SOMAIYA
VIDYAVIHAR

K J Somaiya Institute of Engineering and Information Technology
An Autonomous Institute Permanently Affiliated to the University of Mumbai

DEPARTMENT OF INFORMATION TECHNOLOGY

Synopsis of Minor Project On

Diagnosis Of Diabetic Retinopathy Using Machine Learning

Prepared By:

Burhanuddin Dilshad (Roll No. 11)
Viraj Gholap (Roll No. 16)
Samridh Gupta (Roll No. 17)

Under the guidance of:

Dr. Mansingh Rathod

Department of Information Technology

Academic Year: 2021-2022

Autonomy Syllabus Scheme-I (2021-22) - SEMESTER VI (TY - IT)

SOMAIYA
VIDYAVIHAR
K J Somaiya Institute of Engineering and Information Technology
An Autonomous Institute Permanently Affiliated to the University of Mumbai
Accredited by NAAC with 'A' Grade (3.21 CGPA), Approved by AICTE, New Delhi

DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that following students:

	Roll No. / Seat No.
Burhanuddin Dilshad	(Roll No. 11)
Viraj Gholap	(Roll No. 16)
Samridh Gupta	(Roll No. 17)

have submitted PBL – Minor Project Lab II (AI-based Minor Project) Report on *"Diagnosis Of Diabetic Retinopathy Using Machine Learning"* as the partial fulfillment for the requirement of Third Year of Engineering (6rd Semester) in T.Y. - Information Technology under my guidance during the academic year 2021-2022.

Dr.Mansingh Rathod
Project Guide
Assistant Professor
Department of Information Technology

Dr. Radhika Kotecha
Head of Department
Associate Professor
Department of Information Technology

Date of Examination: _________________

Signature of Internal Examiner

Signature of External Examiner

Table of Contents

Acknowledgement

We would like to convey our sincere thanks to everyone who guided us throughout this project endeavor, before we present our project entitled "**Diagnosis Of Diabetic Retinopathy Using Machine Learning**". We wish to express our sincere thanks to Dr. Suresh Ukarande, our Principal, and Dr. Sunita Patil, our Vice-principal for providing us with all the necessary facilities for the research. We place on record our sincere thank you to Dr. Radhika Kotecha, Head of Department, for the continuous encouragement and support. We are also grateful to our project guide Dr. Mansingh Rathod for mentoring us.

We are extremely thankful to her for sharing expertise, and sincere and valuable guidance and encouragement extended to us. We take this opportunity to express gratitude to Prof. Nasim Shah, Mini-Project coordinator for her valuable guidance, help and support. We also thank our parents for unceasing encouragement, support and attention.

Abstract

Diabetic Retinopathy (DR) is a common complication of diabetes mellitus, which causes lesions on the retina that effect vision. If it is not detected early, it can lead to blindness. Unfortunately, DR is not a reversible process, and treatment only sustains vision. DR early detection and treatment can significantly reduce the risk of vision loss. The manual diagnosis process of DR retina fundus images by ophthalmologists is time-, effort-, and cost-consuming and prone to misdiagnosis unlike computer-aided diagnosis systems. Recently, deep learning has become one of the most common techniques that has achieved better performance in many areas, especially in medical image analysis and classification. Convolutional neural networks are more widely used as a deep learning method in medical image analysis and they are highly effective. For this article, the recent state-of-the-art methods of DR color fundus images detection and classification using deep learning techniques have been reviewed and analyzed. Furthermore, the DR available datasets for the color fundus retina have been reviewed. Difference challenging issues that require more investigation are also discussed.

Chapter 1: Introduction

In the healthcare field, the treatment of diseases is more effective when detected at an early stage. Diabetes is a disease that increases the amount of glucose in the blood caused by a lack of insulin. It affects 425 million adults worldwide. Diabetes affects the retina, heart, nerves, and kidneys. Diabetic Retinopathy (DR) is a complication of diabetes that causes the blood vessels of the retina to swell and to leak fluids and blood. DR can lead to a loss of vision if it is in an advanced stage. Worldwide, DR causes 2.6% of blindness. The possibility of DR presence increases for diabetes patients who suffer from the disease for a long period. Retina regular screening is essential for diabetes patients to diagnose and to treat DR at an early stage to avoid the risk of blindness. DR is detected by the appearance of different types of lesions on a retina image. These lesions are microaneurysms (MA), haemorrhages (HM), soft and hard exudates

- Microaneurysms (MA) is the earliest sign of DR that appears as small red round dots on the retina due to the weakness of the vessel's walls. The size is less than 125 µm and there are sharp margins. Michael et al. classified MA into six types. The types of MA were seen with AOSLO reflectance and conventional fluorescein imaging.
- Haemorrhages (HM) appear as larger spots on the retina, where its size is greater than 125 µm with an irregular margin. There are two types of HM, which are flame (superficial HM) and blot (deeper HM)
- Hard exudates appear as bright-yellow spots on the retina caused by leakage of plasma. They have sharp margins and can be found in the retina's outer layers.
- Soft exudates (also called cotton wool) appear as white spots on the retina caused by the swelling of the nerve fiber. The shape is oval or round.

Red lesions are MA and HM, while bright lesions are soft and hard exudates (EX). There are five stages of DR depending on the presence of these lesions, namely, no DR, mild DR, moderate DR, severe DR and proliferative DR

1.1 Motivation

For Diabetic Retinopathy (DR) diagnosis there exist multiple techniques, an ocular manifestation of diabetes that affects more than 75% of patients with longstanding diabetes and is the leading cause of blindness for the age group 20-64. Researches shows that it contributes around 5% of total cases of blindness. WHO estimates that 347 million of world population is having the disease diabetes and about 40-45% of them have some stage of the disease. By seeing below image one can differentiate between image produced by normal eye and DR eye.

(a)Normal Eye

(b) DR affected eye

Fig : Difference between image produced by normal eye and DR eye

There are various factors affecting the disease like age of diabetes, poor control, pregnancy but Researches shows that if we can detect DR in early stage of the diseaseprogression to vision impairment can be slowed or averted.

So the aim of our project is to provide a automated, suitable and sophisticated model using image processing technic we can detect DR at early levels easily so that damage to retina can be minimized.

1.2 Problem Analysis

Diabetic retinopathy(DR) is a disease which causes blindness in people having diabetes. Currently, to detect DR, medical staff has to thoroughly examine images of the retina manually taken by the technique of Fundus photography. This is time consuming. We proposed a model to detect DR using machine learning techniques such as Neural networks to make the detection process automated as well as accurate.

1.3 Objectives

- To make a user friendly interface for users
- To make an early detection of diabetic retinopathy enables medication to be performed to prevent delay visual loss
- To implement an automated detection of diabetic retinopathy using digital fundus images
- To make it easy for patients to diagnosis their eye related disease on their own

1.4 Scope

- The use of AI has high potential to facilitate the way eye disease is diagnosed
- Modified versions can help detect and diagnosis any eye related problem
- Cost effective solutions can be provided for a social cause of blindness in people
- In time of crisis like covid-19 pandemic, this system could turn out to be a boon

Chapter 2: Literature Review

2.1 Related Work

Previous work has been done in using machine learning and various models for automated DR screening.For development of our method and result analysis, we have conducted a literature survey describing DR features and past work done to detect DR.

GiriBabuKandeet al. represented Segmentation of Vessels in Fundus Images using Spatially Weighted Fuzzy c-Means Clustering an algorithm for the extraction of Blood Vessels from Fundus images.They used a set of linear filters sensitive to vessels of different thickness and orientation.A vessel detection methods recently reported in the literature is simple and an experimental evaluation demonstrates excellent performance over global thresholding. Their algorithm were expected to be applicable to a variety of other applications due to its simplicity and general nature.

Bob Zhang et al described a mothod for Detecting Diabetes Mellitus and Nonproliferative Diabetic Retinopathy. In their mathod, theGeometry Features, Color and Texture utilized for detection. The image is captured and preprocessed. Then color features are extracted from pre-processed image. In order to represent the texture feature, the image is separated into eight blocks. The 2-D Gabor filter is applied on surface for compute texturefeatures. Finally the geometry features such as distances, areas are extracted from foreground image. In order to separate NPDR samples from healthy sample, the support vector machine classifier is used. However, it only achieves 80.52% accuracy.

Yuji Hatanaka et al presented the improvement of automatic hemorrhages detection methods using brightness correction on fundus Images. They indicates the importance of developing several automated models for finding out the abnormalities in fundus images. The purpose of this paper was to improve their automated hemorrhage detection model to diagnose diabetic retinopathy. They represented a new method for preprocessing and false positive elimination. They removed false positives by using a 45-feature analysis. To verify their new method, they examined 125 fundus images, including 35 images with hemorrhages and 90 normal images. The sensitivity and specificity for the detection of abnormal cases was were 80% and 88%, respectively. These verified results indicate that their new method may effectively improve the performance of their diagnosis system for hemorrhages.

Vujosevic et al. build a binary classifier on a dataset of 55 patients by explicitly forming single lesion features.The scope of this study is limited in that the dataset.

Wang et al. used a CNN(LeNet-5 architecture) as a feature extractor for addressing blood vessel segmentation. The model has three heads at different layers of the convnet which then feed into three random forests. The final classifier essembledthe random forests for a final prediction achieving an accuracy and AUC on 0.97/0.94 using a standard dataset for comparing models addressing vessel segmentation

Lim et al. where the authors represent building a convolutional neural network for lesion-level classification and then use the learned feature representations for image-level classification. This scope of the study is limited in that the dataset which contains 200 images.

The authors in introduced morphological image processing techniques to extract blood vessel, microaneurysm, exudate, and hemorrhage features and then train an SVM on a data set of 331 images achieving sensitivity 82% and specificity 86%.

Bhat P. S. Acharya et al. reported sensitivity of 90% and accuracy of 90% with a dataset of 140 images using image processing techniques for extracting the area of blood vessels, area of exudates, and texture features which are then fed into a small Neural Network.

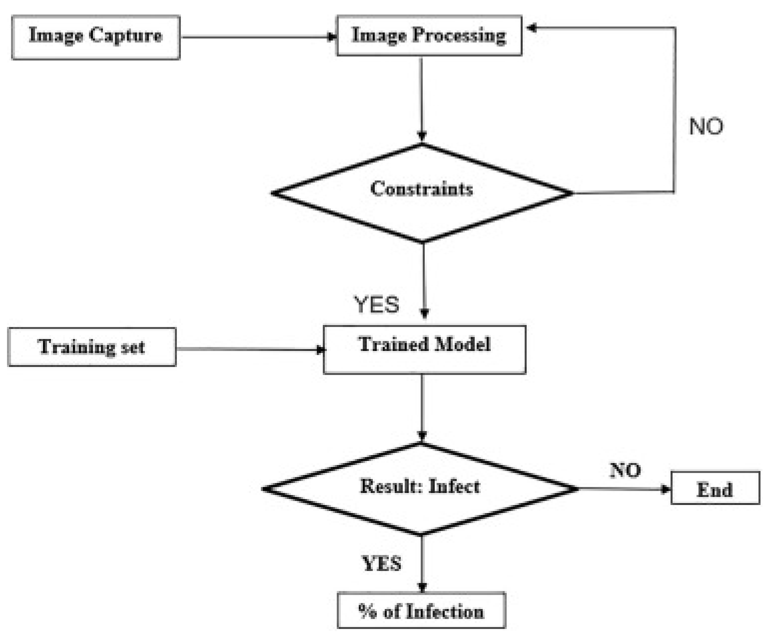
2.2 Existing System

The Automatic Diabetic Retinopathy detection system involves detection and segmentation of the abnormal features from the fundus images. The input image from the retinal image database is first preprocessed to remove noise and to enhance the contrast of the image for further processing. The next step is to remove optic disc and blood vessels. After that, lesions such as microaneurysms, hemorrhages, excudates and cotton wool spots are detected using calculation of different features. These features are analyzed with different techniques to perform severity classification of the disease as normal, mild, moderate, severe Non proliferative Retinopathy and Proliferative Retinopathy. This paper is organized as follows: Section 2 gives a review of various preprocessing methods. In Section 3, techniques for detection and extraction of blood vessels and optic disc are discussed. In section 4 lesion detection and feature extraction methods are discussed. Section 5 gives review of different classifiers used for severity As a result of the external lighting source and usage of different cameras for image acquisition, color fundus images experience variation in illumination change, contrast, shading and resolution. In the preprocessing stage, above mentioned problems are rectified. Different authors have used different techniques according to their requirements and characteristics of images in Datasets they have used. Hoover et al.has applied illumination equalization to the image. Mendonça et al. has performed the image normalization by subtracting an estimate of the background from the monochromatic representation of the original color image followed by thin vessel enhancement by processing with a set of line detection filters, corresponding to the four orientations. The set of convolution kernels used in this operation. For each pixel, the highest filter response is kept and added to the normalized image.

Chapter 3: Proposed System

Here, we are introducing a novel approach for detecting and classifying the diabetic retinopathy disease at the earliest stage by extracting the features such as Hemorrhages and Microaneurysms from the retinal images using Convolutional Neural Network method. So, we attempt to reduce the number of features to two including Microaneurysms which is the earliest symptom of the disease. We exclude the blood vessel feature since it shows symptoms only after Microaneurysms. The main motive of our project is to detect the disease at a stage where we still have possibilities to cure the disease and to locate the symptom showing feature in the retinal images. Since, the patients realizes no symptoms until their vision turns to a little blinder and have only minimum chance of getting cured, the best way to detect the disease is as early as possible which has been covered.

3.1 Proposed Approach and Details



1) Resizing the images The need for resizing is to convert all the varying dimensional (height and width) images to a single dimension. Because the training algorithm takes images which are of same height and width as input for better processing and thus reduces the processing time. Here in our dataset we have random dimensionality images such as 228x221, 998x630 and so on. Hence we resize the images to a height of 96 and width of 96 with depth as 3 that is three channels since we are using color images.

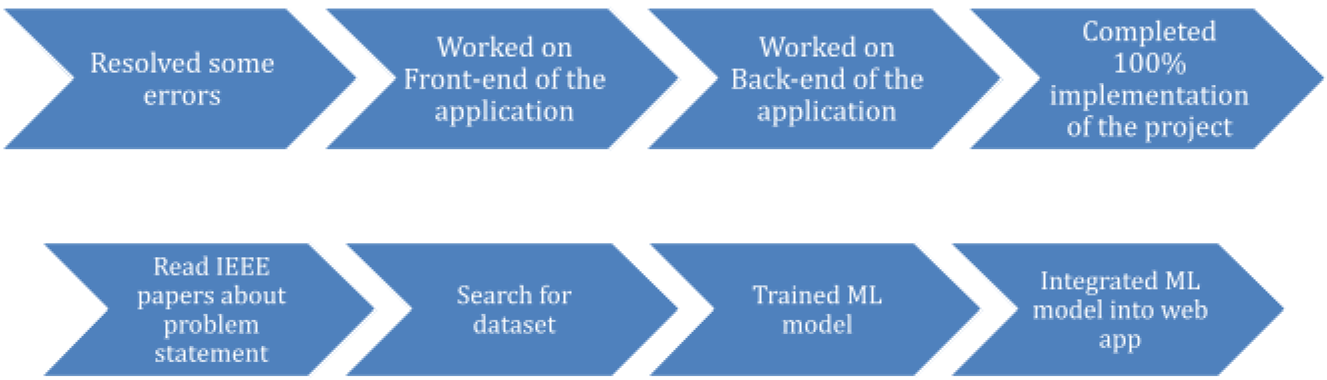
2) Training the model On the whole, we first update the input shape of image as channel first. Then for Smallervggnet we pass parameters such as the convolution window size as 3x3, the number of filters it should built out of it as 32,for the next pass its 64 and for the next pass its 128 and the padding value is marked 'same' to preserve the dimension of the output as same as the input size. Next, the activation function used here ReLu, the Rectified Linear Unit, takes values which are either positive or zero. It also increases the non-linearity of the image. Then, we do batch normalization followed by max pooling of pool size 3x3. At last, the dropout is marked as 0.25 for convolution and pooling layer and it is 0.5 for fully connected layer. Dropout helps to control overfitting. Finally we end up with softmax classifier to return our predicted probabilities for each class label.

3) Testing the model The same process starting from preprocessing till the last layer of Tensorflow is followed again for the test images. Now an operation is performed with the output of the testing image values from the fully connected layer with the values of all filters that was constructed previously in the training part to calculate the probability. The highest probability value is taken as the correct match. Finally we segment the two features from the abnormal images.

3.2 Innovation in Idea

- Higher accuracy test on our web application
- Best in class AI models for users
- Majority of working class population need this technology available at finger tips

3.3 Timeline



3.4 Roles and Responsibilities

Project Team Member Responsibilities

Project team member responsibilities include:

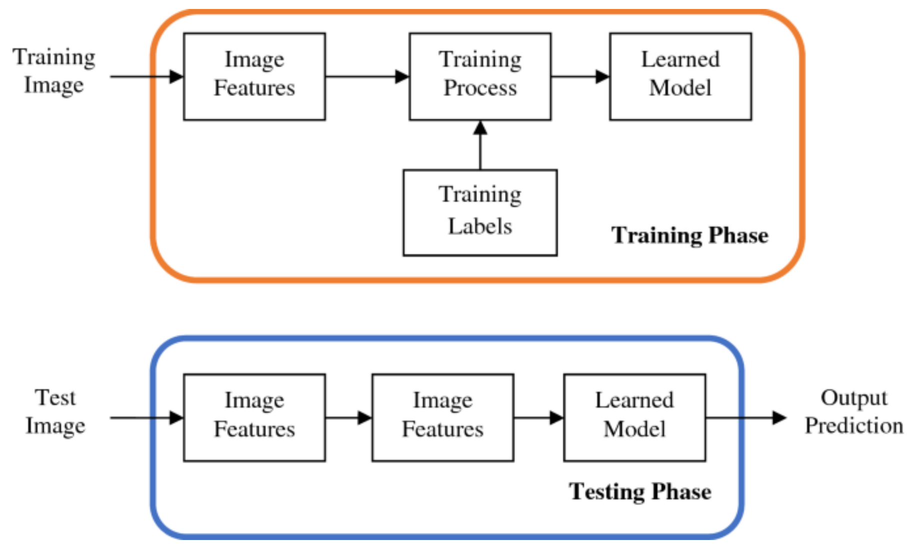
- Contributing to overall project objectives
- Completing individual deliverables
- Providing expertise
- Working with users to establish and meet business needs

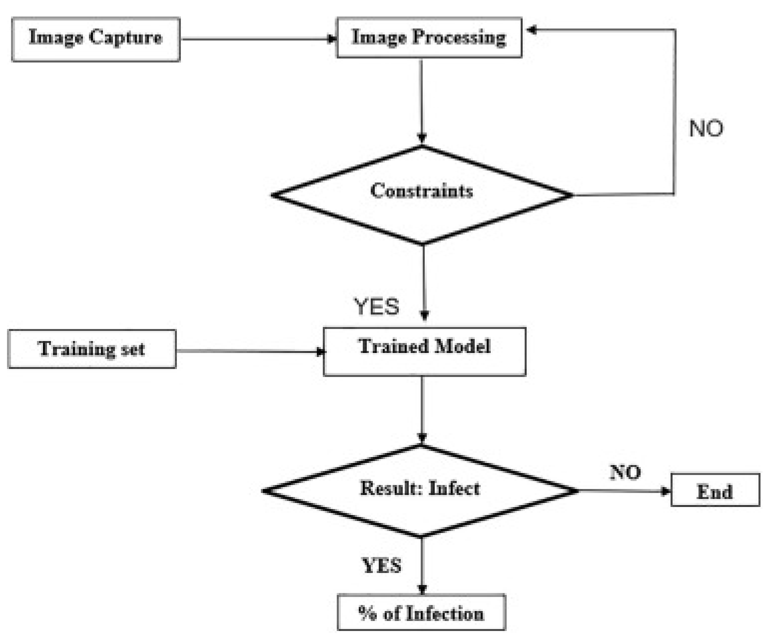
Chapter 4: Implementation Details and Results

4.1 Technology Stack

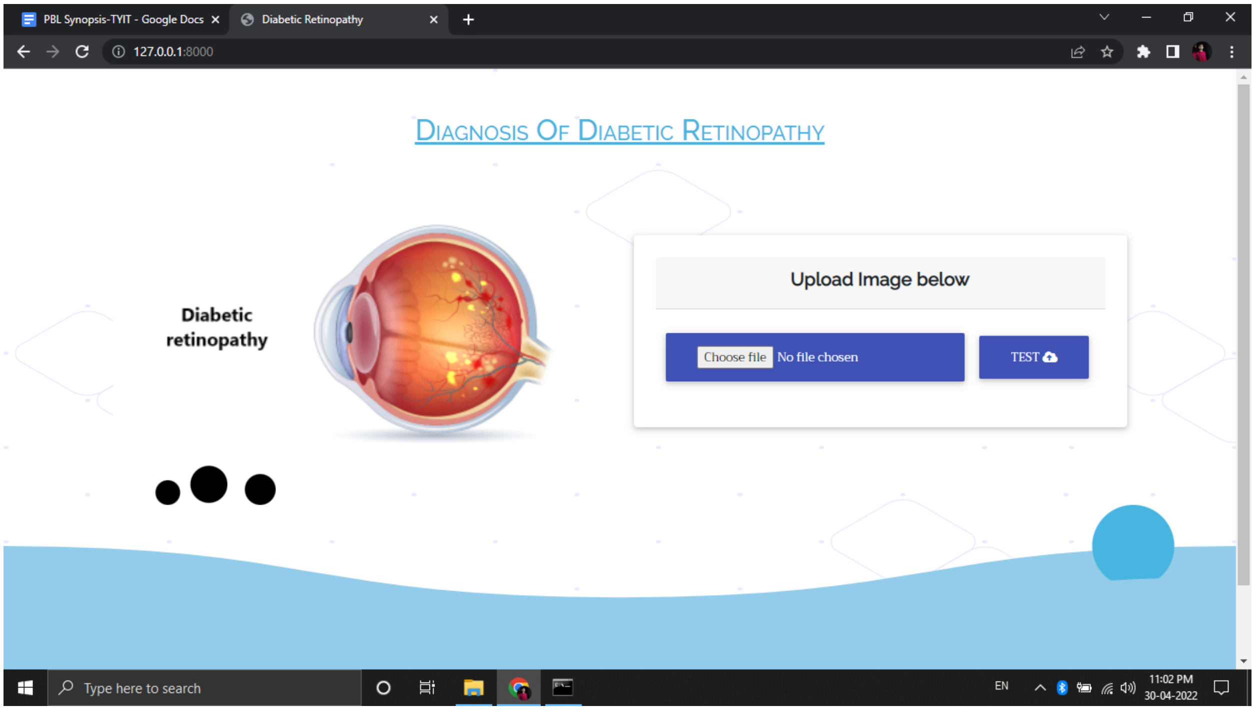
- **Django** - Django is a high-level Python web framework that enables rapid development of secure and maintainable websites. Built by experienced developers, Django takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel. It is free and open source, has a thriving and active community, great documentation, and many options for free and paid-for support.
- **TensorFlow -** TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications.
- **SQLite -** SQLite is a C-language library that implements a small, fast, self-contained, high-reliability, full-featured, SQL database engine. SQLite is the most used database engine in the world.
- **Python: Pillow (a fork of PIL) -** Python Imaging Library (expansion of PIL) is the de facto image processing package for Python language. It incorporates lightweight image processing tools that aids in editing, creating and saving images.
- **NumPy -** NumPy is the fundamental package for scientific computing in Python.

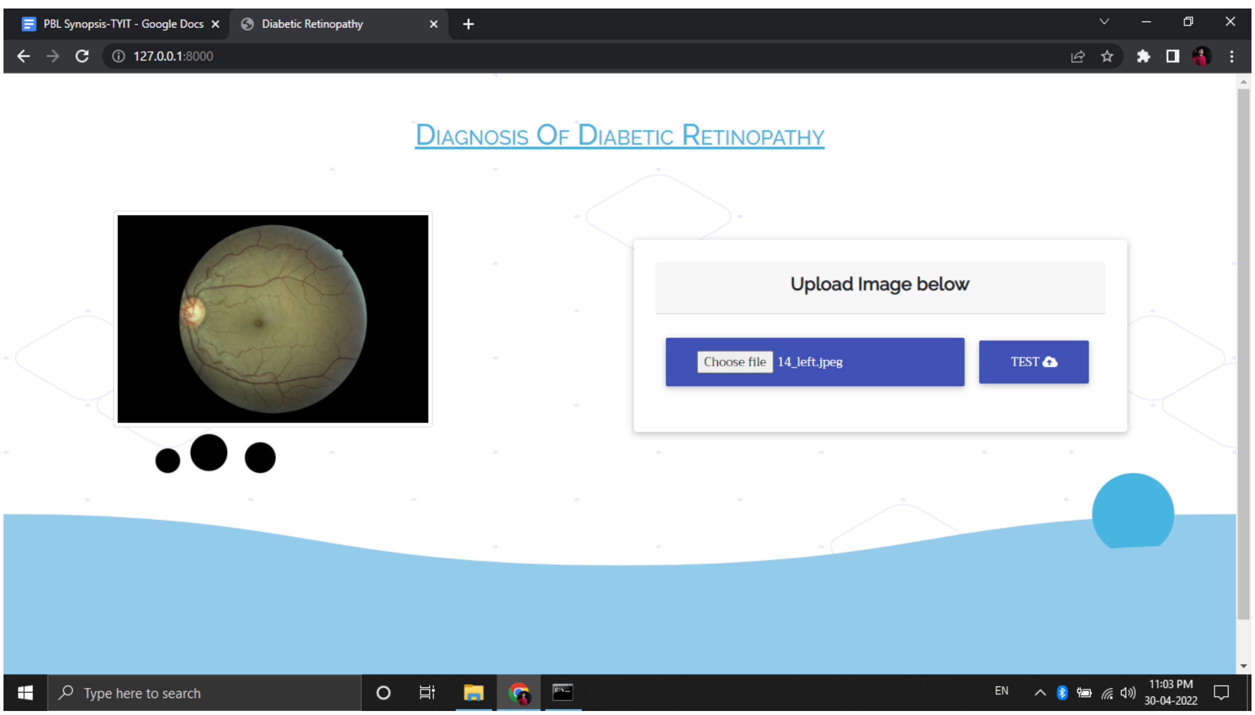
4.2 Implementation Parameters

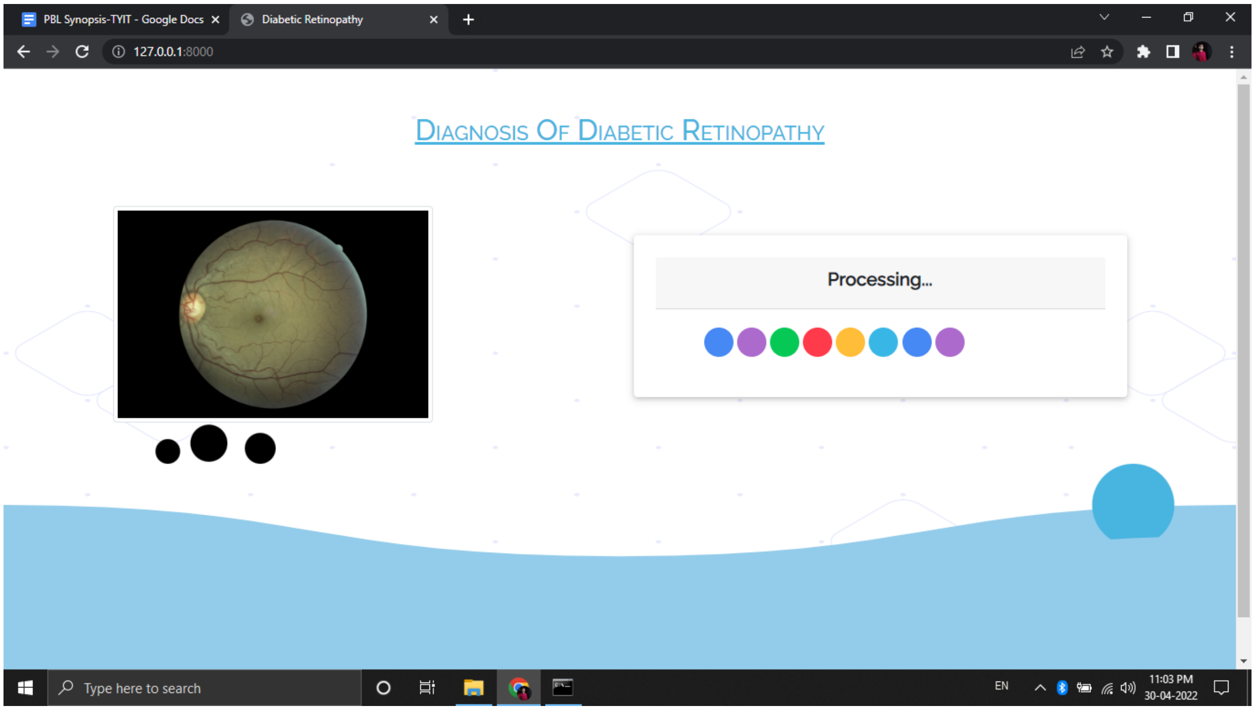




4.3 Preliminary Results







PBL Synopsis-TYIT - Google Docs
Diabetic Retinopathy
127.0.0.1:8000
Diagnosis Of Diabetic Retinopathy
Processing...
Type here to search
EN
11:03 PM
30-04-2022

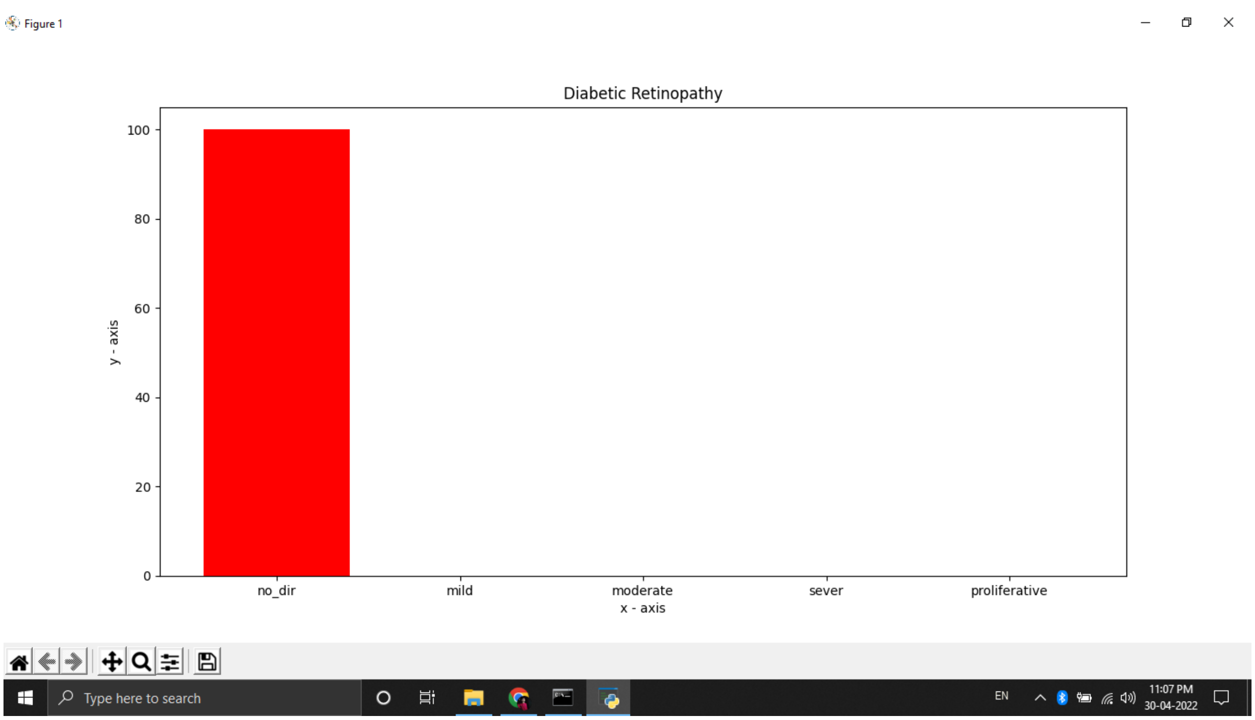
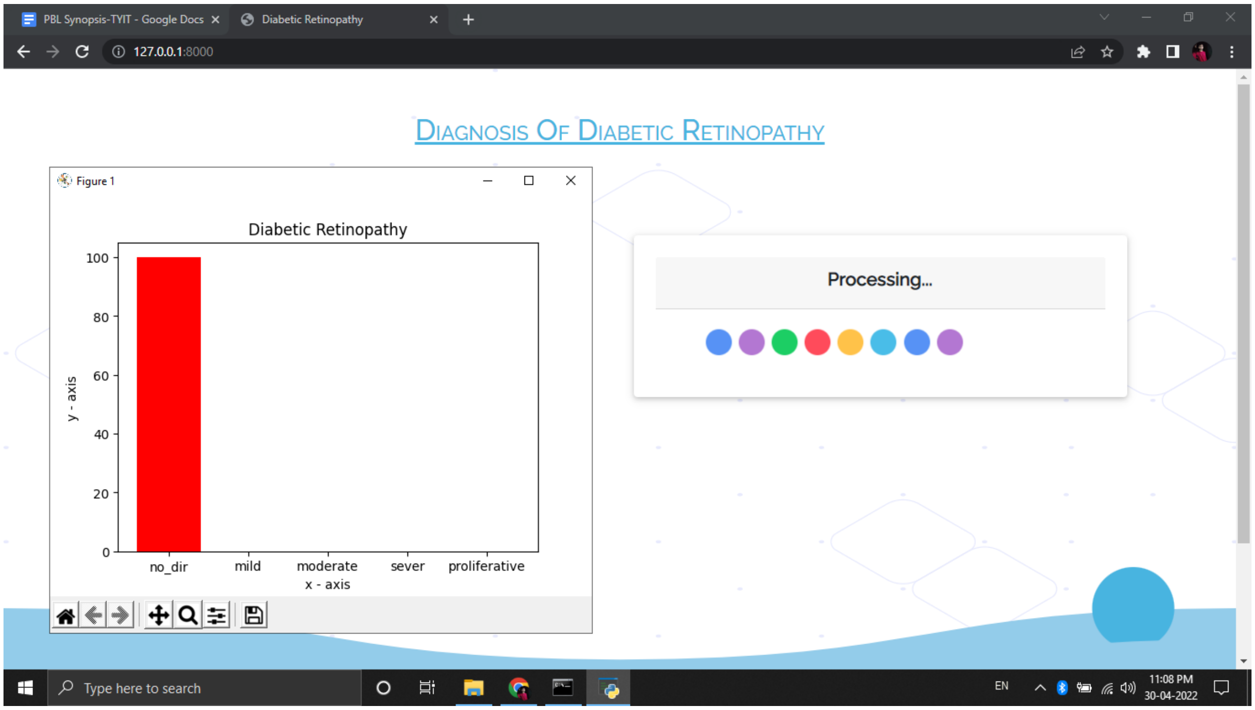
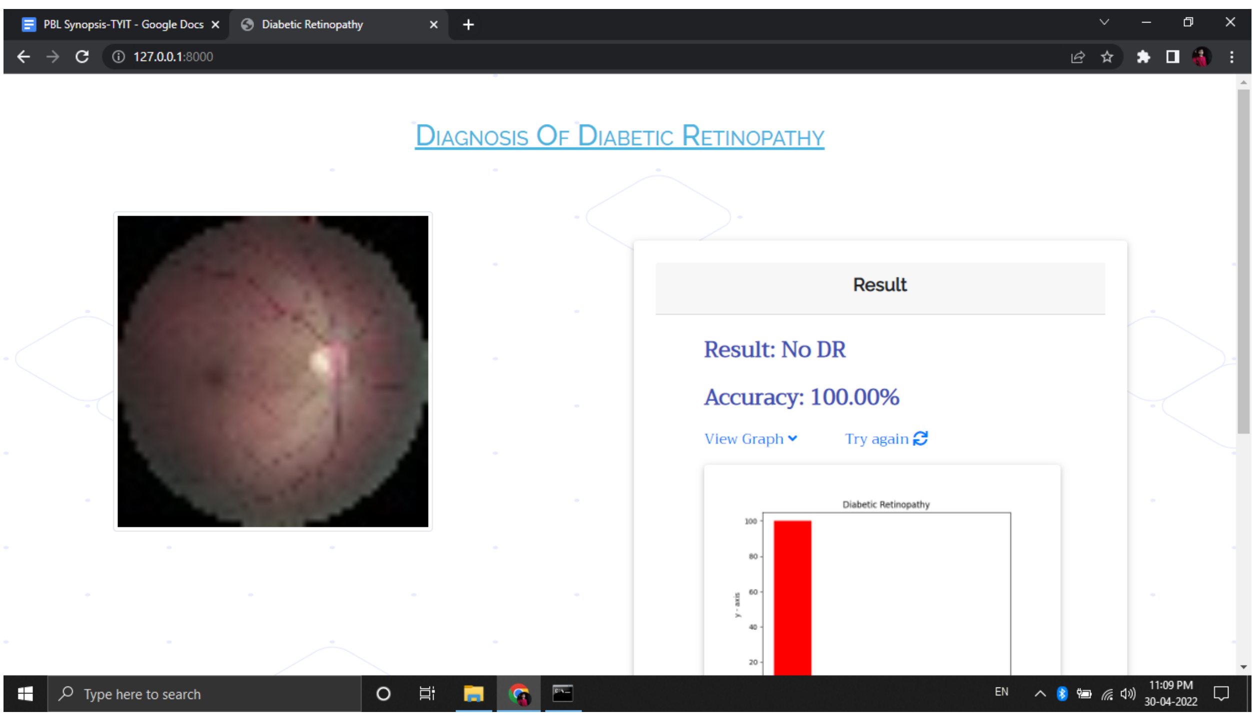
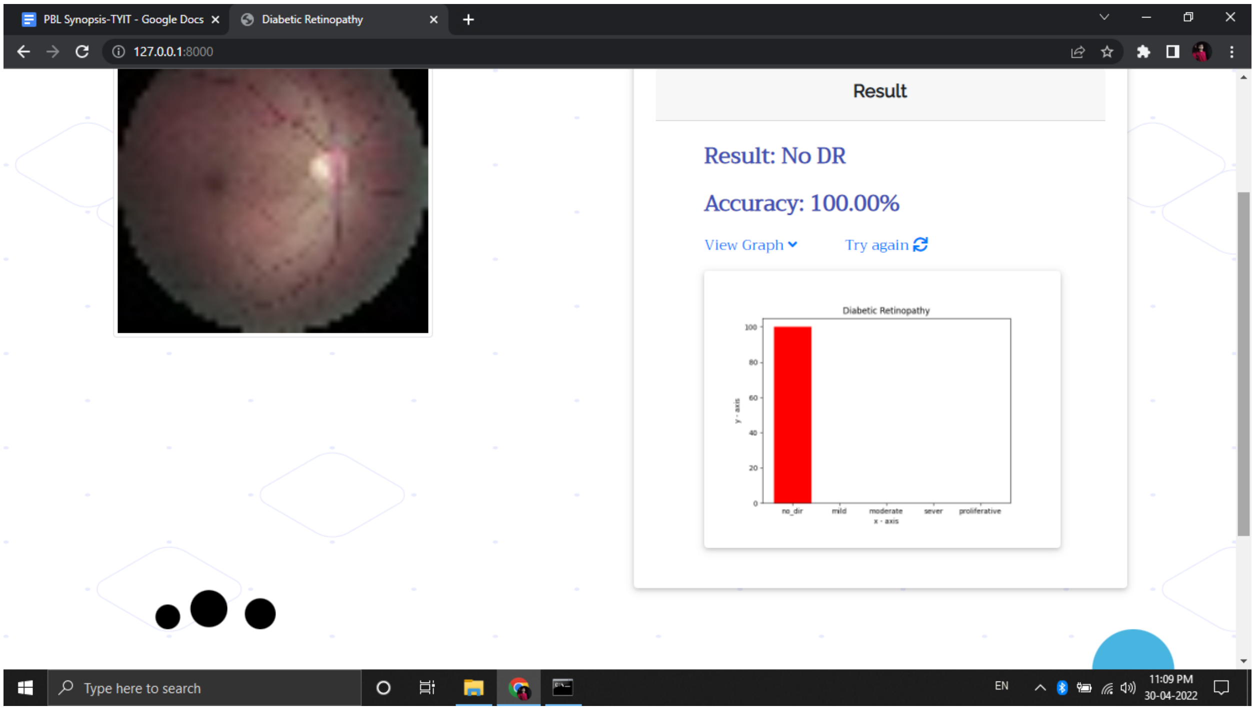


Figure 1
Diabetic Retinopathy
100
80
60
40
20
0
y - axis
no_dir
mild
moderate
sever
proliferative
x - axis
Type here to search
EN
11:07 PM
30-04-2022


PBL Synopsis-TYIT - Google Docs
Diabetic Retinopathy
127.0.0.1:8000
Diagnosis Of Diabetic Retinopathy
Figure 1
Diabetic Retinopathy
100
80
60
40
20
0
y - axis
no_dir
mild
moderate
sever
proliferative
x - axis
Processing...
Type here to search
EN
11:08 PM
30-04-2022


PBL Synopsis-TYIT - Google Docs
Diabetic Retinopathy
127.0.0.1:8000
Diagnosis Of Diabetic Retinopathy
Result
Result: No DR
Accuracy: 100.00%
View Graph
Try again
Diabetic Retinopathy
Type here to search
EN
11:09 PM
30-04-2022


PBL Synopsis-TYIT - Google Docs
Diabetic Retinopathy
127.0.0.1:8000
Result
Result: No DR
Accuracy: 100.00%
View Graph
Try again
Diabetic Retinopathy
y - axis
x - axis
no_dir
mild
moderate
sever
proliferative
Type here to search
EN
11:09 PM
30-04-2022

Chapter 5: Conclusion

Every input image will be classified with the highest accuracy assuring no misclassification happens. We first trained the model to recognize lines, edges, corners etc. from the retinal image followed by recognizing small parts of a single feature which is then followed by recognizing a whole feature. By this way it learnt to identify Microaneurysms and hemorrhage features from the retinal image and thus classified the images accordingly to normal, beginning, mild and severe stages. Since we are extracting a feature which shows signs of the disease at the earliest stage, recovery chances will be high, thus vision of many people could be saved. After this classification, to help the doctors locate those features showing symptoms in the image, we automate marking those features in every classified abnormal retinal image. The proposed idea is not only helpful for Diabetic retinopathy affected people but can also be used by the melanoma and myeloid leukemia disease affected people. One more advantage of this system is, it can be implemented to achieve higher accuracy even with less number of dataset

Chapter 6: References

[1]Meindert Niemeijer,"Retinopathy Online Challenge: Automatic Detection of Microaneurysms in Digital Color Fundus Photographs",IEEE Transactions on Medical Imaging vol.29,no.1, January 2010.

[2] D. Klein, B. E Klein, S. E Moss et al "The Wisconsin epidemiologic study of diabetic retinopathy VII. Diabetic non proliferative retinal lesions", Br. J Ophthalmol, vol. 94, 1986.

[3]Michael Goldbaum,Saied Moezzi, Adam Taylor, Shankar Chatterjee, Jeff Boyd, Edward Hunter, and Ramesh Jain, "Automated diagnosis and image understanding with object extraction, object classification, and differencing in retinal images", Br. J Ophthalmol, vol 83, august 1999.

[4]C.Sinthanayothin, J.Boyce,H. Cook, and T. Williamson, "Automated localisation of optic disc, fovea, and retinal blood vessels from digital colour fundus images", Br. J Ophthalmol, vol. 83, august 1999.

[5]Anil K. Jain Michigan State University, Jianchang Mao IBM Almaden Research Centre.," Artificial neural networks: A tutorial", 1996.

[6] Rosenblatt, Frank. x.," Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms", Spartan Books, Washington DC, 196.

[7] Rumelhart, David E., Geoffrey E. Hinton, and R. J. Williams. "Learning Internal Representations by Error Propagation", David E. Rumelhart, James L. McClelland, and the PDP research group. (editors), Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: Foundations. MIT Press, 1986.

[8] Cybenko,”Approximation by superpositions of a sigmoidal function Mathematics of Control, Signals, and Systems”, 2(4), 303–314.

[9]S.JeraldJeba Kumar, Madheswaran,“Extraction of Blood Vascular Network for Development of an Automated Diabetic Retinopathy Screening System”, International Conference on Computer Technology and Development IEEE 10.1109/Icctd.2009.212, 2009.

[10]María García, Carmen Valverde, “Comparison of Logistic Regression and Neural Network Classifiers in the Detection of Hard Exudates in Retinal Images”, 35th Annual International Conference of the IEEE EMBS Osaka, Japan, 3 - 7 July, 2013.

[11] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, ““Breast cancer diagnosis with transfer learning and global pooling,” arXiv preprint arXiv:1909.11839, 2019.

[12] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, R. Deters et al, “A hybrid deep learning architecture for leukemic blymphoblast classification,” arXiv preprint arXiv:1909.11866, 2019.

[13] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, , “Classification of histopathological biopsy images using ensemble of deep learning networks,” arXiv preprint arXiv:1909.11870, 2019.

[14] Xiaomin Zhou, Chen Li, Md Mamunur Rahaman, Yudong Yao et al. "A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks", IEEE Access, 2020

[15] Varun Gulshan, Subhashini Venugopalan, Rajiv Raman, “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” JAMA. 2016;316(22):24022410. doi:10.1001/jama.2016.17216.

[16] Kele Xu, Dawei Feng, and Haibo Mi, “Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy Using Fundus Image,” Received: 10 November 2017; Accepted: 22 November 2017; Published: 23 November 2017.

[17] Sheikh Muhammad Saiful Islam, Md Mahedi Hasan, and Sohaib Abdullah, “Deep Learning based Early Detection and Grading of Diabetic Retinopathy Using Retinal Fundus Images,” arXiv:1812.10595v1 [cs.CV] 27 Dec 2018.

[18] Lam C, Yi D, Guo M, Lindsey T., “Automated Detection of Diabetic Retinopathy using Deep Learning,” AMIA Jt Summits Transl Sci Proc. 2018 May 18;2017:147-155. PMID: 29888061; PMCID: PMC5961805.

[19] Sara Hosseinzadeh Kassani, Peyman Hosseinzadeh Kassani, Reza Khazaeinezhad, Michal J. Wesolowski et al. "Diabetic Retinopathy Classification Using a Modified Xception Architecture", 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2019

[20] Anuj Jain, Arnav Jalui, Jahanvi Jasani, Yash Lahoti, Ruhina Karani. "Deep Learning for Detection and Severity Classification of Diabetic Retinopathy", 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019