MINOR PROJECT REPORT

OCCULIST

(A portal for Diabetic Retinopathy detection)

Submitted By:

Pious Annie	(20103049)
Karanpreet Singh	(20103065)
Rijul Singla	(20103080)
Armaan Badhaan	(20103102)

Under the Supervision of:

Prof. Amandeep Kaur

Department of Computer Science and Engineering

Punjab Engineering College

DECLARATION

We hereby declare that the project work entitled **“OCCULIST”** is an authentic record of our own work carried out at **Punjab Engineering College (Deemed to be University),** as a requirement of **Minor Project**, under the guidance of **Prof. Amandeep Kaur** (Department of Computer Science and Engineering). We further declare that the information has been collected from genuine and authentic sources and we have not submitted this project report to this or any other university for award of diploma or degree of certificate examination.

-Pious Annie (20103049)

-Karanpreet Singh (20103065)

-Rijul Singla (20103080)

-Armaan Badhaan (20103102)

CERTIFICATE

Certified that the Project work entitled **"OCCULIST"** submitted by Pious Annie, Karanpreet Singh, Rijul Singla and Armaan Badhaan for the fulfilment of Minor Project offered by Punjab Engineering College, Chandigarh during the academic year 2022-23 is an original work carried out by the students under my supervision and this work has not framed any basis for the award of and Degree, Diploma or such other titles.

Certified that the above statement made by the students is correct to the best of my knowledge and belief.

Prof. Amandeep Kaur

(Faculty Mentor)

Department of Computer Science and Engineering

Punjab Engineering College

19th December 2022

ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely fortunate to have got this all along the completion of our project work. Whatever we have done is only due to such guidance and assistance and we would not forget to thank them. We respect and thank **Prof. Amandeep Kaur** for providing us with this opportunity to do the project work and providing us with all support and guidance which made us complete the project on time. We are extremely grateful to her for providing such nice support though she had a busy schedule managing the college affairs.

This project gave us the opportunity to learn and explore several fields of Computer Science, which has opened up a plethora of opportunities to carry out further development. We acknowledge gratefully the help and suggestion of fellow batch mates who were always eager to help us with their warm attitude and technical knowledge, and provide critical feedback when required.

-Pious Annie (20103049)

-Karanpreet Singh (20103065)

-Rijul Singla (20103080)

-Armaan Badhaan (20103102)

ABSTRACT

Diabetic retinopathy (DR) is the result of a complication of diabetes affecting the retina. It can cause blindness, if left undiagnosed and untreated. An ophthalmologist performs the diagnosis by screening each patient and analysing the retinal lesions via ocular imaging using a fundus camera. In practice, such analysis is time-consuming and cumbersome to perform. Our solution is intended for rural areas which have small dispensaries and can afford a fundus camera, but don't have enough expert medical professionals available. Occulist aims at making healthcare affordable and accurate.

Our work is an attempt to speed up preliminary screening of DR to cater to the future requirement of such a huge number of diabetic patients. We have trained and validated robust classification models on publicly available datasets for early detection of DR. We have applied state-of-the-art deep learning models based on Convolutional Neural Networks (CNN), to exploit data-driven machine learning methods for the purpose. We framed the problem as a binary classification for the detection of DR of any grade (Grade 1 – 4) vs No-DR (Grade 0). The ensemble of models has evaluated on separate datasets other than the dataset used for training; i.e., EyePACS test set an MESSIDOR-2 test sets. Our algorithm scored an AUC of 0.92 on MESSIDOR-2. We have used the shortest distance as the operating point selection strategy for the ROC curve. We obtain Sensitivity and Specificity for MESSIDOR-2 as 81.02% and 86.09%, respectively. For the EyePACS test dataset from Kaggle (14,210 images), our algorithm has scored an AUC of 0.927, Sensitivity and Specificity of 83.74% and 89.65%, respectively at the mentioned operating point. For binary classification, the modified ResNet50 model shows an accuracy of 89.56%. The Multistage classification accuracy obtained using modified ResNet50 is 91.73. Further, a comparison of the proposed model with the models in literature shows an overall accuracy improvement by at least 0.83% and more than 5% for binary and multistage classification respectively. The developed preliminary automated screening system will act as an aid to the manual diagnostic process by referring DR patients to an ophthalmologist for further examination (if detected positive) well in time to reduce the risks of vision loss.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

INTRODUCTION

Diabetic retinopathy (DR) is a type of ocular disease caused by high levels of blood glucose and high blood pressure, which can damage the blood vessels in the back of the eye (retina) and lead to blindness. One-third of people living with diabetes have some degree of diabetic retinopathy, and every person who has diabetes is at risk of developing it. Almost 40% of diabetic patients have DR, from which about 5% of patients face vision-threatening complications. Accurately grading diabetic retinopathy is time-consuming for ophthalmologists and can be a significant challenge for beginner ophthalmology residents. Therefore, developing an automated diagnosis system for diabetic retinopathy has significant potential benefits.

According to international protocol, the severity of DR can be graded into five stages (0-4):

(0) No Retinopathy
(1) Mild Non-Proliferative DR (NPDR)
(2) Moderate NPDR
(3) Severe NPDR
(4) Proliferative DR

The grading usually depends on the number and size of different related lesion appearances and complications. Figure 1 provides two examples, comparing a normal and a diabetic retinopathy retina containing multiple lesions. For example, microaneurysms (MAs) are the earliest clinically visible evidence of DR. Hard Exudate Microaneurysms Soft Exudate Neovascularization Intra-retinal microvascular abnormality Haemorrhage. The left image shows a normal retina, while the right one is a DR-4 retina. are local capillary dilatations that appear as small red dots. Moderate NPDR contains ‘dot’ or ‘blot’ shaped haemorrhages (HEs) in addition to microaneurysms. Hard exudates (EXs) are distinct yellow-white intra-retinal deposits which can vary from small specks to larger patches. They are principally observed in the macular region, as the lipids coalesce and extend into the fovea. Soft exudates (SE), also sometimes referred to as ‘cotton-wool spots’ (CWS), are greyish-white patches of discoloration in the nerve fibre layer, or precapillary arterial occlusions. They usually appear in severe DR stages. Moreover, intra-retinal microvascular abnormalities (IRMAs) are areas of capillary dilatation and new intra-retinal vessel formation. A pre-proliferate DR stage can be predicted once IRMA is present in numbers. Neovascularization (NV) is a significant factor of proliferate DR. As the retina becomes more ischaemic, new blood vessels may arise from the optic disc or in the periphery of the retina. Therefore, identifying these related regions can be helpful for DR grading.

As much as 95% of the cases of vision loss and blindness can be prevented with regular screening and appropriate management. Despite the importance of a timely diagnosis fewer than a half of the patients with DR are aware of their condition. Due to the asymptomatic nature of the disease in its early stages, patients should be screened with a fundus exam, to look for early signs of the disease and make a timely detection of the illness. However, access to specialized care by an ophthalmologist is limited for some populations, given that most of the ophthalmologists in the world are concentrated mainly in urban areas and big cities. The development of an automated detection system for DR could improve access to specialized care by reducing the time, cost, and effort of screening. In addition, the diabetic population is expected to increase 54% by 2030, while the projected increase of ophthalmologists is only 2%. Thus, the need to integrate methods to automate the screening process responds to these challenges in the present and future diabetic retinopathy panorama. Ophthalmologists perform DR screening by manually inspecting fundus photographs, which may be complex, biased (dependent on the doctor's expertise), and prone to error. These manual errors have generated a need for an automated system based on image processing and machine learning algorithm that can help in accurate identification of DR. Dealing with the previous issues, we propose a model for the automatic detection of DR.

The proposed method is different from state of-the-art models based on Convolutional Neural Networks (CNN), because it follows the clinician workflow, this approach has the advantage of providing additional information regarding the lesions found in the input image that give an improved interpretability. The experimental results also showed that this strategy also improves the overall classification performance of the system. The main three contributions of this paper are as follows: a fine-grained lesions annotation for microaneurysms, haemorrhages, cotton wool spots, venous beading, neovascularization and exudates; global labels for ocular diseases and referable conditions for 3209 images from the Kaggle EyePACS dataset and the complete MESSIDOR-2 dataset; and a strategy for the detection of DR that relies on the identification of its related retinal abnormalities.

Over the past decade, computer vision and deep learning-based algorithms have been largely explored to contribute to the medical imaging research community. With successful developments in deep convolutional neural networks (CNNs), image classification, object detection, semantic segmentation, and image synthesis frameworks, have all been investigated to analyse medical images for addressing different tasks. To study diabetic retinopathy, most previous works can be coarsely categorized into three important branches. First, the most valuable task is to predict diabetic retinopathy progression.

We first adopted the Inception-v3 architecture to train a DR grading model, which aims to directly learn the local features rather than explicitly detecting lesions. An automated image-level DR grading system was built on an ensemble of multiple well-trained deep learning models. Some of these deep models were also combined with the AdaBoost to reduce the bias of each individual model. Lesion based diabetic retinopathy detection has also been investigated.

Secondly, we proposed to integrate lesion detection and grading by designing two-stage deep convolutional neural networks. Specifically, a local network is first trained to classify the patches into different lesions, and then the second network predicts the severity grades of DR. A zoom-in-net was proposed to learn attention maps which highlight abnormal regions, and then provides the grading levels of DR in both global and local manners.

Thirdly, we used image generation methods for synthesizing retinal images. This technique can be used for data augmentation to address imbalances in DR training data. We proposed to synthesize fundus images given the pathological descriptors and vessel segmentation masks. We attempted to generate high-resolution retinal images with different grade levels by manipulating arbitrary grading and lesion information.

Currently, the two biggest obstacles to the progress of computer-aided diagnosis systems for DR are limited amounts of training data and inconsistent annotations. While there are a few public DR databases, but most of them only contain image-level labels, and annotations are often inaccurate. Constructing a large dataset with high-quality and fine-grained annotations would significantly contribute to research in DR diagnosis. For example, pixel-level annotations of DR-related lesions are highly beneficial for developing lesion-based segmentation models, as well as for training more interpretable grading models for ophthalmologists. Moreover, if fine-grained annotations of numerous lesions are provided, this rich information can be used to improve the ability of representation learning, as well as enable the models to be transferred for other ocular disease identification tasks without annotations. Therefore, we propose a new benchmark for studying diabetic retinopathy diagnosis systems.

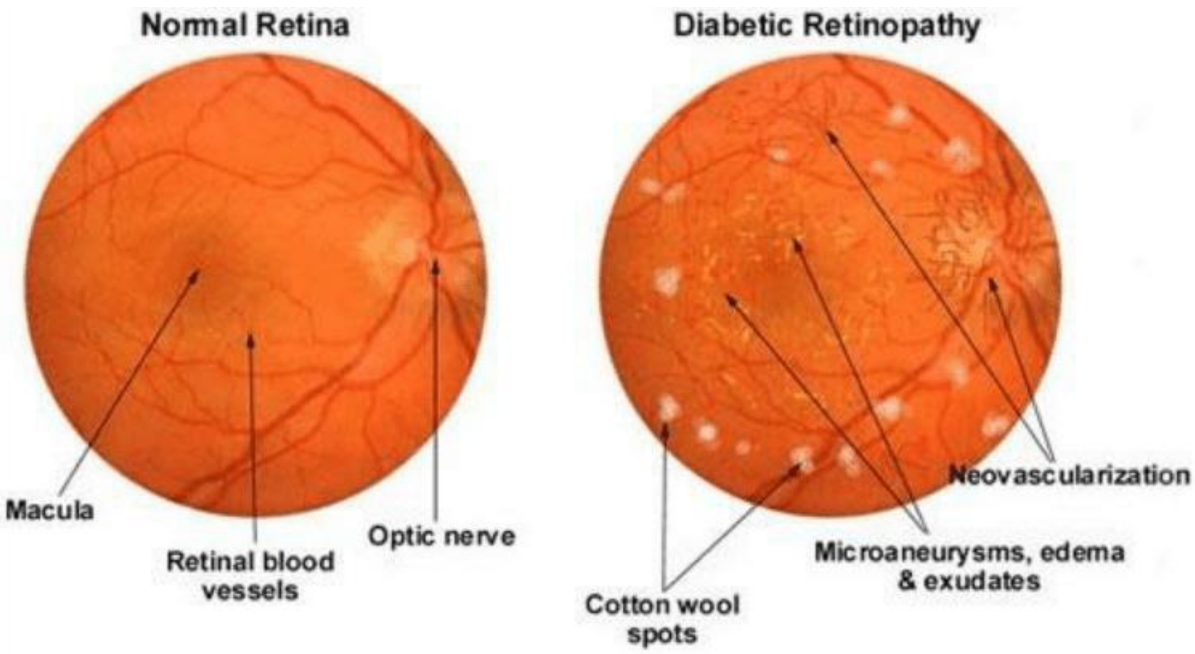


Figure 1. Illustration of diabetic retinopathy retina

1.1 MOTIVATION

India is the diabetes capital with home to 69.1 million people with DM, the second highest number of cases after China. Recent epidemiological evidence indicates a rising DM epidemic across all classes, both affluent and the poor in India. The prevalence of diabetes was 14.3 and 15.1% in Punjab and Haryana, respectively. In Punjab 34.2% of diabetics aware of their condition, 28.2% are on treatment while only 14.2% of the cases are controlled. 29.5% of the respondents were aware of their condition, 22.4% are on treatment while only 13.8% of the cases of diabetes were controlled. Prevalence of diabetes was higher in rural areas in both Punjab and Haryana. Furthermore, in both the states, higher proportion the urban population were on treatment than the rural population. As per the government-led study published in The Lancet, diabetes contributed to 3.1 per cent of the total deaths in 2016.

There is high prevalence of diabetes, especially of undiagnosed cases amongst the adult population, most of whom have uncontrolled blood sugar levels. This indicates the need for systematic screening and awareness program to identify the undiagnosed cases in the community and offer early treatment and regular follow up.

One-third of people living with diabetes have some degree of diabetic retinopathy, and every person who has diabetes is at risk of developing it. Almost 40% of diabetic patients have DR, from which about 5% of patients face vision-threatening complications. As much as 95% of the cases of vision loss and blindness can be prevented with regular screening and appropriate management. Despite the importance of a timely diagnosis fewer than a half of the patients with DR are aware of their condition.

Due to the asymptomatic nature of the disease in its early stages, patients should be screened with a fundus exam, to look for early signs of the disease and make a timely detection of the illness. However, access to specialized care by an ophthalmologist is limited for some populations, given that most of the ophthalmologists in the world are concentrated mainly in urban areas and big cities. The development of an automated detection system for DR could improve access to specialized care by reducing the time, cost, and effort of screening.

In addition, the diabetic population is expected to increase 54% by 2030, while the projected increase of ophthalmologists is only 2%. Thus, the need to integrate methods to automate the screening process responds to these challenges in the present and future diabetic retinopathy panorama. Ophthalmologists perform DR screening by manually inspecting fundus photographs, which may be complex, biased (dependent on the doctor’s expertise), and prone to error. These manual errors have generated a need for an automated system based on image processing and machine learning algorithm that can help in accurate identification of DR. Accurately grading diabetic retinopathy is time-consuming for ophthalmologists.

1.2 PROJECT OVERVIEW

A formal view to the progress of the project was taken and resulted in a systematic development of the required components. In order to meet the stated aims, the following steps were undertaken:

1. **Initial Design** – The use cases for the web site were listed and the control flow was analysed. A detailed database design was laid out to build an optimal website, which runs optimized search queries. A document for planned features was prepared and we brainstormed the best ideas to implement them.

   a) Image acquisition and pre-processing: The system should be able to acquire high-quality retinal images and pre-process them to ensure that they are in a suitable format for analysis. We have chosen publicly available Kaggle EyePACS and MESSIDOR 2 dataset for the purpose of our project.

   b) Machine learning algorithm: The system uses a machine learning algorithm, i.e., Convolutional Neural Networks (CNNs), to analyse the images and detect DR. The algorithm should be trained on a large dataset of labelled images to learn the features that are characteristic of the disease. We present a model for automatic DR classification on eye fundus images by proposing modified deep residual networks for binary and multistage classification of DR.

   c) User interface: The system should have a user-friendly interface that allows users to easily upload and view images, and to view the results of the analysis. The interface should be accessible from a web browser, allowing users to access the system from any device with an internet connection.

   d) Data storage and management: The system should have a secure and efficient system for storing and managing the images and analysis results. This may involve the use of a database or cloud storage.

   e) Security: The system should have robust security measures in place to protect the privacy and confidentiality of patient data. This may involve the use of encryption and secure authentication protocols.

2. **Specification** – The interfaces between all the areas were specified precisely, and the requirements for each of the sections were elicited. The best possible and most relevant technology stack was chosen. We had to keep in mind their compatibility with each other.

3. **Implementation** – The front end is implemented, mostly in React, a JSX-based framework, with additional libraries provided by bootstrap. The back end is implemented in Flask and the Spring Boot framework. Front end and Back end are connected using RESTful. PostgreSQL is used for database querying. All of the implementation was done in Spring Tool Suite and Visual Studio Code and Microsoft Workbench for creating and managing databases.

4. **Testing** – The implementation from each of the parts were then tested as a single system. The app is still in alpha phase. Further testing will be carried out as the project proceeds. We plan to perform unit tests for any new features, integration tests, and finally penetration tests to ensure data privacy as it is highly confidential.

5. **Evaluation** – The system was evaluated through comparisons to existing systems and through qualitative analysis of the optimality of the implementation. The system should be tested and evaluated to ensure that it performs accurately and reliably. This may involve conducting user studies or comparing the results of the system to those of expert ophthalmologists.

1.3 REPORT OVERVIEW

This report fully describes the project undertaken. However, in order to control the length of the report, the reader is on occasion referred to a bibliographical reference if particular details are required in an area. The report is split into five main sections:

1. **Introduction** – This gives an introduction to the project, its motivation and aim, a procedural overview of the work undertaken in the project and an overview of this report for a comprehensive understanding.

2. **Background Research** – This provides an in-depth analysis of current projects available in this area. It also briefly discusses the technologies that we can use to implement this idea. Further, possible extensions to the basic design requirements are proposed.

3. **Working and Implementation** – This section provides a brief overview of the various features and functions that have been included in the project. It discusses the APIs created, and the functionalities they add. It also provides insight to the database design and optimization of the same. Finally, it has various diagrams for a better understanding of the workflow of the project. Discussion of the implementation choices taken and the software that was developed. Detailed reports are given for the implementation of the API along with a sample application which uses the API and analysis of each of the implemented device access methods.

4. **Results and discussions** – Analysis of the successes and failures of the project, and discussion of the advances made. A discussion of the commercial viability of the implementation is also given. Possible extensions and potentials, future scope and further work that could be undertaken are then discussed.

CHAPTER 2
BACKGROUND RESEARCH

Due to the asymptomatic nature of the disease in its early stages, patients should be screened with a fundus exam, to look for early signs of the disease and make a timely detection of the illness. However, access to specialized care by an ophthalmologist is limited for some populations, given that most of the ophthalmologists in the world are concentrated mainly in urban areas and big cities. The development of an automated detection system for DR could improve access to specialized care by reducing the time, cost, and effort of screening. There are several shortcomings with the present methods of detecting diabetic retinopathy. One major issue is the reliance on subjective interpretation by healthcare professionals. The diagnosis of diabetic retinopathy often involves the analysis of fundus images, which can be difficult to interpret accurately without specialized training and experience. As a result, the accuracy of the diagnosis can vary greatly depending on the skill and expertise of the healthcare professional.

Another shortcoming of current methods is the time and resources required for the diagnosis process. The analysis of fundus images can be time-consuming and requires specialized equipment, which can be costly and may not be readily available in all healthcare settings. This can lead to delays in diagnosis and treatment, potentially compromising patient outcomes.

Automatic diabetic retinopathy detection using machine learning (ML) can help to address these issues by improving the accuracy and efficiency of the diagnosis process. ML algorithms can be trained to recognize and classify various features in fundus images, allowing for a more objective and consistent diagnosis. These algorithms can also process images much faster than humans, reducing the time and resources required for the diagnosis process. Automatic computer-aided screening of DR is a highly investigated field. The motivation for creating reliable automatic DR screening systems is to reduce the manual effort of mass screening, which also raises a financial issue. Regarding decision making, automatic DR screening systems either partially follow clinical protocols. Automated grading of diabetic retinopathy has potential benefits such as increasing efficiency, reproducibility, and coverage of screening programs; reducing barriers to access; and improving patient outcomes by providing early detection and treatment. To maximize the clinical utility of automated grading, an algorithm to detect referable diabetic retinopathy is needed.

In addition, the diabetic population is expected to increase 54% by 2030, while the projected increase of ophthalmologists is only 2%. Thus, the need to integrate methods to automate the screening process responds to these challenges in the present and future diabetic retinopathy panorama. Ophthalmologists perform DR screening by manually inspecting fundus photographs, which may be complex, biased (dependent on the doctor's expertise), and prone to error. These manual errors have generated a need for an automated system based on image processing and machine learning algorithm that can help in accurate identification of DR. Dealing with the previous issues, we propose a model for the automatic detection of DR.

Currently, the two biggest obstacles to the progress of computer-aided diagnosis systems for DR are limited amounts of training data and inconsistent annotations. While there are a few public DR databases, but most of them only contain image-level labels, and annotations are often inaccurate. Constructing a large dataset with high-quality and fine-grained annotations would significantly contribute to research in DR diagnosis. For example, pixel-level annotations of DR-related lesions are highly beneficial for developing lesion-based segmentation models, as well as for training more interpretable grading models for ophthalmologists. Moreover, if fine-grained annotations of numerous lesions are provided, this rich information can be used to improve the ability of representation learning, as well as enable the models to be transferred for other ocular disease identification tasks without annotations. Therefore, we propose a new benchmark for studying diabetic retinopathy diagnosis systems.

Our Solution- In this work, modified residual networks with 50 layers have been used for multistage DR classification on the Kaggle EyePACS and MESSIDOR-2 dataset. The network first performs extraction of features from the images (in the initial layers) and thereafter classifies them (in the last layers). Residual networks provide high accuracy gains and are easy to optimize due to the presence of short connections. These networks also overcome the degradation problem, which usually occurs in deeper neural networks.

2.1 DATASET DESCRIPTION

Most of the existing DR datasets only have image-level grading labels, with providing few pixel-level lesion-based annotations. A summary of some commonly used datasets related to DR is provided in Table I. Models trained on these datasets can only be used to predict a severity grade without providing any interpretability for ophthalmologists as to why a fundus image is graded as a certain level. Therefore, one of the main goals of our benchmark is to introduce a large fine-grained annotated dataset for more explainable diagnosis of DR. Detailed information of existing datasets and our proposed dataset are as follows:

1. Kaggle-EyePACS: This consists of 35,126 training images and 53,576 testing images only containing grading labels. The images are collected from different sources with various lighting conditions and weak annotation quality. The presence of DR in each image is rated on a scale of 0 to 4. In this dataset, some images contain artifacts, are out of focus, underexposed, or overexposed.

2. Kaggle-APTOS2019: This consists of 3,662 training images and 1,928 testing images, also with grading labels only. This dataset also suffers from noise in both images and labels.

3. ODIR-5K: This is a structured ophthalmic dataset of 5,000 patients. Multi-label image-level annotations for eight eye disease categories, including diabetes, glaucoma, cataract, age-related macular degeneration (AMD), hypertension, myopia, normal, and other diseases, are provided. Each patient may contain one or more disease labels. We adopt this dataset in the last task to explore transfer learning from DR to ocular multi-disease identification.

4. MESSIDOR: This contains 1,200 eye fundus images but its DR grading scale is different from those of previous datasets, having only four levels (0 to 3). In addition to DR grading, the risk of macular edema is also provided for each image with grading labels 0 to 2.

5. IDRiD: This dataset provides expert annotations of typical diabetic retinopathy lesions and normal retinal structures. The full set contains 516 images, but only 81 of them are labelled with pixel-level binary lesion masks. Abnormalities associated with DR, such as microaneurysms, haemorrhages, soft exudates and hard exudates, are provided.

6. DRIVE: This dataset is used for evaluating the segmentation of blood vessels in retinal images, and contains pixel-level binary vessel masks. The 40 images are divided into a training and a testing set, each containing 20 images.

7. FGADR: This is a fine-grained annotated diabetic retinopathy (FGADR) dataset, which consists of two sets. The first set, named Seg-set, contains 1,842 images with both pixel-level lesion annotations and image-level grading labels. The lesions include microaneurysms (MA), haemorrhages (HE), hard exudates (EX), soft exudates (SE), intra-retinal microvascular abnormalities (IRMA), and neovascularization (NV).

The proposed models have been evaluated on publicly available following datasets:

1. The **Kaggle EyePACS** is a free real-world set of high-resolution photos of the posterior pole of the retina. The Kaggle EyePACS dataset comprises 88702 macula-centred eye fundus images acquired with different types of cameras and under a variety of imaging conditions. These images were labelled using a scale of 0, 1, 2, 3, 4, which stand for no DR, mild, moderate, severe and proliferative DR respectively.

2. The **MESSIDOR-2** dataset is a public dataset that contains 1748 macula-centred eye fundus images that were acquired with a 45-degree field of view and the sizes are ranged between 1440 × 960 and 2304 × 1536 pixels.

These public datasets were used to create two new individual datasets with fine-grained labels of six ocular lesions used for ophthalmologists to diagnose DR. The construction details are explained as follows:

Dataset construction: an ophthalmologist with supra-specialty in the retina from the Fundación Oftalmologica Nacional of Colombia selected 3209 retinal images from the Kaggle EyePACS dataset. The ophthalmologist performed the manual choice of these images based on the quality of fundus images and the presence of DR-related lesions. The MESSIDOR-2 dataset was also analysed by the specialist who determined that 1689 images are suitable for lesion level annotation. The expert labelled the images with six ocular findings: aneurysms, haemorrhages, cotton wool spots, venous beating, neovascularization, and exudates. In addition, these images were manually annotated with: binary labels for referable and non-referable patients and the grade of diabetic macular edema (DME) in a scale 0 to 3 for no DME, mild, moderate, and severe.

Annotation criteria: the whole images were evaluated in their entire area as follows: first, a meticulous analysis looking for any lesions on the optic nerve and the macular region is performed. Then, the extra-macular retina and outside the vascular arch areas are examined in detail to find possible DR-related lesions. Finally, each image was classified in a fine-grained way as positive for that specific finding if at least one ocular lesion was identified. Regardless of whether it was a single injury or many, or whether it affected the centre of the macula or its periphery. In addition, the images were labelled as DME in an image-level way according to the ETDRS scale. Also, images were classified as a referable image, if at least an ocular lesion was present in all image areas, regardless of which one.

	Ocular lesions labels						Image-level labels		
Set of images	MA	H	CWS	VB	NV	EX	DR	REF	DME
Training set	1719 53.55%	1418 44.17%	653 20.34%	720 22.42%	140 4.36%	1812 56.44%	2016 62.80%	2550 79.44%	1812 56.44%
Test set	196 51.86%	443 26.22%	133 7.87%	47 2.78%	15 0.89%	196 11.60%	377 22.32%	893 52.87%	196 11.60%

Table 1. Distribution of DR-related ocular lesions and image-level labels from the customized Kaggle EyePACS (training test) and MESSIDOR-2 (test set) data sets.

A summary with the number of images and percentage per finding and image-level for the customized Kaggle EyePACS dataset which is used as training set and for the MESSIDOR-2 dataset, which is used as test set is presented in Table 1.

Dataset Statistics:

(a) Most images in the Seg-set contain one or more kinds of lesions annotated. We observe that microaneurysms, haemorrhages, and hard exudates are the three most common lesions in DR images, while intra-retinal microvascular abnormalities, neovascularization, laser marks, and proliferate membranes rarely appear.

(b) Since all the samples in the Seg-set are coarsely selected through a pre-trained grading model, the ratios of grade 0 and 1 are low. More specifically, Seg-set has 1,842 images ([‘grade’: the number of images] ‘0’: 101, ‘1’: 212, ‘2’: 595, ‘3’: 647, ‘4’: 287), and Grade-set has 1000 images (‘0’: 143, ‘1’: 125, ‘2’: 566, ‘3’: 105, ‘4’: 61).

(c) We also illustrate various lesion distributions related to the five grading levels with normalization. As shown, microaneurysms are the first DR lesions to appear usually starting in the early stages (grade-0 or grade-1). Moreover, the number of all lesions generally grows as the DR grading level increases. Although it is difficult to differentiate stages 3 and 4, only based on lesion distributions, we observe that neovascularization, laser marks, and proliferate membranes are good factors for further discrimination.

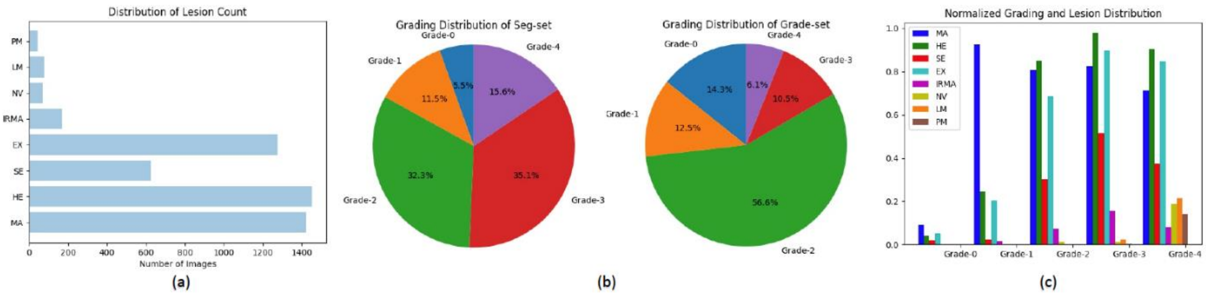


Figure 2. Statistics of our FGADR dataset. (a) Number of images for each pixel-level annotated lesion. (b) The left and right pie charts illustrate the grading distribution of the Seg-set and Grade-set, respectively. (c) Lesion distribution normalized by the number of images for different grades.

The DR-related findings of MA, H, and CWS are the most common lesions as shown in Table 1. Moreover, these ocular findings are related to appear in the initial stages of subjects with DR. On the other hand, VB and NV findings presented few examples during the screening, but also these findings are scarce and rare to find because they are common in an advanced grade of DR. Finally, the Referable condition was the most common image-level labels among the datasets. The comparison of our dataset with others publicly available with fined grained labels of DR related lesion is presented in Table 2. It can be noticed that the available datasets have fewer examples of DR related lesions, this is due to labelling images at a lesion level is costly, tedious and time consuming. We hope that this new set of labels open opportunities to improve and develop new approaches for the detection of retina lesions caused by DR and for the detection of DR based on them, which leads to more relatable models for the clinicians.

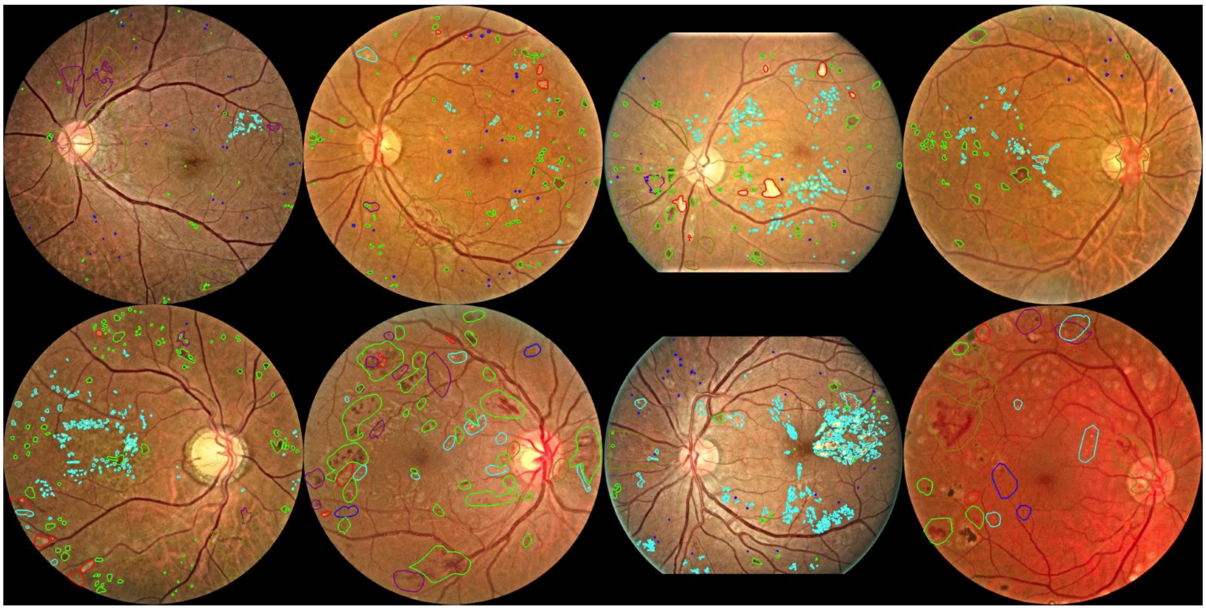


Figure 2. Pixel-level annotation examples from our FGADR dataset, including six different lesions. The blue, green, red, cyan, purple, and olive denote microaneurysms, haemorrhages, soft exudates, hard exudates, intra-retinal microvascular abnormalities, and neovascularization, respectively

Database	MA	H	CWS	VB	NV	EX	Others
DRIVE		X				X	X
DIARETDB0	X	X	X		X	X	
DIARETDB1	X	X	X			X	
STARE	X				X		X
IDRiD	X	X	X			X	
FGADR	X	X	X		X	X	X
e-optha	X					X	
Our dataset	X	X	X	X	X	X	

Table 2. Comparison of DR-related lesion labels from public databases. The DR-related lesion are: microaneurysm (MA), haemorrhages (H), cotton wool spots (CWS), venous beading (VB), neovascularization (NV), exudates (EX), and other lesions.

CHAPTER 3
PROPOSED WORK

3.1 SOLUTION APPROACH

In residual networks each building block is of two layer or three-layer depth. Small networks like ResNet (18, 34) are usually made of two-layer blocks, whereas large networks like ResNet (50, 101, 152) are made up of three-layer blocks. This model bypasses a few convolution layers at a time and the shortcut connections between the layers create a residual block, where the convolutional layer's output is added to the input tensor of the block.

In order to solve the problem of the vanishing/exploding gradient, this architecture introduced the concept called Residual Blocks. In this network, we use a technique called skip connections. The skip connection connects activations of a layer to further layers by skipping some layers in between. This forms a residual block. ResNets are made by stacking these residual blocks together.

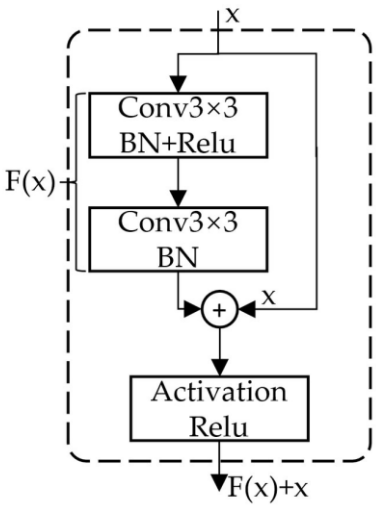


Figure 3. Residual block

The advantage of adding this type of skip connection is that if any layer hurt the performance of architecture, then it will be skipped by regularization. So, this results in training a very deep neural network without the problems caused by vanishing/exploding gradient.

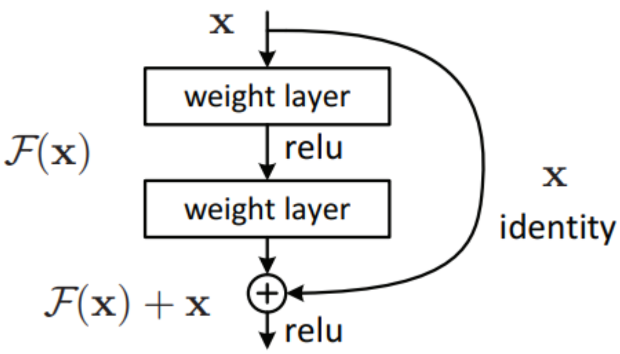


Figure 4. Residual block with a skip connection

Network Architecture: This network uses a 34-layer plain network architecture inspired by VGG-19 in which then the shortcut connection is added. These shortcut connections then convert the architecture into a residual network.

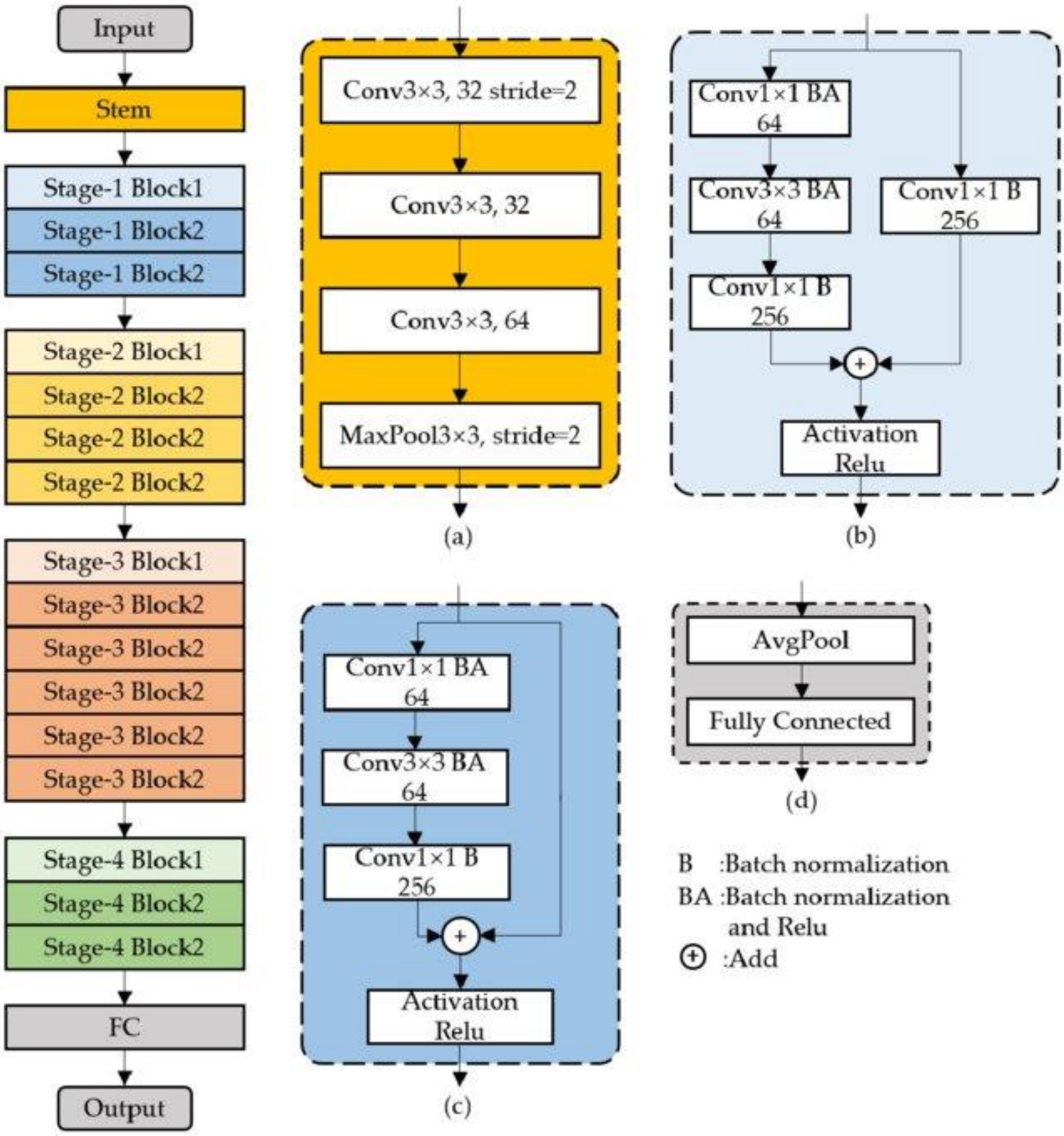
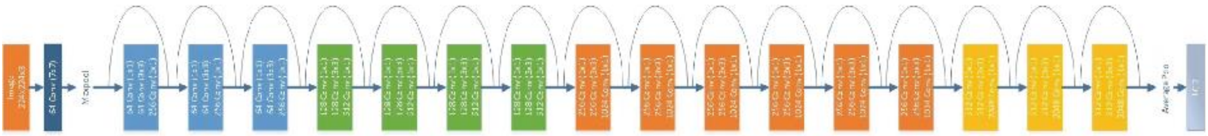
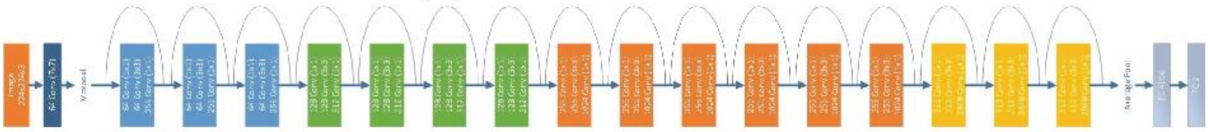


Figure 5. The architecture of ResNet-50-vd. (a) Stem block; (b) Stage1-Block1; (c) Stage1-Block2; (d) FC-Block.

The study presents modified residual network architecture for classification of DR using fundus images.



(a) Resnet50 Architecture



(b) Modified Resnet50 Architecture

Figure 6. ResNet-50 Architecture

In the second stage the predictions of the lesions are concatenated to train multiple classifiers for the DR classification task. The best results are reported and compared with state-of-the-art methods using the MESSIDOR-2 dataset as test set. The models and datasets used in this work will be released in a public repository to ensure reproducibility and encourage other researchers to create new methods. The classifiers and the parameters explored in both stages are:

1. **Support vector machine (SVM):** the parameters explored are the regularization parameter (C), different kernels and the gamma (γ) kernel coefficient.

2. **Gaussian Process (GP):** the kernels 'RBF' and 'Mattern' are evaluated. The kernel parameters explored are lower length scale bound, upper length scale bound, lower noise level and upper noise level.

3. **Multilayer Perceptron (MLP):** the number of layers, number of neurons, activations and learning rate are explored.

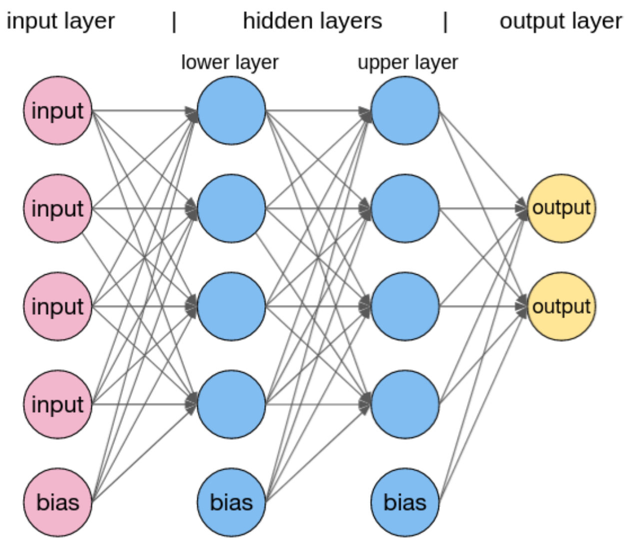


Figure 7. Multilayer Perceptron

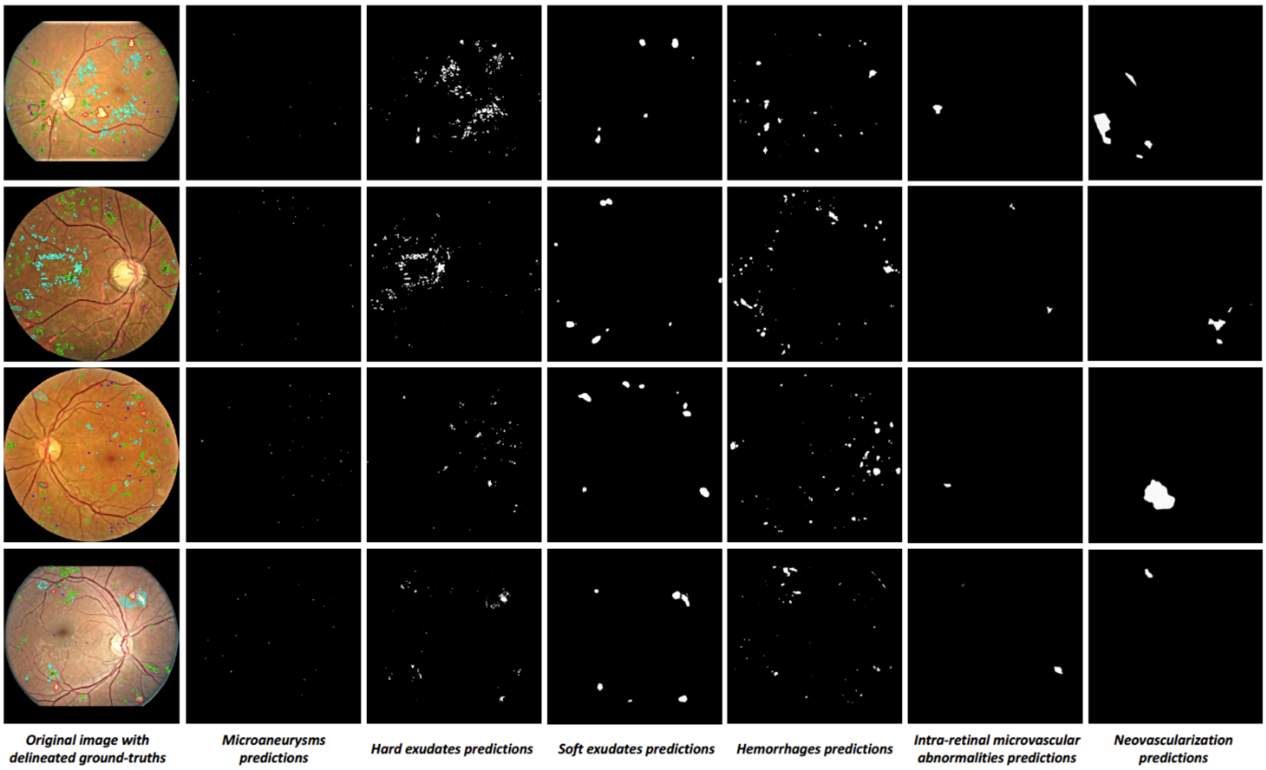


Figure 8. Qualitative segmentation performance of the ResNet50. All the mask outputs are binarized using a threshold of 0.25 for visualization

3.2 TECHNOLOGY STACK

3.2.1 REACT

React is a free and open-source front-end JavaScript library for building user interfaces based on UI components. React can be used as a base in the development of single-page or mobile applications. In our project, we used React as the main framework. It defines the structure of the application, and all other technologies are chosen around it.

3.2.2 CSS

Cascading Style Sheets (CSS) is a style sheet language used for describing the presentation of a document written in a markup language such as HTML. CSS is a cornerstone technology of the World Wide Web, alongside HTML and JavaScript.

CSS is designed to enable the separation of presentation and content, including layout, colours, and fonts. This separation can improve content accessibility; provide more flexibility and control in the specification of presentation characteristics; enable multiple web pages to share formatting by specifying the relevant CSS in a separate .css file, which reduces complexity and repetition in the structural content; and enable the .css file to be cached to improve the page load speed between the pages that share the file and its formatting.

3.2.3 HTML

The Hypertext Markup Language or HTML is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript.

Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document.

3.2.4 TENSOR FLOW

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. TensorFlow can be used in a wide variety of programming languages, including Python, JavaScript, C++, and Java. This flexibility lends itself to a range of applications in many different sectors.

3.2.5 KERAS

Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Keras contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code.

3.2.6 NUMPY

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

3.2.7 FLASK

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions.

3.2.8 PILLOW

Python Imaging Library is a free and open-source additional library for the Python programming language that adds support for opening, manipulating, and saving many different image file formats. Python pillow library is used to image class within it to show the image. The image modules that belong to the pillow package have a few inbuilt functions such as load images or create new images, etc.

3.2.9 CV2

CV2 or OpenCV (Open-Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

CHAPTER 4
IMPLEMENTATION DETAILS

4.1 FRONTEND IMPLEMENTATION

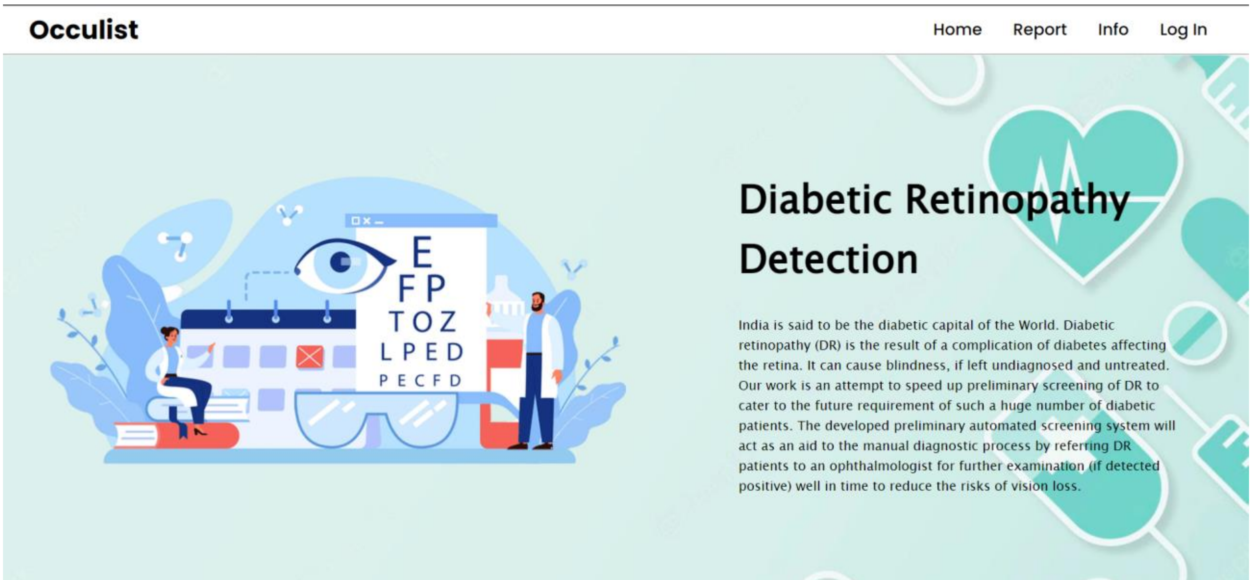
The Front End of this application was developed using React Library in JavaScript, HTML and CSS. React is a free and open-source front-end JavaScript library which makes it possible to write our own components.

HTML is the standard Markup language for web pages and CSS is used to add styling, animations, responsiveness and other media queries to the page.

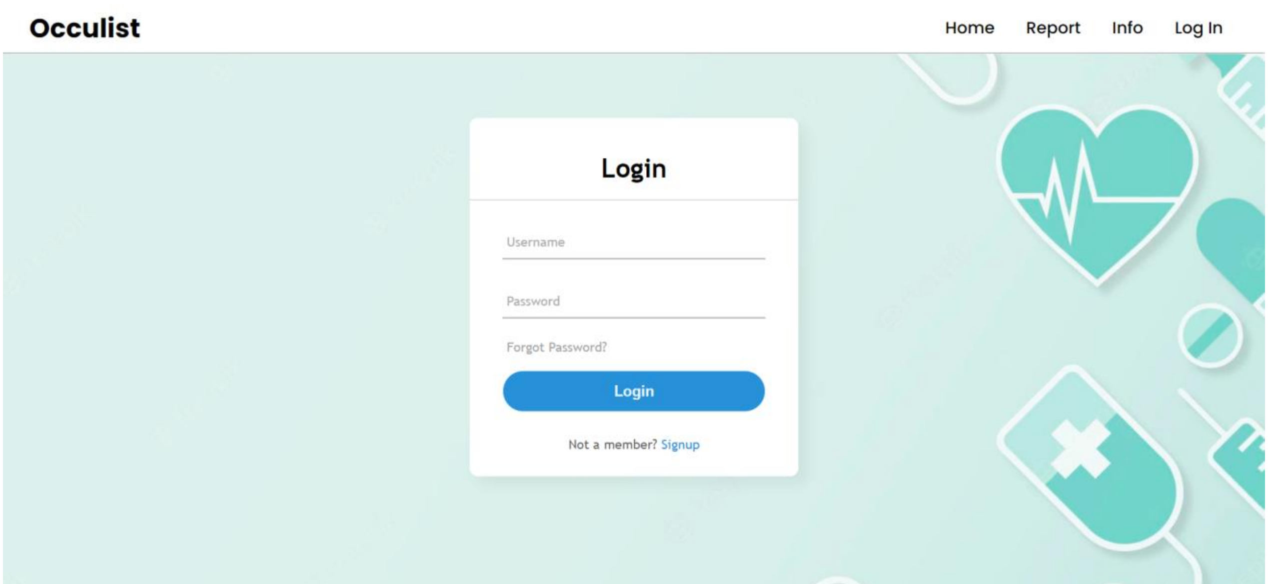
The main components include the NAVBar, forms, buttons which have responsive UI, on hove UI and flexboxes.

JavaScript makes web pages dynamic. Using JavaScript enabled us to change the static elements of the website.

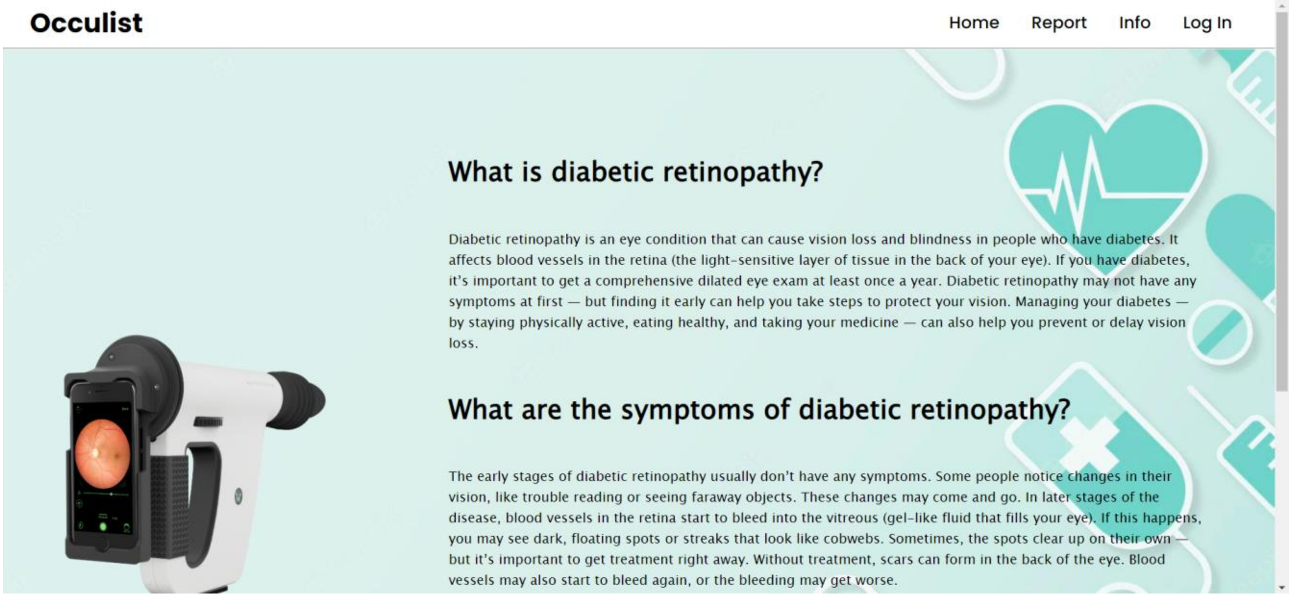
4.2.1 HOME PAGE



4.2.2 LOGIN PAGE

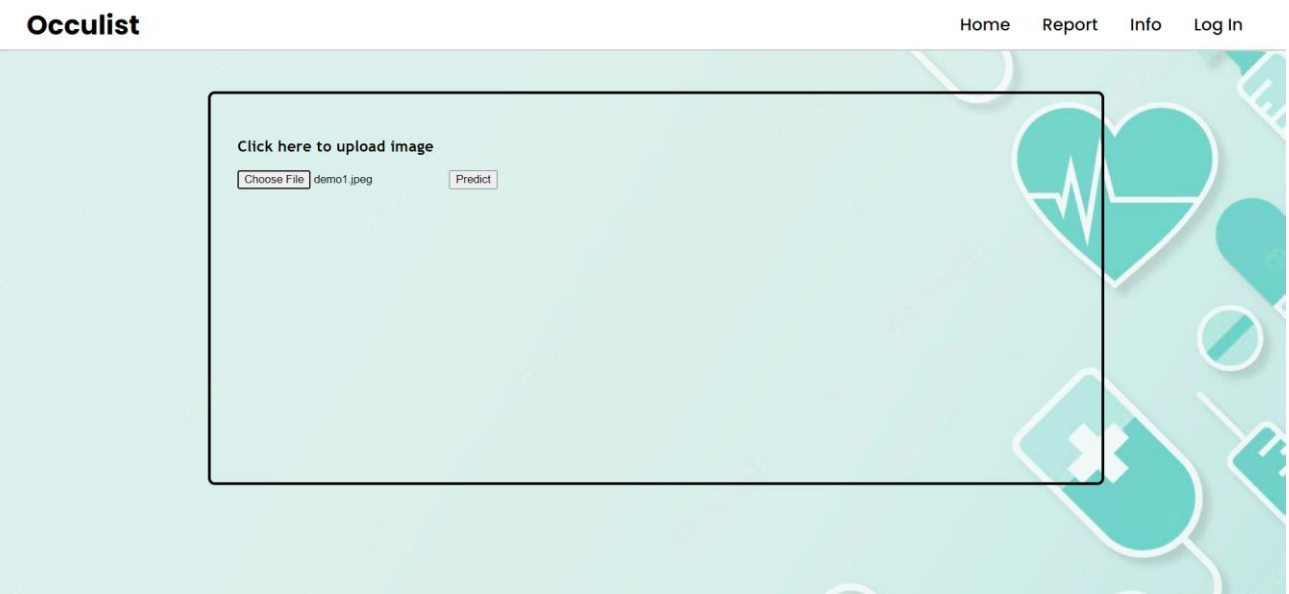


4.1.3. INFORMATION PAGE

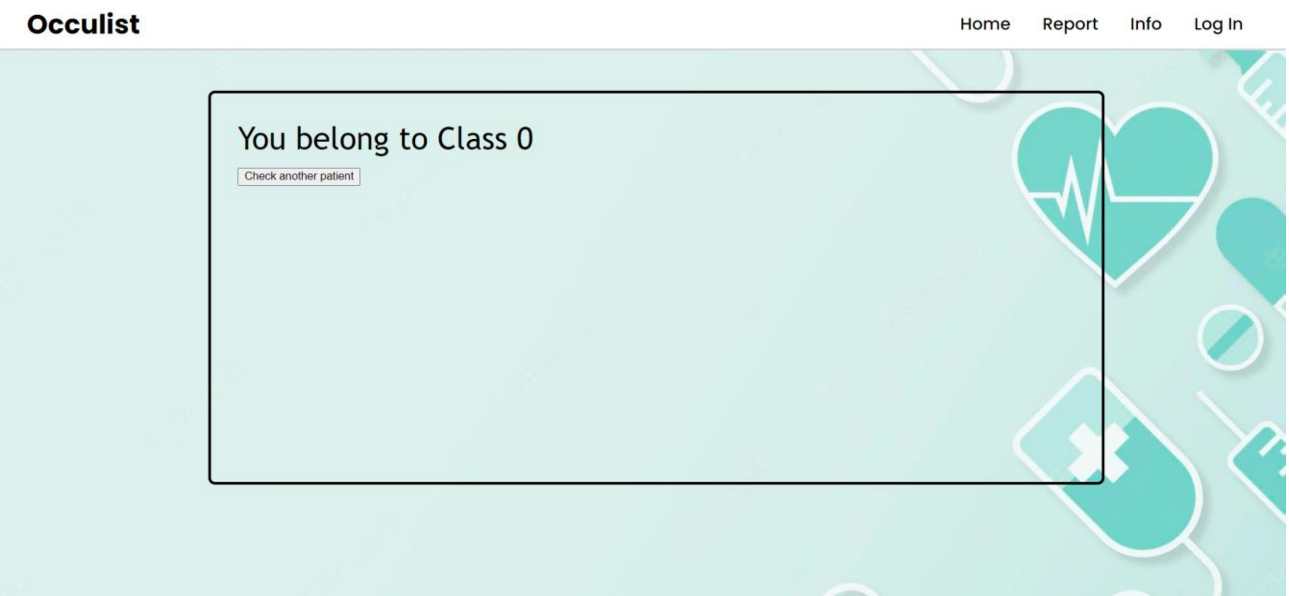


4.1.4 REPORT PAGE

The upload option in report page:



The prediction page after uploading:



4.2 BACKEND IMPLEMENTATION

The backend was done using python and flask, which is a micro-web Python web framework that enables rapid development of secure and maintainable websites.

Passwords generated during signup were secured using PBKDF2 with SHA256 hash. SQLite3 was used to store the data.

In order to run the machine learning model at the backend several python modules were imported. The models were pickled using python pickle library and the later loaded to load the already trained model.

The image received from the front end was base64 encoded, which was converted to a python ‘PIL.Image’ object using NumPy and OpenCV.

4.3 INTEGRATION

API stands for application programming interface. APIs are mechanisms that enable two software components to communicate with each other using a set of definitions and protocols. API acts as a bridge between our frontend and our ML-Model. The flowchart below shows the sequence of actions taking place.

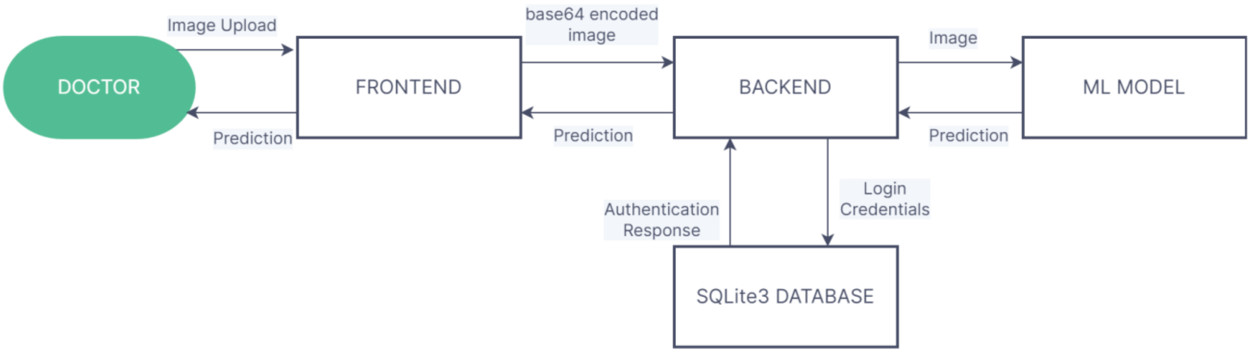


Figure 9. Process Flow Diagram

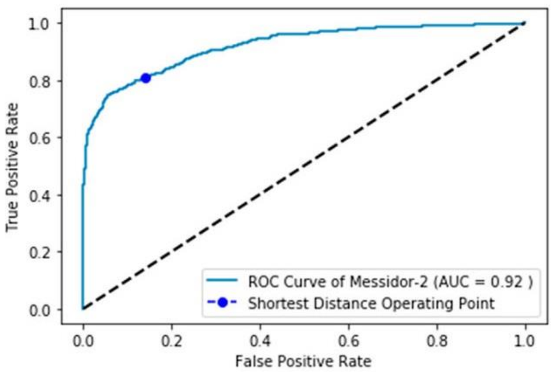
CHAPTER 5
RESULTS AND DISCUSSION

We trained multiple models on the EyePACS train set. Each model has trained on a different train (80%) and validation sets (20%), selected from the EyePACS train set using random splitting. We combined the results of these models by using ensemble averaging methods. The ensemble of models has evaluated on separate datasets other than the dataset used for training; i.e., EyePACS test set an MESSIDOR-2 test sets. Our algorithm scored an AUC of 0.92 on MESSIDOR-2. We have used the shortest distance as the operating point selection strategy for the ROC curve. We obtain Sensitivity and Specificity for MESSIDOR-2 as81.02% and 86.09%, respectively. For the EyePACS test dataset from Kaggle (14,210 images), our algorithm has scored an AUC of 0.927, Sensitivity and Specificity of 83.74% and 89.65%, respectively at the mentioned operating point. The ROC curves for the same are available in Figure 5. The dotted line in Figure 5 represents a random chance classifier while the curve represents the model's trade-off between True Positive Rate and False Positive Rate. The dot on the curve represents the shortest distance threshold.

Test Dataset	Sensitivity	Specificity	Area under the Receiver Operating Characteristic Curve
EyePACS	84.7%	89.65%	0.927
MESSIDOR-2	81.02%	86.09%	0.92

Table 3. represents the results of the studies (including our study) that used non-adjudicated datasets for training

(A)



(B)

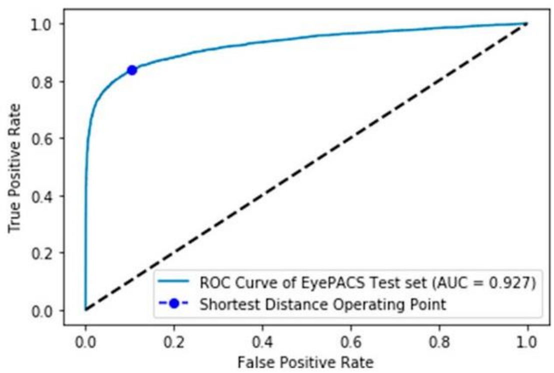


Figure 10. Shows the ROC curves (A) MESSIDOR-2 and (B) EyePACS test sets

CHAPTER 6
CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

We have engineered, trained and validated deep learning-based models for the development of a preliminary automated detection system for the detection of Diabetic Retinopathy in colour fundus images of the retina. We structured the models based on Convolutional Neural Networks and trained them over publicly available datasets. Our models achieved notably better results as compared to the previous studies on the same public datasets. It makes Convolutional Neural Networks a leading choice for tasks like automated detection of ailments in digital medical images. Further, less model inference time and batch processing of images make the said automated system a suitable choice for preliminary screening of a large number of patients. It makes the screening process efficient, addresses the requirement of frequent screening of patients and in-time treatment which prevents vision loss. There is a scope of improvement regarding the quality of datasets. The quality of public datasets is not-so-good because they may contain noise in images and labels. The images may contain artifacts, are out of focus, over-/under-exposed, etc. Noise in the labels comprises of missing/wrong/inconsistent information. With the help of consensus (through adjudication), one can make the labels consistent and free from noise. However, pre-processing and visual inspection helps to remove noise from images. Datasets used in most of the studies were cleaned and graded by multiple ophthalmologists (consensus). It makes them free from noise in images and labels, thereby enabling classifiers to train and score better on them. We have experimented with multiple pre-processing methods and hyper-parameters. We found that pre-processing methods like intelligent cropping (remove black border and centring), random brightness change (~12–13%) and random contrast change (~20%) give better results. Also, the optimal image size is nearly 512 x 512 pixels. Lastly, we have found that residual connections in Inception based models significantly improve the results.

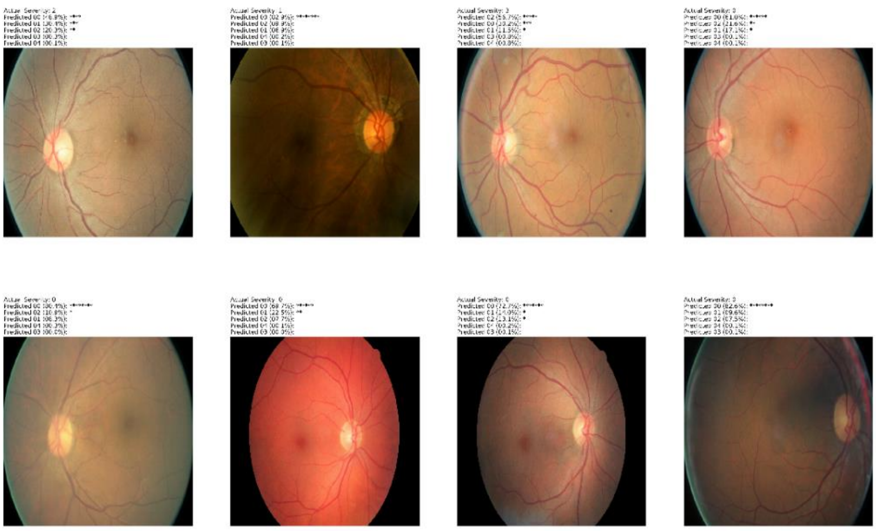


Figure 11. Final image output

6.2 FUTURE WORKS

We can enhance the performance of classifiers by making custom models built for specific purposes. Our study uses general-purpose models. The model was designed to work excellent for the conditions where the image contains varying size features. This is not the case with digital medical images. The size of the features in these images is almost the same across datasets. Here, a general-purpose model may or may not perform better under limited resources. So initially, one can use general-purpose models to get optimum size kernels and later can build custom models using those kernels. The depth of the model can also be optimized similarly. This will reduce the number of parameters(weights) to be trained, enhance targeted learning, increase the model's performance and provide greater insights into the problem. We claim that our classifier worked as a preliminary automated detection system for early detection of DR. We have achieved this by making a classifier to distinguish the images with DR grades 1, 2, 3, 4 vs grade 0 (binary classification). We can enhance the classifier and strengthen the claim in two ways. First, we can train a binary classifier that will distinguish the DR grades 1 and 2 vs grade 0. This will capture the features for early detection in the image more prominently. Here, early detection is the key to prevent vision loss. Second, we can train a multi-class classifier to predict the DR grade. This will make the system usable in conditions other than early detection. We conducted experiments for binary classification of DR grade 1 and 2 vs grade 0, and multi-class classification, and presents our preliminary results on the MESSIDOR-2 data set. For binary classification, we obtained an AUC value of 0.91, and for multi-class classification, an AUC value of 0.95 has been obtained. The experimental details are presented in the Appendix. Lastly, we think that the generation of heat maps can bring out the most relevant portions of the image that cater to the classification decision. This will help novice users and patients to better understand the problems in images. This will also help in the training of medical practitioners.

REFERENCES

1. Modified U-Net architecture for semantic segmentation of diabetic retinopathy images retrieved from: https://www.sciencedirect.com/science/article/pii/S0208521620300747
2. Modified residual networks for severity stage classification of diabetic retinopathy retrieved from: https://www.researchgate.net/publication/358731247_Modified_residual_networks_for_severity_stage_classification_of_diabetic_retinopathy
3. A Benchmark for Studying Diabetic Retinopathy: Segmentation, Grading, and Transferability retrieved from: https://arxiv.org/pdf/2008.09772.pdf
4. Aggregated residual transformation network for multistage classification in diabetic retinopathy retrieved from: https://onlinelibrary.wiley.com/doi/10.1002/ima.22513
5. https://towardsdatascience.com/blindness-detection-diabetic-retinopathy-using-deep-learning-on-eye-retina-images-baf20fcf409e

Datasets retrieved from:

6. https://paperswithcode.com/dataset/drive
7. https://paperswithcode.com/dataset/stare
8. https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid
9. https://www.it.lut.fi/project/imageret/diaretdb1/
10. https://www.adcis.net/en/third-party/e-ophtha/
11. https://www.kaggle.com/datasets/mariaherrerot/aptos2019
12. https://paperswithcode.com/dataset/chase-db1
13. https://paperswithcode.com/dataset/messidor-1
14. https://www.adcis.net/en/third-party/messidor2/
15. https://www.kaggle.com/datasets/vidheeshnacode/stare-dataset
16. https://www.kaggle.com/datasets/andrewmvd/drive-digital-retinal-images-for-vessel-extraction