iteration during model training. Larger batch sizes provide more stable and accurate gradient estimates but require more memory. Smaller batch sizes reduce memory requirements but can result in noisy gradients [77].

- **Dropout rate**

Dropout is a hyperparameter, determines the probability of dropping out a neuron and should be carefully selected through experimentation to strike the right balance between regularization and model performance, it is a regularization technique used in neural networks to prevent overfitting during training. By randomly dropping out neurons, dropout reduces overfitting and enhances the model's ability to generalize to unseen data. The dropout rate [78].

3.5 Deep learning Models Implementation and Deployment

In this section, we will present the tools and technologies used to develop our system.

3.5.1 VGG16 Model

The model consists of several sequential layers:

- **Convolutional layers:** There are 13 layers that extract critical features from the input image using various filters. Each layer captures different levels of abstraction.

- **Pooling layers:** There are 5 layers that reduce the spatial dimensions of the feature maps generated by the convolutional layers. Max pooling is used, selecting the highest value within each pooling window.
- **Additional convolutional layers:** There are 4 layers that reveal more complex features, uncovering hidden depths and providing further insights.
- **Batch normalization layers:** There are 4 layers that stabilize the activations of the previous layers, enhancing the training process and improving generalization.
- **Global average pooling layer:** This layer calculates the average value for each feature map, resulting in a fixed-length vector regardless of the input image size. It standardizes the output for universal interpretation.
- **Fully connected layers:** There are 2 layers that serve as the backbone of traditional neural networks. The first dense layer consists of 512 units, followed by a dropout layer that helps prevent overfitting. The final dense layer has 2 units, representing the output classes for binary classification.

Overall, this model employs a series of convolutional, pooling, batch normalization, and fully connected layers to extract and process features from input images, leading to binary classification outputs.

3.5.2 MobileNetV1 Model

The model's architecture is a complex web of layers:

- **Input layer:** Captures the variable-shaped image.

- **Convolutional layer:** Uses 2D convolution operation with filters to extract low-level features from the image.
- **Batch normalization layer:** Standardizes the activations of the previous layer, enhancing training efficiency.
- **ReLU activation layer:** Introduces non-linearity using the Rectified Linear Unit activation function.
- **Depthwise separable convolution layer:** Splits the convolution operation into a depthwise convolution and a pointwise convolution, reducing computational complexity.
- **Batch normalization layer:** Normalizes the activations of the depthwise separable convolution.
- **ReLU activation layer:** Applies ReLU activation to the output of the depthwise separable convolution.
- **Pointwise convolution layer:** Applies a 1x1 pointwise convolution to increase the number of channels in the feature maps.
- **Batch normalization layer:** Normalizes the activations of the pointwise convolution.
- **ReLU activation layer:** Applies ReLU activation to the output of the pointwise convolution.

The model continues with additional depthwise separable convolution and pointwise convolution layers, increasing the number of filters and channels. This pattern repeats to capture higher-level features.

The architecture also includes the following layers towards the end:

- **Global average pooling layer:** Computes the average value for each channel of the feature maps, resulting in a fixed-length vector.
- **Fully connected layer:** A traditional neural network layer with 512 units.
- **Dropout layer:** Randomly drops a fraction of the input units during training to prevent overfitting.
- **Final fully connected layer:** Has 2 units, representing the output classes for binary classification.

3.5.3 EfficientNetB0 Model

EfficientNet-B0 is a formidable tool that can deftly classify retinal images and identify telltale signs of the ailment, which stems from damage to the blood vessels in the retina. By training EfficientNet-B0 with a vast dataset of labeled retinal images, the model can acquire the ability to recognize patterns and features that portend diabetic retinopathy. The model can subsequently classify new, unseen retinal images and aid in diagnosing and monitoring the disease [61].

- **Input:** The model takes retinal images as input, which can be grayscale or RGB.
- **Convolutional Layers**: Multiple convolutional neural networks are used to extract relevant features from the input images, capturing patterns, edges, textures, and shapes associated with diabetic retinopathy.
- **Bottleneck Blocks:** EfficientNet-B0 includes bottleneck blocks that combine convolutional, pooling, and activation layers. These blocks reduce the dimensions of the feature maps while retaining important features. This compression helps optimize computational efficiency and memory usage.
- **Global Average Pooling:** After the bottleneck blocks, the model applies global average pooling. This operation condenses the feature maps into a single value per channel, summarizing the learned features across the entire image.
- **Fully Connected Layers:** The output of global average pooling is fed into fully connected layers. These layers learn to associate the extracted features with specific classes or categories, such as different stages of diabetic retinopathy.
- **Output:** The final layer of EfficientNet-B0 produces predicted probabilities for each class. In the case of diabetic retinopathy detection, the model's output indicates the probability of diabetic retinopathy presence and its severity level for a given retinal image.

EfficientNetB0's architecture, combining convolutional layers, bottleneck blocks, global average pooling, and fully connected layers, enables efficient and accurate detection of diabetic retinopathy from retinal images.

3.6 Evaluation, experimentation and testing

3.6.1 The environments used and libraries employed

3.6.1.1 Virtual environments and hardware environments

Virtual environment, we used:

- **Python language**

  We choose to engage with Python (v3.11) since it is the open-source programming language of choice for computer gurus. Python's attraction stems from its numerous features, which allow developers to focus on their art rather than the mechanics of it, Python has gained to prominence in infrastructure management, data analysis, and software development. It has improved processes and freed developers from the restraints of outdated, often slow languages. As a result, writing and developing in Python is well-known for its quick execution speed in compared to its competitors [59].
- **Visual Studio Code**

  Visual Studio Code (VS Code) is a free and lightweight, cross-platform source code editor developed by Microsoft. It has allowed and granted us to realize and create HTML and CSS web pages [60].
- **Google Colab**

  Google Colab, also known as Google Colaboratory, is an interactive cloud-based environment for developing and sharing interactive notebooks, enabling users to write, ecute, and collaborate on Python code without additional software installation**.** It has use of several libraries [63].
- **Kaggle Notebook**

  Kaggle Notebook is an online computational environment for data science and machine or deep learning, enabling code creation, sharing, and collaboration in Python, popular libraries and datasets [64].

Hardware environments, we utilized two the first one to execute the Python code on both Google Colab and Kaggle Notebook /

- RAM: 8 GB.
- GPU: GTX 1050 TI GPU.
- CPU: Intel(R) Core (TM) i5-7300Q CPU @2.50GHz.

The second environment was used to create the interface:

- RAM: 2 GB

- CPU: Intel(R) Pentium(R) CPU B950 @ 2.10GHz, 2100 MHz, 2 core(s), 2 logical processor(s).

3.6.1.2 The libraries employed

- **Tensorflow**

Tensorflow is a library developed by Google specialized in training and developing deep learning models [56]. Tensorflow offers an API that allows the implementation of very complex deep architectures in a simple and fast manner. We use Tensorflow to implement and test different configurations of the deep learning model [65].

Fig.3.7: Tensorflow [69].

- **Keras**

Keras is the most widely used framework. It is an open-source library built on Python that allows for easy and fast construction of neural networks and machine learning models, leveraging major frameworks such as Tensorflow and Pytorch. Keras was required for implementing the LSTM and GRU prediction models, as well as for data preprocessing [66].

Fig.3.8: Keras [70].

- **NumPy**

NumPy is an open-source Python library used for numerical computing. It provides a powerful multidimensional array object and a wide range of functions for manipulating arrays and performing mathematical operations [53]. NumPy is commonly used in scientific and data analysis applications due to its efficiency in handling large arrays of data. It offers optimized numerical operations, broadcasting capabilities, and integration with other libraries. NumPy was required in the preprocessing for feature Extraction and data analysis [68].

Fig.3.9: NumPy [71].

- **OpenCV2**

OpenCV2, or Open-Source Computer Vision Library version 2, is an open-source computer vision and machine learning software library [54]. It provides a wide range of functions and algorithms that can be used for tasks such as image and video processing, object detection and tracking, feature extraction. OpenCV2 was required for Image Preprocessing, Feature Extraction and Object Detection [67].

Fig.3.10: OpenCV [72].

- **Flask**

Flask is a lightweight web framework written in Python, known for its simplicity and flexibility [55]. It allows developers to build web applications and APIs with minimal code, it is a popular choice for developing web applications [83].

Fig.3.11: Flask [73].

3.6.2 The evaluation metrics

There are several established evaluation metrics in the literature that are based on the confusion matrix for each class, offering four important pieces of information in terms of recall and precision:

- **Confusion table**

Confusion table is a table used to assess the performance of a classification model. It is a matrix with rows and columns representing the actual and anticipated categories. Each cell in the matrix represents the number of occurrences where the actual and anticipated classes are the same. It shows the number of "True Positive (**TP**) ", "False Positive (**FP**)", "True Negative (**TN**) ", and "False Negative **(FN**) "as indicated below:

		Actual	
		True	False
Expected	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

Tab 3 .1: Structure of table confusion.

- **Accuracy**

A proportion of correctly predicted and classified to total observations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad \textbf{(3.1)}$$

- **Error rate (ER)**

The error rate is the fraction of incorrectly categorized cases in a classification model that represents the classification's overall error.

$$\text{ER} = 1\text{-Accuracy} \quad \textbf{(3.2)}$$

- **Recall (Sensitivity):**

Sensitivity or true positif rate it is the proportion of genuine positive predictions to total positive cases in the dataset.

$$\text{Recall} = \frac{TP}{FP+} \quad \textbf{(3.3)}$$

- **Precision**

It calculates the proportion of accurately anticipated positive events among all predicted positive events.

$$\text{Precision}=\frac{TP}{TP+F} \quad \textbf{(3.4)}$$

- **Test and interpretations**

In our approach, we divide the 1st dataset into 70% for training phase,15% for validation and 15% for testing phase. Additionally, we generate RGB retinal images using VGG16 and MobileNetV1.

Model	**VGG16**			**MobileNetV1**		
	Train	Val	Test	Train	Val	Test
K=1*1 S={0.25,0.5,1,1.5}	10%	9.77%	9.77%	10.18%	9.91%	9.91%
K=1*1 S=1.75	40%	39.84%	39.84%	40.20%	39.95%	39.95%
K=3*3 S={0.25,0.5,1,1.5}	60%	59.78%	59.78%	60.22%	59.80%	59.80%
K=5*5 S={0.25,0.5,1,1.5}	60%	59.78%	59.78%	60.22%	59.80%	59.80%
K=5*5 S={2,2.5,3}	10%	9.80%	9.80%	10.15%	9.93%	9.93%
K=7*7 S={0.25,0.5,1,1.5}	60%	59.79%	59.79%	60,22%	59.80%	59.80%
K=7*7 S={2.2,5.3.5}	10%	9.77%	9.77%	10.16%	9.90%	9.90%

Table 3.2: Test Results on the Dataset using VGG16 and MobileNetV1

Given that: **K**: Kernel; **S**: Sigma; **Ep**: Epochs.

In our experiment, in order to reduce noise, sought to examine how different sigma values (0, 0.25, 0.5, 1, and 1.5) affected a fundus picture. Exploring the connection between rising sigma values and classification accuracy was the main goal. Our research showed that the accuracy substantially dropped as the sigma values rose. This finding demonstrates a proportionate link between accuracy and sigma.

It is significant to remember that bigger smoothing effects, which lessen noise, are produced by higher sigma values. It could also result in the blurring of details and characteristics that are necessary for precise classification. This more pronounced smoothing effect may result in the loss of important data, which will ultimately reduce accuracy.

Additionally, we looked at how altering the kernel might affect accuracy within a single epoch and discovered a 1x1 kernel is limited to capturing highly specific local information, potentially resulting in the loss of essential information necessary for accurate data classification, and by utilizing kernels of 3x3, 5x5 and 7x7 sizes, the accuracy of the model increases to 60% and 60.22%, respectively for VGG16 and MobileNetV1. This is attributed to the capability of the larger kernels to capture broader and more comprehensive features of the image, to be provided in Table 3.3, more analysis of our model's performance with additional epochs is required.

Model	**VGG16**			**MobileNetV1**		
K=3*3 ;S=0.5 ;Ep=1	70%	70%	69.90%	71%	70.97%	70.10%
K=3*3 ;S=0.5 ;Ep=10	70%	70%	69.90%	71%	70.97%	70.10%
K=3*3;S=0.5 ; Ep=50	70%	70%	69.90%	71%	70.97%	70.10%
K=5*5 ;S=1 ;Ep=1	70%	70%	69.90%	71%	70.97%	70.10%
K=5*5 ;S=1 ;Ep=10	70%	70%	69.90%	71%	70.97%	70.10%
K=5*5 ;S=1 ;Ep=50	70%	70%	69.90%	71%	70.97%	70.10%

Table 3.3: Test Results on the first Dataset using VGG16 and MobileNetV1

According to the results' interpretation, a pretrained Convolutional Neural Network (CNN) may identify more accurately by using more training epochs throughout its training phase. For MobileNetV1 and VGG16, accuracy rose from 60% to 70% and 60.2% to 71% respectively, when the number of epochs was raised from a starting value. This indicates the beneficial effect of extensive training on the CNN models' capacity to accurately recognize DR.

Further increasing the number of epochs did not result in appreciable accuracy, this shows that the models attained a steady performance level and that further training did not result in appreciable accuracy gains after that.

Based on our observations and analysis of the table 3.4 We conclude that the VGG16 and MobileNetV1 gives a good accuracy but we want to enhance the accuracy using RGB images and add layers numbers of CNN also, change the learning rate in the table below:

Model	**VGG16**			**MobileNetV1**		
	Train	Val	Test	Train	Val	Test
Ep=1 Lr=1e -5	72.86%	73.37%	73.30%	73.10%	72.21%	72.50%
Ep=10 Lr=1e – 5	74.68%	74.11%	74.00%	75.10%	74.70%	74.11%
Ep=50 Lr=0.001	87.34%	73.56%	73.47%	88.20%	74.09%	74.00%
Ep=110 Lr=1e -5	95.21%	73.60%	73.49%	95.40%	73.50%	73.40%

Table 3.4: Test Results on the first Dataset using VGG16 and MobileNetV1.

According to the current study, using extra convolutional neural network (CNN) layers in RGB image processing results in more accuracy than using Gaussian filter-filtered pictures. This demonstrates how well CNN architectures work at immediate learning complex patterns and features from RGB input. After 110 training epochs, the performance of the MobileNet and VGG16 architectures were evaluated, and the results showed outstanding accuracy rates of 95.40% and 95.21%, respectively. This result demonstrates a positive relationship between the number of epochs and the models' performance, since more epochs allowed the models to go through more training iterations and enhance their accuracy in picture classification. Also, the changing of value of learning rate from 0.001 to 1e-5 improve the accuracy this suggests that a slower weight decrease and a more exact correction during training were beneficial to the model.

So, utilizing CNN architectures on RGB images and adding extra layers produced results that were more accurate than those obtained by using filtered images. The trials using MobileNetV1 and VGG16 showed that more epochs had a favorable effect on accuracy, enabling the models to continue learning and enhance their classification ability.

And in the 2nd dataset,we allocate 80% training and 20% validation using EfficientNetB0 to generate filtered Gaussian images with batch size =32.

Model	**EfficientNetB0**	
	Train	Val
Ep=1	67.59%	61.42%
Ep=10	85.23%	77.84%
Ep=40	97.54%	79.62%

Table 3.5: Test Results on the Dataset using VGG16 and MobileNetV1

The accuracy results for the EfficientNetB0 model for various numbers of epochs are shown in the supplied table 3.5. On the dataset (training and validation), the model starts at epoch 1 with a reasonable level of accuracy. All datasets see a considerable improvement in accuracy as the number of epochs rises, reaching high levels by epoch 40.

3.6.3 Final results and evaluations

For evaluate the models, we will base our evaluation on a set performance metrics using accuracy. During the evaluation phase, we have selected the best result obtained for each table from the previous test's results.

The neural networks we employed were infused with pretrained VGG16 and MobilenetV1 models for the first database. It's crucial we utilize RGB images and alter the learning rate value to 1e-5, which we'll tackle during the hyperparameter optimization phase. As for the second database, we opted for images filtered through a Gaussian process.

Result on the entire dataset: The following table presents the accuracy of each model on the training, validation and test datasets.

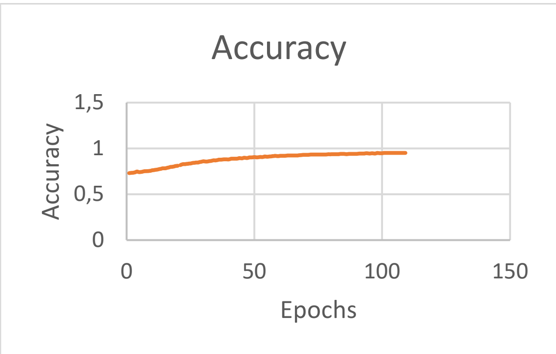
Model	Accuracy		
	Train	Val	Test
VGG16	95.21%	73.60%	73.50%
MobileNetV1	95.40%	73.71%	73.55%
EfficientNetB0	97.54%	80.74%	/

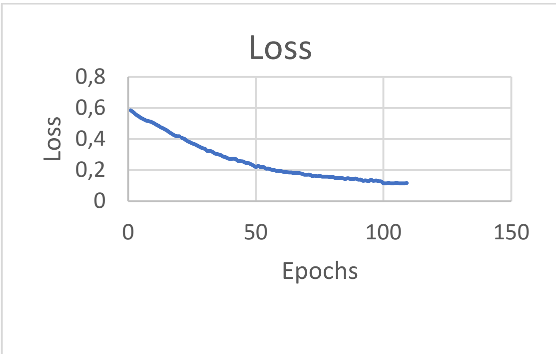
Table 3.6: Accuracy of the models.

According to the results in the table above, we observe that the Efficientnetb0 model outperformed VGG16 and Mobilenet with a staggering 97.54% accuracy. However, for the first database, Mobilenetv1 and VGG16 were neck and neck in terms of accuracy. But, Mobilenetb0 triumphed with a precision percentage of 95.40%.

3.6.4 Curves of Accuracy

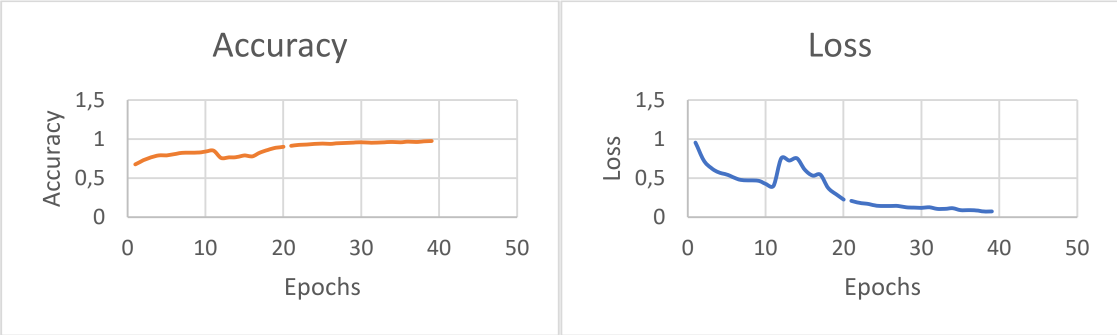
Bellow, we will find the accuracy curves with respect to loss (a) and accuracy curves with respect to epochs (b) for three models: VGG16, MobileNetV1 and EfficientNetB0.





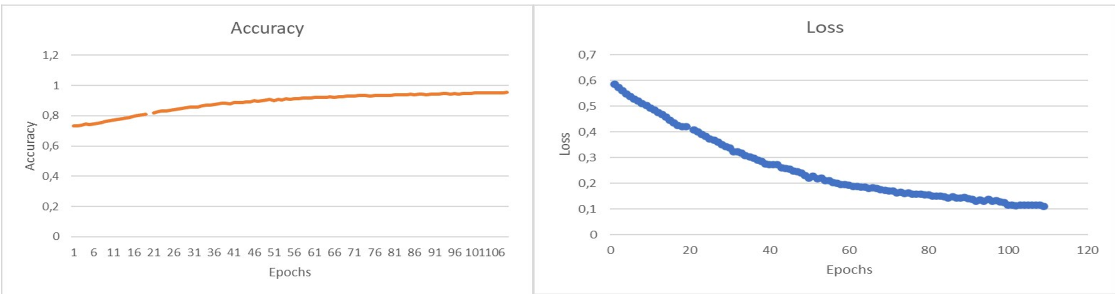
(a) **(b)**

Fig.3.12: (a) Accuracy and (b) loss obtained using VGG16.



(a) **(b)**

Fig.3.13: (a) Accuracy and (b) loss obtained using EfficientNetB0.



(a) **(b)**

Fig.3.14: (a) Accuracy and (b) loss obtained using MobileNetV1.

VGG16 and MobileNetV1 achieved an accuracy of 95.21% and 95.40% for RGB images after 110 epochs, with a loss of 0.1161 and 0.11. On the other hand, EfficientNetB0 achieved an accuracy of 97.54% for gray images after 40 epochs, with a loss of 0.0743. VGG16, MobileNetV1 and EfficientNetB0 models performed well in detecting and classifying the images. Their low loss values demonstrate their ability to minimize errors and make precise predictions, making them valuable tools for diagnosing diabetic retinopathy.

Comparative analysis:

Models	**Train loss**	**Train Accuracy**	**Test loss**	**Test Accuracy**
VGG16	2.04	0.22	1.81	0.25
ResNet50	2.16	0.70	1.52	0.70

Table 3.7: The results of the article of reference [87].

On the other hand, the article of reference that we worked on, reported lower accuracies of 25% and 70%, respectively. Our results suggest that their models may have faced challenges in achieving higher accuracy with the given dataset. However, with our model, we achieved better accuracy.

3.7 Presentation of the application

After developing our deep learning models and making them functional for users, we created a user-friendly interface. Our application is a two-tier web application, with a server-side based on Flask, a micro-framework written in Python. It allows for the processing of medical images submitted by users and provides a diagnosis. The application primarily consists of two interfaces:

3.7.1 Graphical user interface

- The fundus image uploading interface: the user can upload an images through this interface.
- The results interface: This interface displays the results of the prediction made on the submitted images, providing relevant diagnostic information.

To better understand the functionality of our application, we will present some screenshots illustrating its interface and operating principle.

- **Uploading and checking page**

Our graphical home page consists of two buttons: one for uploading an image and the other for checking whether or not diabetic retinopathy is present. The following figure illustrates what we have mentioned.

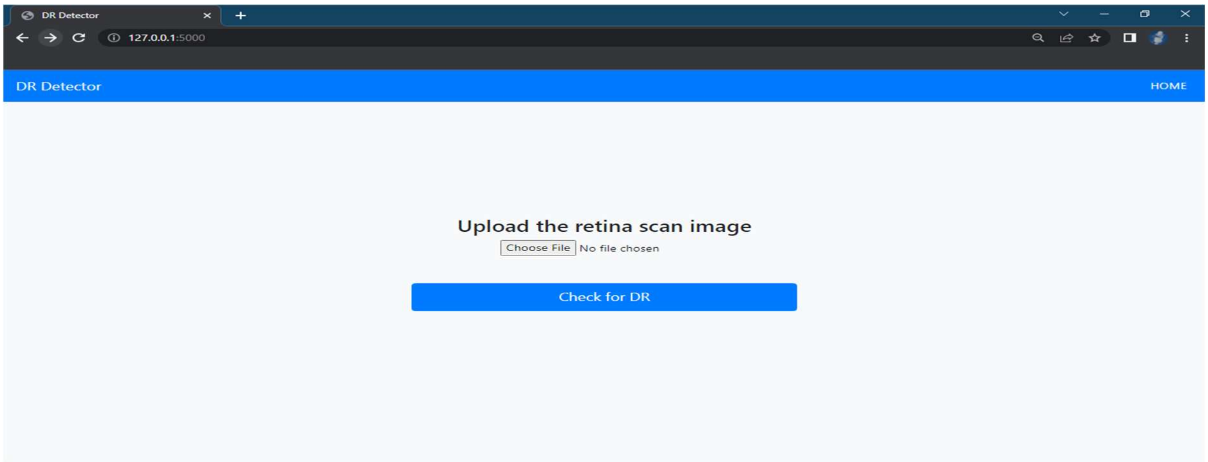


Fig.3.15: The interface home page.

As it's shown, no file is chosen, by clicking on choose file a window will be opened where we classified the retinas into DR and No DR just to make sure it working correctly, the window allows you to choose any retina scan image, as it's shown in figure 3.16 bellow:

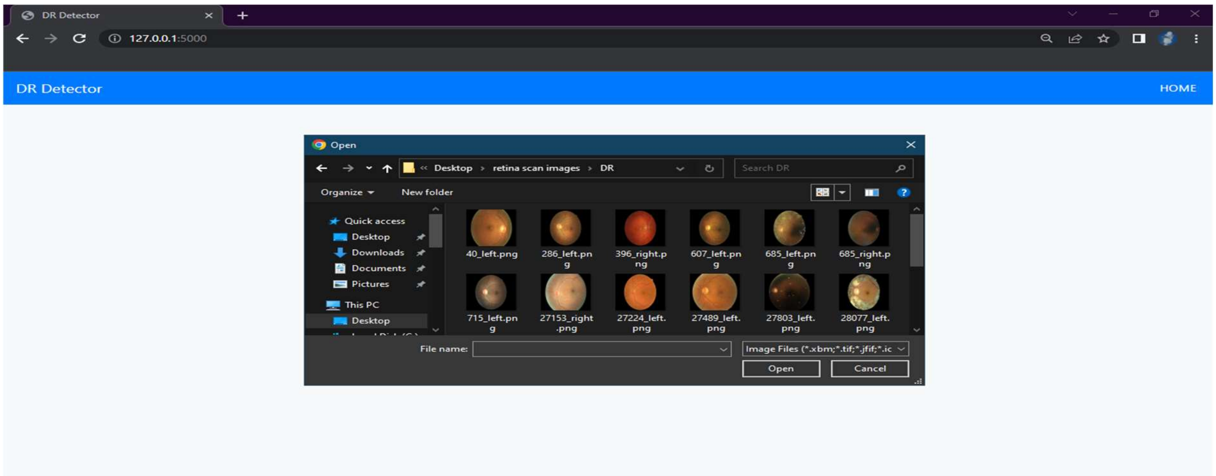


Fig.3.16: The window for choosing a file by the user.

If u click on check for DR button without firstly selecting any image an error text will appear to tell you to choose file, as it's shown in figure 3.17 bellow:

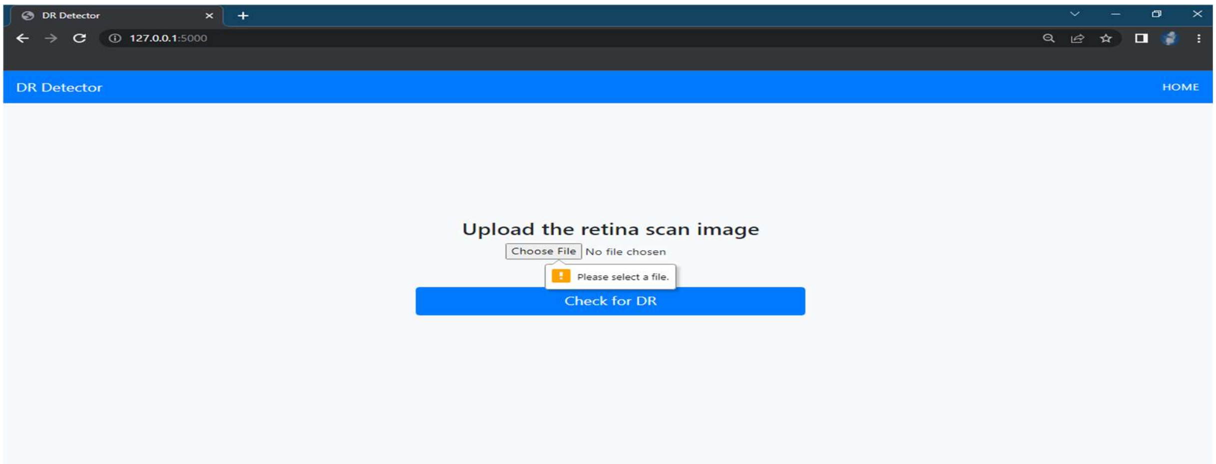


Fig.3.17: The error text.

❖ Results interface

In the 2nd page of our application, the outcome based on the selected RGB image of the 1st dataset will be displayed, indicating whether or not DR is present.

In figure 3.18 DR is detected:

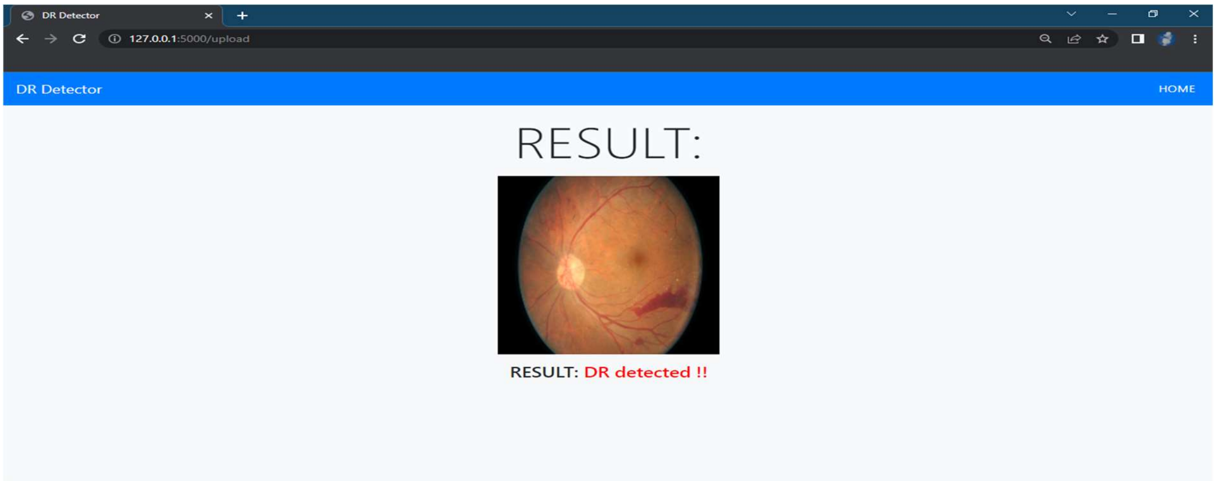


Fig.3.18: The result of detecting DR.

In figure 3.19 No DR detected:

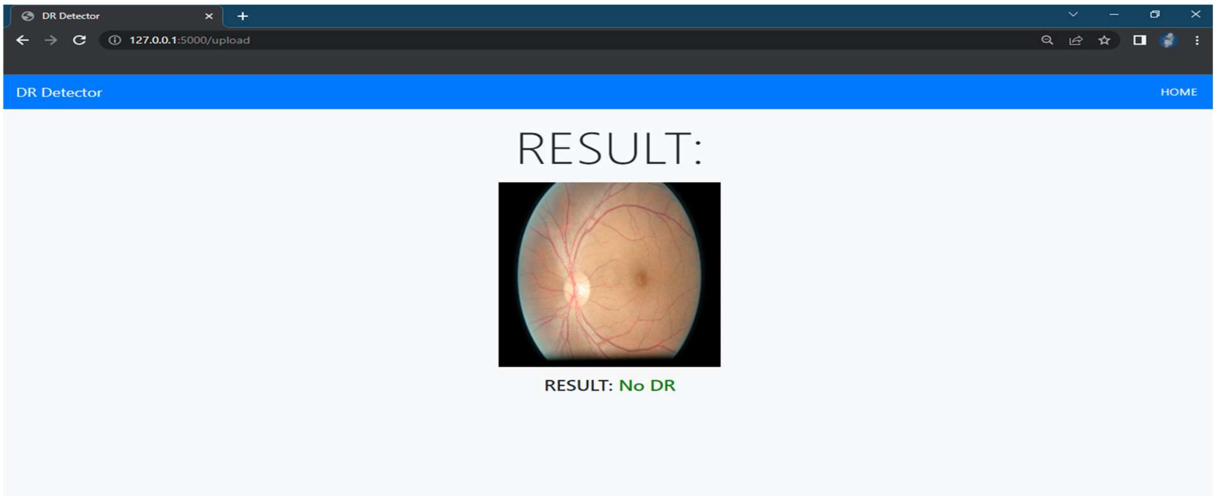


Fig.3.19: The result of No DR detection.

Also, we present the result of the 2nd dataset grayscale images:

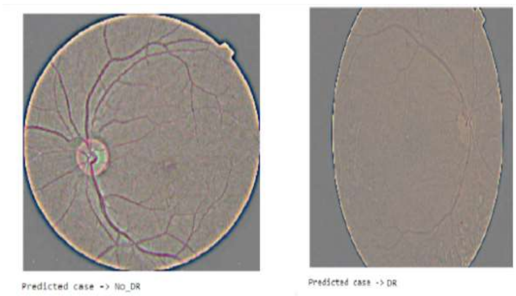


Fig.3.20: The result of grayscale images.

3.8 Conclusion

In this final chapter, our focus has been on the various tools and technologies utilized to develop our diabetic retinopathy detection system, as well as the architecture of the three deep learning models we constructed. These models consist of VGG16 and MobileNetV1 for the initial RGB image dataset, and EfficientNetB0 for grayscale images. We conducted numerous experiments and tests, exploring diverse configurations of these architectures. Alongside these tests, we provided our observations and interpretations to compare the accuracy performance of the architectures in the final testing phase. Lastly, we showcased our developed application, highlighting the libraries employed in the process.

General Conclusion

The development of this project has proven to be extremely beneficial, allowing us to apply our knowledge and gain valuable experience. With the constant evolution of technology, the prevalence of diabetic retinopathy has increased, posing challenges in its detection. Throughout our project, we focused on addressing this issue, resulting in the development of our application.

Our findings indicate that both the MobileNetV1 and VGG16 architectures yield similar results, with MobileNetV1 exhibiting better accuracy for RGB images and EfficientNetB0 performing well with grayscale images. This project provided us with an opportunity to explore various architectures used in convolutional neural network (CNN) processing, keeping us abreast of the latest trends and techniques in the field.

The application remains open for future improvements and can be enhanced with more advanced features. Utilizing a larger dataset and implementing more sophisticated algorithms could potentially yield better results. Additionally, the application could be expanded to detect and diagnose other ocular diseases. One possible future direction could be to optimize the project by modifying and improving the algorithm used, and even exploring the possibility of deploying the model on mobile phones for on-the-go detection using the phone's camera.

Having achieved our initial goal, we conclude this project by presenting a cohesive body of work that leverages the means and knowledge we have acquired throughout our academic journey.

References

[1] Center of Disease Control and Prevention, "What is diabetes? | CDC," [Online]. Available: https://www.cdc.gov/diabetes/basics/diabetes.html . [Accessed: Feb. 12, 2023].

[2] World Health Organization, "Diabetes (who.int)," [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes . [Accessed: Feb. 16, 2023].

[3] Center of Disease Control and Prevention, "What Is Type 1 Diabetes? | CDC," [Online]. Available: http://www.cdc.gov/ . [Accessed: Mar. 11, 2023].

[4] Cleveland Clinic, "Eyes - Cleveland Clinic," [Online]. Available: https://my.clevelandclinic.org/health/body/21823eyes#:~:text=Your%20eyes%20are%20organs%20that,lead%20to%20permanent%20vision%20loss . [Accessed: Feb. 20, 2023].

[5] National Institutes of Health, "National Eye Institute, N,IH Publication No: 06-2171 diabeticretino.pdf," [Online]. Available: https://www.nei.nih.gov/ . [Accessed: Feb. 10, 2023].

[6] National Institutes of Health (NIH). "Turning Discovery Into Health." [Online]. Available: https://www.nih.gov/. [Accessed: Feb. 10, 2023].

[7] Mayo Clinic, "Diabetic retinopathy - Symptoms and causes - Mayo Clinic," [Online]. Available: https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/ . [Accessed: Feb. 10, 2023].

[8] "Symptoms of Diabetic Retinopathy," [Online]. Available: https://kadrmaseyecare.com/ . [Accessed: Feb.11,2023].

[9] American Academy of Ophthalmology Retina/Vitreous Panel, "Preferred Practice Pattern® Guidelines. Diabetic Retinopathy," Sec. for Quality-of-Care Anne L. Coleman, MD, PhD, Academy Staff Nicholas P. Emptage, MAE, Doris Mizuiri, Shannon Kealey, MLS, Flora C. Lum, MD, Med. Editor Susan Garratt, Design Socorro Soberano, Approved by Board of Trustees, Sept. 20, 2014.

[10] "Algérie : le nombre de diabétiques représente environ 15% de la population," [Online]. Available: https://www.algerie-eco.com/ . [Accessed: Feb.20,2023].

[11] A. UzZaman and S. K. Bashir, "Diabetic Retinopathy Detection Using Image Processing," Supervisor: Prof. Dr. Md. Haider Ali, Co-supervisor: Mohammad Hammad Ali, Department of Computer Science and Engineering, School of Engineering and Computer Science, BRAC University, Submitted on 20th April 2016.

[12] Y. Tokuda et al., "Automatic Diagnosis of Diabetic Retinopathy Stage Focusing Exclusively on Retinal Hemorrhage," Medicina, vol. 58, p. 1681, Nov. 2022. [Online]. Available: https://doi.org/10.3390/medicina58111681 . [Accessed: June. 12, 2023].

[13] "Diabetic Retinopathy Stages," [Online]. Available: https://www.healthline.com/health/diabetes/diabetic-retinopathy-stages#stages . [Accessed: Apr. 19, 2023].

[14] "Diabetic Retinopathy Pathophysiology," [Online]. Available: https://eyewiki.aao.org/Diabetic_Retinopathy_Pathophysiology#:~:text=3%20Conclusions-,Introduction,in%20working%20middle%2Daged%20adults . [Accessed: Apr. 2, 2023].

[15] "Anatomy of the Eye," [Online]. Available: https://www.scottpautlermd.com/anatomy-of-the-eye/ . [Accessed: Apr. 2, 2023].

[16] "Difference between Normal Retina and Diabetic Retinopathy," [Online]. Available: https://www.researchgate.net/profile/Srinivasa-Rao-Avanapu/publication/282609747/figure/fig2/AS:281548759814145@1444137863086/Difference-between-Normal-Retina-and-Diabetic-Retinopathy.png . [Accessed: Apr. 19, 2023].

[17] "Diabetic Retinopathy Image," [Online]. Available: https://miro.medium.com/max/1400/1*Om8kniLZCpQbf1oznWCvFA.png .

[Accessed: Apr. 2, 2023].

[18]URL:https://www.europarl.europa.eu/news/en/headlines/society/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used . [Accessed: May 3, 2023].

[19] URL: https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence#:~:text=Artificial%20intelligence%20is%20the%20simulation,speech%20recognition%20and%20machine%20vision. [Accessed: May 3, 2023].

[20] URL: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained . [Accessed: May 3, 2023].

[21] URL: https://www.coursera.org/articles/types-of-machine-learning . [Accessed: May 3, 2023].

[22] URL: https://addepto.com/blog/what-are-the-top-10-challenges-of-machine-learning/ . [Accessed: May 6, 2023].

[23] URL: https://www.provintl.com/blog/5-common-machine-learning-problems-how-to-beat-them . [Accessed: May 6, 2023].

[24] URL: https://addepto.com/blog/what-are-the-top-10-challenges-of-machine-learning/ . [Accessed: May 6, 2023].

[25] URL: https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms . [Accessed: May 6, 2023].

[26] Antoine Crochet-Damais. "Deep learning: définition et principes de l'apprentissage profond." URL: https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501333-deep-learning-definition-et-principes-de-l-apprentissage-profond/. [Accessed: May 8, 2023].

[27] "Neural Network: Architecture, Components & Top Algorithms." URL: https://www.upgrad.com . [Accessed: May 6, 2023].

[28] URL: https://www.lebigdata.fr/reseau-neurones-artificiels-definition . [Accessed: May 11, 2023].

[29] URL: https://deepai.org/machine-learning-glossary-and-terms/activation-function . [Accessed: May 7, 2023].

[30] URL: https://deepai.org/machine-learning-glossary-and-terms/loss-function. [Accessed: May 7, 2023].

[31] URL: https://www.javatpoint.com/gradient-descent-machine-learning. [Accessed: May 7, 2023].

[32] "Optimizers in Deep Learning." URL: https://medium.com/mlearning-ai/optimizers-in-deep-learning-7bf81fed78a0. [Accessed: May 13, 2023].

[33] "Optimizers in Deep Learning." URL: https://ruder.io/optimizing-gradient-descent/index.html#nesterovacceleratedgradient. [Accessed: May 13, 2023].

[34] URL: https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures/. [Accessed: May 9, 2023].

[35] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," CoRR, abs/1409.1556, 2015.

[36] "VGG16 – Convolutional Network for Classification and Detection." URL: https://neurohive.io/en/popular-networks/vgg16/. [Accessed: May 10, 2023].

[37] "Great learning by G.Rohini." URL: https://medium.com/great-learning/everything-you-need-to-know-about-vgg16-2f64347b01ae. [Accessed: May 10, 2023].

[38] "VGGNet-16 Architecture: A Complete Guide." URL:https://www.kaggle.com. [Accessed: May 10, 2023].

[39] S. Patel, "Diabetic Retinopathy Detection and Classification using Pre-trained Convolutional Neural Networks," International Journal on Emerging Technologies, vol. 11, no. 3, pp. 1082–1087, 2020.

[40] "Understanding VGG16: Concepts, Architecture, and Performance." URL: https://www.datagen.tech. [Accessed: May 9, 2023].

[41] A. M. Mutawa, S. Alnajdi, and S. Sruthi, "Transfer Learning for Diabetic Retinopathy Detection: A Study of Dataset Combination and Model Performance," Applied Sciences, vol. 13, no. 9, p. 5685, 2023. https://doi.org/10.3390/app13095685.

[42] URL: https://www.researchgate.net/figure/Depthwise-separable-convolution-block_fig1_341090667. [Accessed: May 10, 2023].

[43] K. Bapat, "Introduction to MobileNet v1 using Depth Wise Separable Convolution – Krutika Bapat – Engineering at IIIT-Naya Raipur | 2016-2020." URL:

https://krutikababapat.medium.com/introduction-to-mobilenet-v1-using-depth-wise-separable-convolution-8676b3a716e9. [Accessed: May 10, 2023].

[44] D. (DJ) Sarkar, "A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning." URL: https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a.

[45] URL: https://cs231n.github.io/transfer-learning. [Accessed: May 9, 2023].

[46] K. Kamal, R. A. Mohamed, A. Darwish, and A. E. Hassanien, "Deep Learning with Class Imbalance for Detecting and Classifying Diabetic Retinopathy on Fundus Retina Images," Research Square, August 12th, 2022. [Accessed: May 16, 2023].

[47] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 2019. [Accessed: May 16, 2023].

[48] "EfficientNet [model architecture]." URL: https://www.opengenus.org. [Accessed: May 16, 2023].

[49] G. Liang and L. Zheng. "A transfer learning method with deep residual network for pediatric pneumonia diagnosis." Computer Methods and Programs in Biomedicine, Vol. 187, pp. 1-9, 2020. [Accessed on May 16, 2023].

[50] "Squeeze-and-Excitation Block." URL: https://paperswithcode.com/method/squeeze-andexcitationblock#:~:text=The%20Squeeze%2Dand%2DExcitation%20Block,convolutional%20block%20as%20an%20input. [Accessed on May 17, 2023].

[51] Sandler, Mark, Howard, Andrew, Zhu, Menglong, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. [Accessed on May 16, 2023].

[52] "Keras MobileNet Documentation." URL: https://keras.io/api/applications/mobilenet. [Accessed on May 18, 2023].

[53] NumPy: A fundamental package for scientific computing with Python." [Online]. Available: https://numpy.org/.[Accessed on May 7,2023].

[54] "OpenCV: Open-Source Computer Vision Library." [Online]. Available: https://opencv.org/. [Accessed on May 8,2023].

[55] Flask Project. "Flask: a lightweight Python web framework." [Online]. Available: https://flask.palletsprojects.com/. [Accessed on June 9,2023].

[56] URL : https://www.tensorflow.org/. [Accessed on June 9,2023].

[57] "APTOS 2019 Blindness Detection Kaggle Competition." URL: https://www.kaggle.com/c/aptos2019-blindness-detection/overview.

[Accessed on June 1, 2023].

[58] "Diabetic Retinopathy Detection Kaggle Competition." URL: https://www.kaggle.com/c/diabetic-retinopathy-detection. [Accessed on June 1, 2023].

[59] "Python Documentation." URL: https://www.python.org/doc/essays/blurb/. [Accessed on June 3, 2023].

[60] URL: https://www.computerhope.com/jargon/v/visual-studio.htm. [Accessed on June 10, 2023].

[61] "EfficientNet - Papers with Code." URL: https://paperswithcode.com/method/efficientnet. [Accessed on June 3, 2023].

[62] "Transfer Learning Guide Examples for Images and Text in Keras." URL: https://neptune.ai/blog/transfer-learning-guide-examples-for-images-and-text-in-keras. [Accessed on June 3, 2023].

[63] "What is Google Colab? - Education Ecosystem Blog." URL: https://educationecosystem.com/blog/what-is-google-colab/. [Accessed on June 3, 2023].

[64] "Kaggle Getting Started Guide." URL: https://www.kaggle.com/getting-started/13245. [Accessed on June 3, 2023].

References

[65] "TensorFlow - TechTarget." URL: https://www.techtarget.com/searchdatamanagement/definition/TensorFlow#:~:text=TensorFlow%20is%20an%20open%20source,statistical%20and%20predictive%20analytics%20workloads. [Accessed on June 3, 2023].

[66] "KERAS - Definitions.net." URL: https://www.definitions.net/definition/KERAS. [Accessed on June 3, 2023].

[67] "OpenCV - Introduction." URL: https://docs.opencv.org/4.x/d1/dfb/intro.html. [Accessed on June 3, 2023].

[68] "NumPy Introduction - W3Schools." URL: https://www.w3schools.com/python/numpy/numpy_intro.asp#:~:text=NumPy%20stands%20for%20Numerical%20Python . [Accessed on June 3, 2023].

[69] TensorFlow - https://www.tensorflow.org/. [Accessed on May 5, 2023]

[70] Keras API - https://keras.io/api/. [Accessed on May 5, 2023]

[71] NumPy - https://numpy.org/. [Accessed on May 5, 2023]

[72] OpenCV Logo file:///C:/Users/Acer/Desktop/OpenCV_Logo_with_text_svg_version.svg. [Accessed on May 5, 2023].

[73] Pocoo Flask Logo - https://www.vectorlogo.zone/util/preview.html?image=/logos/pocoo_flask/pocoo_flask-ar21.svg . [Accessed on May 5, 2023].

[74] Hyperparameters Optimization - https://towardsdatascience.com/hyperparameters-optimization-526348bb8e2d . [Accessed on June 7, 2023].

[75] Understanding Learning Rate - https://towardsdatascience.com/https-medium-com-dashingaditya-rakhecha-understanding-learning-rate-dd5da26bb6de#:~:text=Learning%20rate%20(%CE%BB)%20is%20one,move%20towards%20the%20optimal%20weights . [Accessed on June 7, 2023].

References

[76] Introduction to Neural Network - https://medium.com/geekculture/introduction-to-neural-network-2f8b8221fbd3#:~:text=Number%20of%20Neurons%20and%20Number%20of%20Layers%20in%20Hidden%20Layer&text=The%20number%20of%20hidden%20neurons,size%20of%20the%20output%20layer.n . [Accessed on June 7, 2023].

[77] Batch Size in Deep Learning - https://www.accountingtools.com/articles/batch-size#:~:text=Batch%20size%20is%20the%20number,setup%20cost%20over%20more%20units .[Accessed on June 7, 2023].

[78] Dropout Regularization in Deep Learning - https://www.analyticsvidhya.com/blog/2022/08/dropout-regularization-in-deep-learning/#:~:text=What's%20Dropout%3F,are%20codependent%20with%20one%20another . [Accessed on June 7, 2023].

[79] Gaussian Blur in Image Processing - https://www.southampton.ac.uk/~msn/book/new_demo/gaussian/#:~:text=It%20is%20used%20to%20remove,blurring%20for%20larger%20template%20sizes . [Accessed on June 7, 2023].

[80] RGB and CMYK Color Models - https://guides.lib.umich.edu/c.php?g=282942&p=1885348#:~:text=RGB%20and%20CMYK-,RGB,for%20web%20and%20presentation%20purposes . [Accessed on June 7, 2023].

[81] Introduction to Batch Normalization - https://www.analyticsvidhya.com/blog/2021/03/introduction-to-batch-normalization/#:~:text=Normalization%20is%20a%20data%20pre,values%20to%20a%20balanced%20scale . [Accessed on June 7, 2023].

[82] Convolutional Layer in Deep Learning - https://www.sciencedirect.com/topics/engineering/convolutional-layer#:~:text=A%20convolutional%20layer%20is%20the,and%20creates%20an%20activation%20map . [Accessed on June 7, 2023].

[83] Flask Documentation - https://pymbook.readthedocs.io/en/latest/flask.html . [Accessed on June 7, 2023].

References

[84] Diabetic Retinopathy - https://www.hopkinsmedicine.org/health/conditions-and-diseases/diabetes/diabetic. [Accessed on June 7, 2023].

[85] Fundus Stock Photos - https://www.istockphoto.com/ . [Accessed on June 17, 2023].

[86] OCT (Optical Coherence Tomography) -OCT (Optical Coherence Tomographie) – Centre Ophtamologique Antony (antony-ophtalmologie.fr) .[Accessed on June 17, 2023]

[87] A. Mustapha, M. Lachgar, H. Hrimech, and A. Kartit, "Diabetic Retinopathy Classification Using ResNet50 and VGG-16 Pretrained Networks," International Journal of Computer Engineering and Data Science, vol. 1, no. 1, pp., July 2021.