Ocular Disease Intelligent Recognition

Author

SYEDA GHINA SAHAR

273566

Dr. Syed Omer Gilani

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

Ocular Disease Intelligent Recognition (ODIR)

Author

SYEDA GHINA SAHAR

273566

A thesis submitted in partial fulfillment of the requirements for the degree of

MS BIOMEDICAL Engineering

Thesis Supervisor:

Dr. Syed Omer Gilani

Thesis Supervisor's Signature: _____

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD

July 2022

FORM TH-4

National University of Sciences & Technology

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by:

(Student Name & Regn No.): Syeda Ghina Sahar 00000273566

Titled: Ocular Disease Intelligent Recognition be accepted in partial fulfillment of the requirements for the award of MS degree.

Examination Committee Members

1. Name: Dr. Asim Waris Signature: _____

2. Name: Dr. Adeeb Shahzad Signature: _____

Supervisor's name: Dr. Syed Omer Gilani Signature: _____

Date: _____

_____ _____

Head of Department Date

COUNTERSINGED

Date: _____

_____

Dean/Principal

Thesis Acceptance Certificate

It is certified that final copy of MS thesis written by **Ms. Syeda Ghina Sahar**, Registration No. **00000273566** of **MS Biomedical Sciences and Engineering** (SMME) has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfilment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have been incorporated in the said thesis.

Supervisor: _____

Dr. Syed Omer Gilani

HOD: _____

Dr. Syed Omer Gilani

Principal: _____

Dr. Javed Iqbal


بسم الله

Declaration

I certify that this research work titled "*Ocular Disease Intelligent Recognition"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

Signature of Student

S.Ghina

2022-NUST-Ms-BME-000273566

Proposed Certificate for Plagiarism

It is certified that MS Thesis Titled **Ocular Disease Intelligent Recognition** by **Syeda Ghina Sahar** has been examined by us. We undertake the follows:

a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.
b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.
c. There is no fabrication of data or results which have been compiled/analyzed.
d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.
e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

Name & Signature of Supervisor

Name: Dr. Syed Omer Gilani

Signature

Copyright Statement

Acknowledgements

I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought which You setup in my mind to improve it. Indeed I could have done nothing without your priceless help and guidance. Whosoever helped me throughout the course of my thesis, whether my parents or any other individual was your will, so indeed none be worthy of praise but you.

I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout in every department of my life.

I would also like to express special thanks to my supervisor Dr. Omer Gilani for his help throughout my thesis and also for Signals and Images course which he has taught me. I can safely say that I haven't learned any other engineering subject in such depth than the ones which he has taught.

I would also like to pay special thanks to Mr. Talha Anwar for his tremendous support and cooperation. Each time I got stuck in something, he came up with the solution. Without his help I wouldn't have been able to complete my thesis. I appreciate his patience and guidance throughout the whole thesis.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study.

Dedicated to my exceptional parents and adored siblings whose tremendous support and cooperation led me to this wonderful accomplishment.

Abstract

To record anatomical details of the eye and anomalies, fundus imaging has proved very efficient. The most effective way to see and diagnose a wide range of eye diseases is through fundus imaging. Conditions that affect the blood vessels and areas surrounding it include diabetes-related retinopathy, glaucoma, AMD, myopia, cataract and hypertension. It's possible for the patient to have more than one ophthalmological problems that can be seen in one or both of his eyes. The dataset provided by ODIR is used in this study. The data has eight different categories for the diseases to be detected. By using transfer learning, two simultaneous models are described for solving the multi label problem for both the eyes (left and right). For the convolutional network, two synchronous efficient net models are implemented which are used with ADAM optimizers for better detection and results outcome. On the ODIR data set, B7 Efficient net along with focal loss outperformed the other approaches with an accuracy rate of 0.96%.

Key Words: *Fundus imaging; Ocular Disease detection; Convolutional Neural Networks; Ocular disease detection*

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Background, Scope and Motivation

A cheap and effective method of preventing blindness brought on by diabetes, glaucoma, cataracts, macular degeneration, and other disorders is fundus screening.

Early diagnosis of many eye conditions is difficult since few symptoms of diseases may be seen in their early stages. For persons who have had diabetes for a long period are more prone to diabetic retinopathy and macular edema. [1]

To develop methods for automatically categorizing eye diseases is the aim of this challenge. Categorizing the patients into following 8 categories is our agenda in ocular disease classification. [2]

- Normal
- Diabetes
- Glaucoma
- Cataract
- AMD
- Hypertension
- Myopia
- Other

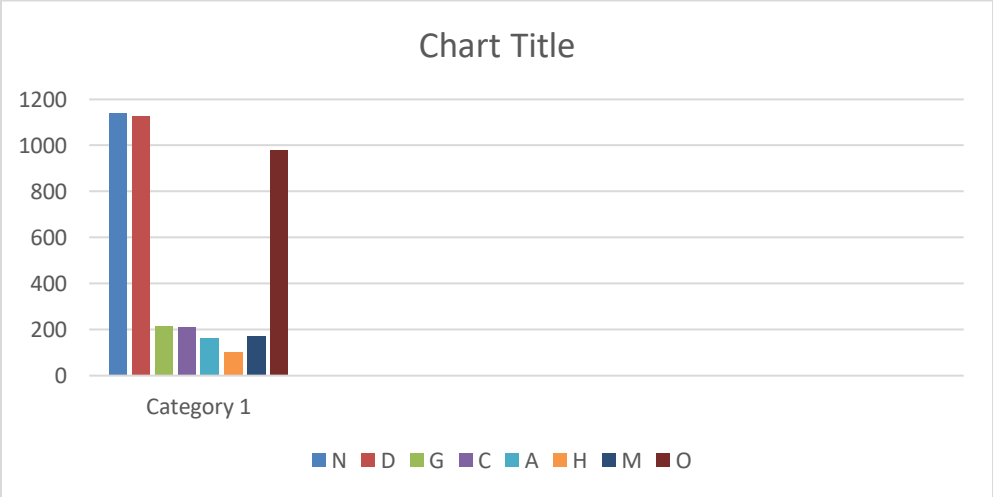


Figure 1.1: Uneven Distribution of Data

1.2 FUNDUS SCREENING

Retinal imaging and screening programs have been significantly impacted by the development of fundus photography. Fundus cameras are essential in the fight against blindness that can be prevented. A retinal camera also termed as a fundus camera is used to capture the details of the retina and other structures surrounding it. [2]

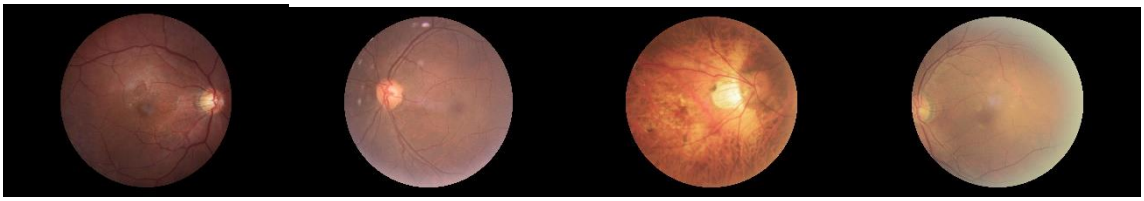


Figure 1.2: Fundus Screening

1.3 RETINA

The area which receives light and is responsible for converting it into chemical energy is known as retina. The nerves are activated by that chemical energy and as a result they conduct messages into the higher regions of the brain. Retina is an extension of the forebrain and is a complex nervous structure. The light which enters the eye focuses on the retina which is also light sensitive, and it ultimately results in vision. [2]

1.3.1 RETINA AND ITS PARTS

The retina has three different layers which are functional,

- The outer layer is the photoreceptor layer which constitutes of rods and cones. The vision is generated when the converted light energy in the form of electric signals gets processed and is then transmitted to the brain.
- The central & detailed vision is produced in the macula which is also coined as the central part of retina. [3]

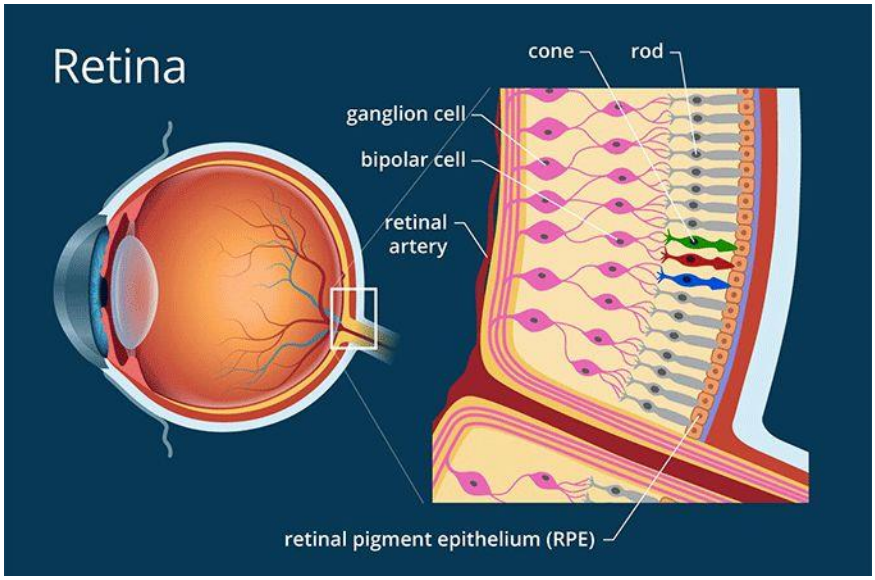


Figure 1.3: Retina And Its Parts

1.3.2 ORDERLY ARRANGEMENT OF RETINAL CELLS

The outer core layer, which is made up of the rod and cone nuclei, is created by the retinal cells arranged in a systematic way. Two layers: the inner nuclear layer and ganglion layer. The connections that neurons make. As a result, the bipolar cells' dendritic processes are connected to the rod and conical projections in the outer plexiform layer via these rod balls and cone stems, which also serve as a conduit for the alterations brought about by light in the rods and cones.

Axons are the projections that the cell uses to transmit impulses. They enable communication between the bipolar cells and the ganglion cells, where it is passed via the ganglion cells' axons as communication from the optic nerve, located in the inner plexiform layer.

1.3.3 PHOTORECEPTOR CELLS

Rods and cones are the names for the two categories of photoreceptor cells found in the retina of humans and the majority of animals. Cones tend to be thicker than rods, but they are all constructed on the same plane. The outer portion, which is present in the pigment epithelium, contains the photosensitive pigment. The bipolar and horizontal cells get the effects of light through the other end, known as the synaptosome. The restricting film that wrapped the outside of the rod and cone appeared to have been folded, as seen under an electron microscope, to form a stack of discs on

the outside. Because of this, the visual pigments on the surface of these discs are dispersed over a relatively large area, which enhances the visual unit's ability to absorb light more effectively. [3]

Light must pass through the light-insensitive layer in order to reach the photosensitive rod and cone due to the structure of the retina. The development of the fovea, a portion of the retina that is close to the optical axis of the eye but lacks the inner layer of the retina, substantially mitigates the optical flaws of this arrangement. As a result, the photoreceptor cell develops a depression, or centre well, through which light can enter relatively freely. Basically, the eye is guided toward the object of interest such that its image fits within the defined area using this region of the retina to produce an accurate vision. [4]

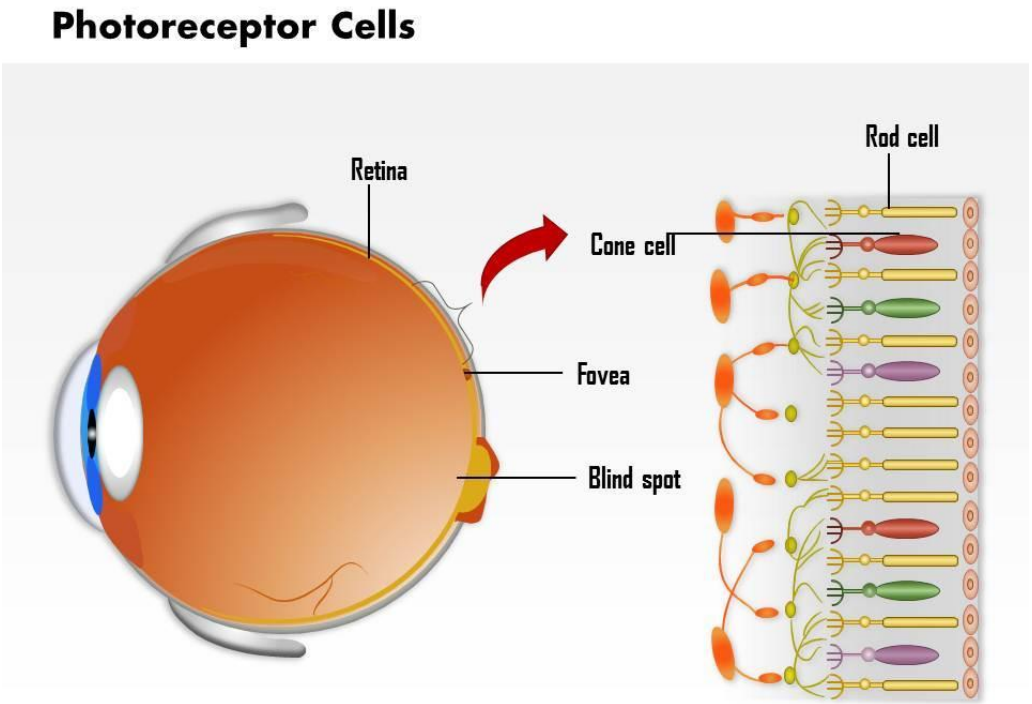


Figure 1.4: Photoreceptor Cells

CHAPTER 2: LITERATURE REVIEW

2.1 Functioning Of Retina

The lens together with cornea serve as two prime bodies that focus light onto the retina. Some people need glasses to function at some point in their lives as their vision deteriorates. To allow light to reach the retina, the vitreous body and the center of the eye must be clear. Cataracts are one medical disorder that prevents this process and causes the opacification of the eye. The retina must be in good condition for individuals to see under typical circumstances.''

2.1.1 Retinal Diseases

The retina receives light to create images that are then converted into electric signals, much like a digital card in a camera. Similar to how retina requires nutrients to work properly, everything else does as well. Additionally, human retinal blood vessels carry out this function. The diabetic retinopathy that affects these blood vessels. A person will become blind if the normal blood flow to the retina is interrupted, either by vein or artery obstruction.

Another retinal illness, known as age-related macular degeneration (AMD), causes edema and bleeding from choroid blood vessels. Diseases of the choroid can also result in blindness. Vision loss results from retinal detachment.
If the macula, the retina's core region, is uneven, vision becomes blurred and not clear.
The macula become moist due to AMD, diabetes, and excessive blood pressure. Vision blur could result from this.

2.2 Diabetic retinopathy classification using deeply supervised ResNet

In this paper, a well-trained ResNet approach was proposed for categorizing the severity of DR. Here the convolutional neural networks design is extended to three sets of side output layers. Multi-

scale learning can be implemented by incorporating predictions from the intermediate supervised layer, improving the overall performance. [5]

2.3 Machine learning applied to retinal image processing for glaucoma detection

A variety of technological tools can detect glaucoma. [6]. Many internal eye structures can be easily detected using the technique of Fundus imaging. [7] A wavelet feature extraction method in 2016 was given by Singh et al. [10]. In addition, in the segmented OD, the methodology of extracting wavelet feature was used.

2.4 Machine learning applied to retinal image processing using disc segmentation algorithm

An algorithm coined as 'cup and disc segmentation' was suggested in 2019 by Mohamed et al [11]. The super pixel method was used to group the given pixels of available data. During preprocessing the PSNR and CNR were used for evaluation purposes, quantitatively. After segmentation, the mean, variance, kurtosis, and skewness features were extracted using SPL. For classification step of the SVM used both the RBF and the linear function kernel. The RIM-ONE database was used for the approach testing. [12].

2.5 Machine learning applied to retinal image processing using Inception-V3

In 2018, Li et al. [13] developed a technique for evaluating a deep learning system's efficacy for GON screening. To achieve this, a 32-size mini-batch gradient descent was employed throughout the training process together with the Inception-v3 [14] in the network design. Then, Adam Optimizer was used to complete the convergence procedure [15]. The learning rate of 0.002 was termed as the best.

2.6 Machine learning applied to retinal image processing using U-net Convolutional Network

A picture fractionalization and classification technique was suggested by Dos Santos. [16]. It started off with the training of a U-net CNN before the completion of segmentation of the OD. Currently, training uses 80 percent of the data, while testing uses 20 percent.

2.7 Classification of retinal images based on convolutional neural network

A crucial step in obtaining high-accuracy results for ophthalmologists to detect and treat the disease early is the automatic diagnosis of maculopathy disease. Ophthalmologists must put in a lot of effort and time manually identifying diabetic maculopathy. In retinal pictures, exudates can be seen and used to diagnose maculopathy. The fundamental architecture for retinal image classification that is suggested in this study starts with fuzzy preprocessing to enhance the original image and the contrast between the foreground and background. After segmenting the image using binarization to separate the optic disc and blood vessels, the original image is then cleaned up. After that, a gradient method is applied to the retinal image to help distinguish between normal and pathological conditions. [17]

CHAPTER 3: "OVERVIEW OF DATASET"

3.1 ODIR DATASET

The data set was acquired from the ODIR 2019 Grand challenge which was hosted by Peking University and organized by Institute of Artificial Intelligence at Peking University and Shang gong Medical Technology Co. Ltd. The dataset has an ophthalmic data of 5000 patients from both the eyes, left and right, correspondingly. The images have varied resolutions since they're captured by using Kiowa, Canon and Zeiss. The data set is divided into eight categories.

Normal (N)	Diabetes (D)	Glaucoma (G)	Cataract (C)	AMD (A)	Hypertension (H)	Myopia (M)	Other (O)

Table 3.1: Eight Categories of Ophthalmological Diseases

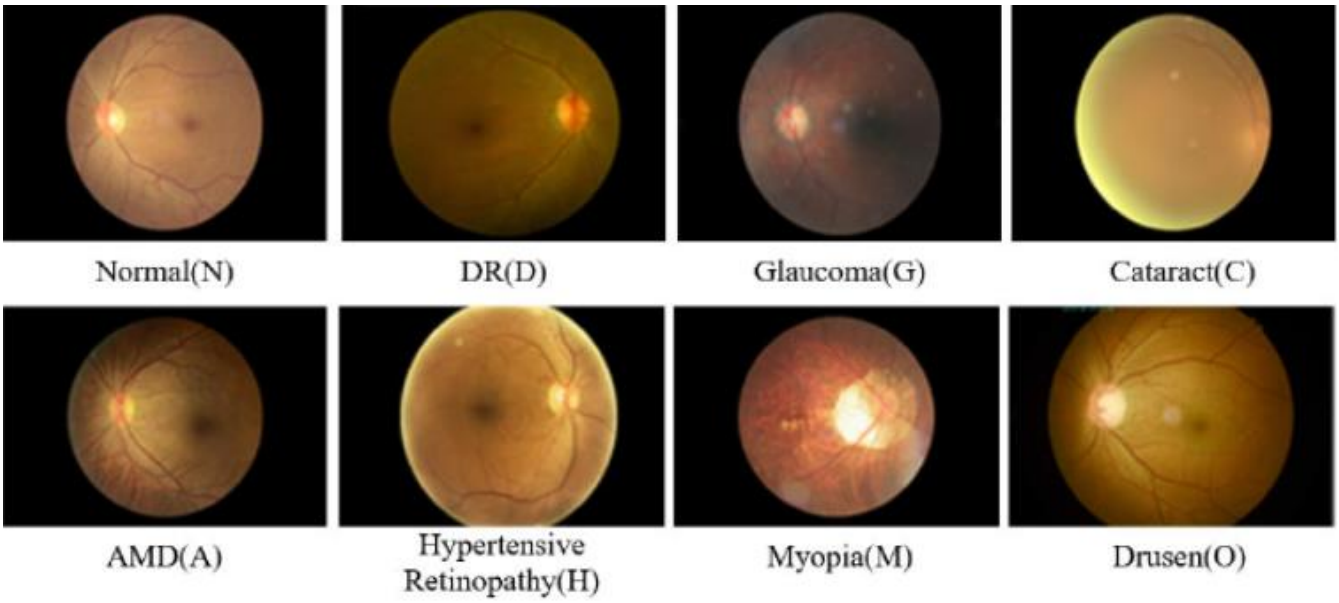


Figure 3.1: Image Sample Corresponding to Eight labels

Since the data set is multi-label, so there are eye-data sets which obviously has more than one disease.

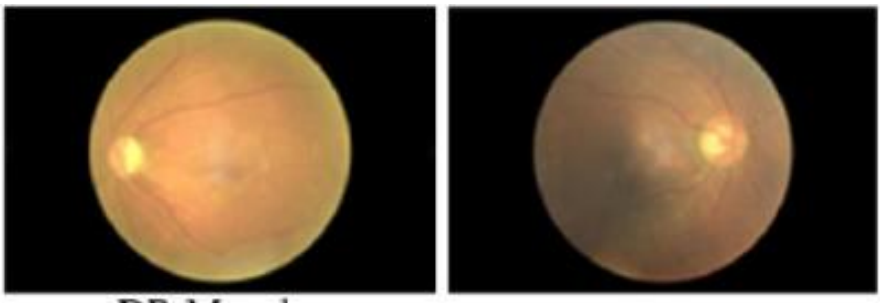


Figure 3.2: Multi-label Ophthalmological Images

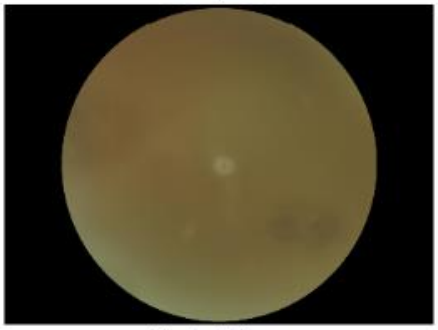
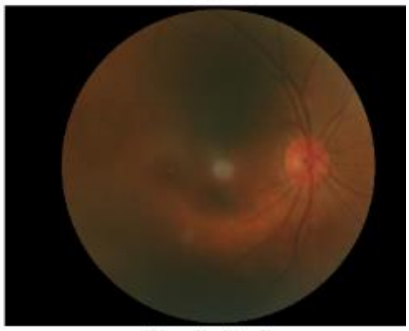
Basic Info.	*Patient Sex*	Female			*Patient Age*	69		
Fundus Images	 0_left.jpg				 0_right.jpg			
Laterality	Left				Right			
Disease Labels	*N*	*D*	*G*	*C*	*A*	*H*	*M*	*O*
	0	0	0	1	0	0	0	0
Diagnostic Keywords	Cataract				Normal fundus			

Figure 3.3: Structured ophthalmic record in ODIR Database

The 5000 patient's data set is divided into testing data set and training data set both on-site and off-site. Roughly, 4000 of these ophthalmological eye data set is in the trainig data and the rest in in the testing data.

No.	Labels	Training Cases	Off-site Testing Cases	On-site Testing Cases
1	N	1,135	161	324
2	D	1,131	162	323
3	G	207	30	58
4	C	211	32	64
5	A	171	25	47
6	H	94	14	30
7	M	177	23	49
8	O	944	134	268

Figure 3.4: Distribution of Dataset: Training And testing

CHAPTER 4: METHODOLOGY

4.1 DATA PRE-PROCESSING

Firstly, the Colab notebook is connected to the drive. The command used for this is *google.colab import drive.* Next the gpu info is extracted using *gpu_info = !nvidia-smi.*

```
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 440.48.02    Driver Version: 418.67       CUDA Version: 10.1     |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|===============================+======================+======================|
|   0  Tesla T4            Off  | 00000000:00:04.0 Off |                    0 |
| N/A   36C    P8     9W /  70W |      0MiB / 15079MiB |      0%      Default |
+-------------------------------+----------------------+----------------------+
```

Next the libraries are installed to unpack 7z files and to perform other functions. Importing pandas as pd and numpy as np and import cv2. Then a folder named as data is created.

Path where the data is saved.

Path= '/content/drive/My drive/thesis/'

Next the 7zfile is extracted in the data folder we created before. This will not only extract the data but move it from drive to colab for efficient processing.

Next the annotations file is read and first 5 rows of annotation file are read and printed.

	ID	Patient Age	Patient Sex	Left-Fundus	Right-Fundus	Left-Diagnostic Keywords	Right-Diagnostic Keywords	N	D	G	C	A	H	M	O
0	0	69	Female	0_left.jpg	0_right.jpg	cataract	normal fundus	0	0	0	1	0	0	0	0
1	1	57	Male	1_left.jpg	1_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0
2	2	42	Male	2_left.jpg	2_right.jpg	laser spot，moderate non proliferative retinopathy	moderate non proliferative retinopathy	0	1	0	0	0	0	0	1
3	3	66	Male	3_left.jpg	3_right.jpg	normal fundus	branch retinal artery occlusion	0	0	0	0	0	0	0	1
4	4	53	Male	4_left.jpg	4_right.jpg	macular epiretinal membrane	mild nonproliferative retinopathy	0	1	0	0	0	0	0	1

The most popular data analysis library is Pandas. It is also the most widely used manipulation library. It's versatile and powerful functions makes it so popular.

Testing data and Training data (given) is also read into the code and is displayed.

```
train_df = pd.read_excel('../multilabel/ODIR/ODIR-5K_Training_Annotations(Updated)_V2.xlsx')
train_df.head()
```

4.2 IMAGE AUGMENTATION

Image augmentation is used to expand the available date set. This technique alters the existing data to create some more data for training. Abominations is a library in Python for fast image augmentations. It's a part of the Pytorch Ecosystem. It's also greatly compatible with Pytorch and TensorFlow.

```
from albumentations import *

from albumentations.pytorch import ToTensorV2

aug= Compose([
            Resize(256, 256),
            CenterCrop(224,224),
            HorizontalFlip(p=0.5),
            VerticalFlip(p=0.5),
            ShiftScaleRotate(0.05,0.05,5),
            RandomBrightnessContrast(p=0.5),
            CLAHE(),
            Cutout(p=0.5),
            Normalize(mean=(0,0,0), std=(1, 1, 1)),
            ToTensorV2(p=1.0),
        ], p=1.0,additional_targets={'image0': 'image'})
```

Image Augmentation starts by resizing all the images to a size of 256,256. After that, the boundary of the images is cropped to a size of 224,224. This is done because the boundary information is less as compared to the information residing in the middle. Given that, no information is lost. Other augmentations given are Horizontal flip, vertical flip and random brightness contrast by a value of 0.5. ShiftscaleRotate by a dimension of 0.05,0.05,5.

CLAHE [Contrast Limited Adaptive Histogram Equalization]

At times, when global equalization generates a lot of noise in the images, this technique may not be the ideal one to use. In such a scenario Adaptive Histogram Equalization is used along with contrast limiting to reduce the chances of noise being generated. CLAHE is a variant of AHE and it handles the over-amplification of contrast very well. This technique works by operating on small regions of the image.

RandomBrightnessContrast(p=0.5),

CLAHE()

CUTOUT

Cutting off arbitrary portions of input images during training is a straightforward regularization6 technique for convolutional neural networks. This method simulates occluded cases and encourages the model to base its conclusions on a larger number of minor features as opposed to a small number of big attributes.

```
from albumentations import *

from albumentations.pytorch import ToTensorV2

aug= Compose([
            Resize(256, 256),
            CenterCrop(224,224),
            HorizontalFlip(p=0.5),
            VerticalFlip(p=0.5),
            ShiftScaleRotate(0.05,0.05,5),
            RandomBrightnessContrast(p=0.5),
            CLAHE(),
            Cutout(p=0.5),
            Normalize(mean=(0,0,0), std=(1, 1, 1)),
            ToTensorV2(p=1.0),
        ], p=1.0,additional_targets={'image0': 'image'})
```

Data reader is used to send the data batch by batch. When one batch completes we get one epoch.

def __getitem: [index, 0] Randomly picks one row but column first.

Python pillow library: We used CV2 bgr to rgb color

If self-transform: Both augmentation on both left and right options sees both eyes the same way.

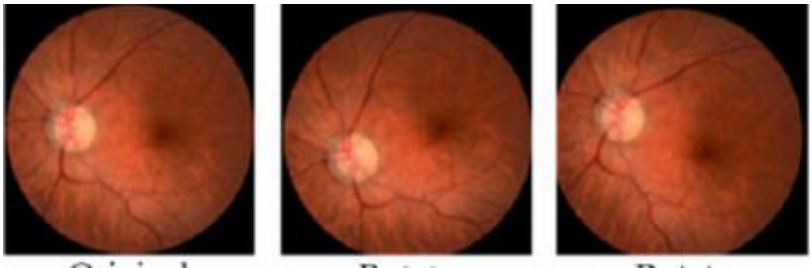


Figure 4.1: Data Pre-processing

```
from PIL import Image
import cv2
from torch.utils.data import DataLoader, Dataset
class DataReader(Dataset):
  def __init__(self,df,path,transform=None):
    super(DataReader,self).__init__()
    self.df=df
    self.path=path
    self.transform=transform
  def __len__(self):
    return len(self.df)

  def __getitem__(self,index):
    image_path=self.df.iloc[index,0]
    image_label=self.df.iloc[index,-8::]
    #read data
    left=cv2.imread(self.path+str(image_path)+'_left.jpg')
    left=cv2.cvtColor(left,cv2.COLOR_BGR2RGB)
    right=cv2.imread(self.path+str(image_path)+'_right.jpg')
    right=cv2.cvtColor(right,cv2.COLOR_BGR2RGB)


    if self.transform:
        image=self.transform(image=left,image0=right)
        left=image['image']
        right=image['image0']

    return left,right,torch.tensor(image_label,dtype=torch.float32)
```

Giving data frame and augmentation in the next block.

Data loader: Pre-defined function. Images are not yet loaded so we iterate.

```
train=DataReader(train_df,train_path,aug)
train_loader = DataLoader(train,shuffle=True,num_workers=0,batch_size=16)
left,right,label=next(iter(train_loader))
print(left.shape,right.shape,label.shape)
```

4.3 PLOTTING:

```
from torch.utils.data.dataloader import DataLoader
import torchvision
import matplotlib.pyplot as plt
plt.figure(figsize=(16,16))
grid_img=torchvision.utils.make_grid(left,8,4)
plt.imshow(grid_img.permute(1, 2, 0))
plt.figure(figsize=(16,16))
grid_img=torchvision.utils.make_grid(right,8,4)
plt.imshow(grid_img.permute(1, 2, 0))
```

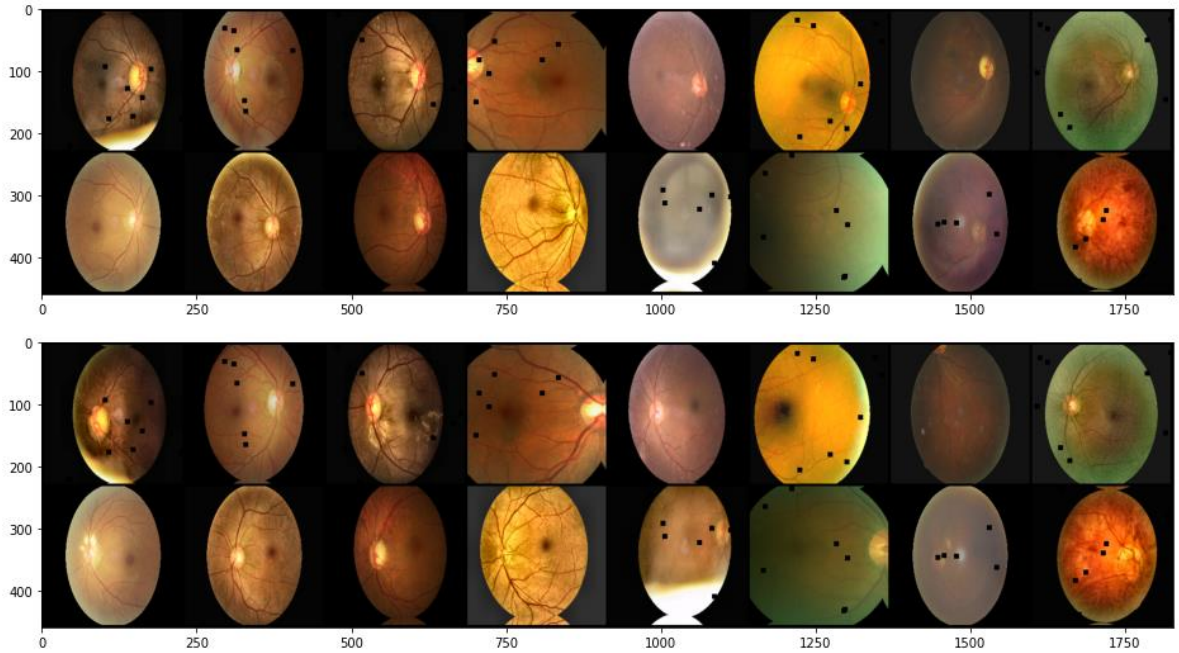


Figure 4.2: Processed Retinal Images

4.4 SPLIT DATA [TRAIN TEST]

```
from sklearn.model_selection import train_test_split
from sklearn import metrics
from torchvision.models import densenet121,resnet50
train_df,val_df=train_test_split(train_df,test_size=0.2,random_state=21)
```

We used scheduler,cosine Annealing starts. If learning rate is 1*10-3, we want it to change since we want to achieve global minimum.

Learning Rate monitor:

```
lr_monitor = LearningRateMonitor(logging_interval='epoch')

early_stop_callback = EarlyStopping(monitor='val_loss',  patience=10, verbose=True, mode='min')
checkpoint_callback = ModelCheckpoint(monitor='val_loss',dirpath='',
                                      filename='savefile' )
model = OurModel()
trainer = Trainer(max_epochs=10,
                  gpus=-1,precision=16,
                  callbacks=[checkpoint_callback,early_stop_callback,lr_monitor],
                  )
```

Precision 16 is used.

4.5 MODEL ARCHITECTURE

4.5.1 CONVOLUTIONAL NEURAL NETWORK

In this approach two Efficient-net models are used simultaneously for left and right eye. It's a convolutional neural network architecture that uniformly and efficiently scales up all the dimensions and models by incorporating a coefficient that's compound.

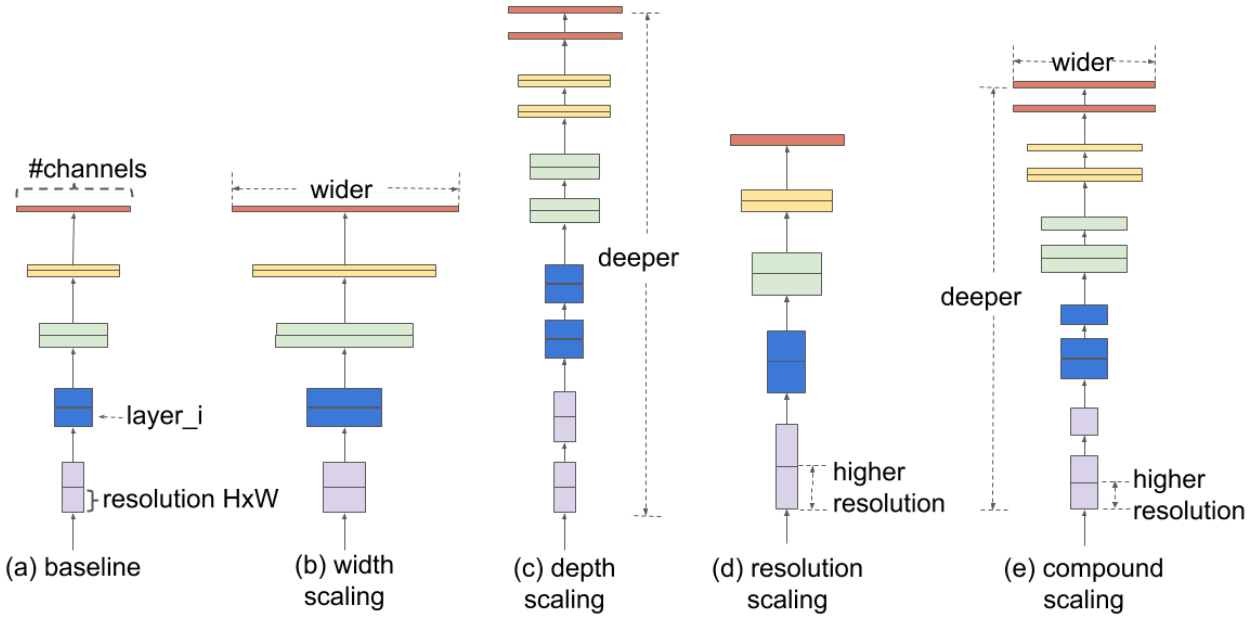


Figure 4.3: Efficient Net Model Mapping

Two models with two input layers is created as we have two images simultaneously [Left eye and right eye]. Both models are combined at latter stage the features are combined to make one model. The output of model has 8 neuron, one for each class. Here Efficient-Net is used as our architecture. The inputs from the left and right eye are fed to two Efficient-net models. The outputs are then concatenated. The classification layers are:

- Global Average Pooling
- Dense Layer
- Dense layer
- Dense layer

The output layer constitutes of 8 neurons. Each one for the classification of eight different eye diseases.

Output Layer	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	N	**D**	**G**	**C**	**A**	**H**	**M**	**O**

Table 4.1: Output Layer Eight Neurons Labels

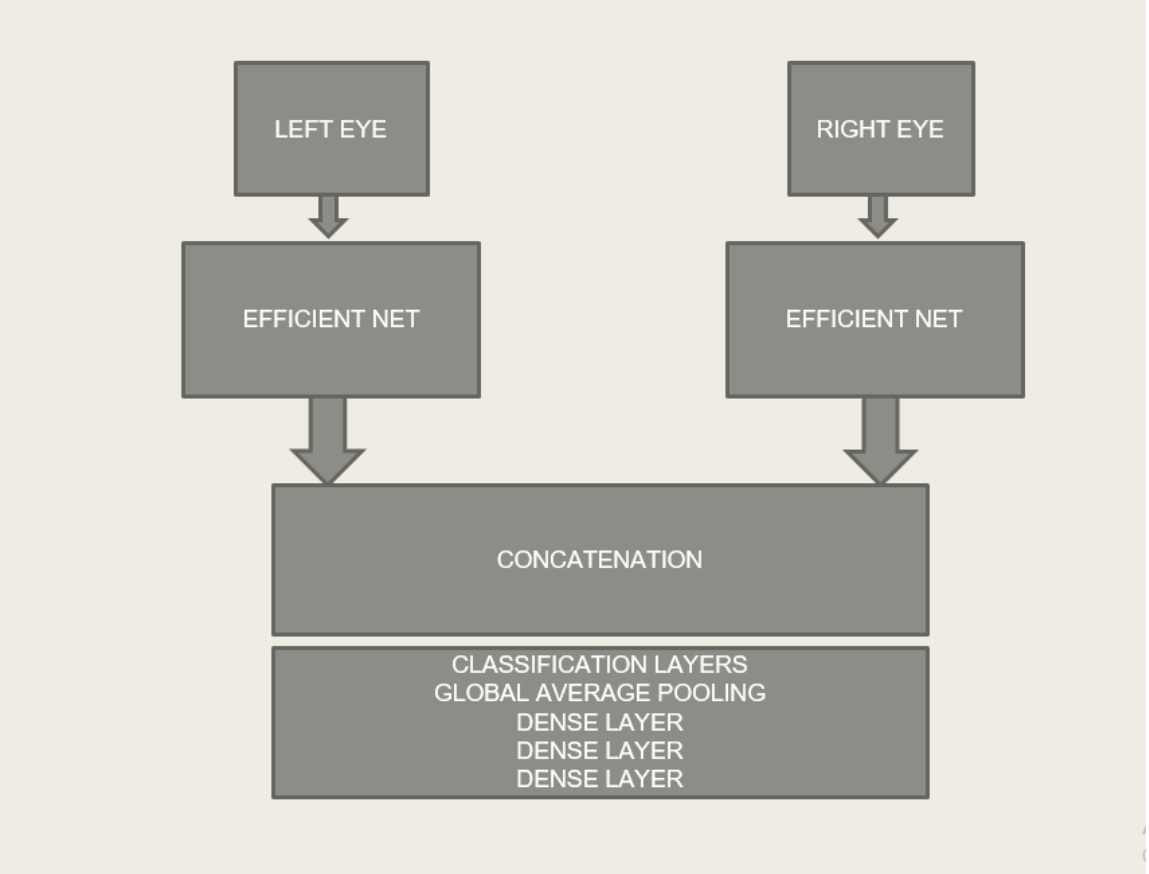


Figure 4.4: Implementation of The CNN Model

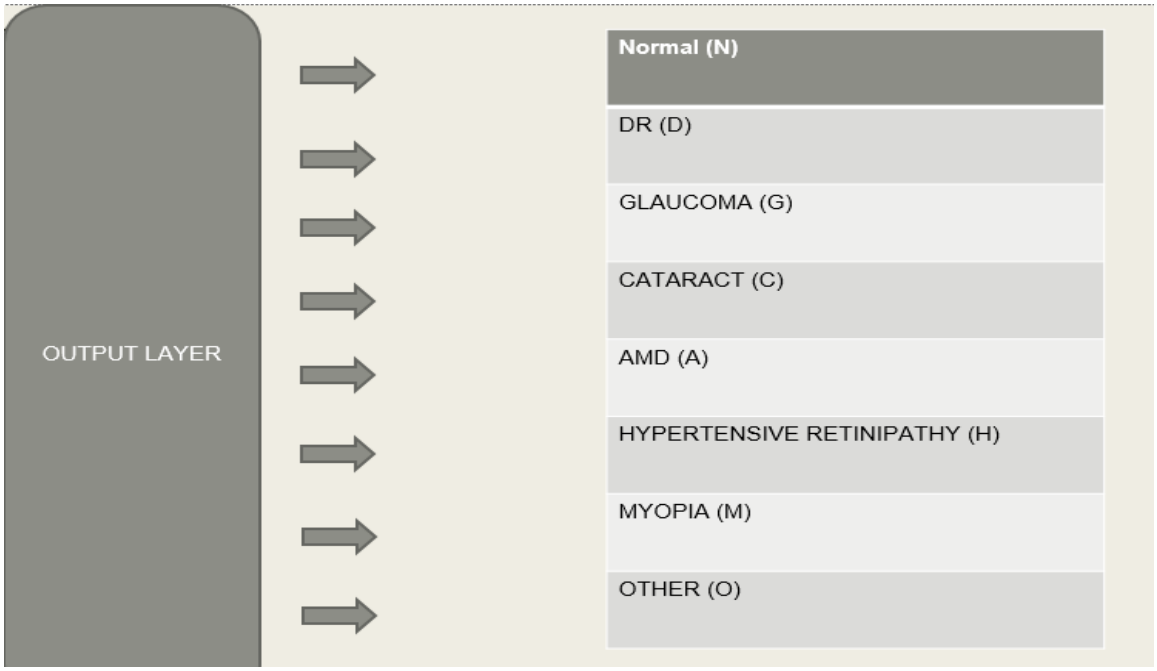


Figure 4.5: Output Layers with 8 neurons

4.5.2 Loss function:

Loss Functions in classification problems are the computed values which tells us the values at the stake wrong predictions in our model. It's a standard to evaluate how well the implemented model in working on our data given the predictions. The values; low or high, determines what type of predictions we've got. A lower value signifies better predictions while a comparatively higher value signifies that the predictions on th given model are off.

We tried two different loss functions

Binary Loss (0-1): This approach is used in binary classification problems. This approach is used where binary classification is required. Like the options of yes or no, 1 & 0 etc

Focal Loss (give more weightage to low occurring): This focuses on addressing the class imbalance problem. It helps in downsizing the loss by emphasizing more on the less-classified examples.

4.5.3 OPTIMIZER:

In order to reduce the loss predicted and for the improvement of our accuracy, optimizer is used. It does so by modifying and improving the weights used and improving the overall learning rate.

4.5.4 ADAM OPTIMIZER

A stochastic gradient used for optimization which uses the approach of adaptive estimation, with momentum. It updates the weights of the given network based on the given training data t hand.

FEATURES OF ADAM OPTIMIZATION

- Doesn't requires memory
- Computing is effective
- Suitable for sparse data and that which has noise
- Easy to customize
- Incorporating the properties of best algorithms i.e. AdaGrad and RMSProp
- Suitable for problems relating to size of parameter

Figure 4.6: Adam Optimizer Curve

4.5.5 IMPLEMENTATION OF EFFICIENT-NET

```
def parallel_model():
    input_layer = Input(shape=(2, img_shape,img_shape,3))
    left_input, right_input = Lambda(lambda x: tf.split(x, 2, axis=1))(input_layer)

    left_input = Reshape([img_shape, img_shape, 3])(left_input)
    right_input = Reshape([img_shape, img_shape, 3])(right_input)



    left_model = efn.EfficientNetB7(input_shape =  (img_shape,img_shape,3), include_top = False, weights = None,input_tensor=left_input)
    x1 = left_model.output
    x1 = GlobalMaxPooling2D()(x1)
    out1 = Dense(5, activation='softmax')(x1)
    left_model=tf.keras.Model(inputs=left_model.input, outputs=out1)
    left_model.load_weights('../input/aptos-model/aptos_model.hdf5')
    left_model=tf.keras.Model(left_model.input, left_model.layers[-3].output)



    right_model = efn.EfficientNetB7(input_shape =  (img_shape,img_shape,3), include_top = False, weights = None,input_tensor=right_input)
    xr = right_model.output
    xr = GlobalMaxPooling2D()(xr)
    outr = Dense(5, activation='softmax')(xr)
    right_model=tf.keras.Model(inputs=right_model.input, outputs=outr)
    right_model.load_weights('../input/aptos-model/aptos_model.hdf5')
    right_model=tf.keras.Model(right_model.input, right_model.layers[-3].output)

    for layer in right_model._layers:
        layer._name = layer._name + '_right'
    for layer in left_model._layers:
        layer._name = layer._name + '_left'

    left_model._name="left_eff"
    right_model._name="right_eff"


    con = concatenate([left_model.output, right_model.output])
    GAP= GlobalAveragePooling2D()(con)
    fc1=Dense(256, activation = 'relu')(GAP)
    fc2=Dense(128, activation = 'relu')(fc1)
    fc3=Dense(64, activation = 'relu')(fc2)
    out = Dense(8, activation = 'sigmoid')(fc3)

    model = tf.keras.Model(inputs=input_layer, outputs=out)
    return model
```

```
from tensorflow.keras.optimizers import Adam
from focal_loss import BinaryFocalLoss

def get_model():
    opt = Adam(lr=0.0003, decay=1e-3)

    with strategy.scope():
        model = parallel_model()
        #model.summary()

    model.compile(optimizer=opt, loss=BinaryFocalLoss(gamma=2),metrics=['accuracy'])
    return model
```

The above model depicts and describes the implementation of two efficient net models which are later concatenated to get results for the left and right eye classification.

4.5.6 Score improvement

For improving the score, first the model was pre-trained on Diabetic Retinopathy Detection from kaggle. We are provided with high-resolution images, for both the left and right eye, given the labels.

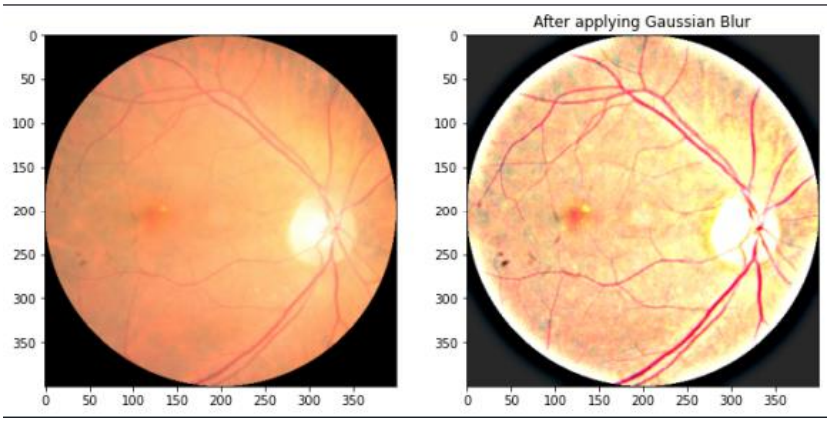


Figure 4.7: Gaussian Blur on DR Dataset

CHAPTER 5: OTHER APPROACHES TO ODIR DATA

ResNet50: Image shape (224, 224) was used and batch size was given 50.

	precision	recall	f1-score	support
0	0.29	0.15	0.19	350
1	0.30	1.00	0.46	299
2	1.00	0.02	0.03	62
3	0.09	0.34	0.15	62
4	0.00	0.00	0.00	41
5	0.02	0.55	0.05	22
6	0.00	0.00	0.00	53
7	0.27	1.00	0.42	266
micro avg	0.22	0.56	0.32	1155
macro avg	0.25	0.38	0.16	1155
weighted avg	0.28	0.56	0.29	1155
samples avg	0.24	0.53	0.31	1155

Figure 5.1: Resnet50 Model Behavior

InceptionResNetV2:

20 epoch and a batch size of 15 was given.

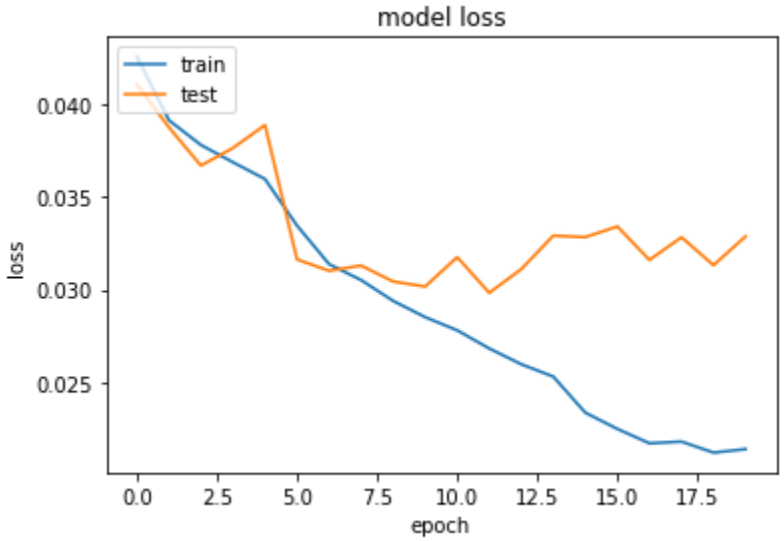


Figure 5.2: Inception ResNet Model Loss

VGG16

Activation sigmoid and a batch size of 50 was given.

	precision	recall	f1-score	support
N	0.30	1.00	0.47	303
D	0.35	1.00	0.52	349
G	0.07	1.00	0.13	70
C	0.06	1.00	0.11	58
A	0.05	1.00	0.10	50
H	0.03	1.00	0.06	29
M	0.05	1.00	0.09	49
O	0.28	1.00	0.44	281
micro avg	0.15	1.00	0.26	1189
macro avg	0.15	1.00	0.24	1189
weighted avg	0.26	1.00	0.40	1189
samples avg	0.15	1.00	0.26	1189

Figure 5.3: VGG16 Model Behavior

CHAPTER 6: RESULTS

6.1 Evaluation Metrics

Evaluation metrics are metrics defined to determine the success of a trained model. Themetrics we used in this study are Accuracy, the Area under Curve (AUC) and F1-score. Accuracy (ACC) is characterized as the ratio of correctly predicted values and total valuespredicted by the model. True positives (TP) are the positives predicted correctly whereaspositives predicted as negatives are termed as false negatives (FN). True negatives (TN)are the negatives predicted correctly and the negatives predicted as positives are calledfalse positives (FP). Accuracy (ACC) is given by the equation (1):

ACC= TP +TN / TP + TN+ FP +FN…… ……(1)

F1-score is defined as the harmonic mean of precision and recall and is defined below:

F1-score = 2 x (precision x recall / precision + recall)……..(2)

Where precision is given in equation (3):

Precision = TP/ TP+FP……(3)

And recall is given in equation (4):

Recall = TP/ TP +TN……(4)

Finally, AUC is a metrics that defines how well the model can distinguish between differentclasses. The classification metrics are summarized in Table 6.1.

Metrics	Definition	Formula
Accuracy (ACC)	the ratio of correctly predicted values and total values predicted by the model	$\frac{TP + TN}{TP + TN + FP + FN}$
AUC	It determines the model'sability to differentiate among classes	$\int_{x=0}^{1} TPR(FPR{-}1(x))dx$
F1- Score	The harmonic mean of precision and recall	$\frac{2TP}{2TP + FP + FN}$

Precision (Specificity)	Ratio of correctlyidentified TN	$\frac{TP}{TP + FP}$
Recall	Ratio of correctlyidentified TP	$\frac{TP}{TP + TN}$

Table 6.1: Classification Metrics

Following table demonstrates the results obtained by implementing various CNN models on our data set. The parameters as AUC and F1 were computed and a final score was calculated as their mean.

MODEL	**AUC**	**F1**	**FINAL SCORE**
Vgg 16	0.64	0.81	0.725
Inception V3	0.65	0.85	0.75
ResNet 50	0.63	0.81	0.72
DenseNet	0.63	0.83	0.73
B7 Binary Loss	0.90	0.88	0.89
B7 Focal Loss	0.95	0.97	0.96
B5 Focal Loss	0.88	0.84	0.86
B5 Binary Loss	0.83	0.81	0.82

Table 6.2: Comparative Evaluation of Implementation Of Different CNN Models

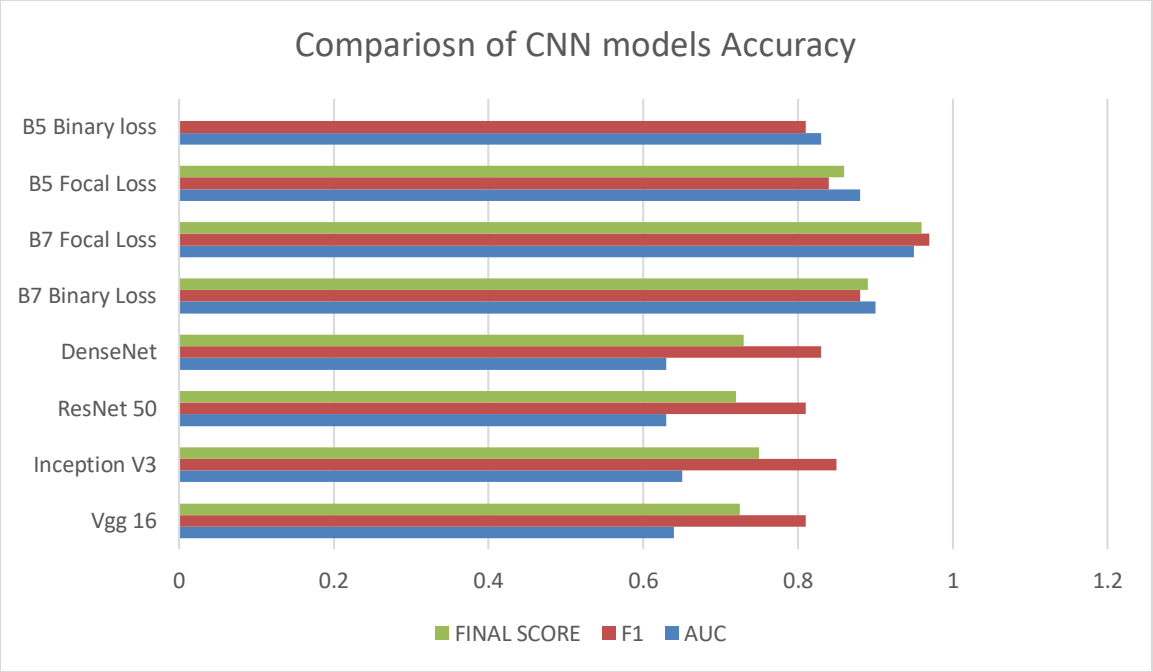


Figure 6.1: Comparison of CNN Models Accuracy

Our approach to the data set was using Efficient Net B5 and B7 in conjunction with binary loss and focal loss respectively. From the results below, it is evident that B7 out performed on the dataset given any loss approach. Further, the results and accuracy ratio was better while focal loss was incorporated, as compared to Binary Loss entropy.

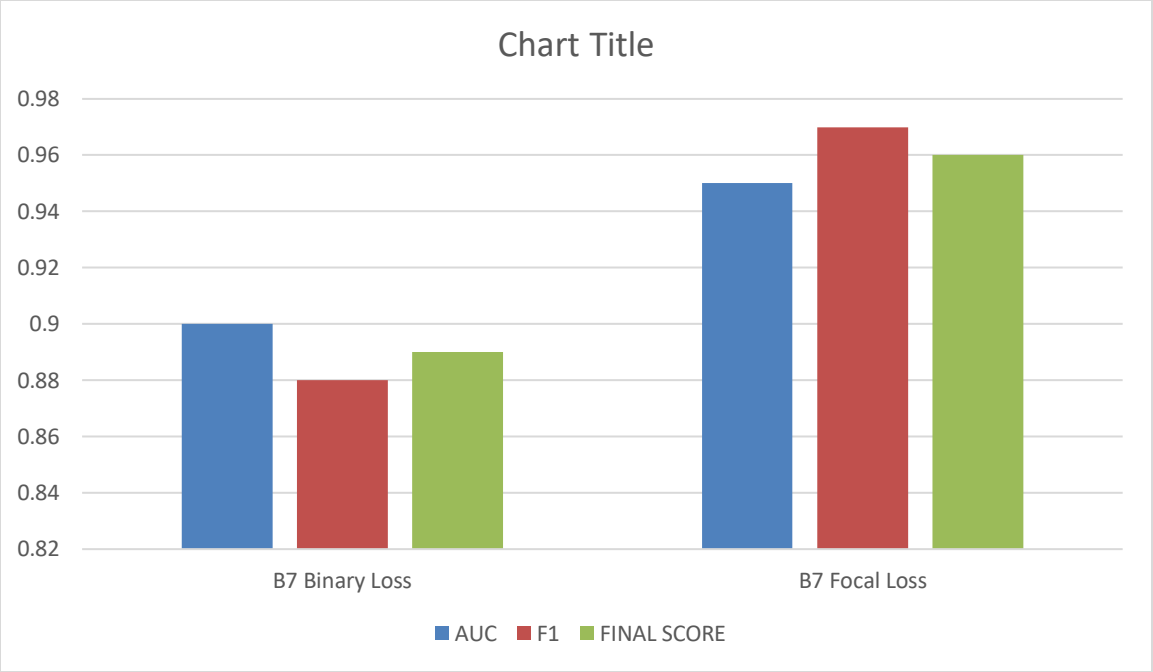


Figure 6.2: B7 Approach using Binary Loss and Focal Loss

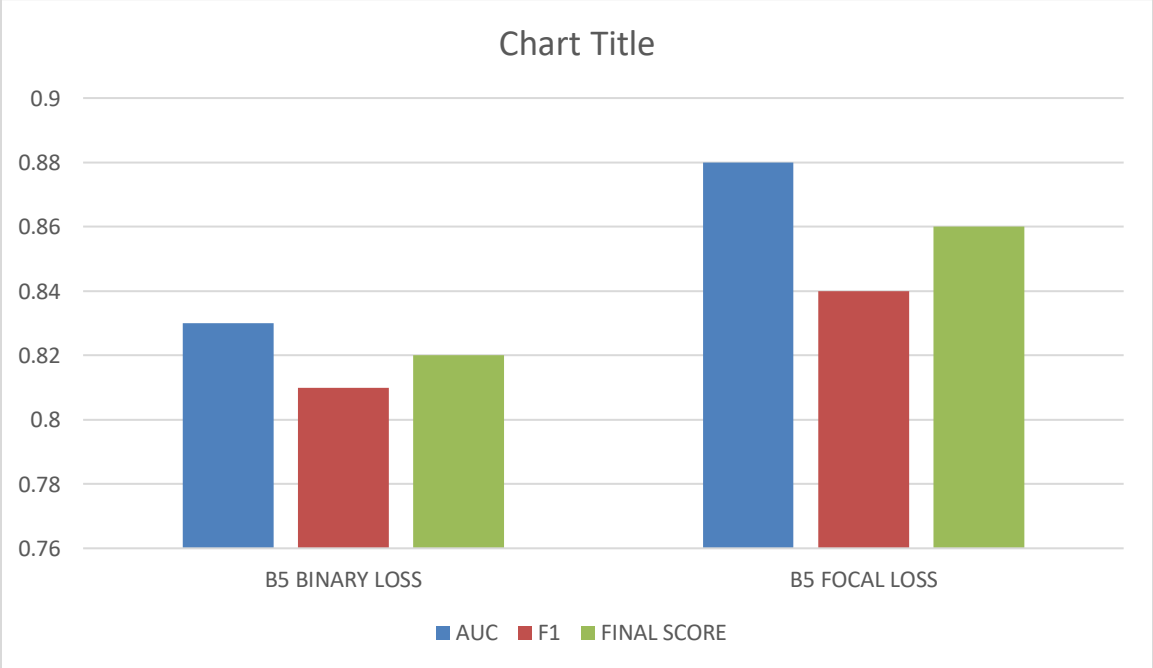


Figure 6.3: B5 Approach Using Binary Loss and Focal Loss

CHAPTER 7: CONCLUSION AND LIMITATIONS

Two parallel Efficient-net model architectures were used along with Focal loss and Adam optimizer to achieve the required accuracy. Further, the data set provided by ODIR is greatly unbalanced. The category having the labels as 'Other (O)', contains a stack of diseases which are unknown. Also, the dataset for some of the eye diseases is very less as compared to the other given categories. All of this combined, makes it challenging to acquire the required accuracy and precision.

Another limitation to this approach comes from the nature of deep learning. The black box to this approach is that we never know that which specific features are being learned by our architecture.

CHAPTER 8: FUTURE ASPECTS

For future work, more diverse data from various hospitals, taken with different resolution cameras, is a requirement. This will help in gaining a dataset that is diverse in its attributes and features. This will not only help in generalization, but also in improving the overall accuracy of the results.

Further, apart from the image data and features in itself, other provided data such as the patient's age, gender, sex etc. can also be considered for detecting the probability of the retinal diseases. The deep learning model's acceptance is the state of the art requirement and it need to be improved, along with understanding its various aspects and diversity.

REFERENCES

[1] Jackman WT, Webster JD. On photographing the retina of the living human eye. *Philadel Photogr* 1886**;**23:340–341

[2] Donaldson DD. A new camera for stereoscopic fundus photography. *Trans Am Ophthalmol Soc* 1964**;**62:429–458

[3] Behrendt T, Wilson LA. Spectral reflectance photography of the retina. *Am J Ophthalmol* 1965**;**59:1079–1088

[4] LaRocca F, Nankivil D, Farsiu S, Izatt JA. Handheld simultaneous scanning laser ophthalmoscopy and optical coherence tomography system. *Biomed Opt Express* 2013**;**4:2307–2321

[5] https://ieeexplore.ieee.org/abstract/document/8397469

[6] Son J, Shin JY, Kim HD, Jung K-H, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. Ophthalmology. 2019;127(1):85–94.

[7] Al-Bander B, Al-Nuaimy W, Williams BM, Zheng Y. Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc. Biomed Signal Proc Control. 2018;40:91–101.

[8] Zhu C, Zou B, Zhao R, Cui J, Duan X, Chen Z, Liang Y. Retinal vessel segmentation in colour fundus images using extreme learning machine. Comput Med Imag Graph. 2017;55:68–77.

[9] Tobin KW, Chaum E, Govindasamy VP, Karnowski TP. Detection of anatomic structures in human retinal imagery. IEEE Trans Med Imag. 2007;26(12):1729–39.

[10] Singh A, Dutta MK, ParthaSarathi M, Uher V, Burget R. Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image. Comput Methods Prog Biomed. 2016;124:108–20.

[11] Mohamed NA, Zulkifley MA, Zaki WMDW, Hussain A. An automated glaucoma screening system using cup-to-disc ratio via simple linear iterative clustering superpixel approach. Biomed Signal Process Control. 2019;53:101454.

[12] Fumero F, Alayón S, Sanchez JL, Sigut J, Gonzalez-Hernandez M. Rim-one: An open retinal image database for optic nerve evaluation. In: 2011 24th International Symposium on Computer-based Medical Systems (CBMS), pp. 1–6 (2011). IEEE

[13] Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology. 2018;125(8):1199–206.

[14] Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-First AAAI conference on artificial intelligence (2017)

[15] Zhang Z. Improved adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th International symposium on quality of service (IWQoS), pp. 1–2 (2018). IEEE

[16] dos Santos Ferreira MV, de Carvalho Filho AO, de Sousa AD, Silva AC, Gattass M. Convolutional neural network and texture descriptor-based automatic detection and diagnosis of glaucoma. Expert Syst Appl. 2018;110:250–63.

[17] Son J, Shin JY, Kim HD, Jung K-H, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. Ophthalmology. 2019;127(1):85–94.