
NETRA

Ansh Tiwari

Government Boys Senior Secondary School - Rajokari
New Delhi
ansschh@gmail.com

Arshroop Singh Saini

The Emerald Heights International School
Indore
arshroopsaini@gmail.com

ABSTRACT

Diabetic Retinopathy (DR) is a common complication of diabetes mellitus, which causes lesions on the retina that affect 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.[1] Convolutional neural networks are more widely used as a deep learning method in medical image analysis and they are highly effective.[1] Netrascopy is a more efficient system for Diabetic Retinopathy detection, which consists of a low cost, Camera, “DIYretCAM Netrascopy FUNDUS Camera V1”, An Android Application and Web Application which aims to help patients and doctors detect diabetic retinopathy at early stages by taking 30-Second video of patient’s retina and passing each frame as an individual test case to a Convolutional Neural Network to detect probability of a patient having diabetic retinopathy.

Keywords Convolutional Neural Network · Deep Learning · ResNet50 · FUNDUS Image · Diabetic Retinopathy (DR) · Validation Accuracy · Validation Loss · Data Optimization · Contrast-limited adaptive histogram equalization · Otsu thresholding · Binary thresholding · Python · Heroku · Cloud

1 Introduction

1.1 Diabetic Retinopathy

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. [2] 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. [3]

1.2 Severity

Worldwide, DR causes 2.6% of blindness. Diabetic retinopathy is the leading cause of blindness in the working-age population of the developed world[4]. It is estimated to affect over 93 million people. The US Centre for Disease Control and Prevention estimates that 29.1 million people in the US have diabetes and the World Health Organisation estimates that 347 million people have the disease worldwide. Around 40% to 45% of Americans with diabetes have some stage of the disease[5].

1.3 Detection Method

Progression to vision impairment can be slowed or averted if DR is detected in time, however this can be difficult as the disease often shows few symptoms until it is too late to provide effective treatment. 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

Netra:

appearance of different types of lesions on a retina image. These lesions are microaneurysms (MA), haemorrhages (HM), soft and hard exudates (EX)[6] [7] [8].

- 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 $125 \mu\text{m}$ [9] and there are sharp margins.
- Haemorrhages (HM) appear as larger spots on the retina, where its size is greater than $125 \mu\text{m}$ 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 fibre. The shape is oval or round.

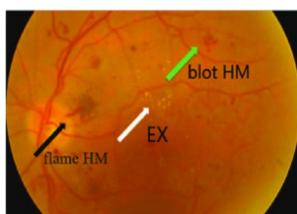
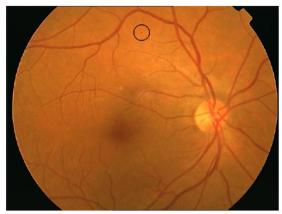


Figure 1: Microaneurysms (MA)

Figure 2: Haemorrhages (HM)

Figure 3: Soft exudates

Figure 4: Hard exudates

1.4 Current Methods.

Currently, detecting DR is a time-consuming and manual process that requires a trained clinician to examine and evaluate digital colour fundus photographs of the retina. By the time human readers submit their reviews, often a day or two later, the delayed results lead to lost follow up, miscommunication, and delayed treatment.

Clinicians can identify DR by the presence of lesions associated with the vascular abnormalities caused by the disease. While this approach is effective, its resource demands are high. The expertise and equipment required are often lacking in areas where the rate of diabetes in local populations is high and DR detection is most needed. As the number of individuals with diabetes continues to grow, the infrastructure needed to prevent blindness due to DR will become even more insufficient [10].

1.5 Previous Deep Learning Algorithms

The need for a comprehensive and automated method of DR screening has long been recognized, and previous efforts have made good progress using image classification, pattern recognition, and machine learning [11]. Current research on artificial intelligence based DR diagnosis shows promise but still has room for improvement before clinical implementation in terms of diagnostic accuracy of each stage and most of the deep learning algorithms were trained on small datasets of a few thousand one channel black and white Images due to which the algorithms were not able to give correct predictions for 5 different classes which were different stages of diabetic retinopathy [12].

1.6 Research Methods

In order to tackle the problem, we met with Dr. Ravi Jasuja, a Professor of Medicine at HMS (Harvard Medical School). We also conducted online interviews with scientists at Plaksha University, and studied peer-reviewed scientific papers to understand research into the problem we faced. In this study, deep learning was used to train an algorithm to detect referable diabetic retinopathy and assess the performance of the algorithm in 2 clinical validation sets.

1.7 Our Solution

To build our deep learning algorithm, ResNet50, VGG-16, InceptionV3, and EfficientNetB7 were first explored, and ResNet50 was selected for further optimizations. Upon testing convolutional neural network based ResNet50 on a dataset of 35,000 Images, ResNet50 gave the highest accuracy of 95% and a validation loss score of 0.1 (1 percentage), which is higher than the recent best algorithms trained on the same dataset of >35,000 with total size of 88 GigaBytes by California Healthcare Foundation.

2 Related Work.

Sinthaniyothin [13] uses maximum variance to obtain the optic disk center and a region growing segmentation method to obtain the exudates. [14] tracks the optic disk through a pyramidal decomposition and obtains disk localization from a template-based matching that uses the Hausdorff distance measure on the binary edge image. However, the above methods will fail if exudates similar in brightness and size to the optic disk are present. [15] combines matched-filter responses, confidence measures and vessel boundary measures to obtain blood vessels robustly. But the paper doesn't extend it to identify diabetic retinopathy in images. citearticle used blood vessel intersection property to obtain the optic disk. However, they use the whole blood vessel network which can lead to wrong or inconclusive results because of noise from the fringe blood vessels. In contrast, we use only the main blood vessels, which is more robust. Statistical classification techniques have been very popular lately for the problem of lesion classification. Exudates have color properties similar to the optic disk while Microaneurysms are difficult to segment due to their similarity in color and proximity with blood vessels. In order to classify detected features, typically, candidate regions are detected using color/morphological techniques and then classification is done on these regions using some classifier. STARE is a complete system for various retinal diseases [16]. The optic disk is detected using blood vessel convergence and high intensity property. In order to determine the features and classification method to be used for a given lesion, a Bayesian probabilistic system is used.

3 Convolutional Neural Network

3.1 Data

For algorithm development, Dataset which was used to make our deep learning algorithm was obtained from a competition organised in 2015 on kaggle organised by California Healthcare Foundation. We also tested our model using the public MESSIDOR 2[17] and E-Ophtha databases[18] for external validation. Below are samples of each dataset.

It was 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 (e.g. 1_left.jpeg is the left eye of patient id 1).

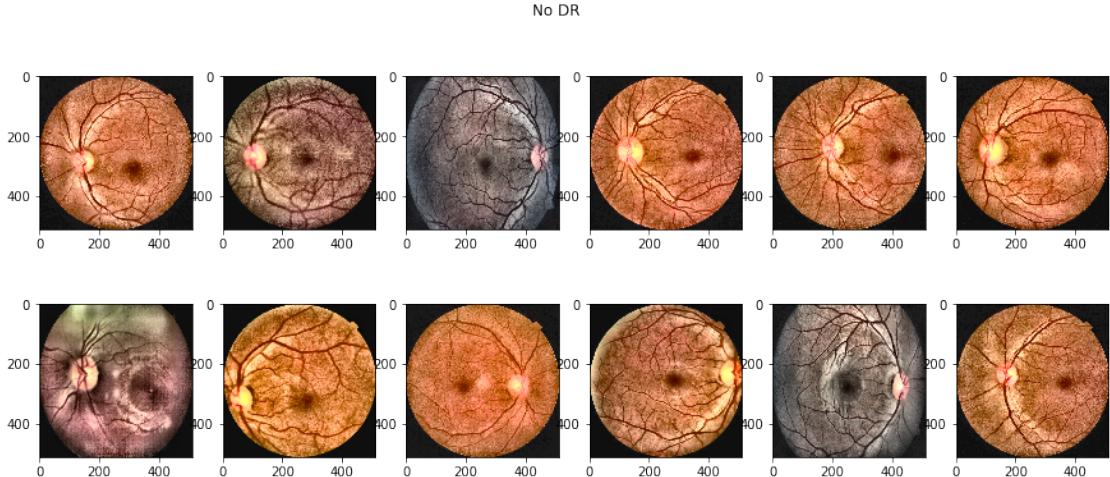


Figure 5: No DR Data Samples

Netra:

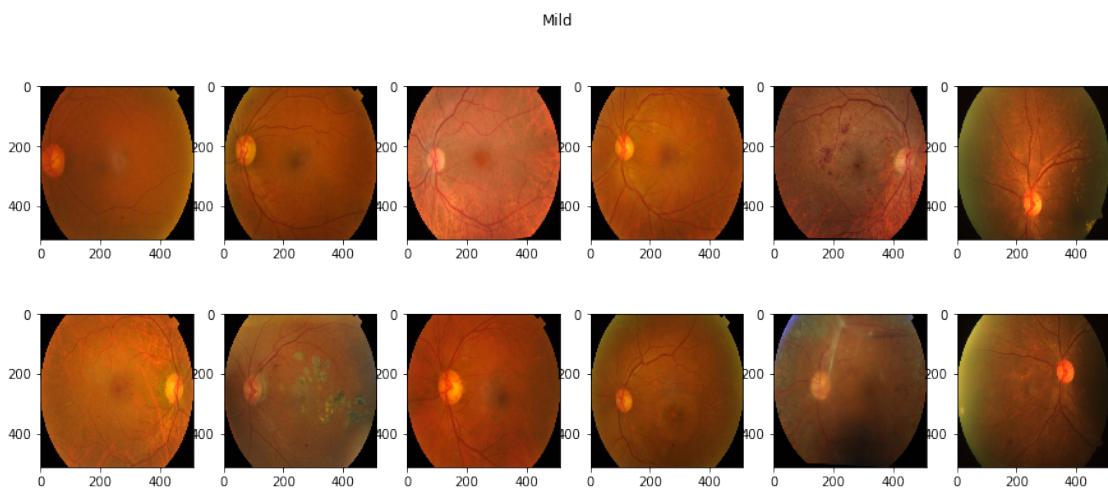


Figure 6: Mild DR Data Samples

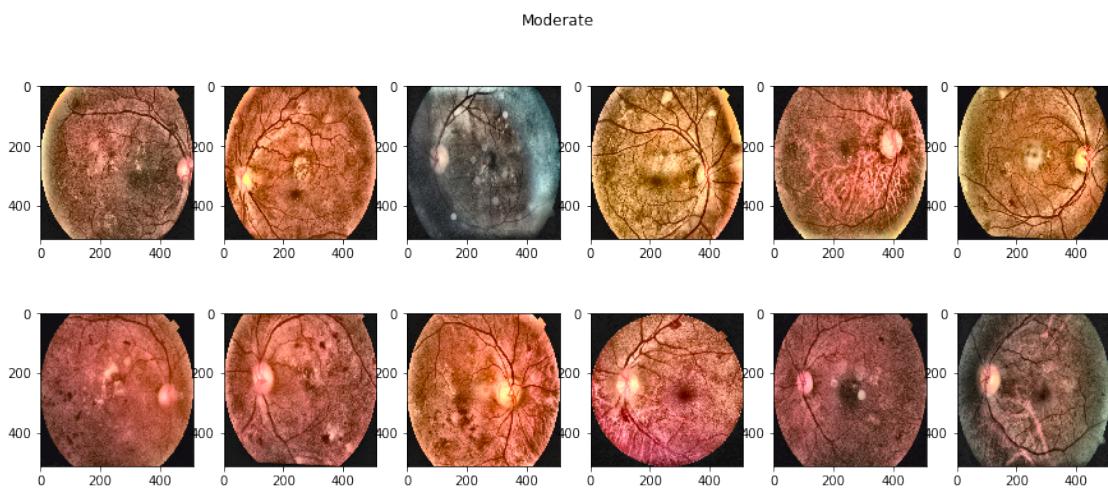


Figure 7: Moderate DR Data Samples

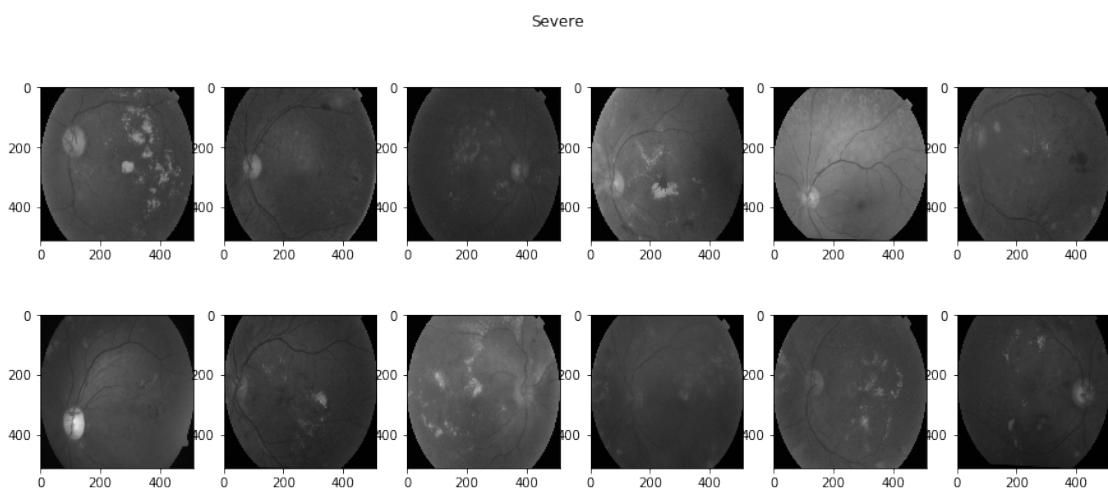


Figure 8: Severe DR Data Samples

Netra:

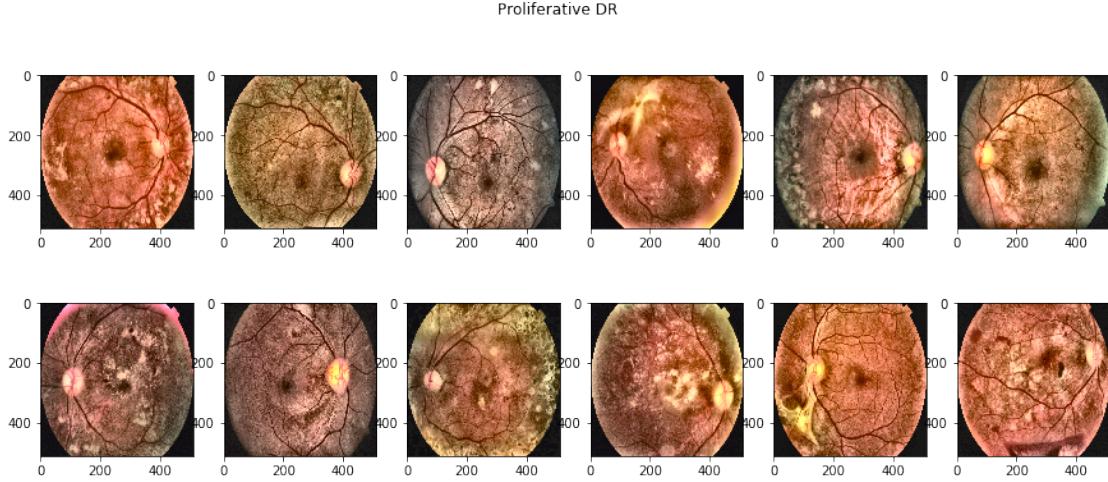


Figure 9: Proliferative DR Data Samples

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 DR
- 3 - Severe
- 4 - Proliferative DR

. The images in the dataset come from different models and types of cameras, which can affect the visual appearance of left vs. right. Some images are shown as one would see the retina anatomically (macula on the left, optic nerve on the right for the right eye). Others are shown as one would see through a microscope condensing lens (i.e. inverted, as one sees in a typical live eye exam). There are generally two ways to tell if an image is inverted:

- It is inverted if the macula (the small dark central area) is slightly higher than the midline through the optic nerve. If the macula is lower than the midline of the optic nerve, it's not inverted.
- If there is a notch on the side of the image (square, triangle, or circle) then it's not inverted. If there is no notch, it's inverted.

. The second data set was the publicly available Messidor-2 [17] data set which has been used by other groups for benchmarking performance of automated detection algorithms for diabetic retinopathy. The images were obtained between January 2005 and December 2010 at 3 hospitals in France using a Topcon TRC NW6 non mydriatic camera and 45° fields of view centred on the fovea. Approximately 44% of images were acquired with pupil dilation [17]. We can summarise all these works of DR classification problem into two groups. The first one is the binary classification of DR i.e. either the patient has DR or not. The problem in this method is we can not identify the severeness of the disease even after knowing that a patient has DR. The solution to this problem is multi-class classification. In multi-class classification, we classify DR into five different classes or stages as discussed in the introduction section. But most of the related work is unable to properly classify all the stages of DR, especially the early stages. It is important to predict the DR at early stages for the cure as in later stages it is difficult to cure the disease and can lead to blindness. To the best of our knowledge, no other work has detected the mild stages of DR, by using the Kaggle dataset which we have used for our research. Our model can detect the mild stage and performs better than the current state of the art. Moreover, in the related work, no one has shown the effect of a balanced dataset. The imbalanced dataset can lead to bias in the classification's accuracy. If the samples in the classes are equally distributed as in the case of a balanced dataset then the network can learn the features properly, but in case of unequal distribution the network outperforms for highly sampled classes.

3.2 Grading.

All images in the development and clinical validation sets were graded by ophthalmologists for the presence of diabetic retinopathy, diabetic macular edema, and image quality using an annotation tool . Diabetic retinopathy severity (none, mild, moderate, severe, or proliferative) was graded according to the International Clinical Diabetic Retinopathy scale. Referable diabetic macular edema was defined as any hard exudates within 1 disc diameter of the macula, which is a proxy for macular edema when stereoscopic views are not available. Image quality was assessed by graders using the rubric in the “Grading Instructions” section in the Supplement. Images of excellent, good, and adequate quality were considered gradable [19].

3.3 Data Preprocessing

We use the Kaggle dataset which contains 35126 color fundus images, each is of size 3888×2951 . It contains the images from five different classes based on the severity of diabetic retinopathy (DR).

	Class-0 (Normal)	Class-1 (Mild)	Class-2 (Moder- ate)	Class-3 (Severe)	Class-4 (PDR)	Total
Original	25,810	2,443	5,292	873	708	35,126
Up Sample	25,810	25,810	25,810	25,810	25,810	1,29,050
Down Sample	708	708	708	708	708	3,540

Table 1.

Table 1, shows the distribution of sample images in different classes of the Kaggle dataset. The distribution of different classes is shown in the first row of Table 1, which is perfectly imbalanced. Training of deep networks with imbalanced data leads to classification biasness. In the first preprocessing step, we resize each input image shown in Figure 5 (a) to 512×512 shown in Figure 6 (b) by maintaining the aspect ratio to reduce the training overhead of deep networks. Moreover, for balancing the dataset we performed up-sampling and down-sampling. The up-sampling (The Table 1, second row) is performed with augmentation of minority classes by randomly cropping patches, of size 512×512 as shown in Figure 7 (c). The dataset is divided into three parts: training, testing, and validation sets with ratio 64% and 20% and 16% respectively. During training, the validation set is used to check and reduce overfitting.

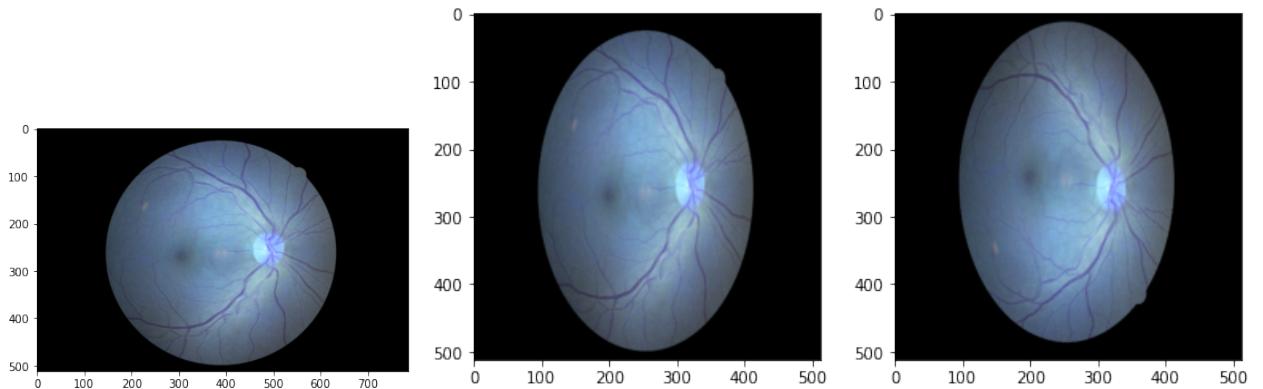


Figure 10: Original Image

Figure 11: Resized Image

Figure 12: Augmented Image

Netra:

3.4 Data Optimization.

3.4.1 Circular Image Cropping.

Computational Efficiency is one of the biggest feature of our model. And to make the CNN model more accurate and efficient, We first remove the uninformative black background of FUNDUS Images, so that model only has to do feature extraction from informative and crucial parts only.

$$dst(x, y) \Rightarrow src(x, y) \text{ if } src(x, y) > \text{threshold, otherwise, 0.}$$

We cropped our image by first taking out the black part using binary thresholding method, masking method and then resize the image into 512 x 512.

$$\begin{aligned} I(i, j) &= 5 * I(i, j) - [I(i - 1, j) + I(i + 1, j) + I(i, j - 1) + I(i, j + 1)] \\ &\quad i^j \backslash j \quad -1 \quad 0 \quad +1 \\ \iff I(i, j) * M, \text{ where } M &= \begin{matrix} -1 & 0 & -1 & 0 \\ 0 & -1 & 5 & -1 \\ +1 & 0 & -1 & 0 \end{matrix} \end{aligned}$$

Masking method matrix

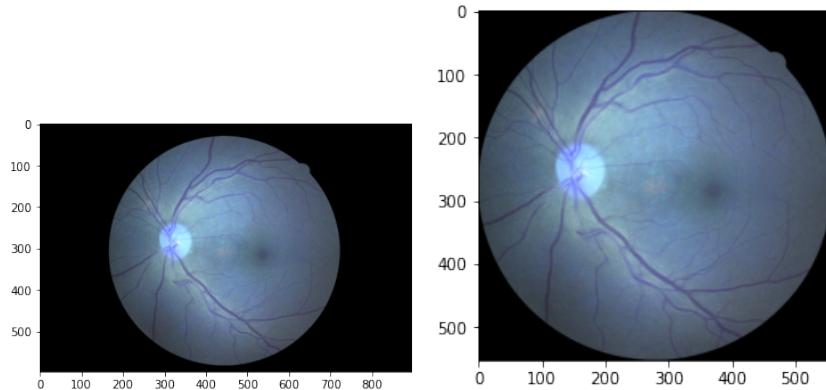


Figure 13: Original Image

Figure 14: Cropped Image

3.4.2 Green Channel Extraction.

After using various methods, we found out that green channel extraction and then doing principal component analysis gave the highest accuracy and computational efficiency for FUNDUS Images.

Netra:

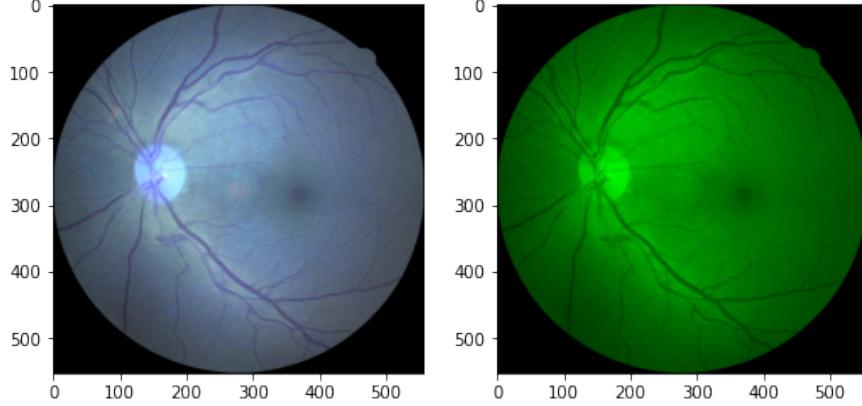


Figure 15: Original Image

Figure 16: Green Channel Image

3.4.3 Green Channel to Grayscale.

Grayscale conversion to increase computational efficiency and better feature extraction.

$$\text{RGB[A] to Gray: } Y \leftarrow 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \quad (1)$$

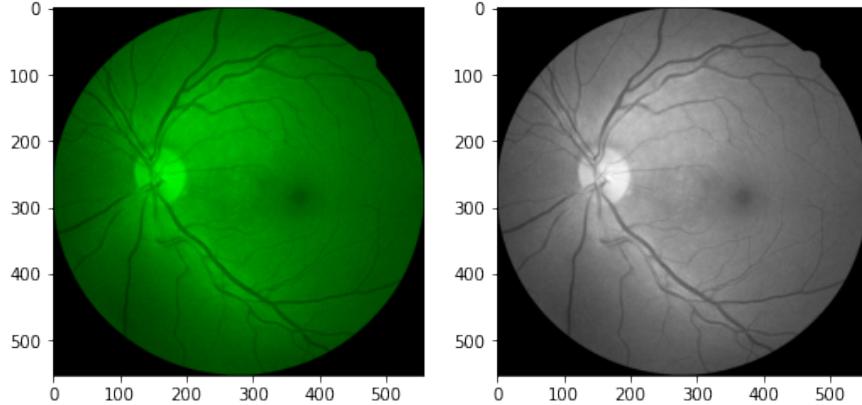


Figure 17: Original Image

Figure 18: Grayscale Image

3.4.4 Implementation of CLAHE on Grayscale Image

Below, it shows, how histogram equalization method can increase visibility of informative parts in FUNDUS Images. Consider a discrete grayscale image x and let n_i be the number of occurrences of gray level i . The probability of an occurrence of a pixel of level i in the image is

$$p_x(i) = p(x = i) = \frac{n_i}{n}, \quad 0 \leq i < L \quad (2)$$

L being the total number of gray levels in the image , n being the total number of pixels in the image, and Px_i being in fact the image's histogram for pixel value i , normalized to $[0,1]$. Let us also define the cumulative distribution function corresponding to i as

$$\text{cdf}_x(i) = \sum_{j=0}^i p_x(x = j) \quad (3)$$

Netra:

which is also the image's accumulated normalized histogram. We would like to create a transformation of the form $y = T(x)$ to produce a new image y , with a flat histogram. Such an image would have a linearized cumulative distribution function (CDF) across the value range, i.e.

$$\text{cdf}_y(i) = (i + 1)K \quad (4)$$

for some constant K . The properties of the CDF allow us to perform such a transform; it is defined as

$$y = T(k) = \text{cdf}_x(k) \quad (5)$$

where k is in the range $[0, L - 1]$. Notice that T maps the levels into the range $[0, 1]$, since we used a normalized histogram of x . In order to map the values back into their original range, the following simple transformation needs to be applied on the result:

$$y' = y \cdot (\max\{x\} - \min\{x\}) + \min\{x\} = y \cdot (L - 1) \quad (6)$$

y is a real value while y' has to be an integer. An intuitive and popular method is applying the round operation:

$$y' = \text{round}(y \cdot (L - 1)) \quad (7)$$

However, detailed analysis results in slightly different formulation. The mapped value y' should be 0 for the range of $0 < y \leq 1/L$ And $y' = 1$ for $1/L < y \leq 2/L$, $y' = 2$ for $2/L < y \leq 3$..., and finally $y' = \text{ceil}(L \cdot y) - 1$

CLAHE is advantageous not to discard the part of the histogram that exceeds the clip limit but to redistribute it equally among all histogram bins [20].

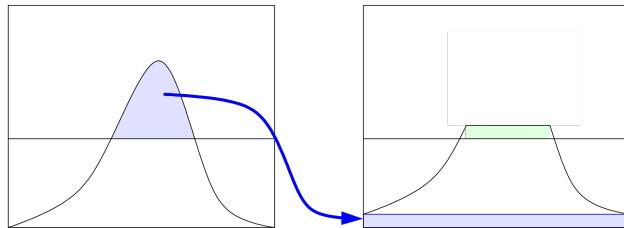


Figure 19: Histogram Equilization

The redistribution will push some bins over the clip limit again (region shaded green in the figure), resulting in an effective clip limit that is larger than the prescribed limit and the exact value of which depends on the image. If this is undesirable, the redistribution procedure can be repeated recursively until the excess is negligible.

3.4.5 Applying Contrast-limited adaptive histogram equalization

To increase the contrast between viens and lesions in the FUNDUS Image(which are the primary representation and indicators of progression/presence of diabetic retinopathy in a patient.) we tried various color contrasting methods and found out Contrast-limited adaptive histogram equalization method was ideale for FUNDUS Images

Netra:

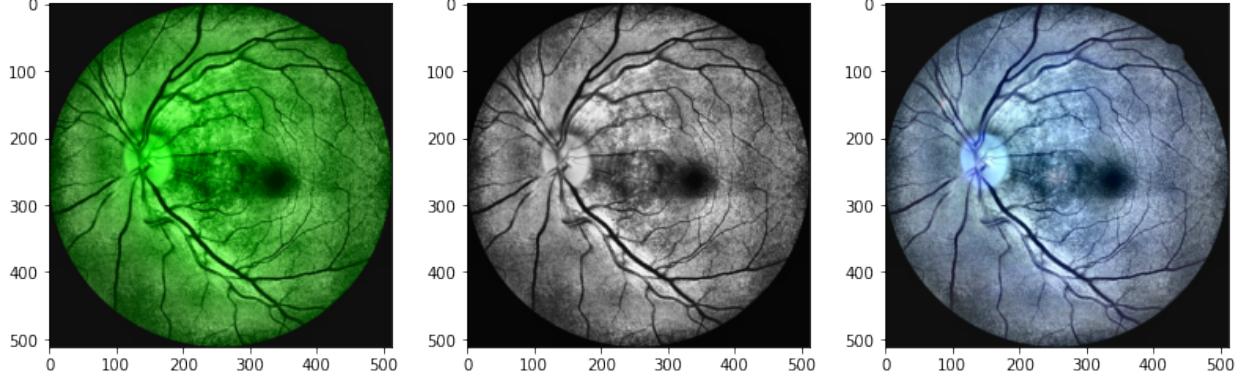


Figure 20: Green Histogram Equalized Image

Figure 21: Grayscale Histogram Equalized Image

Figure 22: Original Histogram Equalized Image

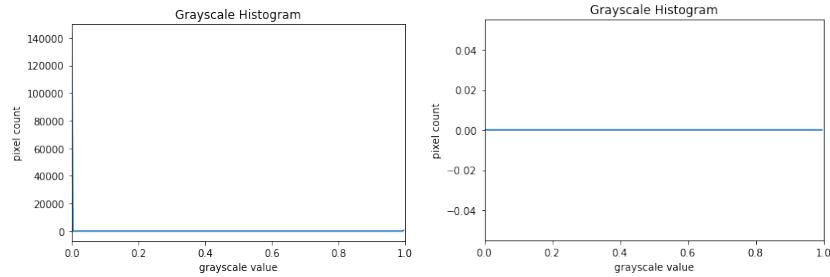


Figure 23: Original Image Histogram

Figure 24: CLAHE Imahe Histogram

3.4.6 Otsu Threshold Implementation.

To increase the visibility of veins in the eye, we used different thresholding methods and came down to otsu thresholding method to further enhance the image to increase the accuracy of CNN model. Since we are working with bimodal images, Otsu's algorithm tries to find a threshold value (t) which minimizes the weighted within-class variance given by the relation:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \quad (8)$$

$$\begin{aligned} q_1(t) &= \sum_{i=1}^t P(i) \quad \& \quad q_2(t) = \sum_{i=t+1}^I P(i) \\ \mu_1(t) &= \sum_{i=1}^t \frac{iP(i)}{q_1(t)} \quad \& \quad \mu_2(t) = \sum_{i=t+1}^I \frac{iP(i)}{q_2(t)} \\ \sigma_1^2(t) &= \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \quad \& \quad \sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)} \end{aligned} \quad (9)$$

Netra:

3.4.7 Otsu Threshold Filtering.

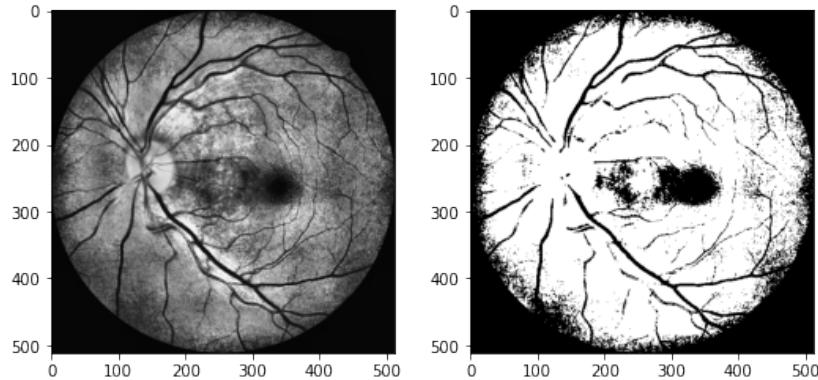


Figure 25: Grayscale Histogram Equalized Image

Figure 26: Threshold Image

3.4.8 Vien Density Extraction.

To increase the visibility of veins in the eye, we subtracted the mean filtered image from threshold image.

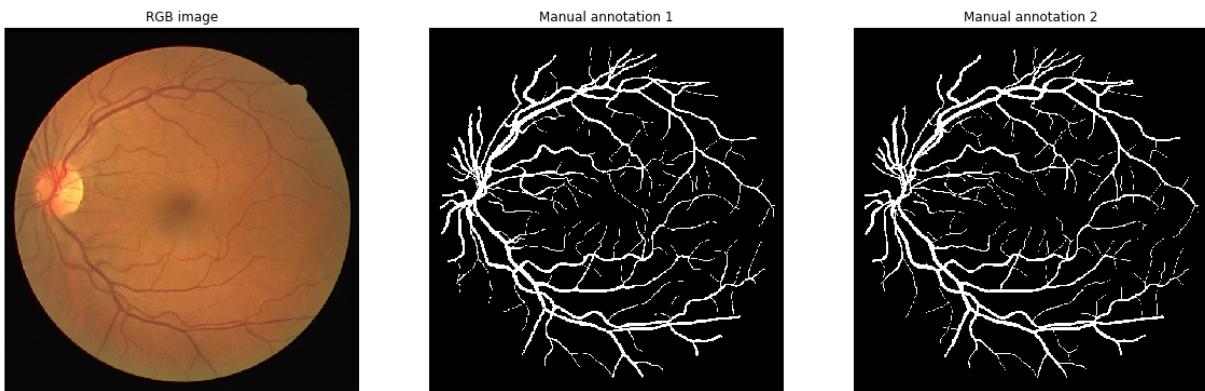


Figure 27: Grayscale Histogram Equilized Image

3.4.9 Train Test Split

After optimizing dataset based on methods described above, we distributed data into training dataset and testing dataset.

Netra:

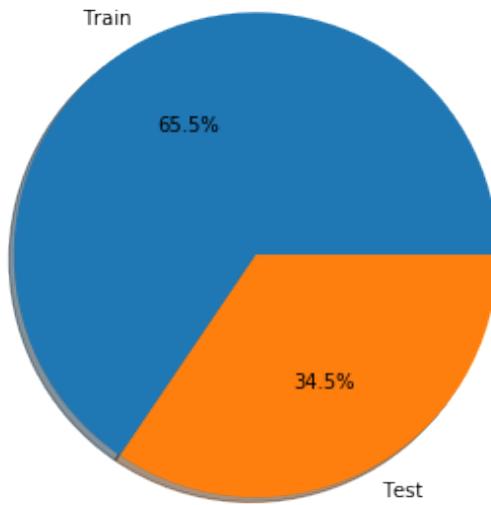


Figure 28: Data Distribution

3.5 Convolutional Layer

Kernel size

The Residual module in the ResNet architecture uses 1×1 and 3×3 filters as a form of dimensionality reduction which helps to keep the number of parameters in the network low (or as low as possible given the depth of the network).

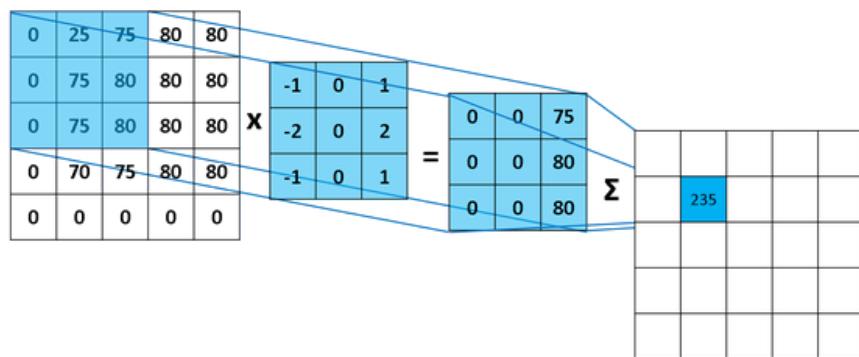


Figure 29: Kernel

Strides

The strides parameter is a 2-tuple of integers, specifying the “step” of the convolution along the x and y axis of the input volume. The strides value defaults to $(1, 1)$, implying that: A given convolutional filter is applied to the current location of the input volume. The filter takes a 1-pixel step to the right and again the filter is applied to the input volume. This process is performed until we reach the far-right border of the volume in which we move our filter one pixel down and then start again from the far left.

Netra:

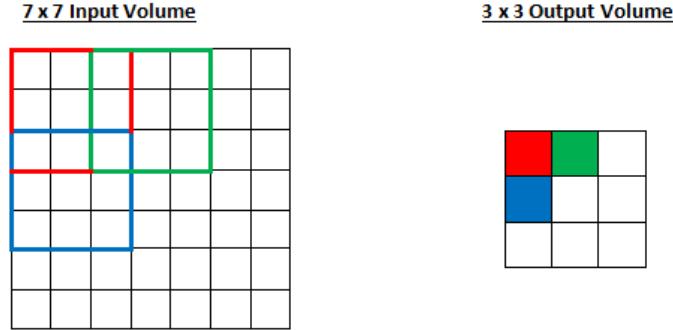


Figure 30: Strides

Padding

The padding parameter to the Keras Conv2D class can take on one of two values: valid or same . With the valid parameter the input volume is not zero-padded and the spatial dimensions are allowed to reduce via the natural application of convolution.

Activation algorithm

We have used the rectified linear unit activation algorithm. Both the ReLU function and its derivative are monotonic. If the function receives any negative input, it returns 0; however, if the function receives any positive value x , it returns that value. As a result, the output has a range of 0 to infinite.

$$f(x) = \max(0, x)$$

The ReLU deep learning function is simple and does not require any heavy processing. As a result, the model can train or operate in less time. Sparsity is another significant quality that we consider to be an advantage of utilising the ReLU activation function. Then flattening and then dropout layer.

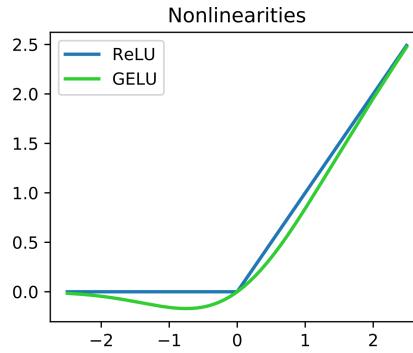


Figure 31: Data Distribution

The Dropout layer randomly sets input units to 0 with a frequency of rate at each step during training time, which helps prevent overfitting. Inputs not set to 0 are scaled up by $1/(1 - \text{rate})$ such that the sum over all inputs is unchanged.

The mechanism was implemented through Keras. We initially treated each individual frame of a video as independent from the others. This type of implementation caused label flickering. We, then, implemented a more advanced neural network, including LSTMs (Long Short Term Memory - this type of model is built to “remember” patterns over short and long durations) and the more general RNNs (Recurrent Neural Networks - these are used for sequential models like our data frame), helped combat this problem and led to much higher accuracy. Adding a rolling window average based on another machine learning trained window size, enabled us to smooth out the predictions and make for a better detection model.

Netra:

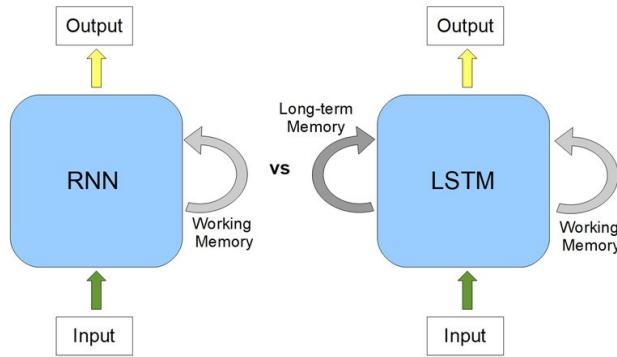


Figure 32: RNN v/s LSTM

Padding

3.6 Python Model

We used flask to host our CNN at heroku, below is the API code.

```
from flask import Flask, request, jsonify
import numpy as np
import cv2
from skimage import io
import json
import matplotlib.pyplot as plt

pickled_model = pickle.load(open('model.pkl', 'rb'))

app = Flask(__name__)

@app.route('/')
def home():
    return "Hello World"

@app.route('/predict', methods=['POST'])
def predict():
    links=[]
    zeros=0
    ones=0
    twos=0
    threes=0
    fours=0
    for i in range(1,5):
        path = request.form.get(f'url{i}')
        links.append(path)
        img = io.imread(path, plugin='matplotlib')
        img = cv2.cvtColor(img, cv2.COLOR_RGB2Lab)
        clahe = cv2.createCLAHE(clipLimit=10,tileGridSize=(8,8))
        img[:, :, 0] = clahe.apply(img[:, :, 0])
        img = cv2.cvtColor(img, cv2.COLOR_Lab2RGB)
        bgr = cv2.resize(img, (512,512))
        image = np.array(bgr) / 255.0
        predict=pickled_model.predict(np.array([image]))
        per=np.argmax(predict, axis=1)
        if per[0]==0:
            zeros+=1
        elif per[0]==1:
            ones+=1
```

Netra:

```
    elif per[0]==2:  
        twos+=1  
    elif per[0]==3:  
        threes+=1  
    elif per[0]==4:  
        fours+=1  
    return jsonify({"zeros":zeros,"ones":ones,"twos":twos,"threes":threes,"fours":fours})  
  
if __name__=='__main__':  
    app.run(debug=True)
```

3.7 Model Summary and Results

We built the CNN model (using Keras with a backend of TensorFlow). For it We had to make the convolution layers, pooling layers, and dense layers and pass my data through them. Convolutional layers do the feature extraction by using Kernels and pooling layers reduce dimensionality. I then flattened the data matrix and prepared it for dense layers. Our model has 16 layers (refer to the next figure for a full model summary and the name and output of each layer in our model). We trained our model to the training images dataset with 20 epochs (1 epoch means that the computer looked through the data 1 time.).

Model Summary

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 512, 512, 3)	0	
conv1_pad (ZeroPadding2D)	(None, 518, 518, 3)	0	input_1[0][0]
conv1 (Conv2D)	(None, 256, 256, 64)	9472	conv1_pad[0][0]
bn_conv1 (BatchNormalization)	(None, 256, 256, 64)	256	conv1[0][0]
activation_1 (Activation)	(None, 256, 256, 64)	0	bn_conv1[0][0]
pool1_pad (ZeroPadding2D)	(None, 258, 258, 64)	0	activation_1[0][0]
max_pooling2d_1 (MaxPooling2D)	(None, 128, 128, 64)	0	pool1_pad[0][0]
res2a_branch2a (Conv2D)	(None, 128, 128, 64)	4160	max_pooling2d_1[0][0]
bn2a_branch2a (BatchNormalizati	(None, 128, 128, 64)	256	res2a_branch2a[0][0]
activation_2 (Activation)	(None, 128, 128, 64)	0	bn2a_branch2a[0][0]
res2a_branch2b (Conv2D)	(None, 128, 128, 64)	36928	activation_2[0][0]
bn2a_branch2b (BatchNormalizati	(None, 128, 128, 64)	256	res2a_branch2b[0][0]
activation_3 (Activation)	(None, 128, 128, 64)	0	bn2a_branch2b[0][0]
res2a_branch2c (Conv2D)	(None, 128, 128, 256)	16640	activation_3[0][0]
res2a_branch1 (Conv2D)	(None, 128, 128, 256)	16640	max_pooling2d_1[0][0]
bn2a_branch2c (BatchNormalizati	(None, 128, 128, 256)	1024	res2a_branch2c[0][0]
bn2a_branch1 (BatchNormalizatio	(None, 128, 128, 256)	1024	res2a_branch1[0][0]
add_1 (Add)	(None, 128, 128, 256)	0	bn2a_branch2c[0][0] bn2a_branch1[0][0]
activation_4 (Activation)	(None, 128, 128, 256)	0	add_1[0][0]

Netra:

```
final_output (Dense)      (None, 5)      5125      dropout_2[0][0]
=====
Total params: 25,691,013
Trainable params: 25,637,893
Non-trainable params: 53,120
```

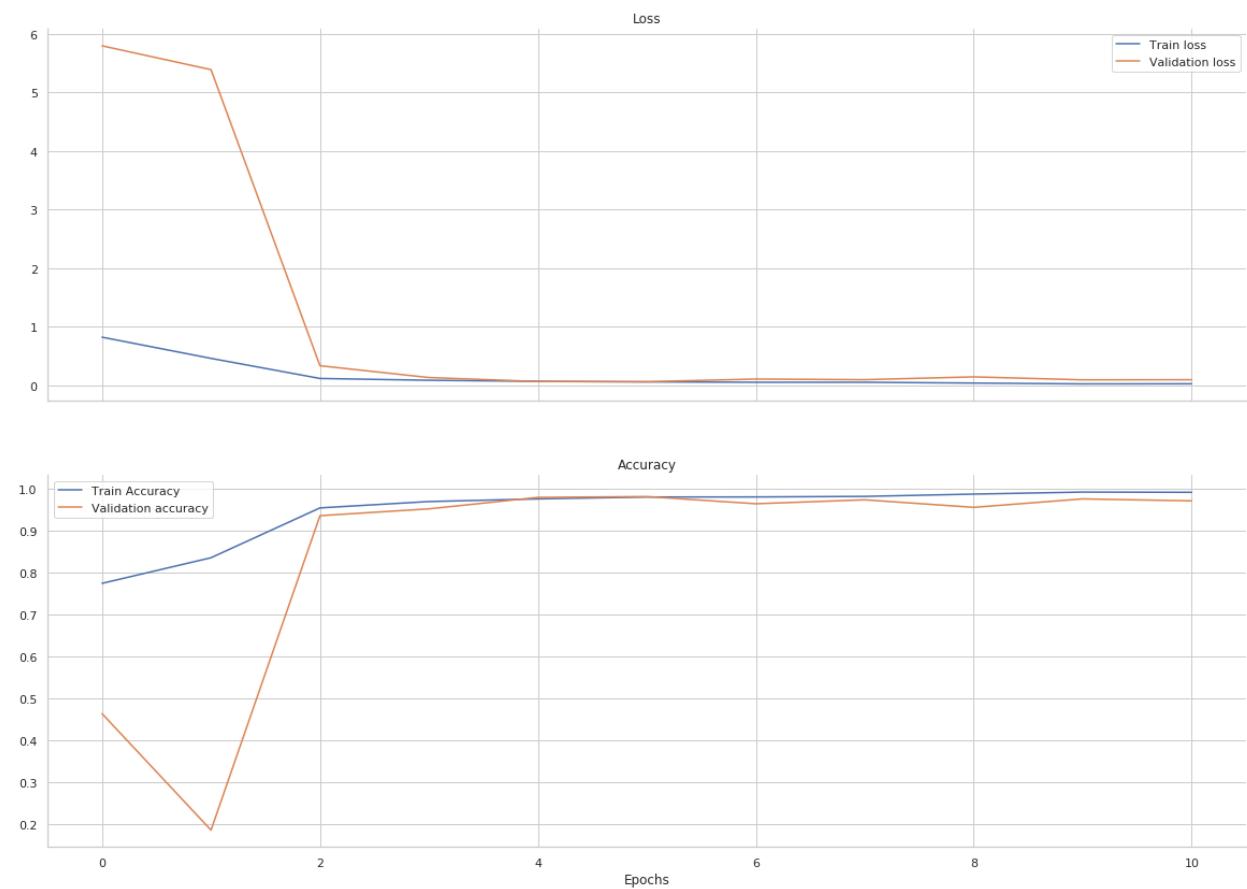


Figure 33: CNN Model Validation Accuracy (VAL_ACC) and Validation Loss (VAL_LOSS)

Netra:

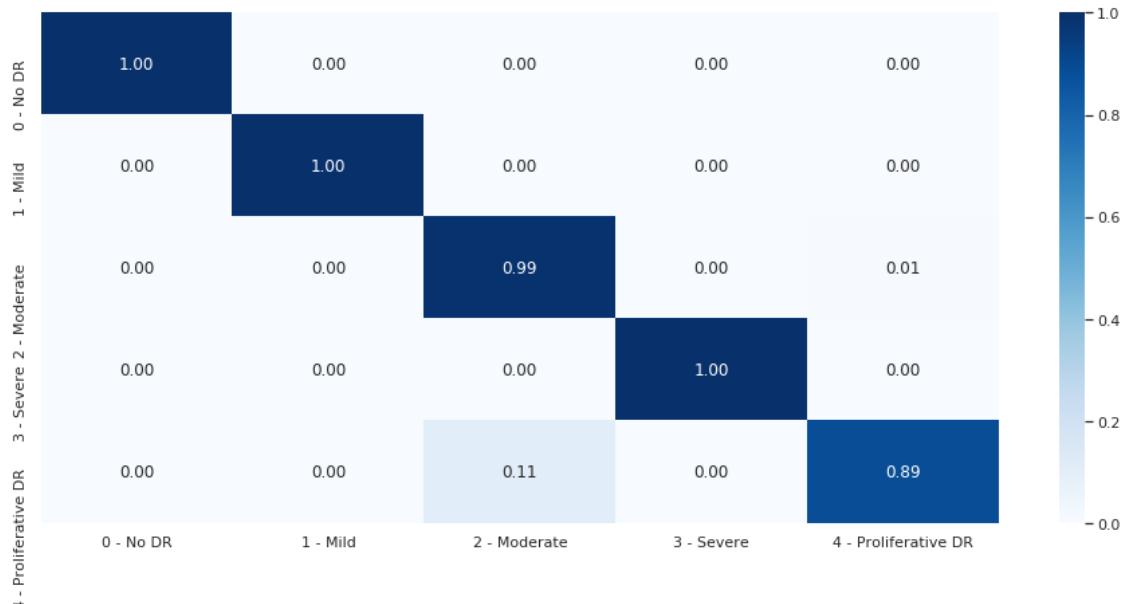


Figure 34: CNN Model Confusion Matrix

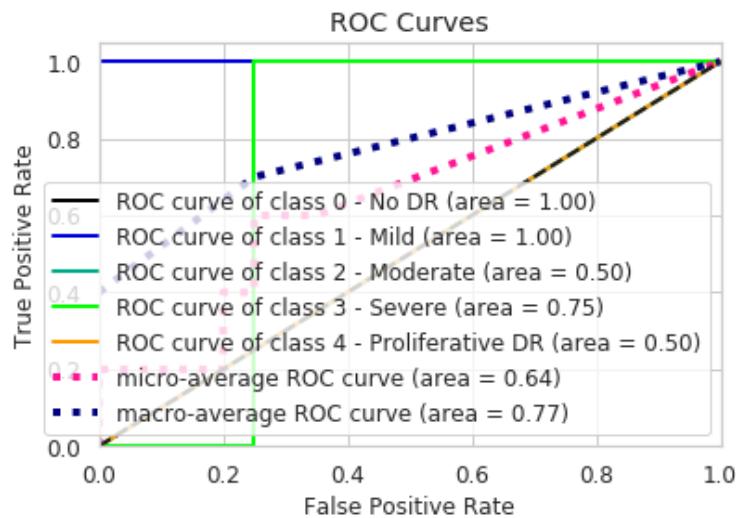


Figure 35: True Positive Rate v/s False Positive Rates

3.8 Testing the CNN Model using Validation datasets

The CNN model was validated using two validation datasets mentioned above. And the results were:

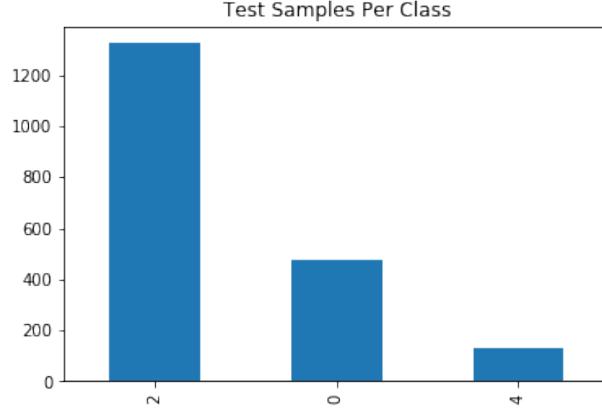


Figure 36: Testing Results

4 DIYretCAM Netrascopy FUNDUS Camera V1

4.1 Problems with Fundus Imaging

Fundus photography plays a key role in monitoring and follow-up of patients. Recently, smartphones have been used for fundus documentation [21]. A fundus camera is a complex optical system used for imaging the retina of the eye. Retinal imaging presents a unique difficulty considering that the retina must be illuminated and imaged simultaneously, a process which forces illumination and imaging systems to share a common optical path [22]. Because the retina is a minimally reflective surface, the power of the back reflections from the shared optics of the illumination and imaging paths is greater than the power reflected by the retina. Retinal imaging is complicated due to a number of factors: the location of the retina forces imaging illumination to share a common path through the pupil of the eye, the retina is minimally reflective, aberrations in the eye blur the image of the retina, the pupil of the eye restricts the amount of light that can illuminate the retina, and back reflections from the cornea and any other optical surface is likely to be significantly greater than the light reflected from the retina [23].

4.2 Principal

The basic design principle of fundus cameras is to provide maximum illumination to the retina while eliminating back reflections [24]. Although current fundus cameras have advanced significantly since their introduction, the traditional tabletop optical design has remained largely unchanged [23]. Complex optical assemblies in current devices provide high-resolution imaging of the fundus but also require dedicated clinical space and high manufacturing costs. Portable cameras have recently become commercially available, but most remain difficult to use in a hand-held manner and often have substandard image quality, compared to their tabletop counterparts [25]. The commercial field of fundus camera equipment stands in unique contrast to consumer digital camera technology, where personal cameras are becoming ever cheaper, smaller, and easier to use. Although other ophthalmic equipment manufacturers have recently incorporated consumer digital single-lens reflex (DSLR) cameras into their fundus camera designs, they do not make full use of the consumer camera's built-in functions or space-saving design. Traditional fundus camera designs are thus ill suited to leverage the significant cost reductions and technological advancements of consumer camera technology [26].

4.3 Our Design

Here, We present the technical modifications needed to transform a consumer digital camera into a mydriatic fundus camera. We also demonstrate the functional feasibility of using such a camera in a variety of clinical settings to produce images of the posterior pole of the eye with pathology detail resolution comparable to that of existing fundus cameras. . We designed an optical system based upon this imaging principle that comprised a two-part modular system: an optical attachment that integrated all of the optical components necessary to produce an image of the fundus, and a camera that was used to compose, capture, and store an image of the fundus.

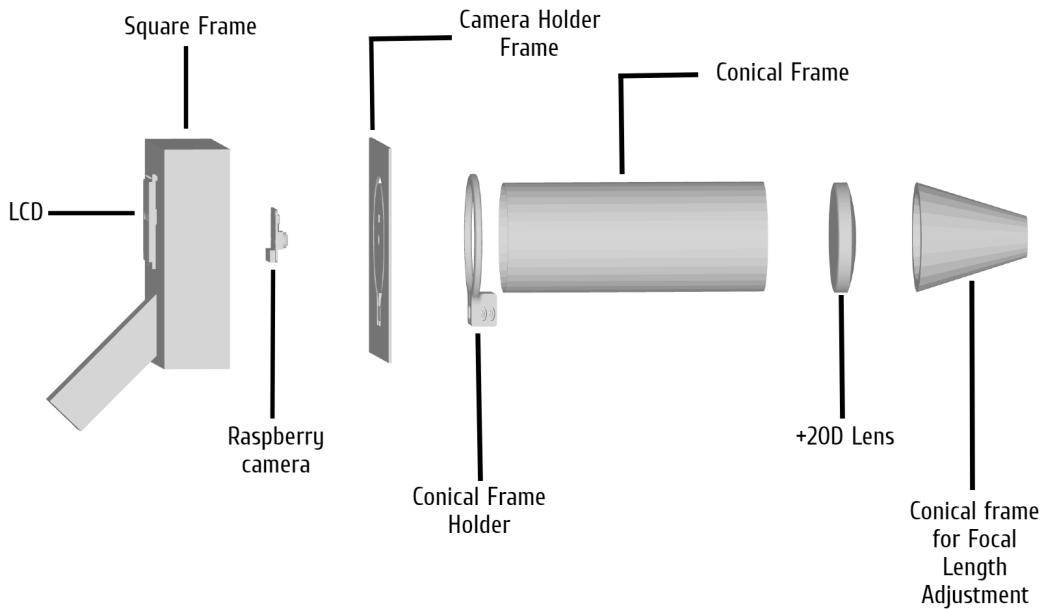


Figure 37: Netrascopy FUNDUS Camera CAD Model

The primary function of the optical attachment is to provide the imaging path necessary for the consumer camera to capture an image of the fundus. A 22-diopter (D) indirect ophthalmic lens was used as the front objective lens of the fundus camera. This lens was selected owing to a number of favourable characteristics including: a 60 Degree field of view, a large clear aperture of 52 mm for improved light transmission, and advanced antireflection coatings. A Raspberry Camera was selected as the camera back . This camera has a number of desirable features including large 12-megapixel (MP) complementary metal-oxide-semiconductor (CMOS) sensor, rapid automatic focus and exposure capabilities, hot-shoe adapter, Live-View imaging, interchangeable lensing, and built-in image stabilisation.

4.4 Components Required

- 50 mm PVC pipe
- 24cm long for the optical tube.
- 120 grit sandpaper.
- Electrical insulation tape
- Raspberry Pi.
- Camera to take photos.
- 3D Printed Frames.
- +22 D lens

Netra:



Figure 38: +22D Lens

Figure 39: Raspberry pi and Raspberry pi camera

Human fundus photography is based upon the principle of indirect ophthalmoscopy. A front objective lens is positioned at a working distance of 5 to 50 mm from the front of the eye. This lens is used to simultaneously relay light rays toward the eye, collect the reflected light, and also provide a magnified view of the fundus.

Dimensions: 100.99 x 243.41 x 130.00 mm

Surface: 115,784.99 mm²

4.5 Circuit

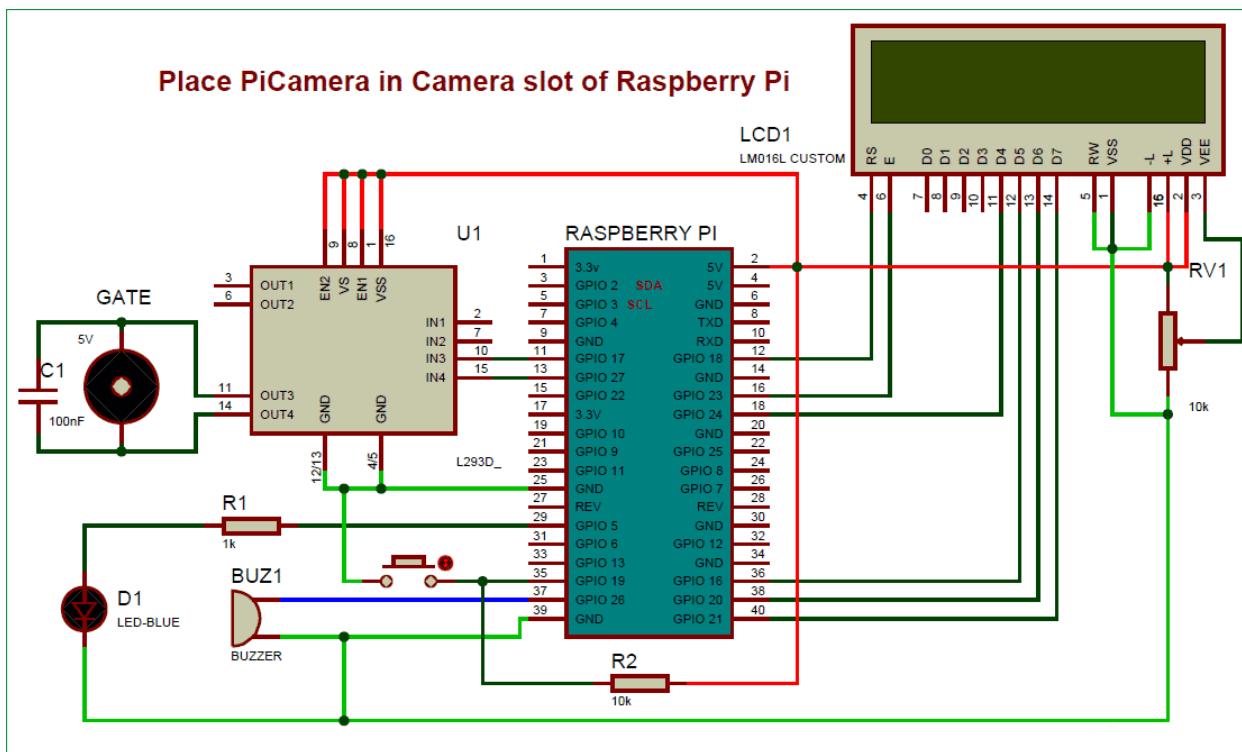


Figure 40: Netrascopy FUNDUS Camera CAD Model Circuit

4.6 Fundus Illumination

A continuous illumination source for image composition and focusing was provided by a light emitting diode (LED) with a colour temperature of 2700*K. The LED was prebuilt with a 10* condensing lens and was powered by a 1000-mA BuckPuck current driver using four AA batteries. Although this LED source could be used for final image capture, the low optical power would necessitate slow exposure settings. Therefore, an external xenon flash was incorporated into

Netra:

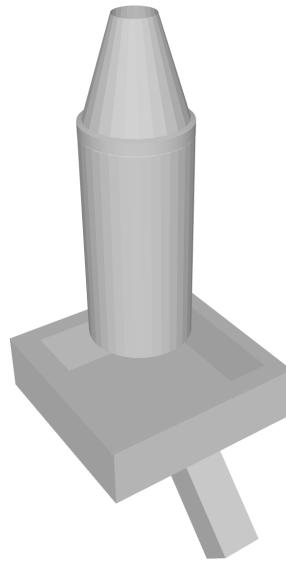


Figure 41: Netrascopy FUNDUS Camera CAD Model

the design to allow high shutter speed exposures. This feature decreases image noise and eliminates artefacts from subject eye drift or camera shake.

5 User End Product

5.1 Using CNN Model

We stored our CNN model as .pkl file and uploaded it to heroku using flask. To start testing user can submit 30-Second Fundus Video and upload it, once uploaded, the video will be sent to heroku where, our first API, FRAMES API, will convert the video into individual frames, and will send these images to CNN model hosted on another heroku pipeline, using a for loop, once all the test results are captured, the user will be redirected to results page. Where, they will be shown out of how many test case, they were tested positive/negative for different stages of diabetic retinopathy.

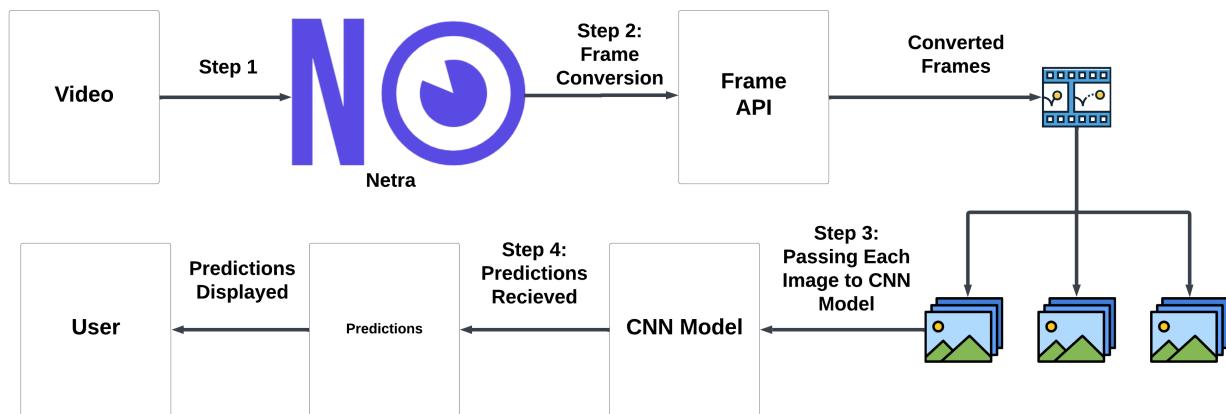


Figure 42: Using CNN Model

Netra:

5.2 Using Web Application

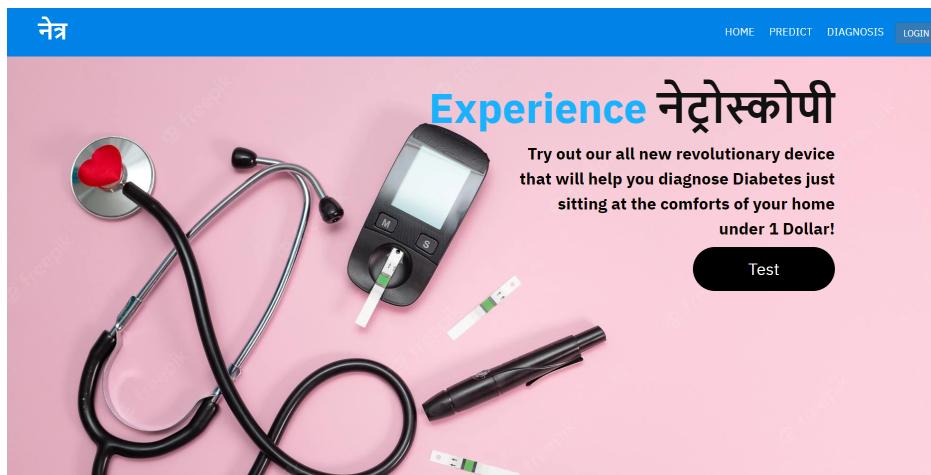


Figure 43: Home Page

The image shows the "Predict Diabetes" form page. At the top, there is a blue header bar with the word "नेत्र" (Netra) in white. To the right of the header are four navigation links: "HOME", "PREDICT", "DIAGNOSIS", and "LOGIN". Below the header, the main content area has a light gray background. The title "Predict Diabetes" is centered at the top. Below the title, a sub-instruction "Please upload a clear Fundus Image using Netrascopy" is displayed. There are several input fields: a file upload field labeled "Choose File: No file chosen", a text input field for "Name", a text input field for "Age", a dropdown menu for "Gender", and a green "Predict" button at the bottom.

Figure 44: Prediction Form Page

The image shows the "Predict Diabetes" form page with data entered. At the top, there is a blue header bar with the word "नेत्र" (Netra) in white. To the right of the header are four navigation links: "HOME", "PREDICT", "DIAGNOSIS", and "LOGIN". Below the header, the main content area has a light gray background. The title "Predict Diabetes" is centered at the top. Below the title, a sub-instruction "Please upload a clear Fundus Image using Netrascopy" is displayed. The input fields have been populated with data: the file upload field shows "Choose File: My Video.mp4", the name field shows "Arshroop", the age field shows "16", the gender field shows "Male", and the "Predict" button is visible at the bottom.

Figure 45: Pridction Form Page

Netra:

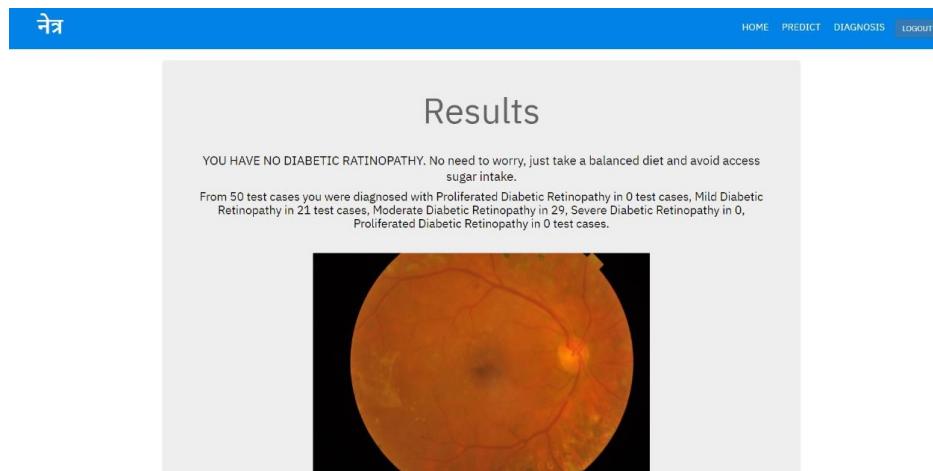


Figure 46: Results Page with test case description

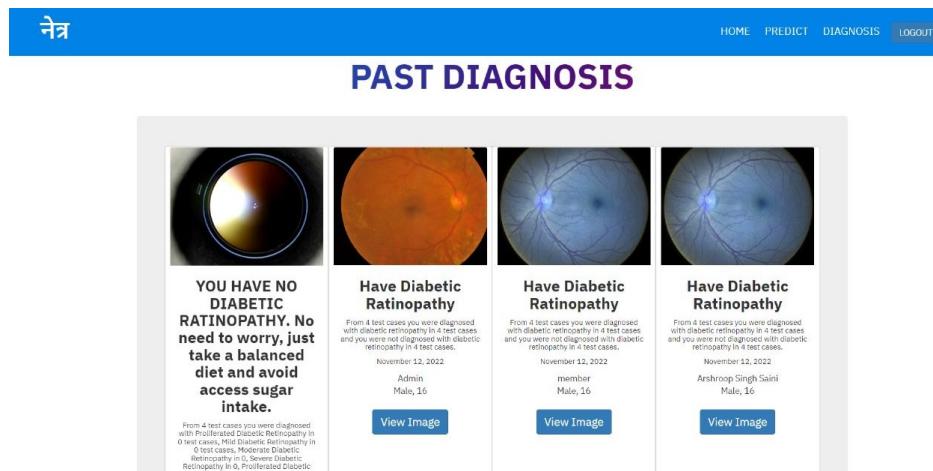


Figure 47: All Test Cases All at one place

6 Test Cases.

6.1 Test Case 1

Title: Fundus Video of Diabetic Retinopathy Patient.

Description: This 30-Second video contained the fundus of a patient who has mild symptoms of diabetic retinopathy. After feeding the video to our web app, the Frame API first extracted 30 frames from the video (1 Frame per second) and passed each frame to our convolutional neural network to reduce the overall validation loss.

Result: It took 20.62 seconds by the Web App to display results of each test case, and the patient was tested positive for Mild Diabetic retinopathy in 26 test cases, No DR in 3 test cases and Moderate diabetic retinopathy in 1 test case, so over all the model suggested that the patient is suffering from Mild DR, which was validated by an ophthalmologist.

6.2 Test Case 2

Title: Fundus Video of a Normal Person.

Description: This 30-Second video contained the fundus of a normal person who does not show any kinds of diabetic retinopathy symptoms. After feeding the video to our web app, the Frame API first extracted 30 frames from the video (1 Frame per second) and passed each frame to our convolutional neural network to reduce the overall validation loss.

Result: It took 16.56 seconds by the Web App to display results of each test case, and the patient was tested negative for Diabetic retinopathy in 29 test cases, Mild DR in 1 test case which was validated by an ophthalmologist. Similarly

Netra:

the model was tested multiple times with multiple people who were tested negative for any stage of diabetic retinopathy, which shows overall computational efficiency, and accuracy of the convolutional neural network, and eligibility of it being clinically usable.

7 Conclusion

In this project, we have developed a computationally efficient, and user-friendly solution for the one of its kind state-of-the-art stage specific early detection of diabetic retinopathy. Through our research we were successful in demonstrating that, data optimization before training the convolutional neural network can drastically reduce validation loss and can increase validation accuracy during training. Diabetic retinopathy doesn't tend to cause any symptoms in the early stages. Different kinds of data optimization methods were tried and tested but, Grayscale converted Green Channel image enhanced by Contrast-limited adaptive histogram equalization (LAB/CLAHE) method and otsu thresholding algorithm gave highest accuracy and reduce validation losses from 0.19 to 0.1, which is over 50% improvement. Diabetic retinopathy can cause permanent blindness if not diagnosed and treated promptly. screening can detect problems in your eyes before they start to affect your vision. The manual diagnosis process of DR retina fundus images by ophthalmologists is time-, effort-, and cost-consuming. Netrascopy is a more efficient system for Diabetic Retinopathy detection, which consists of a low cost, Camera, "DIYretCAM Netrascopy FUNDUS Camera V1", An Android Application and Web Application which aims to help patients and doctors detect diabetic retinopathy at early stages by taking 30-Second video of patient's retina and passing each frame as an individual test case to a convolutional neural network to detect probability of a patient having diabetic retinopathy. The results achieved in this study are significant, as early diagnosis and treatment can decrease the risks of vision impairment by 95%. This research moves AI-driven diagnosis one step closer to clinical applications from the research phase.

References

- [1] Mihaj Bakator and Dragica Radosav. Deep learning and medical diagnosis: A review of literature. *Multimodal Technologies and Interaction*, 2(3), 2018.
- [2] D. McCance, M. Maresh, and D.A. Sacks. *A Practical Manual of Diabetes in Pregnancy*. Practical Manual of Series. Wiley, 2017.
- [3] Kierstan Boyd. Diabetic retinopathy: Causes, symptoms, treatment, Oct 2022.
- [4] Rupert R A Bourne, Gretchen A Stevens, Richard A White, Jennifer L Smith, Seth R Flaxman, Holly Price, Jost B Jonas, Jill Keefe, Janet Leasher, Kovin Naidoo, Konrad Pesudovs, Serge Resnikoff, Hugh R Taylor, and Vision Loss Expert Group. Causes of vision loss worldwide, 1990–2010: a systematic analysis. *Lancet Glob. Health*, 1(6):e339–49, December 2013.
- [5] Helen K. Li, Mark Horton, Sven-Erik Bursell, Jerry Cavallerano, Ingrid Zimmer-Galler, Mathew Tennant, Michael Abramoff, Edward Chaum, Debra Cabrera DeBuc, Tom Leonard-Martin, and Marc Winchester. Telehealth practice recommendations for diabetic retinopathy, second edition. *Telemedicine and e-Health*, 17(10):814–837, 2011. PMID: 21970573.
- [6] Roy Taylor and Deborah Batey. *Handbook of retinal screening in diabetes: diagnosis and management*. John Wiley & Sons, 2012.
- [7] Early Treatment Diabetic Retinopathy Study Research Group et al. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified airlie house classification: Etdrs report number 10. *Ophthalmology*, 98(5):786–806, 1991.
- [8] Peter H Scanlon, Ahmed Sallam, and Peter Van Wijngaarden. *A practical manual of diabetic retinopathy management*. John Wiley & Sons, 2017.
- [9] Michael Dubow, Alexander Pinhas, Nishit Shah, Robert F. Cooper, Alexander Gan, Ronald C. Gentile, Vernon Hendrix, Yusufu N. Sulai, Joseph Carroll, Toco Y. P. Chui, Joseph B. Walsh, Rishard Weitz, Alfredo Dubra, and Richard B. Rosen. Classification of Human Retinal Microaneurysms Using Adaptive Optics Scanning Light Ophthalmoscope Fluorescein Angiography. *55(3):1299–1309*, 03 2014.
- [10] G S Scotland, P McNamee, A D Fleming, K A Goatman, S Philip, G J Prescott, P F Sharp, G J Williams, W Wykes, G P Leese, and J A Olson. Costs and consequences of automated algorithms versus manual grading for the detection of referable diabetic retinopathy. *British Journal of Ophthalmology*, 94(6):712–719, 2010.
- [11] Xue-Wen Chen and Xiaotong Lin. Big data deep learning: challenges and perspectives. *IEEE access*, 2:514–525, 2014.

Netra:

- [12] João Miguel Pires Dias, Carlos Manta Oliveira, and Luís A. da Silva Cruz. Retinal image quality assessment using generic image quality indicators. *Information Fusion*, 19:73–90, 2014. Special Issue on Information Fusion in Medical Image Computing and Systems.
- [13] Guy Hadash, Einat Kermany, Boaz Carmeli, Ofer Lavi, George Kour, and Alon Jacovi. Estimate and replace: A novel approach to integrating deep neural networks with existing applications. *arXiv preprint arXiv:1804.09028*, 2018.
- [14] Langis Gagnon, Marc Lalonde, M. Beaulieu, and Marie Boucher. Procedure to detect anatomical structures in optical fundus images. volume 4322, pages 1218–1225, 07 2001.
- [15] George Kour and Raid Saabne. Fast classification of handwritten on-line arabic characters. In *Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of*, pages 312–318. IEEE, 2014.
- [16] Michael Goldbaum, S. Moezzi, A. Taylor, S. Chatterjee, J. Boyd, E. Hunter, and Ramesh Jain. Automated diagnosis and image understanding with object extraction, object classification, and inferencing in retinal images. pages 695 – 698 vol.3, 10 1996.
- [17] Etienne Decencière, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Béatrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali Erginay, Béatrice Charton, and Jean-Claude Klein. FEEDBACK ON a PUBLICLY DISTRIBUTED IMAGE DATABASE: THE MESSIDOR DATABASE. *Image Analysis and Stereology*, 33(3):231, August 2014.
- [18] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Laÿ, and A. Chabouis. TeleOphta: Machine learning and image processing methods for teleophthalmology. *IRBM*, 34(2):196–203, April 2013.
- [19] Ling Dai, Liang Wu, Huating Li, Chun Cai, Qiang Wu, Hongyu Kong, Ruhan Liu, Xiangning Wang, Xuhong Hou, Yuxing Liu, Xiaoxue Long, Yang Wen, Lina Lu, Yaxin Shen, Yan Chen, Dinggang Shen, Xiaokang Yang, Haidong Zou, Bin Sheng, and Weiping Jia. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature Communications*, 12(1), May 2021.
- [20] Stephen M. Pizer, E. Philip Amburn, John D. Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B. Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3):355–368, 1987.
- [21] R Gutner and D Miller. Inside the fundus camera. *Annals of ophthalmology*, 15(1):13–16, 1983.
- [22] Edward DeHoog and James Schwiegerling. Fundus camera systems: a comparative analysis. *Applied optics*, 48(2):221–228, 2009.
- [23] Kenneth Tran, Thomas A. Mendel, Kristina L. Holbrook, and Paul A. Yates. Construction of an inexpensive, hand-held fundus camera through modification of a consumer point-and-shoot camera. *Investigative Ophthalmology & Visual Science*, 53(12):7600, November 2012.
- [24] Rudolf Kingslake. *Applied Optics and Optical Engineering V6*, volume 6. Elsevier, 2012.
- [25] Christine I Gliss, Jean-Marie A Parel, John T Flynn, Hans Pratisto, and Peter F Niederer. Toward a miniaturized fundus camera. *Journal of Biomedical Optics*, 9(1):126–131, 2004.
- [26] Kanagasingam Yogesan, Ian J Constable, Chris J Barry, Robert H Eikelboom, Ian L McAllister, and Mei-Ling Tay-Kearney. Telemedicine screening of diabetic retinopathy using a hand-held fundus camera. *Telemedicine Journal*, 6(2):219–223, 2000.