Diabetic Retinopathy Detection From Fundus Images: Adjusted Activation Functions

By:
Nisarg Parikh(804018)
Aharnish Solanki
Hansal Dhandha(804065)

A thesis
Presented to **Maharaja Sayajirao University of Baroda**

in partial fulfilment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

At

सत्यं शिवं सुन्दरम्

Department of Computer Science & Engineering

Faculty of Technology & Engineering

The Maharaja Sayajirao University of Baroda

MAY-2021

Declarations

I hereby declare that the project entitled "***Diabetic Retinopathy Detection***" submitted for the B. E. (CSE) degree is my original work and the project has not formed the basis for the award of any other degree, diploma, fellowship or any other similar titles.

Abstract

Diabetic Retinopathy (DR) is an eye disease that leads to blindness when it progresses to proliferative level. The earliest signs of DR are the appearance of red and yellow lesions on the retina called hemorrhages and exudates. Early diagnosis of DR prevents blindness. In this thesis, a method for detecting diabetic retinopathy is presented. Convolution Neural Network(CNN) is used to classify retinal images into normal or abnormal cases of DR including non-proliferative (NPDR) or proliferative diabetic retinopathy (PDR). The proposed method has been tested on fundus images from Kaggle. The implementation of the presented methodology was done in Python. The methodology is tested for sensitivity and accuracy.

Acknowledgements

It is a great pleasure to express our gratitude and indebtedness to our guide Bela Shah for his guidance, encouragement, affection and moral support through the course of our work.

We are also sincerely thankful to Dr. Apurva Shah, Head of the Department of Computer Science & Engineering, for allotment of this work and also for his encouragement.

We are thankful to Dr. Anjali Jivani and Dr. Viral Kapadia for valuable opinions, and continuous assessment, that led to timely completion of our project.

Also, we are thankful to all the teachers of our department, who were always there to solve our doubts and queries during our dissertation.

Certificate

This is to certify that the project titled "Diabetic Retinopathy Detection" is the genuine and original work carried out by Nisarg Parikh,Aharnish Solanki and Hansal Dhandha student/s of B E (CSE) of faculty of Technology & Engineering, The Maharaja Sayajirao University of Baroda during the academic year 2018-19, in partial fulfilment of the requirements for the award of the degree of Bachelor of Engineering (Computer Science and Engineering).

Signature of the Guide

Signature of the Head of the Department

Place:

Date:

Table of contents

Contents

Chapter 3: Results

Chapter 4: Conclusions

List of Figures

List of Tables

List of Abbreviations

DR	Diabetic Retinopathy
SoTA	State of The Art
NPDR	Non-proliferative diabetic retinopathy
PDR	Proliferative diabetic retinopathy
OD	Optic Disk

Chapter 1: Introduction

1.1 Background

Diabetic retinopathy is a serious sight-threatening complication of diabetes. Diabetes interferes with the body's ability to use and store sugar (glucose). The disease is characterized by too much sugar in the blood, which can cause damage throughout the body, including the eyes. Over time, diabetes damages small blood vessels throughout the body, including the retina. Diabetic retinopathy occurs when these tiny blood vessels leak blood and other fluids. This causes the retinal tissue to swell, resulting in cloudy or blurred vision.

Diabetic retinopathy usually affects both eyes. The longer a person has diabetes, the more likely they will develop diabetic retinopathy. If left untreated, diabetic retinopathy can cause blindness. When people with diabetes experience long periods of high blood sugar, fluid can accumulate in the lens inside the eye that controls focusing. This changes the curvature of the lens, leading to changes in vision. However, once blood sugar levels are controlled, usually the lens will return to its original shape and vision improves. Patients with diabetes who can better control their blood sugar levels will slow the onset and progression of diabetic retinopathy.

1.1.1 Causes & risk factors:

Diabetic retinopathy results from the damage diabetes causes to the small blood vessels located in the retina. These damaged blood vessels can cause vision loss:

Fluid can leak into the macula, the area of the retina responsible for clear central vision. Although small, the macula is the part of the retina that allows us to see colors and fine detail. The fluid causes the macula to swell, resulting in blurred vision. In an attempt to improve blood circulation in the retina, new blood vessels may form on its surface. These fragile, abnormal blood vessels can leak blood into the back of the eye and block vision.

Diabetic retinopathy is classified into two types.

Non-proliferative diabetic retinopathy

Non-proliferative diabetic retinopathy (NPDR) is the early stage of the disease in which symptoms will be mild or nonexistent. In NPDR, the blood vessels in the retina are weakened. Tiny bulges in the blood vessels, called microaneurysms, may leak fluid into the retina. This leakage may lead to swelling of the macula.

Proliferative diabetic retinopathy

Proliferative diabetic retinopathy (PDR) is the more advanced form of the disease. At this stage, circulation problems deprive the retina of oxygen. As a result, new, fragile blood vessels can begin to grow in the retina and into the vitreous, the gel-like fluid that fills the back of the eye. The new blood vessels may leak blood into the vitreous, clouding vision.

Other complications of PDR include detachment of the retina due to scar tissue formation and the development of glaucoma. Glaucoma is an eye disease in which there is progressive damage to the optic nerve. In PDR, new blood vessels grow into the area of the eye that drains fluid from the eye. This greatly raises the eye pressure, which damages the optic nerve. If left untreated, PDR can cause severe vision loss and even blindness.

1.2 Literature Review

The research field of retinal image analysis has attracted a lot of interest in the last couple of decades, with the automated detection of diabetic retinopathy having received a considerable share of this interest. Landmark detection is also an area that has received considerable interest. Landmarks consist of blood vessels, the optic disc and the fovea. This section will start with a brief review of the automated segmentation of blood vessels. Most DR detection methodologies use it as a prerequisite before identifying pathological entities, in particular new vessel detection methods. A brief review of the main methodologies used to detect the main DR features (microaneurysms, haemorrhages and exudates) will be provided. This will be followed with a section providing a detailed account on the detection of new vessels (proliferative DR), which is the main focus of this project.

Most methodologies start by pre-processing the images. The main pre-processing steps are applied to correct for poor illumination and poor contrast. Poor illumination is often tackled with a technique called shade correction [1,2], whereby an image approximating the background is subtracted from the original image. The background image is obtained with the application of a median filter whose size is significantly greater than the largest retinal feature. Poor contrast is frequently tackled with contrast limited adaptive histogram equalisation (CLAHE) [3,4]. This is a technique for local contrast enhancement which is preferred to global contrast enhancement. However, pre-processing can only correct up to a certain extent; therefore, it is the task of the screener (photographer) to ensure an adequate standard of photographs are captured. To avoid an overly lengthy literature review, pre-processing steps shall be omitted from further discussions, although a comparative evaluation of pre-processing steps for retinal analysis is provided by Youssif [5].
It should be noted that the word "feature" will be used in two different contexts. The word "feature" refers to those components that make up a feature set used for classification. DR features on the other hand refer to pathological features such as microaneurysms and haemorrhages.

1.3 Challenges and Facts

Research on early detection of diabetic retinopathy was inspired by the desire to save patient's vision. It also was inspired by the need to help doctors to make better decisions during the diagnosis process. However, several factors affect the results of detection. For example, hemorrhages and exudates are two different signs of DR which must be detected separately. Additionally, hemorrhages and exudates are similar in color to the retinal blood vessels and the OD, respectively. Thus, an efficient algorithm should be able to distinguish between the disease indicators and the normal features of fundus images. For instance, blood vessels are different in shape than hemorrhages, and the OD is bigger than exudates. Furthermore, the research also focuses on classification of retinal images into normal or abnormal cases of DR. Here a number of features should be defined so that the algorithm can differentiate between the classes of DR.

According to Jesse and Thomas [6], DR is a serious medical disorder which influences 3.4 percent of the population. The duration of the disease, the type of diabetes mellitus and the patient's age are the three main factors that increase the risk of developing DR. For example, for individuals with type 1 diabetes, the probability of DR is 50% for a 10 year old person; however, by age 30 the risk increases up to 90%. The probability of DR for individuals with type 2 diabetes is 5%. Moreover, according to the World Health Organization, after 15 years of diabetes, 2% of DR patients become blind and 10% suffer from severe vision problems.

In this study, a novel approach for automatic detection and classification of DR from retinal fundus images is proposed. Images from the standard diabetic retinopathy database are considered. EraseReLU as a modification on successful SoTA Model Architectures is applied to increase performance and convergence to a better accuracy sooner as the novelty approaches.

1.4 Dataset

Deep learning requires the well trained network be exposed to at least thousands of images, else the details which differentiate the weights of the network can't capture the classes.
For the project we used data sets of high-resolution retinal color Images from Kaggle data set
It is a large set of high-resolution retina images taken under a variety of imaging conditions. A left and right field is provided for every subject. Images are labeled with a subject id as well as either left or right . This dataset contains image set to predict whether an image contains signs of diabetic retinopathy or not

A clinician has rated the presence of diabetic retinopathy in each image on a scale of 0 to 4, according to the following scale:

0 - No DR

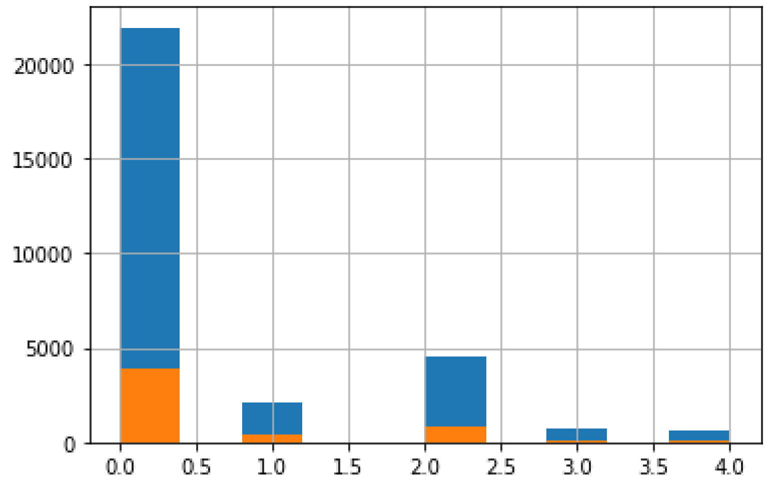
1 - Mild

2 - Moderate

3 - Severe

4 - Proliferative DR

The images in this dataset come from different models and types of cameras and feature very mixed quality.



Fig(1)Class wise distribution of Fundus Images

1.5 Data Augmentation

Most of the time, even after using a proper model we do not get satisfactory results. The problem then lies in the data used to train the network. Having a large dataset is crucial for the performance of the deep learning model. However, we may lack the quantity and diversity of data we wish to train the model for a custom requirement thereby hampering its performance.
Data augmentation is a technique to artificially create new training data from existing training data. It helps us to increase the size of the dataset ,increase the diversity of data and introduce variability in the dataset, without actually collecting new data.The neural network treats these images as distinct images .So, to get more data, we need to make minor alterations to our existing training data.

As you can see we don't have much images of 2, 3 and 4th class as compared to that of 0 class . Which shows oversampling of 0 class images and undersampling of other class images. So, to get more data, we did minor alterations to our existing training data. These alterations include flipping the image horizontally, vertically, rotating and few other translations.

```
from keras.preprocessing.image import ImageDataGenerator, array_to_img, img_to_array ,load_img

datagen = ImageDataGenerator(
     featurewise_center=False,
    rotation_range=25,
    width_shift_range=0.1,
    height_shift_range=0.1,
    shear_range=0.1,
    zoom_range=0.1,
    horizontal_flip=True,
    vertical_flip=True,
    fill_mode='constant'
)
```

1.6 Software Description

1.6.1 Python

Python is a multi-paradigm programming language. Object-oriented programming and structured programming are fully supported, and many of its features support functional programming and aspect-oriented programming (including by metaprogramming and metaobjects. Many other paradigms are supported via extensions, including design by contract and logic programming

Python's developers strive to avoid premature optimization, and reject patches to non-critical parts of CPython that would offer marginal increases in speed at the cost of clarity. When speed is important, a Python programmer can move time-critical functions to extension modules written in languages such as C, or use PyPy, a just-in-time compiler. Cython is also available, which translates a Python script into C and makes direct C-level API calls into the Python interpreter. An important goal of Python's developers is keeping it fun to use. Python's design offers some support for functional programming in the Lisp tradition. It has filter, map, and reduce functions, list comprehensions, dictionaries, sets, and generator expressions.

Benefits of Python

· Presence of Third-Party Modules

· Extensive Support Libraries

· Open Source and Community Development

· Learning Ease and Support Available

· User-friendly Data Structures

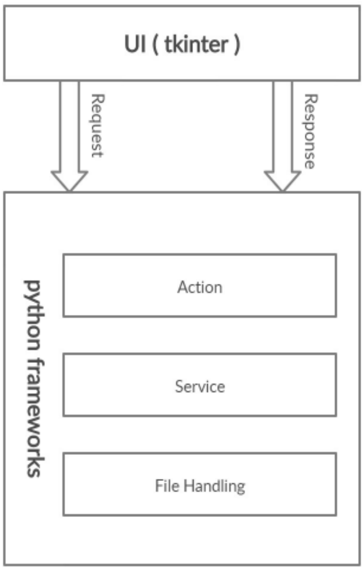
· Productivity and Speed

· Highly Extensible and Easily Readable Language.

- Operating System : Windows 8.1/10, Linux ,Mac
- Front End : Python Tkinter
- Back End : Pytorch, Sklearn, OpenCV, SkImage, Pandas, Numpy

1.6.2 Tkinter Interface

Tkinter is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit and is Python's de facto standard GUI. Tkinter is included with standard Linux, Microsoft Windows and Mac OS X installs of Python . As with most other modern Tk bindings, Tkinter is implemented as a Python wrapper around a complete Tool Command Language (TCL) interpreter embedded in the Python interpreter.Python offers multiple options for developing GUI (Graphical User Interface). Out of all the GUI methods, tkinter is the most commonly used method. It is a standard Python interface to the Tk GUI toolkit shipped with Python. Python with tkinter outputs the fastest and easiest way to create the GUI applications.

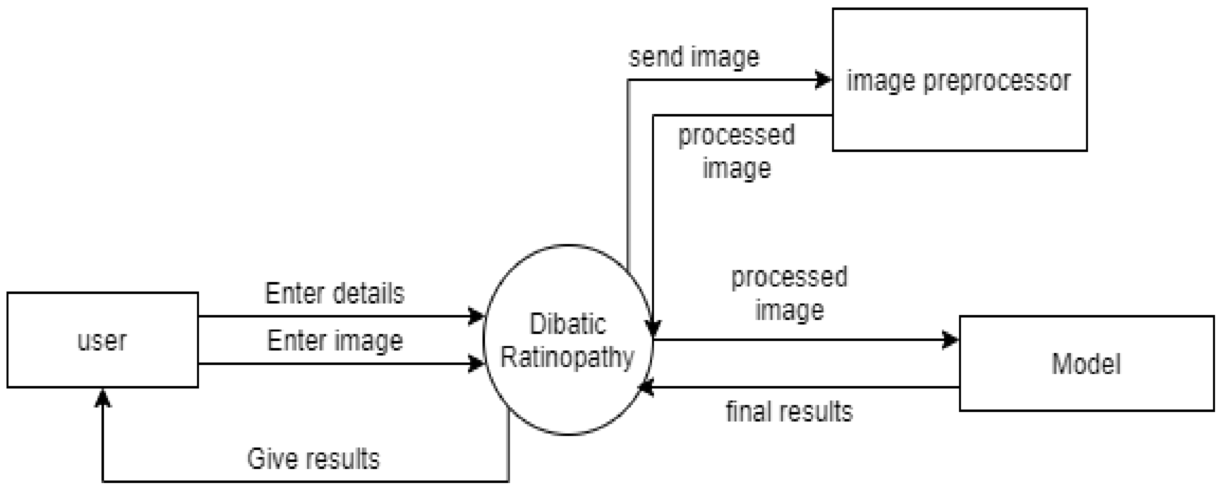


Fig(2)UI interface

1.7 Hardware Specifications

- Ram: 4 GB or above
- Memory: 1 GB for dependency software

1.8 UML Diagram



Fig(3)DFD Diagram

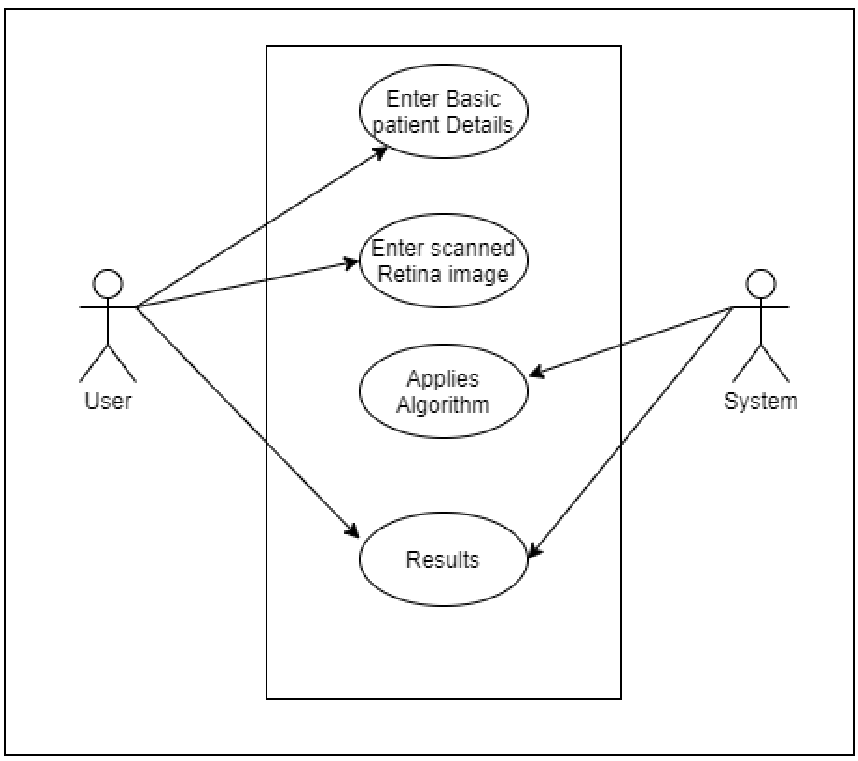
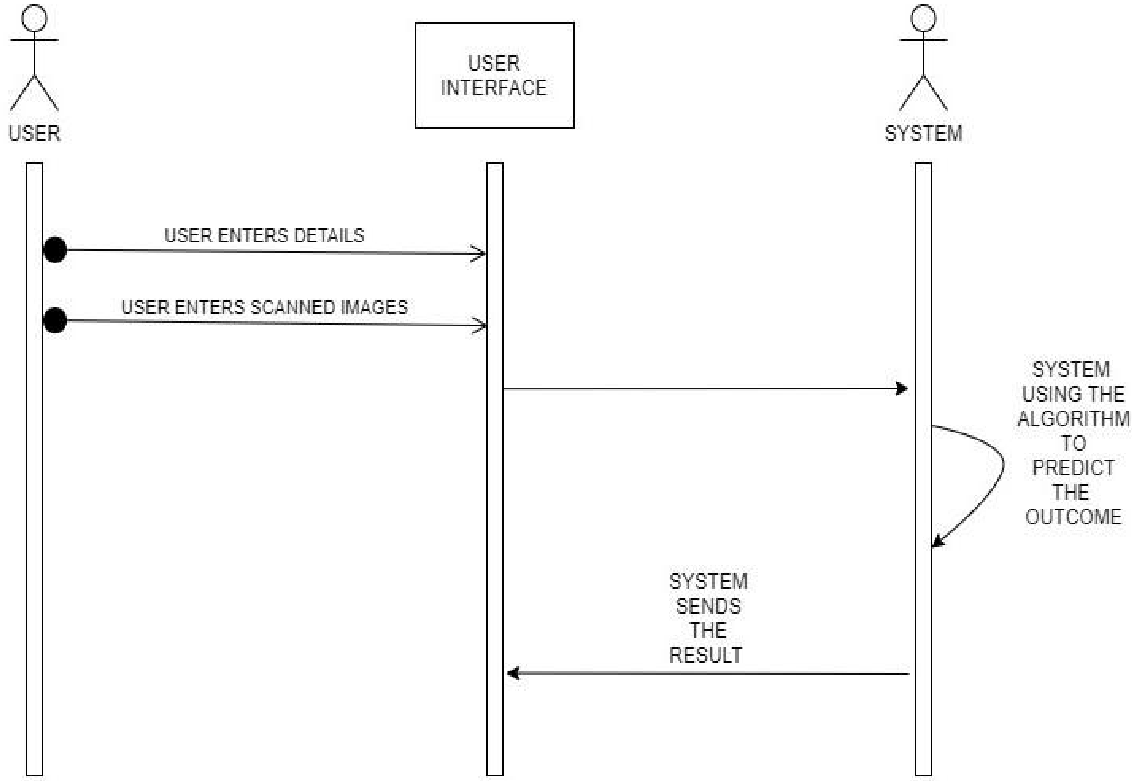


Fig (4) Use Case Diagram



Fig(5)Sequence Diagram

Chapter 2: Implementation and Approach

2.1 Prospective Classification Methods:

1. CNN
2. SVM
3. Random Forest method with DL classification
4. Fuzzy C-Means (FCM) clustering and Morphological Restructuring

Comparative analysis of image classification algorithms based on traditional machine learning and deep learning

[7]SVM is a very powerful classification model in machine learning. CNN is a type of feedforward neural network that includes convolution calculation and has a deep structure. It is one of the representative algorithms of deep learning. Taking SVM and CNN as examples, this paper compares and analyzes the traditional machine learning and deep learning image classification algorithms. This study found that when using a large sample mnist dataset, the accuracy of SVM is 0.88 and the accuracy of CNN is 0.98; when using a small sample COREL1000 dataset, the accuracy of SVM is 0.86 and the accuracy of CNN is 0.83. The experimental results in this paper show that traditional machine learning has a better solution effect on small sample data sets, and deep learning framework has higher recognition accuracy on large sample data sets.

[8]This study compares the classification accuracy of convolution neural networks artificial neural networks and support vector machines on thirteen forest-vegetation species. SVM classifier is experimented with linear, RBF and polynomial kernels. CNN has higher overall accuracy as compared to the three SVM types and artificial neural network. We can conclude that CNN is better than conventional morphological methods for feature extraction in plant species. The data taken from

Bruker VERTEX 70 FTIR spectrometer (Bruker Optics GmbH, Ettlingen, Germany) was used to acquire the Directional Hemispherical Reflectance (DHR) spectrum of each leaf. Experimental results indicate that CNN based approach is significantly effective with an overall accuracy of about 99%. Upon increasing the training data, the classification accuracy of both SVM and CNN improves.
RF (random forest) measures of variable importance are used to detect factors that affect classification performance. Principal Findings Both types of data were informative when discriminating participants with or without DR. RF based models produced much higher classification accuracy than those based on logistic regression. Then there were automatic methods of exudates detection on low-contrast images taken from non-dilated pupils. The process had two main segmentation steps which are coarse segmentation using Fuzzy C-Means clustering and fine segmentation using morphological reconstruction.

We chose to work with CNN architecture as that approach gave historically the most success with accurate detection.

2.2 Chosen Classification Architecture

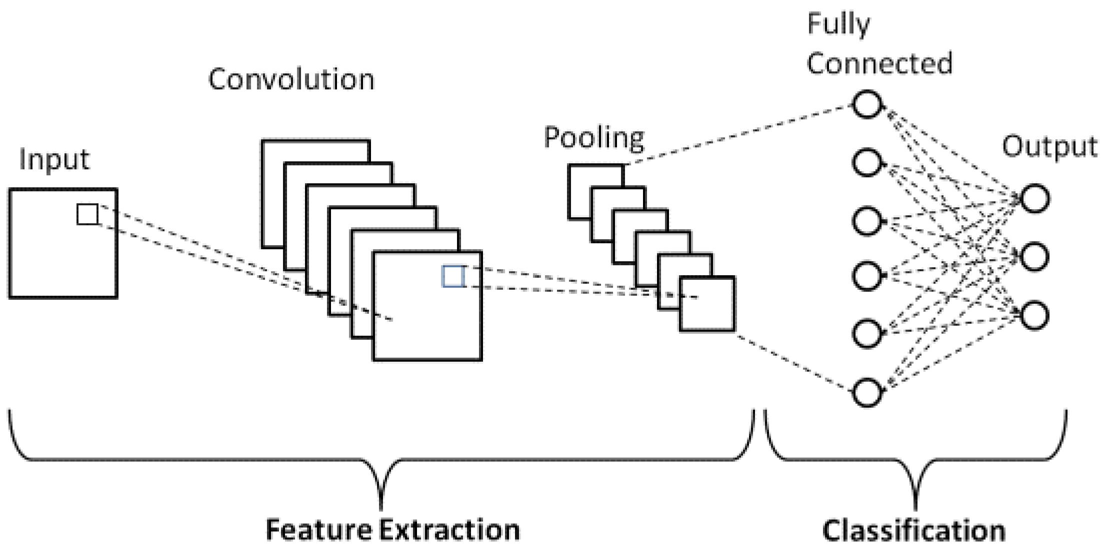
2.2.1 CNNs

In recent years usage of convolution neural networks or CNNs for image classification has become quite popular.
The success of LeNet, a CNN used for classifying digits showed that CNNs could be used in solving various vision tasks [11]. Large-scale image recognition tasks were made easy using CNNs [9]. A deep convolutional neural network was trained to classify 1.2 million images present in ImageNet [10],[11]. Also, CNNs computationally have fewer connections as compared to fully connected architectures and thus are easier and faster to train [12]. To prevent the networks from overfitting, dropout notion was introduced [13] and to address the internal covariate shift,batch normalization was added to the network designing [14], which led to good performing CNNs. CNNs training usually takes a lot of time which can be reduced by using a few high-end GPUs (Graphics Processor Units) . The building components of a CNN are the convolutional, subsampling, max pooling, and batch normalization layers. We will explore a few convolutional neural network architectures as transfer learning is becoming quite popular. Transfer learning is where features defined by all layers are intact, but the final few layers are re-trained with minimal problem-specific dataset which reduces both training and network designing time.

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

CNN Architecture

A convolutional neural network consists of an input layer, hidden layers and an output layer. In any feed-forward neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final convolution. In a convolutional neural network, the hidden layers include layers that perform convolutions. Typically this includes a layer that performs a dot product of the convolution kernel with the layer's input matrix. This product is usually the Frobenius inner product, and its activation function is commonly ReLU. As the convolution kernel slides along the input matrix for the layer, the convolution operation generates a feature map, which in turn contributes to the input of the next layer. This is followed by other layers such as pooling layers, fully connected layers, and normalization layers.



Fig(6) CNN Architecture

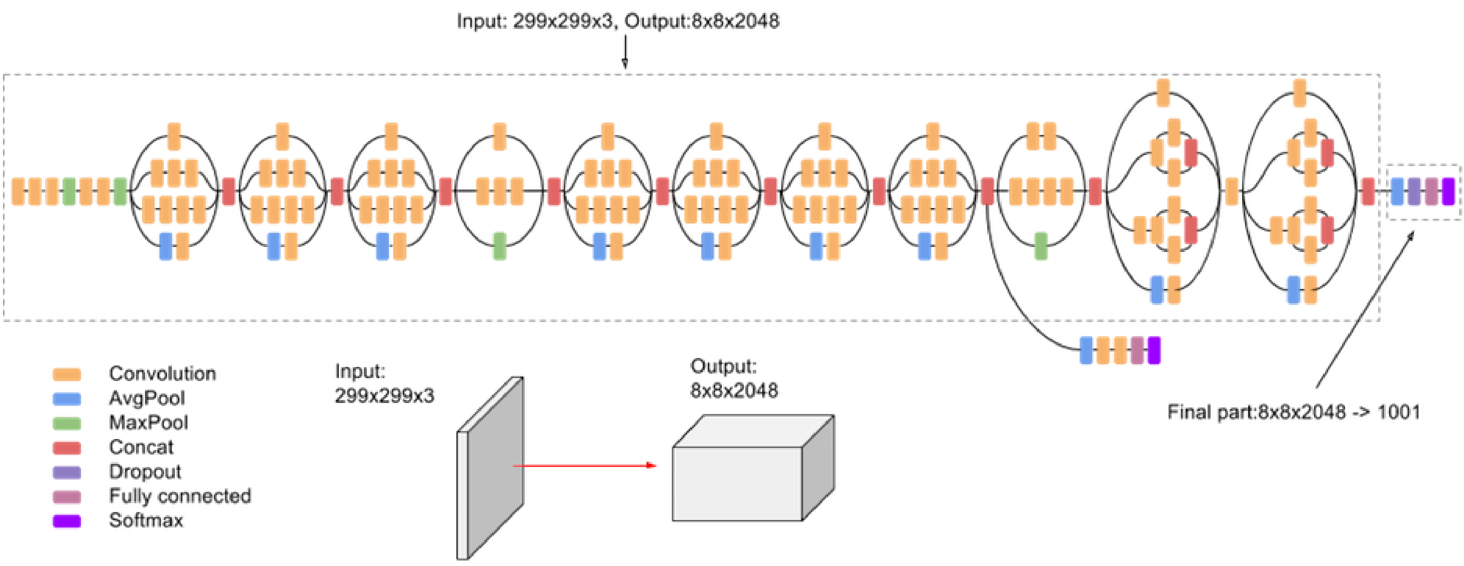
Convolutional Neural Networks (CNNs) is the most popular neural network model being used for image classification problems. The big idea behind CNNs is that a local understanding of an image is good enough. The practical benefit is that having fewer parameters greatly improves the time it takes to learn as well as reduces the amount of data required to train the model. Instead of a fully connected network of weights from each pixel, a CNN has just enough weights to look at a small patch of the image

2.2.1.1 Inception v3

Inception v3 is a convolutional neural network for assisting in image analysis and object detection

Inception v3 is a widely-used image recognition model that has been shown to attain greater than 78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years. It is based on the original paper: "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et. al[29].

The model itself is made up of symmetric and asymmetric building blocks, including convolutions, average pooling, max pooling, concats, dropouts, and fully connected layers. Batchnorm is used extensively throughout the model and applied to activation inputs. Loss is computed via Softmax.



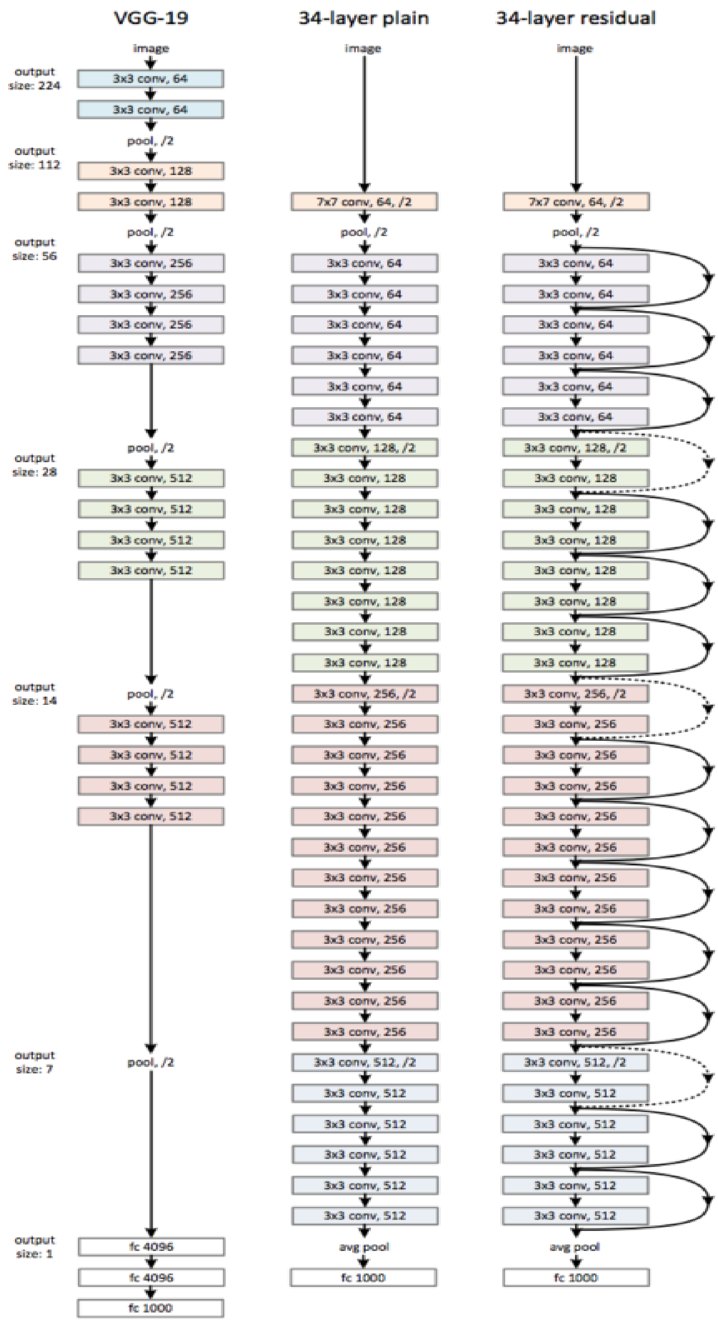
Fig(7)Inception Model Architecture

2.2.1.2 Resnet

A residual neural network (ResNet) is an artificial neural network (ANN)
Residual Network is a specific type of neural network that was introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun in their paper "Deep Residual Learning for Image Recognition"[30]
Typical *ResNet* models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between. An additional weight matrix may be used to learn the skip weights; these models are known as *HighwayNets*.] Models with several parallel skips are referred to as *DenseNets*.
The fundamental breakthrough with ResNet was it allowed us to train extremely deep neural networks with 150+layers successfully. Prior to ResNet training very deep neural networks was difficult due to the problem of vanishing gradients.

ResNet architecture

ResNet 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 the residual network as shown in the figure below:



Fig(8)Resnet Model Architecture

2.3 Preprocessing

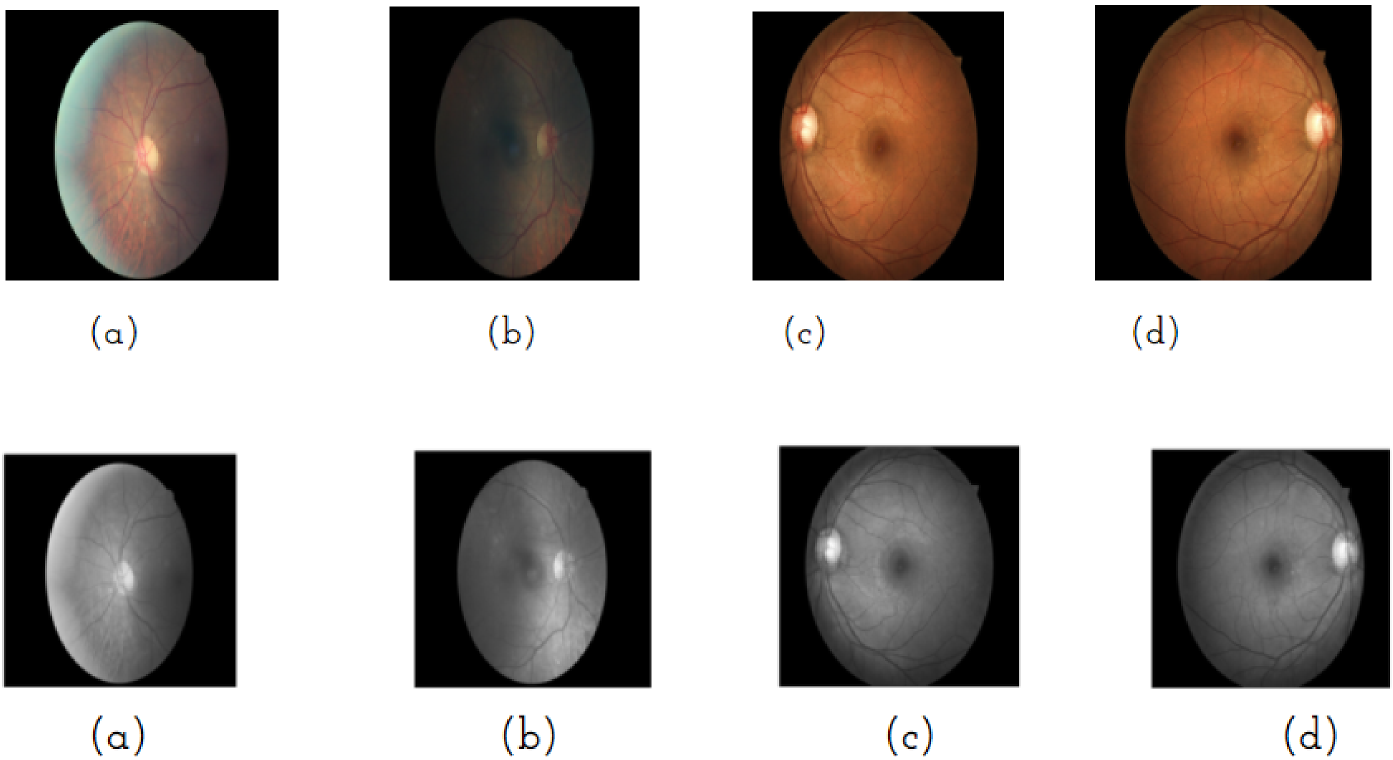
In detecting abnormalities associated with fundus image, the images have to be pre-processed in order to correct the uneven illumination, not sufficient contrast between exudates and Image background pixels and the presence of noise in the input fundus Image . Preprocessing is required to ensure that the dataset is consistent and displays only relevant features.

The techniques for preprocessing include ,Green Channel Extraction ,Gray scale Conversion, Histogram Equalization, Sharpening of Images ,Gamma correction ,Better Contrast, Median Blur , Contrast Limited Adaptive Histogram Equalization (CLAHE) .

2.3.1 Gray Scale Conversion

converting a coloured image to a grayscale image will not matter, because eventually the model will be learning from the geometry present in the image. The image-binarization will help in sharpening the image by identifying the light and dark areas.
Each of these R, G and B values usually vary from 0 to 255 for each pixel. However, in gray-scaling, a certain pixel value will be one-dimensional instead of three-dimensional, and it will just vary from 0 to 255. So, yes, there will be some information loss in terms of actual colours, but that is in tradeoff with the image-sharpness.
So, there can be a combined score of R, G, B values at each point (probably their mean (R+G+B)/3), which can give a number between 0 to 255, which can eventually be used as their representative. So that, instead of specific colour information, the pixel just carries the intensity information. [15]

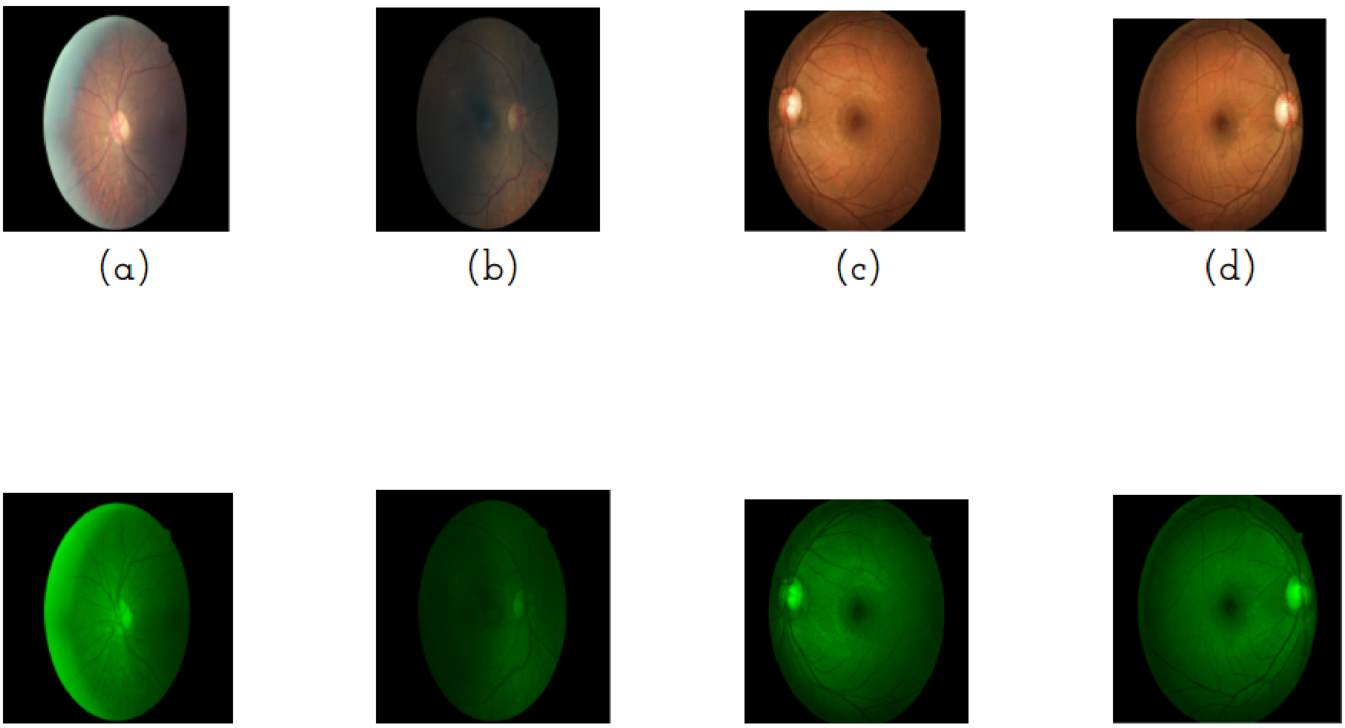
The color image i.e RGB image of an eye is taken as input image and is converted to a grayscale image as shown in Fig.(A) and Fig.(B) Respectively.



FIG(9)Gray Scale Conversion

2.3.2 Green Channel Extraction

[16]Green Channel of the three color channels in the image (Red, Green, and Blue) the contrast between the blood vessels, exudates and hemorrhages is best seen in the green channel and these channels neither under- illuminated nor over-saturated like the other two. Hence, we have extracted only the green channel for analysis and classification given as an illustrative example in Figure (10)



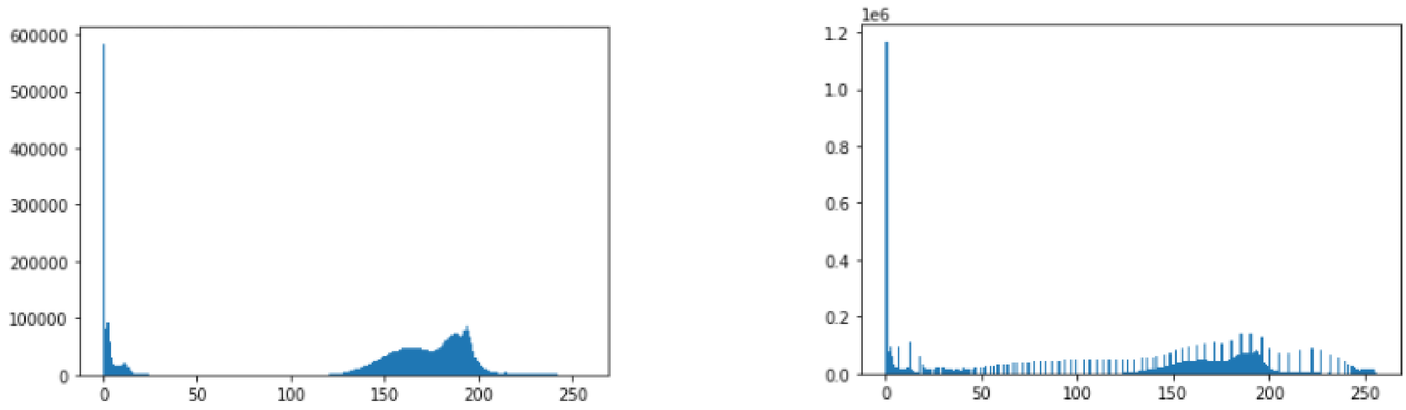
(a) (b) (c) (d)

(a) (b) (c) (d)

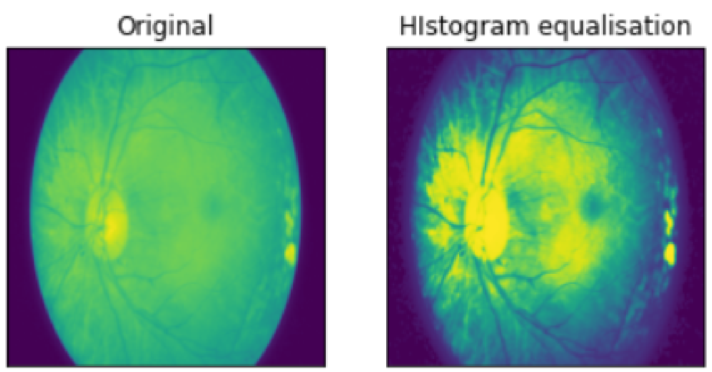
Fig(10)Green Channel Extraction

2.3.3 Histogram Equalization

Histogram Equalization is a technique for adjusting image intensities to enhance contrast. This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values. The method is useful in images with backgrounds and foregrounds that are both bright or both dark.[17][18]



Fig(11) Comparison of Original and Histogram Equalized Histograms



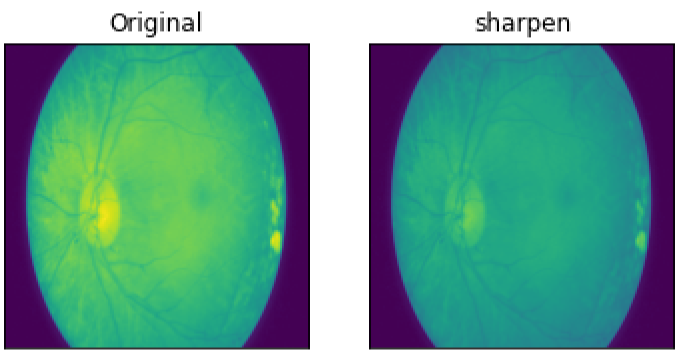
Fig(12) Histogram Equalised Image

2.3.4 Sharpening Of Images

[19]Sharpening an image increases the contrast between bright and dark regions to bring out features . Basically to enhance line structures or other details in an image .The sharpening process is basically the application of a high pass filter to an image. The following array is a kernel for a common high pass filter used to sharpen an image.

$$\begin{bmatrix} -1/9 & -1/9 & -1/9 \\ -1/9 & 1 & -1/9 \\ -1/9 & -1/9 & -1/9 \end{bmatrix}$$

Line structures and edges can be obtained by applying a difference operator (=high pass filter) on the image.

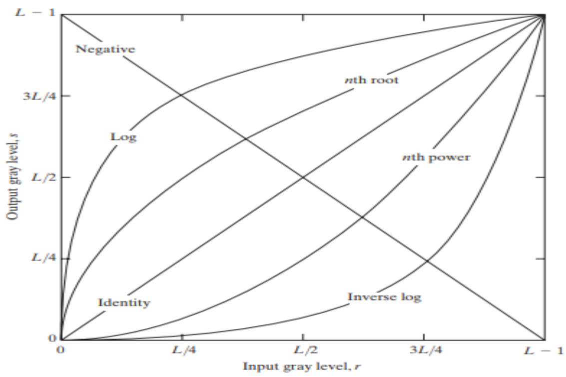


Fig(13)Sharpened Image

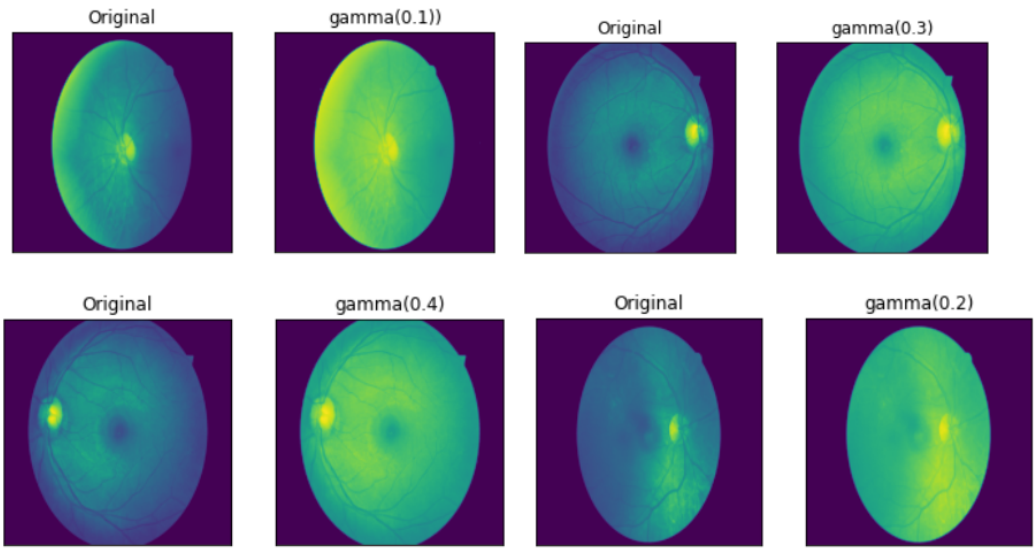
2.3.5 Gamma Correction

Gamma transformations can be mathematically expressed as $s = cr^{\gamma}$. Gamma correction is important for displaying images on a screen correctly, to prevent bleaching or darkening of images when viewed from different types of monitors with different display settings. This is done because our eyes perceive images in a gamma-shaped curve, whereas cameras capture images in a linear fashion.

gamma>1 (indicated by the curve corresponding to 'nth power' label on the graph), the intensity of pixels decreases i.e. the image becomes darker. On the other hand, gamma<1 (indicated by the curve corresponding to 'nth root' label on the graph), the intensity increases i.e. the image becomes lighter.[20]



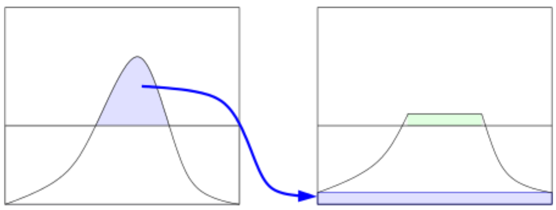
Fig(14)Log Transformation



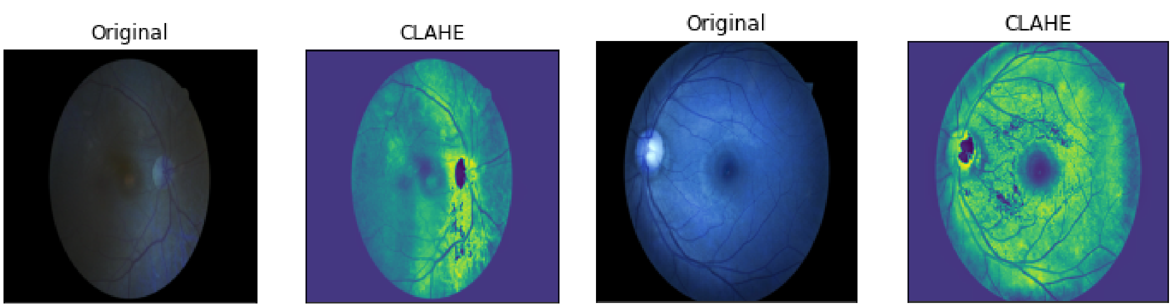
Fig(15)Images after Gamma Transformation

2.3.6 CLAHE (Contrast Limited Adaptive Histogram Equalization)

Contrast Limited AHE (CLAHE) is a variant of adaptive histogram equalization in which the contrast amplification is limited, so as to reduce this problem of noise amplification . CLAHE operates on small regions in the image, called tiles, rather than the entire image. The neighboring tiles are then combined using bilinear interpolation to remove the artificial boundaries.
In CLAHE, the contrast amplification in the vicinity of a given pixel value is given by the slope of the transformation function. This is proportional to the slope of the neighbourhood cumulative distribution function and therefore to the value of the histogram at that pixel value. CLAHE limits the amplification by clipping the histogram at a predefined value before computing the CDF. This limits the slope of the CDF and therefore of the transformation function. The value at which the histogram is clipped, the so-called clip limit, depends on the normalization of the histogram and thereby on the size of the neighbourhood region. Common values limit the resulting amplification to between 3 and 4.[21]



Fig(16)CLAHE Limit Clipping



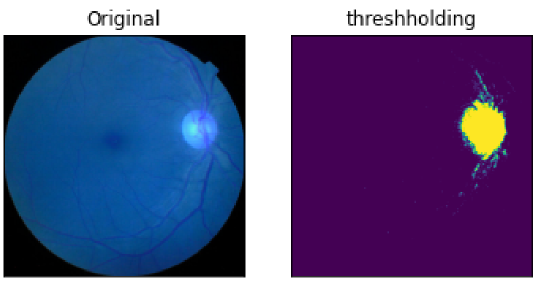
Fig(17) Before and After CLAHE implementation

2.4 Feature Extraction

Optic Disk Elimination and Exudate Detection The main objective of exudate detection is the removal of the optic disc before the onset of the process. It is essential because it appears with similar intensity, color and contrast to the other attributes of the fundus image.

2.4.1mage Thresholding For Optic Disk Detection

The optic disc can be separated out by the presence of high contrast circular shape areas. It should be noted that vessels also show with high contrast. However, they are distinctively smaller in area and number.[22]



Fig(18)Optic Disk Detection

2.4.2 Haemorrhages Detection Via Morphological Processing

Morphological operations are used to extract image components that are useful in the representation and description of region shape. Morphological operators take an input image and a structuring component as input and these elements are then combines using the set operators.
Opening is similar to erosion as it tends to remove the bright foreground pixels from the edges of regions of foreground pixels. The impact of the operator is to safeguard

foreground region that has similarity with the structuring component, or that can totally contain the structuring component while taking out every single other area of foreground pixels. Opening operation is used for removing internal noise in an image.[23]

Opening is erosion operation followed by dilation operation.

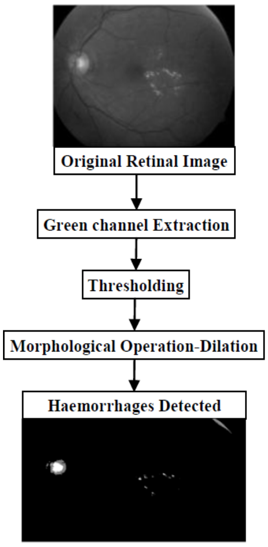
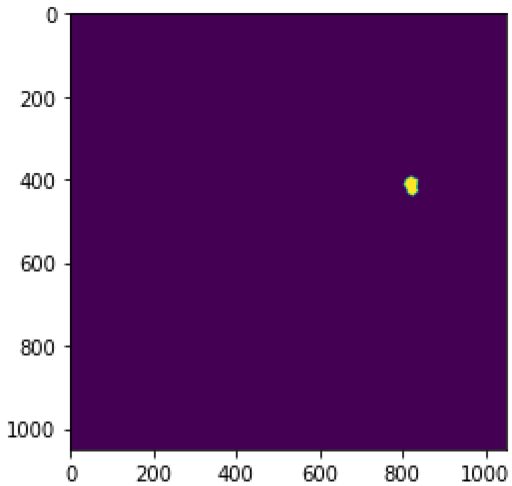


Fig.(19) Block diagram for Haemorrhages Detection



Fig(20) HAEMORRHAGES DETECTION

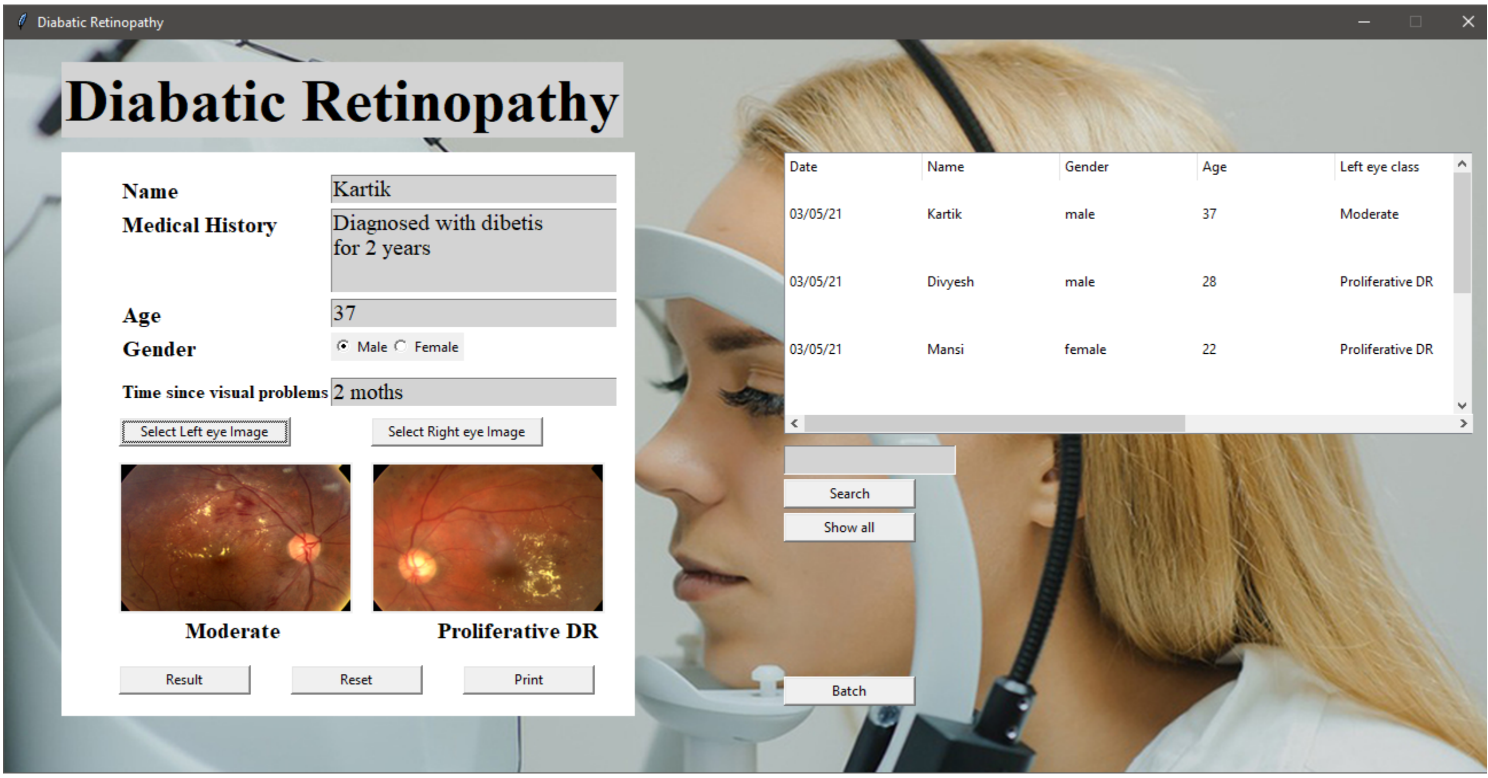
Here Dilation is used after thresholding and the exudates or Haemorrhages are detected.

2.5 GUI

The user interface of this system consists of python's library interface called tkinter.

Then it goes into the framework model where all the actions and services are combined and then the result is processed.

It also consists of a file system where all the user related information is stored and pdf is generated at the end and results are declared in it



Fig(21)GUI

Diabatic Retinopathy Analysis Report

Patient Information

Name: Kartik

Age: 37

Gender: Male

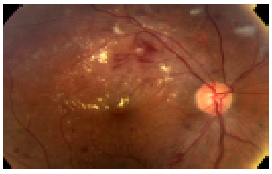
History: Diagnosed with dibetis for 2 years

Duration: 2 moths

Date: 2021-05-03

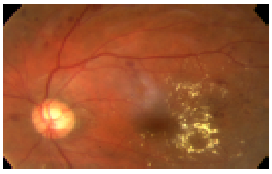
Results:

Left Eye Result:



Moderate

Right Eye Result:



Proliferative DR

Fig (22)Results Generated in PDF format

2.6 Novelty

2.6.1 EraseReLU

In most architectures, the network consists of multiple stacked core modules. Therefore, a network can be formulated as :
$F(\mathbf{x}) = fn \;\; fn\,1 \;\; ::: \;\; fi \;\; ::: \;\; f2 \;\; f1 \;\; \mathbf{x};$
where fi indicates the i-th basic unit in the network and fix equals fi($\mathbf{x}$). As we stack more such modules in CNN, the network tends to overfit the training set and the optimization becomes more difficult. Residual connection [24] can alleviate this phenomenon. However, these two problems are still unsolved [25] We empirically find that reducing nonlinearities of f can be helpful for ameliorating the overfitting and optimization problems in very deep neural networks.

There are usually three operations in CNN models: convolution, batch normalization, and ReLU. Convolution operation is a linear unit and also essential for the model capacity, we thus do not change the convolution operation. BN is not a linear unit strictly but can be approximately regarded as a linear unit. It can avoid the gradient explosion by stabilizing the distribution and reducing the internal covariate, we thus should also retain this operation. Therefore, there left two directions to reduce the nonlinearities, modifying ReLU or optimizing the module structure. The module structure has thousands of combinations of different operations, which is beyond the scope of our paper. Hereto, we eliminate all choices except modifying ReLU.
We can observe that for some models the last layer is a ReLU layer:
$F(\mathbf{x}) = \text{Module}(\mathbf{x}) = \text{ReLU}(\text{Module0}(\mathbf{x}));$
where most architectures also have this character. If we erase the last ReLU layer, we can preserve the overall structure. On the contrary, if we erase the middle ReLU layer of these modules, it will destroy the module structure and decrease the module capacity, thus be harmful to the performance. We empirically observe erasing the last ReLU layer in each module is capable of easing the training difficulty and considered as a regularization to improve the final performance.

Pre-activation structure [26, 27, 28] is another kind of module. It moves BN+ReLU to the head of the convolution layer, and ReLU thus is not the tail of the module. To apply EraseReLU to these architectures, we first transfer them into the after-activation structure and then apply EraseReLU, because the middle ReLU layer is essential for performance as we discussed before.

Hereto, the locations where ReLU layers should be erased have been discussed. It is still not clear that EraseReLU should be applied to which module. If we arbitrarily choose the module to apply EraseReLU, there exist thousands of choice combinations and some of them are even equivalent. Therefore, we use the proportion of modules which should be applied with EraseReLU for efficiency. The location where ReLU should be erased and the proportion of modules to apply EraseReLU are two key factors for performance improvement.
In practice, we usually use the proportion of 100% for EraseReLU, which means that we apply EraseReLU on all modules. There will usually be more than one ReLU layer in the module of state-of-the-art architectures. Therefore, even the proportion of 100% can maintain enough nonlinearity of these models.

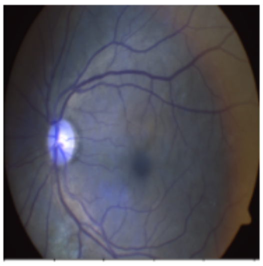
CHAPTER 3: Results

This study aims to develop an approach for automatic detection and classification of DR through developing the main three stages of detection, i.e., processing, segmentation and classification. Therefore, to test the performance of the used algorithms and contributions in retinal images, we performed three different experiments.

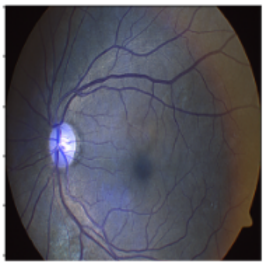
3.1 Preprocessing Results:

As mentioned earlier in this work different preprocessing techniques are used for preprocessing .The following figures indicate various phases different types of retinal images goes through bringing out the better results for processing

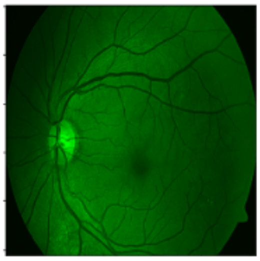
.



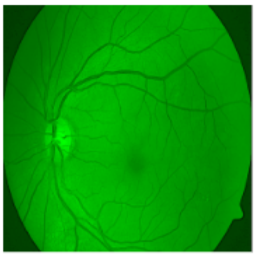
fig(a) Original



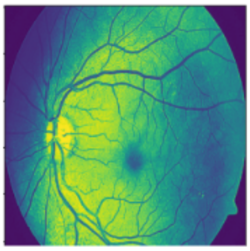
fig(b)Sharpened



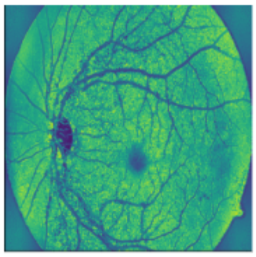
fig(c)Green Channel Extracted



fig(d)Gamma Correction



fig(e)Histogram Equalised



fig(f) CLAHE

Fig(23) Preprocessing Phases of Fundus Images

The original fundus Image as you can see in fig(a) is first sharpened and gives the result as shown in fig(b). Then on that sharpened Image Green Channel Extraction is done as shown in fig(c) .Furthermore that green channel extracted image is given gamma correction which is depicted in fig(d). Doing that histogram Equalisation is performed on gamma corrected image which distributes the frequencies of the intensities as shown respectively in fig(e). Then comes the last preprocessing phase where CLAHE is implemented on the image and the results can be seen in fig(f).

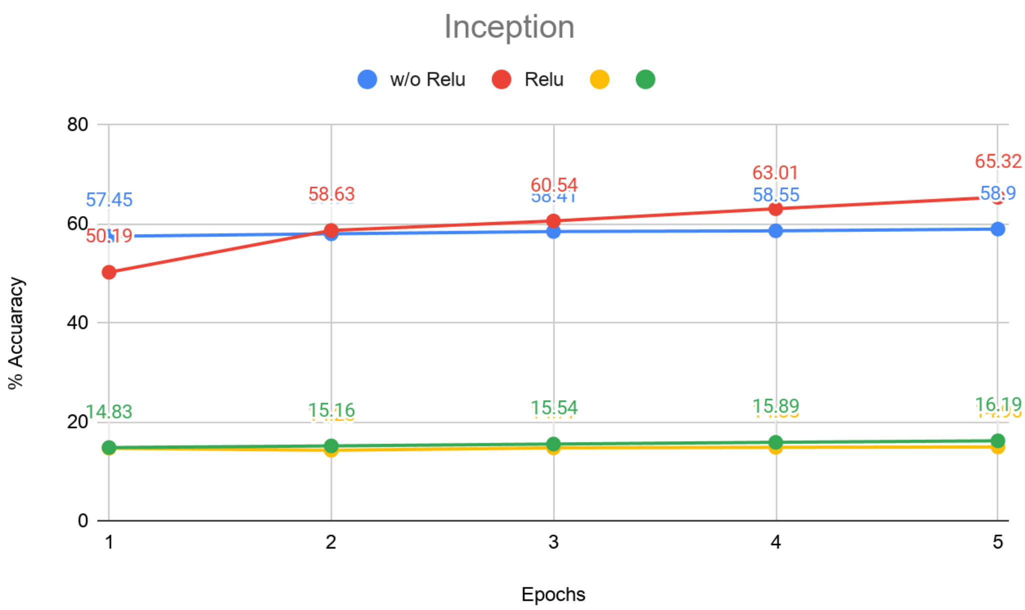
3.2 Training and Validation Results:

Table 3.1 INCEPTION

epochs	training	validation	CLASS	% Accuracy
1	50.19	14.83	NO DR	100
2	58.63	15.16	MILD DR	33
3	60.54	15.54	MODERATE	91
4	63.01	15.89	SEVERE	0
5	65.32	16.19	PDR	0

Table 3.2 INCEPTION with EraseRelu

Epochs	Training	Validation	CLASS	% Accuracy
1	57.45	14.69	NO DR	100
2	57.95	14.28	MILD DR	50
3	58.41	14.77	MODERATE	93
4	58.55	14.88	SEVERE	0
5	58.9	14.96	PDR	50



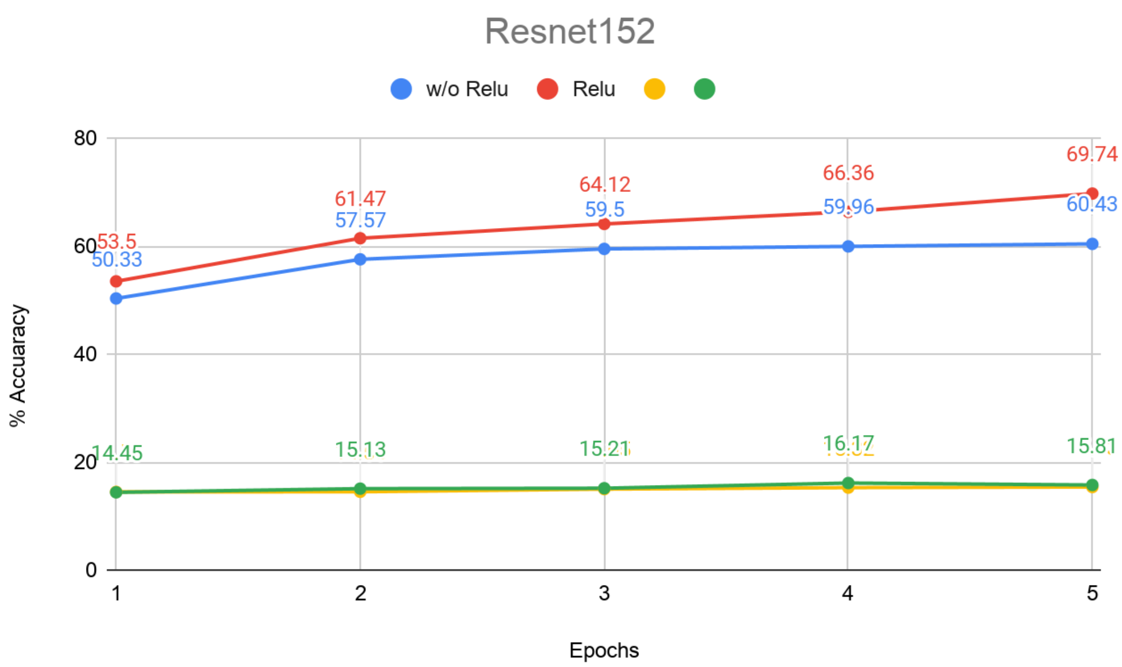
Fig(24)Comparison of Inception v3 with and w/o EraseRelu

Table 3.3 RESNET152

Epoch	training	validation	CLASS	% Accuracy
1	50.33	14.53	NO DR	88
2	57.57	14.55	MILD DR	50
3	59.5	15.05	MODERATE	89
4	59.96	15.32	SEVERE	25
5	60.43	15.43	PDR	0

Table 3.4 RESNET152 with EraseRelu

Epoch	training	validation	CLASS	% Accuracy
1	53.5	14.45	NO DR	95
2	61.47	15.13	MILD DR	40
3	64.12	15.21	MODERATE	92
4	66.36	16.17	SEVERE	33
5	69.74	15.81	PDR	50



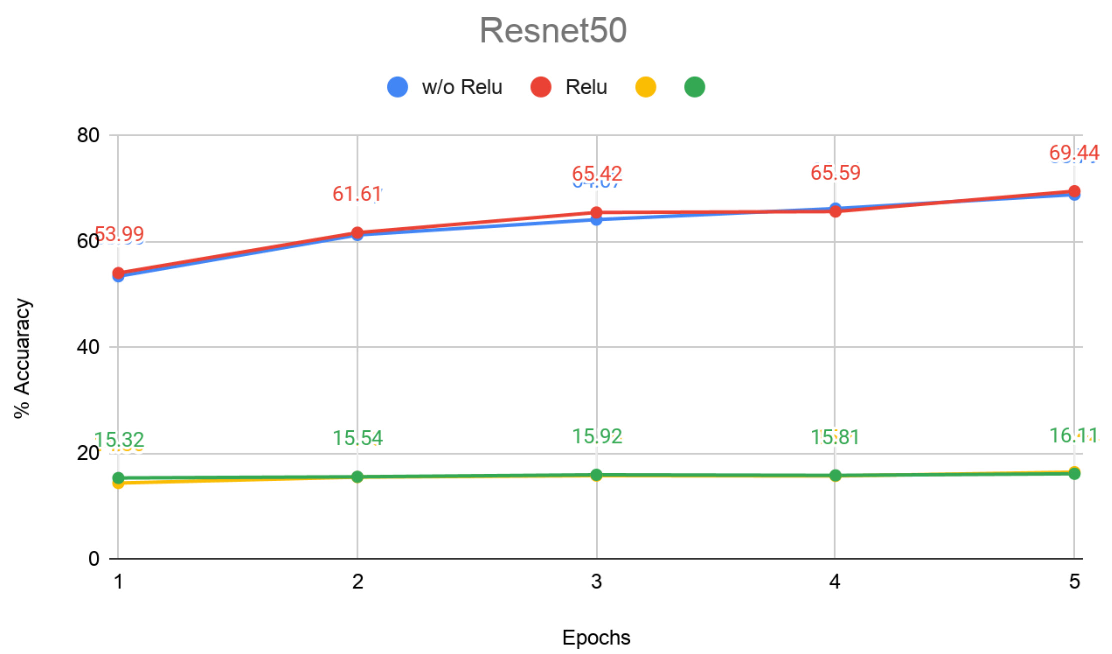
Fig(25) Comparison of Resnet152 with and w/o EraseRelu

Table 3.5 RESNET50

Epochs	training	validation	CLASS	% Accuracy
1	53.38	14.36	NO DR	100
2	61.17	15.48	MILD DR	60
3	64.07	15.76	MODERATE	89
4	66.16	15.7	SEVERE	50
5	68.79	16.42	PDR	0

Table 3.6 RESNET50 with EraseRelu

Epochs	training	validation	CLASS	% Accuracy
1	53.99	15.32	NO DR	100
2	61.61	15.54	MILD DR	60
3	65.42	15.92	MODERATE	89
4	65.59	15.81	SEVERE	50
5	69.44	16.11	PDR	0



Fig(26) Comparison of Resnet50 with and w/o EraseRelu

FOR 5 EPOCHS DATASET : APTOS

MODEL	Training Accuracy	Validation Accuracy
Inception w/o EraseRelu	58.9	14.96
Inception with EraseRelu	65.32	16.19
Resnet152 w/o EraseRelu	60.43	15.43
Resnet152 with EraseRelu	69.74	16.16
Resnet50 w/o EraseRelu	68.79	16.42
Resnet50 with EraseRelu	69.44	16.11

3.7 Training and Validation Accuracy

Chapter 4: Conclusion

4.1 Discussion

Several algorithms have been proposed for automatic detection and classification of DR. However, this work is unique in a very simple way, which is it has augmented SoTA models which show great performance with EraseReLU which shows increased performance and eases the training process.

4.2 Future Work

Future work based on this work can be done in several aspects in these areas:

1. We have only tested the novel approach for 3 models namely Inceptionnet V3, Resnet 50 and Resnet 152, more work can be done on other models to show relevance to current topics of CV in general.
2. More research is to be done with respect to data augmentation and training on a more balanced dataset
3. Another approach can be done by feeding these models only lesions in an image or Microaneurysms to show the percentage of DR a patient has or wether the patient has DR

4.3 Summary

The main contribution of this study is to apply a general approach to improve performance of CNNs to DR Detection, from which both researchers and physicians can benefit. For researchers, EraseReLU can be used as a way to enhance their CNN architecture. On the other hand, doctors can use the algorithm for automatic segmentation of DR lesions as a tool to locate hemorrhages and exudates.

References

[1] T. Spencer, J.A. Olson, K.C. McHardy, P.F. Sharp, J.V. Forrester, An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus, Computers and Biomedical Research 29 (1996) 284-302.

[2] M. Niemeijer, B. van Ginneken, J. Staal, M.S.A. Suttorp-Schulten, M.D. Abramoff, Automatic detection of red lesions in digital color fundus photographs, IEEE Transactions on Medical Imaging 24 (2005) 584-592.

[3] S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, et al., Adaptive histogram equalization and its variations, Computer Vision, Graphics, and Image Processing 39 (1987) 355-368.

[4] G.S. Ramlugun, V.K. Nagarajan, C. Chakraborty, Small retinal vessels extraction towards proliferative diabetic retinopathy screening, Expert Systems with Applications 39 (2012) 1141-1146.

[5] A.A.A. Youssif, A.Z. Ghalwash, A.S. Ghoneim, A comparative evaluation of preprocessing methods for automatic detection of retinal anatomy, Proceedings of the Fifth International Conference on Informatics & Systems (2007) 24-30.

[6] J. Vislisel and T. Oetting, "Diabetic Retinopathy: from one medical student to another," Department of Ophthalmology and Visual Sciences, University of Iowa, September, 2010

[7]]Jun-e Liu, Feng-Ping An, "Image Classification Algorithm Based on Deep Learning-Kernel Function", *Scientific Programming*, vol. 2020

[8]COMPARATIVE ANALYSIS OF SVM, ANN AND CNN FOR CLASSIFYING VEGETATION SPECIES USING HYPERSPECTRAL THERMAL INFRARED DATA Mehmood ul Hasan1,, Saleem Ullah2 , Muhammad Jaleed Khan1 , Khurram Khurshid1

[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," CoRR, vol. abs/1409.1556, 2014.

[10] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, May 24, 2017. [Online].

[11] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, June 2009, pp. 248–255.

[12] S. Nitish, H. Geoffrey, K. Alex, S. Ilya, and S. Ruslan, "Dropout: A simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929–1958, 2014.

[13] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," CoRR,vol. abs/1502.03167, 2015.

[14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan,V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions,"2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 1–9, June 2015.

[15] Wilhelm Burger, Mark J. Burge (2010). *Principles of Digital Image Processing Core Algorithms*. Springer Science & Business Media. pp. 110–111. ISBN 978-1-84800-195-4.

[16]Microaneurysms Detection for Early Diagnosis of Diabetic Retinopathy Using Shape and Steerable Gaussian Features by G. Indumathi, V. Sathananthavathi, in Telemedicine Technologies, 2019
[17]Mengko TR, Handayani A, Valindria VV, Hadi S, Sovani I. Image Processing in Retinal Angiography: Extracting Angiographic Features without the Requirement of Contrast Agents; 2009. p. 451-4.
[18] T Kauppi, V Kalesnykiene, JK Kamarainen, L Lensu... - BMVC, 2007 - it.lut.fi
[19]A Novel Integrated Approach Using Dynamic Thresholding and Edge Detection (IDTED) for Automatic Detection of Exudates in Digital Fundus Retinal Images by Anantha Vidya Sagar ,Balasubramanian S ,Venkatachalam Chandrasekaran, Mar 2007
[20]Diabetic retinopathy retinal image enhancement based on gamma correction
Z Xiao, X Zhang, F Zhang, L Geng, J Wu... - Journal of Medical ..., 2017 - ingentaconnect.com
[21]Methods to enhance digital fundus image for diabetic retinopathy detection
H Ab Rahim, AS Ibrahim, WMDW Zaki... - 2014 IEEE 10th ..., 2014 - ieeexplore.ieee.org

[22]A. Sopharak, B. Uyyanonvara, S. Barman, and T. H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Comput. Med. Imaging Graph.*, vol. 32, no. 8, pp. 720–727, 2008.

[23]S. Gupta and R. Jadhav, "Diabetic Retinopathy using Morphological Operations and Machine Learning," in *Advance Computing Conference (IACC), 2015 IEEE International. IEEE,* 2015, pp. 617–622.

[24]K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016. 1, 2, 3, 5, 6, 7

[25]S. Zagoruyko and N. Komodakis. Diracnets: Training very deep neural networks without skip-connections. arXiv preprint arXiv:1706.00388, 2017. 2, 3

[26]K. He, X. Zhang, S. Ren, and J. Sun. Identity mappings in deep residual networks. In ECCV, 2016. 2, 3, 4, 6, 7

[27] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten. Densely connected convolutional networks. In CVPR, 2017. 1, 4, 5

[28] S. Zagoruyko and N. Komodakis. Wide residual networks. In BMVC, 2016. 2, 4, 5, 6, 7

[29] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna.Rethinking the Inception Architecture for Computer Vision. arXiv:1512.00567 [cs.CV]

[30]Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition. arXiv:1512.03385 [cs.CV]