UNIVERSITY OF DELHI

Undergraduate Thesis

Automated Detection of Diabetic Retinopathy Images using Gabor Filter and Synergic Deep Learning Model

Arnav Mahajan	11910
Ayush Gupta	11912
Gitansh Wadhwa	11920

Supervisor:	Dr. Nirmal Yadav
Department:	Cluster Innovation Centre
Program:	Bachelor of Technology
Specialization:	IT and Mathematical Innovation
Course:	Semester Long Project

April 2021

Declaration of Authorship

We, declare that this thesis titled, 'Automated Detection of Diabetic Retinopathy Images using Gabor Filter and Synergic Deep Learning Model' and the work presented in it are our own. We confirm that:

This work was done wholly or mainly while in candidature for a degree at this University.

Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

Where we have consulted the published work of others, this is always clearly attributed.

Where we have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely our own work.

We have acknowledged all main sources of help.

Where the thesis is based on work done by us jointly with others, we have made clear exactly what was done by others and what we have contributed ourself.

Signed:
Arnav Mahajan
Ayush Gupta
Gitansh Wadhwa

Acknowledgement

We would like to express our deep and sincere gratitude to our research supervisor, Dr. Nirmal Yadav, for giving us the opportunity to do research and providing invaluable guidance throughout this project.

We are thankful to Cluster Innovation Center for providing us with the facilities and the necessary resources to complete our work.
We would also like to express our thanks to various colleagues, who helped us at many stages to overcome a lot of hindrances and helped us in understanding various aspects during the whole period of this project.

Abstract

Automated Detection of Diabetic Retinopathy Images using Gabor Filter and Synergic Deep Learning Model

Submitted by: Arnav Mahajan, Ayush Gupta, Gitansh Wadhwa

Recently, Diabetic Retinopathy(DR) is on the rise as a disease. It's affecting the eye of the diabetic patients because of the drastic increase in the blood sugar level. About 1.5 million people died directly from diabetes in 2019. Diabetic Retinopathy may lead to permanent vision loss under proper guidance and medication. The current research paper tries to identify the Diabetic Retinopathy images using a deep learning model. The method to identify the images involves various steps such as preprocessing the data where unnecessary noises are removed from the background, segmentation where useful regions of the images are taken out and filtering the images using filters such as Gabor Filter. Next, involves image augmentation where the dataset is enlarged to better precision. Then Synergic Deep Learning(SDL) model is applied to identify the DR related images.

Contents

List of Figures

1 Introduction

Diabetes mellitus which is commonly known as diabetes is a disease that raises the blood sugar level in the body.As glucose level grows and continue for a long duration the vessels(blood) get damaged.[1] There are three major types of diabetes- Type 1, Type 2 and Gestational diabetes.The age group and conditions as well as the side effects of each diabetes are different.[2] In the case of Type 1 diabetes, it usually occurs in young children and teenagers although others can also be diagnosed with it. The body does not create the insulin as the immune system attacks and kill the cells with create insulin inside pancreas. Type 2 diabetes is the most common diabetes. Type 2 diabetes usually occurs to middle aged people(45 years+). In this case too the body does not make insulin. Gestational diabetes occurs to women who are pregnant. But most of the time, Gestational diabetes goes away as the baby is born. There a quite a few chances that type 2 diabetes might occur in the future to people with gestational diabetes.

The constant increase in the blood sugar level causes the retina to be affected. It damages the blood vessels of the light-sensitive tissue at the back of the retina (eye).[3]This eye condition is known as diabetic retinopathy and it causes blindness as well as vision loss to those people who suffer from diabetes. At first diabetic retinopathy may not cause any symptoms or mild vision loss but eventually it will cause blindness.[4] This condition have more chances to occur with type 1 and 2 diabetic patients. The longer patient wait and less control they have on their blood sugar level, the more likely they are to develop this eye complication. Fig 1 shows the damage done to the blood vessels.

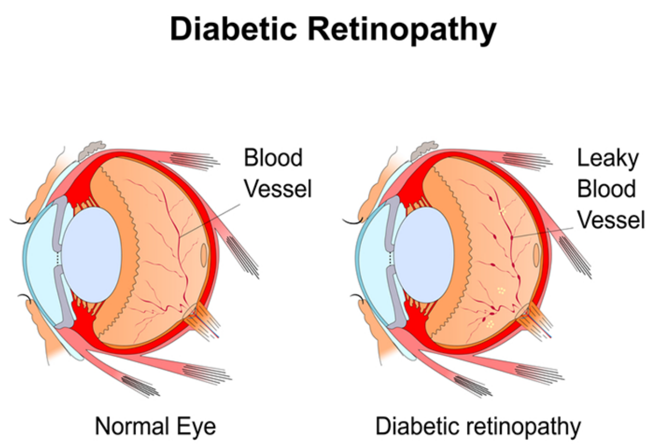

Figure 1: Damage to Blood Vessels in Diabetic Retinopathy

Diabetic Retinopathy(DR) can be classified into 2 categories i.e. Non Proliferative Diabetic Retinopathy (NPDR)' and 'Proliferative Diabetic Retinopathy (PDR).[5] NPDR is the most commonly occurring DR and is considered as early stage of diabetic retinopathy. This is the reason why NPDR is also known as Early Diabetic Retinopathy (EDR). In NPDR the retinal blood vessel's wall weaken and leads protrusion of microaneurysms. Sometimes it causes fluid as well as blood leakage inside the retina. When the condition (DR) increases in severity, it progesses to advanced stage known as proliferative diabetic retinopathy. The blood vessels are closed off and new abnormal vessels appears within the retina which leak fluid and blood inside the retina.Fig2 shows stages of DR.

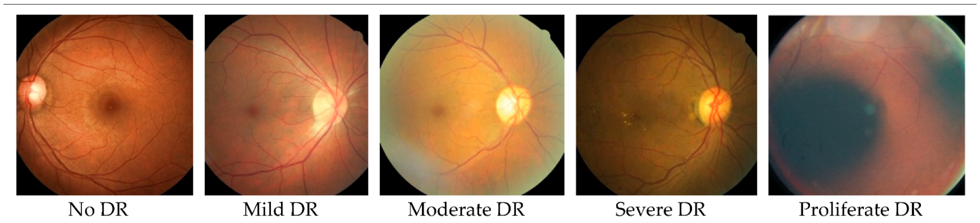

Figure 2: Stages of Diabetic Retinopathy

The current study done by the researchers and our paper focuses on classifying and identifying the DR images. The process starts by preprocessing the image in which all the unncessary noise is cancelled out. It is followed by histogram based segmentation that extracts the important regions from the image. Filters such as Gabor filter are used to analyze the image contents. Augmentation is also applied to increase the dataset and help in better model prediction. At last the Synergic Deep Learning(SDL) model is created and used to identify the DR images with higher precision.

2 Proposed Model

The study intends to classify normal eye images from DR images with maximum accuracy.[6] Since many people are affected by DR it's important to classify the images into various section in the most effective way. So the model proposed and the processes are explained below.

2.1 Preprocessing

The dataset is divided into two classes namely dr(fundus images of eye affected by diabetic retinopathy) and nodr(normal eye fundus images). The first step in preprocessing involves extraction of green channel from the images. The contrast between the relevant features of the retina is clearly visible in the green channel. Contrast limited adaptive histogram equalization is performed to further enhance the features [7]. Fig 3 displays an illustrative example.

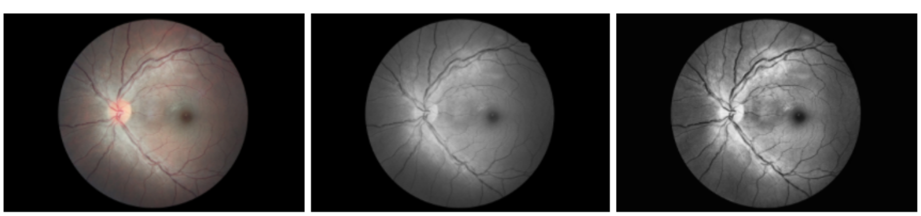

Figure 3: Preprocessing Image

2.2 Segmentation

Blood vessels make up an important feature for the classification process because high blood sugar can damage the blood vessels and lead to abnormal growth of new blood vessels. Segmentation is done using alternate sequential filtering(ASF). Basic morphological processing techniques are used namely opening(erosion followed by dilation) and closing(dilation followed by erosion).

These operations work by taking in the input image along with what is known as a structuring element or a kernel. A structuring element is a set of Euclidean space or its subspace with certain specific geometric shapes [8]. The iterative application of these functions with structuring elements of increasing size results in the image with reduced intensity at the same time preserving the geometrical characteristics of the objects it contained.

The alternating filter is defined as an opening followed by closing or closing followed by the opening [9]. An alternating sequential filter is an iterative application of alternating filters with increasing sizes of structuring elements.

$$\delta_B(f)_x = (f \oplus b)_x = max_{\beta \in B} f(x+\beta) \tag{1}$$

$$\epsilon_B(f)_x = (f \ominus B)_x = max_{\beta \in B} f(x+\beta) \tag{2}$$

$$\gamma_S = \delta_B(\epsilon_B) \tag{3}$$

$$\phi_S = \epsilon_B(\delta_B) \tag{4}$$

The output from ASF is then subtracted from the output of CLAHE which results in the image with hardly detectable blood vessels. The result is then binarized by thresholding to get a segmented image of blood vessels. Thresholding works on a simple principle if the pixel value is greater than the threshold, it assigns one value to the pixel(0 or 1). This operation works by taking in the grayscale image as input along with the second parameter which classifies the pixel value(threshold). The last step is noise removal which is done using erosion. Fig. 4 shows an illustrative example.

Inverted Binary

$$dst(x,y) = \begin{cases} 0 & \text{if src(x,y) ¿ thresh} \\ maxval & \text{otherwise} \end{cases} \tag{5}$$

2.3 Augmentation

Deep learning networks needs a large amount of dataset to achieve good result and performance.[10] Sometimes the number of images in the training set may lead to overfitting which causes irregular accuracy.[11] Issues such as inconsistency in image data(image may be too large or too small,

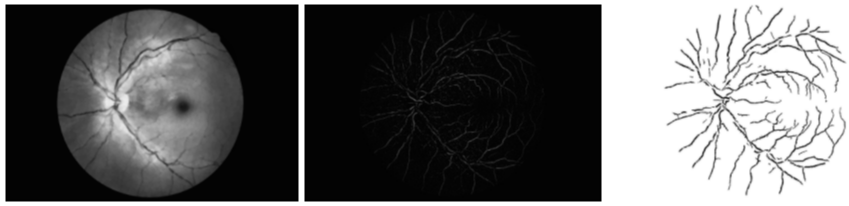

Figure 4: Steps of Segmentation

the shape might be different e.g. one in rectangle other in square) also occur while training a model. To build a proper training set and boost the performance image augmentation is used. Image augmentation basically creates training images with different way of processing such as rotation, reshape, flips, etc.

An augmented image dataset can be developed through various ways. In our case, *imgaug* a library for image augmentation is used. It contains a wide range of augmentation techniques and can be used in random order as per the requirement. There are several augmentation strategies used, including crop, flipping and scaling. The images are cropped upto 30% on each side as well as on top and bottom.[12] The images are flipped both vertically and horizontally by 90%. While scaling the images, a value is uniformally sampled per image from the interval (1.5, 1.0) on both X and Y axis.[13]Augmentation also adds the feature of spatial invariance to the model. If the feature that contributes to the classification is slightly distorted or is on a different side of the image, the model can still detect it. Spatial Invariance adds some level of flexibility to the model. Fig 5 and Fig 6 shows the images before and after augmentation respectively.

Figure 5: Before Augmentation

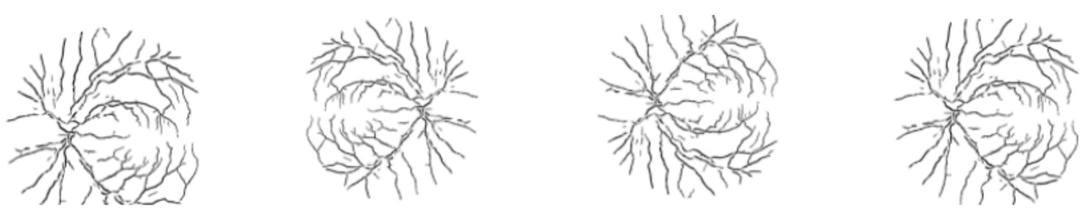

Figure 6: After Augmentation

Applying augmentation, transforms and reduces the errors and improves the accuracy of the model. The dataset is increased by 4X times than the actual one.

2.4 Gabor Filter

A Gabor filter is a linear image filter used in image processing for edge detection, texture analysis, feature extraction, etc. It can be viewed as a Sinusoidal wave of particular frequency and orientation, modulated by a gaussian wave. The frequency and orientation can be set to produce different filters allowing to extract different features from an Image. The 2D Gabor kernel has the following expression

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2+y^2}{2\sigma^2}} e^{[j\omega_x(xcos\theta+ysin\theta)]} \tag{6}$$

where ω_x decides the frequency and θ decides the orientation[14]. We produced a total of 64 gabor kernels having a size of 5x5 with different orientations and frequencies to add into our Convolutional Neural Network

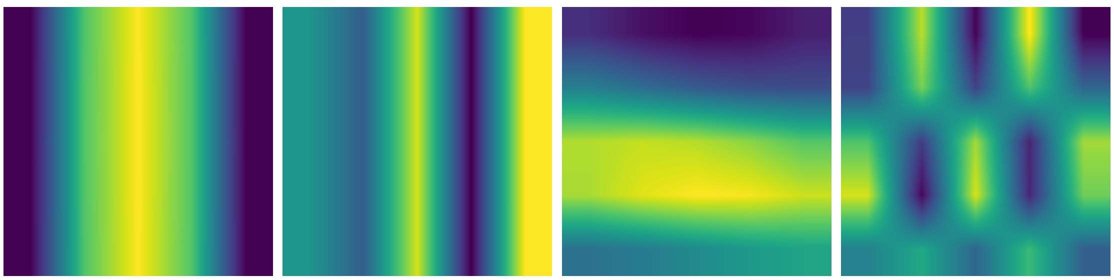

Figure 7: Some of the Gabor kernels resized to visualize

2.5 CNN Classification

Convolutional Neural Network(CNN) also known as ConvNet as specific type of neural network system which focuses on problems such as image recognition.[15] They are deep learning algorithm which can take a image as an input, classify and assign importance to various objects in an image and can differentiate one from the other.

All CNN's are divided into 2 block. The first block of CNN acts as a feature extractor function. It is done by applying convolutional filters and operations. Filters such as Gabor filters are added as a convolutional layer. In return the layers filter the with kernals and return feature maps. The second block. The second block is not CNN but used for classification. The input vector are transformed(due to layers) and gives a different output vector.[16] Fig 8 shows the first and second block of CNN respectively.

All the digital image stored are matrices of number. Each number in each matrix symbolises the brightness of each pixel. In RGB model, the image is classified into 3 such matrices corresponding to channels - red, blue and green. Each matrix values stored ranges from 0 to 255.

There are 4 different types of layers for CNN. First is the convolutional layer then the pooling layer. The third one is ReLU correction layer and the last being the fully connected layer.

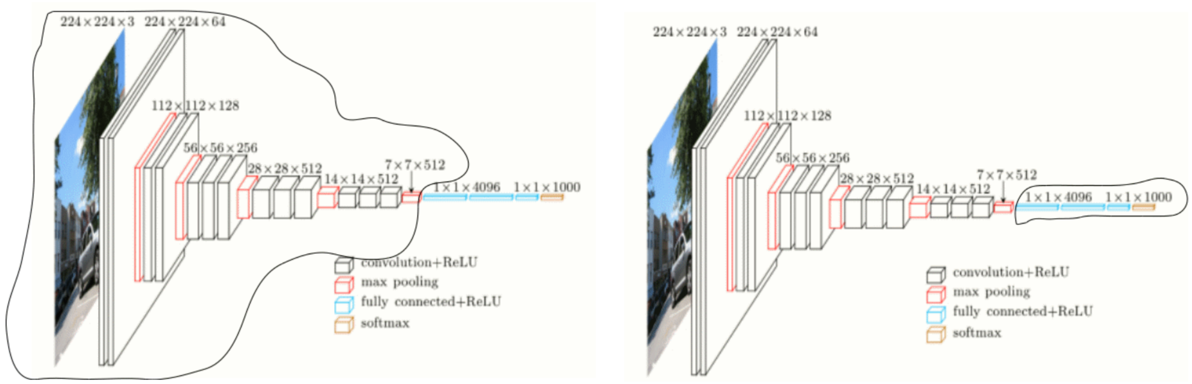

Figure 8: First and Second Block of CNN

2.5.1 Convolutional Layer

Convolutional layer is the first layer to extract features from image. It is the key layer of CNN and always the first layer. It is the process where small matrix(filter or kernal) is passed over the image matrix to transform accordingly the value of filter. Passing the kernal over the image image gives us the feature map. Filters can perform different actions such as blur, edge detection, sharpen, etc.[17] The feature map values are calculated according to the following math formula

$$G[m,n] = (f * h)[m,n] = \sum_j \sum_k h[j,k] f[m-j, n-k] \tag{7}$$

where f denotes the input image, h denotes the kernal filter and m and n represents the row and column of result matrix(feature map).

After placing the filter over the input image, kernal values are taken and multiplied in pair. All the values are sum up to give the output feature map.

2.5.2 Pooling Layer

Pooling layer is often applied between two convolutional layers. It receives several features map and apply pooling layer operations to them. Their primary function is to reduce the size of the tensor to speed up the calculation. This is done by reducing the parameters in the image when the image is too large. This improves the efficiency and avoids overlearning.

The pooling operation is simple. The images are cut into regular smalls cells and the maximum value from each cell is kept. Most commonly 2X2 cells or 3X3 are are used for pooling. We get the same number of feature map as output but they are much smaller. This is known as Max Pool Layers.

2.5.3 ReLU Correction Layer

ReLU stands for Rectified Linear Unit for a non-linear operation.[15] It is defined as

$$f(x) = max(0, x) \tag{8}$$

Graphically it is represented as follow

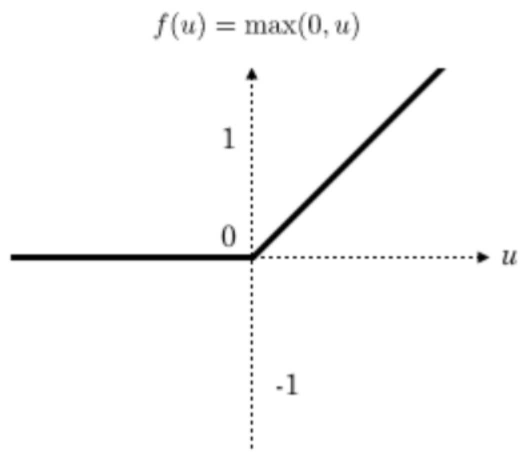

Figure 9: Graphical Representaion of ReLU

The main purpose of ReLU to introduce non-linearity in Convoutional Neural Network. The ReLU correction layer replaces all the negative values to 0. Other functions such as tanh and sigmoid can also be used instead of ReLU but since ReLU is faster as compared to other 2 many scientist prefer to use ReLU.

2.5.4 Fully Connected Layer

The fully connected layer(FC Layer) is always the last layer not only in CNN model but in every deep learning model. This layer receives an input vector and produces output vector.[18] The image matrix is converted to vector and is given to fully connected layer.

To produce a new output from an input vector, FC layer applies a linear combination and then an activation function to the input values which are received to classify outputs in various categories.

2.6 SDL Model

Synergic Deep Learning models are two or more deep learning models which work together to make a prediction. These models can take the same input or different inputs to work together

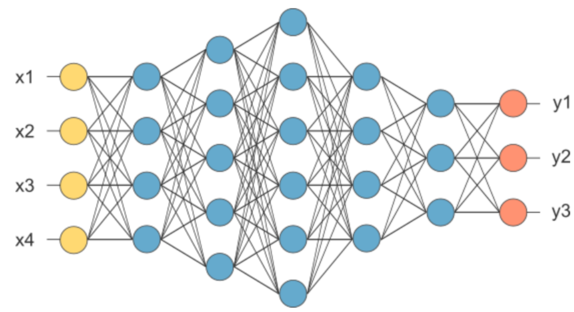

Figure 10: Fully Connected Layer

to get the output. Synergic Deep learning models can also learn from each other. An abstract representation of SDL models is shown in Figure 11. Some SDLs also use a synergic signaling system to supervise the learning from the inputs and bridge the gap between the underlying Deep Learning models as shown in Figure 12.

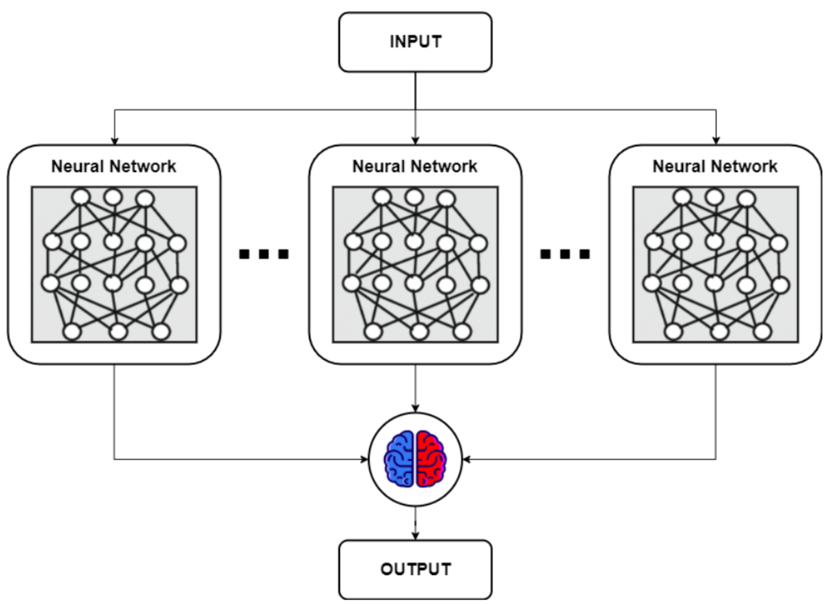

Figure 11: Abstract representation of SDL
[19]

Our SDL model is a concatenation ensemble which concatenates the inputs of the CNNs on a given axis. We have used two Convolutional Neural Networks in our SDL model, first a VGG16 model with a Gabor Convolutional layer and second a VGG19 model. Both these models are first independently trained on the training dataset before being concatenated.

After concatenation of the models, a fully-connected layer is added to the model and then compiled before being trained on the dataset.

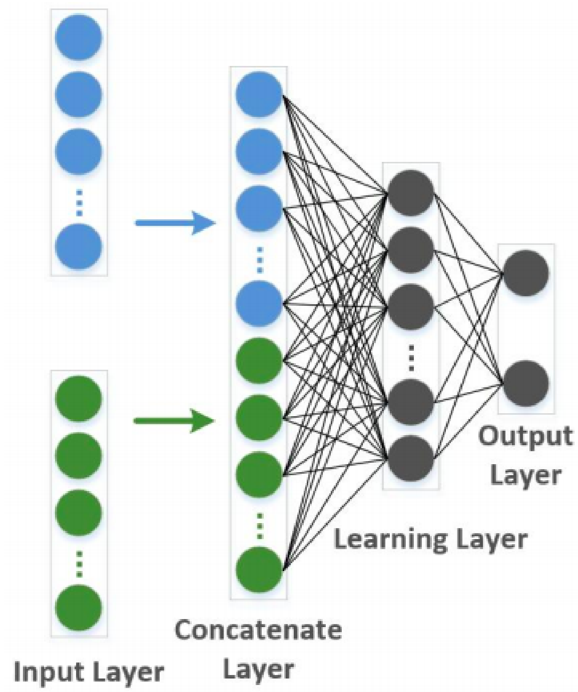

Figure 12: Synergic Signal System
[20]

Our first CNN model for the SDL model, a VGG16 model orginially has 21 layers divided into 6 blocks. The first 5 blocks have some convolutional layers followed by a single pooling layer, the last block only has fully-connected layers. The VGG16 model takes the input in the shape of (224, 224, 3)[21]. We modified the VGG16 model by replacing some convolutional layers in the first block with a Gabor convolutional layer which contains 64 Gabor kernels with different orientations and frequencies and a shape of (5, 5, 3).
Our second CNN model for the SDL model, a VGG19 model originally has 24 layers divided into 6 blocks. The first 5 blocks have convolutional layers followed by a pooling layer, same as VGG16 model, and the last block only has fully-connected layers. The VGG19 model also takes input with the shape of (224, 224, 3)[22].

3 Performance Validation

3.1 Dataset Description

In order to validate the accuracy of the SDL model open source dataset was used. The dataset had approximately 200 images.[6] Preprocessing, segmentaion and augmentation was done to create a new dataset. The new dataset covered around 500 images. The dataset was divided into two parts. One that contained no DR and another containing DR. Table 1 gives information about the dataset

DR Stages	Descriptions	Labels
DR	Diabetic Retinopathic Images	dr
Normal	Normal Images	nodr

Table 1: Dataset Details

3.2 Results

To find the model accuracy of the SDL model we plotted two graph named as Model Accuracy and Model Accuracy. The Model Accuracy graph represent the accuracy of model tested on train and test dataset. Higher the value in the graph better is the prediction. Model loss represents how bad the model prediction is. If the model is perfect then the model loss value is zero. Lower is the value in model loss better is the model. Figure 13 and figure 14 represents model accuracy and model loss.

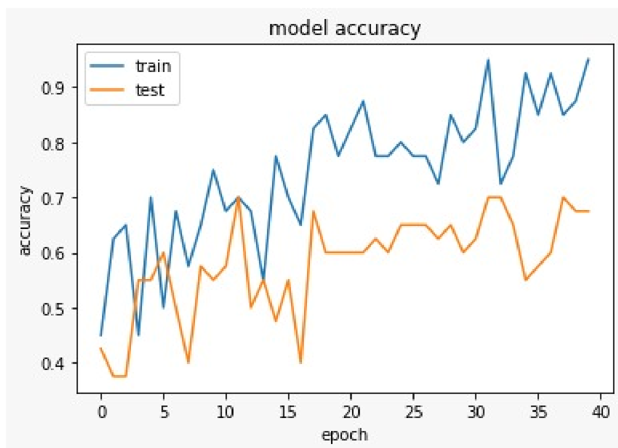

Figure 13: Model Accuracy

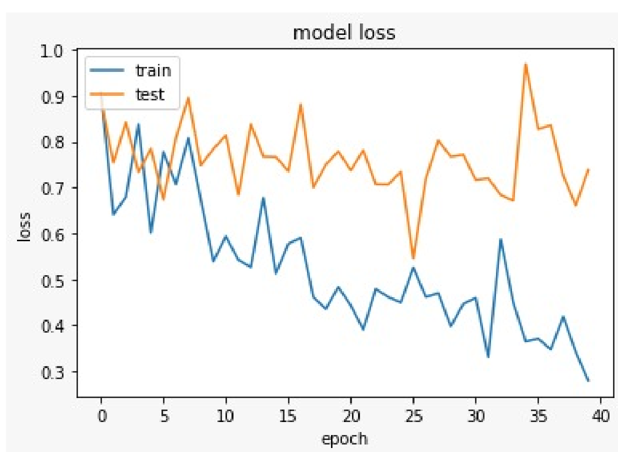

Figure 14: Model Loss

As shown in the graphs the training dataset accuracy is around 95% and the test dataset accuracy is around 75%. The model loss for training dataset is low and around 25%. For test dataset the loss is around 70%.

4 Conclusion

In this paper, a deep learning model named as Synergic Deep Learning(SDL) model was introduced which identified the DR fundus images from normal images. The model gave the accuracy for identifying the images. At the preprocessing stages, all the unwanted noise by the extraction of the green channel from the image. Next segmentaion was done and sugmentation for increasing the dataset. Then two cnn models - VGG16 and VGG19 were used and combined to create the SDL model. After implementing the model, it is observed that the accuracy came around 95% with model loss around 25%. As a part of future scope, the model can be enhanced by improving the dataset and CNN model for better accuracy.

References

[1] What is diabetes? — niddk. https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes#:~:text=Diabetes%20is%20a%20disease%20that,to%20be%20used%20for%20energy., 2017. (Accessed on 04/29/2021).

[2] Diabetes. https://www.who.int/news-room/fact-sheets/detail/diabetes, . (Accessed on 04/29/2021).

[3] Diabetic retinopathy — national eye institute. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy#:~:text=Diabetic%20retinopathy%20is%20an%20eye,at%20least%20once%20a%20year., . (Accessed on 04/29/2021).

[4] Diabetic retinopathy: Causes, symptoms, and treatments. https://www.medicalnewstoday.com/articles/183417, . (Accessed on 04/29/2021).

[5] Are there different types of diabetic retinopathy? — irisvision. https://irisvision.com/different-types-of-diabetic-retinopathy/. (Accessed on 04/29/2021).

[6] K Shankar, Abdul Rahaman Wahab Sait, Deepak Gupta, SK Lakshmanaprabu, Ashish Khanna, and Hari Mohan Pandey. Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recognition Letters*, 133: 210–216, 2020.

[7] Dilip Singh Sisodia, Shruti Nair, and Pooja Khobragade. Diabetic retinal fundus images: Preprocessing and feature extraction for early detection of diabetic retinopathy. *Biomedical and Pharmacology Journal*, 10(2):615–626, 2017.

[8] AP Ilayaraja and MP Gopinath. An analysis of morphological operation using arbitrary structuring element. *International Journal of Modern Engineering Research*, 2(2).

[9] Soo-Chang Pei, Chin-Lun Lai, and Frank Y Shih. An efficient class of alternating sequential filters in morphology. *Graphical Models and Image Processing*, 59(2):109–116, 1997.

[10] Image augmentation for deep learning — by suki lau — towards data science. https://towardsdatascience.com/image-augmentation-for-deep-learning-histogram-equalization-a71387f609b2. (Accessed on 04/29/2021).

[11] Data augmentation in medical images — by cody glickman — towards data science. https://towardsdatascience.com/data-augmentation-in-medical-images-95c774e6eaae. (Accessed on 04/29/2021).

[12] imgaug.augmenters.size — imgaug 0.4.0 documentation. https://imgaug.readthedocs.io/en/latest/source/api_augmenters_size.html#imgaug.augmenters.size.Crop, . (Accessed on 04/29/2021).

[13] imgaug.augmenters.geometric — imgaug 0.4.0 documentation. https://imgaug.readthedocs.io/en/latest/source/api_augmenters_geometric.html#imgaug.augmenters.geometric.Affine, . (Accessed on 04/29/2021).

[14] Gupta Deepak and Khanna Ashish. Gabor filters. https://www.sciencedirect.com/topics/engineering/gabor-filters, 2020. (Accessed on 04/29/2021).

[15] A comprehensive guide to convolutional neural networks — the eli5 way — by sumit saha — towards data science. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53#:~:text=A%20Convolutional%20Neural%20Network%20(ConvNet,differentiate%20one%20from%20the%20other., . (Accessed on 04/29/2021).

[16] Understand the architecture of cnn — by kousai smeda — towards data science. https://towardsdatascience.com/understand-the-architecture-of-cnn-90a25e244c7. (Accessed on 04/29/2021).

[17] Gentle dive into math behind convolutional neural networks — by piotr skalski — towards data science. https://towardsdatascience.com/gentle-dive-into-math-behind-convolutional-neural-networks-79a07dd44cf9. (Accessed on 04/29/2021).

[18] A comprehensive guide to convolutional neural networks — the eli5 way — by sumit saha — towards data science. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53#:~:text=A%20Convolutional%20Neural%20Network%20(ConvNet,differentiate%20one%20from%20the%20other., . (Accessed on 04/29/2021).

[19] Cerliani Marco. Neural networks ensemble. https://towardsdatascience.com/neural-networks-ensemble-33f33bea7df3, July 2020. (Accessed on 04/29/2021).

[20] Zhang Jianpeng, Xia Yong, Wu Qi, and Xie Yutong. Classification of medical images and illustrations in the biomedical literature using synergic deep learning. https://arxiv.org/pdf/1706.09092.pdf. (Accessed on 04/29/2021).

[21] Vgg16 — cnn model. https://www.geeksforgeeks.org/vgg-16-cnn-model/, . (Accessed on 04/29/2021).

[22] Understanding the vgg19 architecture. https://iq.opengenus.org/vgg19-architecture/, . (Accessed on 04/29/2021).