

*Detecting Anemia from retinal images using  
Deep Learning*

---

*Akhil Agrawal*



# Detecting anemia from retinal images using deep learning

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF

Master of Technology  
in  
Computer Science

by

**Akhil Agrawal**

[ Roll No: CS-1919 ]

under the guidance of

**Dr. Sushmita Mitra.**

Professor

Machine Intelligence Unit



Indian Statistical Institute  
Kolkata-700108, India

July 2021

*To my family and my guide*

# CERTIFICATE

This is to certify that the dissertation entitled “**Detecting Anemia from retinal images using deep learning**” submitted by **Akhil Agrawal** to Indian Statistical Institute, Kolkata, in partial fulfillment for the award of the degree of **Master of Technology in Computer Science** is a bonafide record of work carried out by him under my supervision and guidance. The dissertation has fulfilled all the requirements as per the regulations of this institute and, in my opinion, has reached the standard needed for submission.



---

**Dr. Sushmita Mitra**

Professor,  
Machine Intelligence Unit,  
Indian Statistical Institute,  
Kolkata-700108, India.

# Acknowledgments

I would like to show my highest gratitude to my advisor, *Prof. Dr. Sushmita Mitra*, Machine Intelligence Unit, Indian Statistical Institute, Kolkata, for his guidance and continuous support and encouragement. He has literally taught me how to do good research, and motivated me with great insights and innovative ideas.

I would also like to thank *Dr. Rajiv Raman* Senior Consultant, Shri Bhagwan Mahavir Vitreoretinal services, Sankara Nethralaya, Chennai and *Subhashish Bannerjee*, Senior Research Fellow, Machine Intelligence Unit, Indian Statistical Institute, Kolkata, for his valuable suggestions and discussions.

My deepest thanks to all the teachers of Indian Statistical Institute, for their valuable suggestions and discussions which added an important dimension to my research work.

Finally, I am very much thankful to my parents and family for their everlasting supports.

Last but not the least, I would like to thank all of my friends for their help and support. I thank all those, whom I have missed out from the above list.

A handwritten signature in blue ink that reads "Akhil". The signature is written in a cursive style with a horizontal line underneath the name.

**Akhil Agrawal**  
Indian Statistical Institute  
Kolkata - 700108 , India.

# Abstract

Anemia is the condition which formed when there is not enough healthy red blood cells present in body. This ultimately result in the lack of reduced oxygen flow in the body. This causes the fatigue, dizziness, shortness of breath. The main problem with this is it's very common and mostly undetected specially in the Indian population. In general anemia can be detected by the blood test and checking the hemoglobin level with standard WHO standards of men and women to decide anemia or non anemia which is an invasive method. Through some studies, it is been observed that their is an non-invasive way to detect anemia as well, which is by using medical fundus images. It can be used to detect this disease via the Deep learning network. In one of that research [4] medical fundus images with the centered macula region is being used with the huge dataset but in the end their were some evidences in paper that model try to focus on the optic disc region much more and based on this prediction do get affected. We have here the dataset of the medical fundus images provided by the eye hospital and we are picking out images each with the optic disc present, in which we are going to be using for the classification of the anemic and non anemic disease based on the deep learning model. In the end of the experiment we have seen that with AUC 0.6649 in Densenet121, 0.6223 in Resnet50 and 0.6718 in MobilnetV2 we are able to identify the disease.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	What is Anemia? . . . . .	4
1.2	Fundus images . . . . .	4
1.3	Deep Learning . . . . .	5
1.4	Problem Statement . . . . .	6
<b>2</b>	<b>Related Work</b>	<b>8</b>
<b>3</b>	<b>Data Prepration &amp; Visualisation</b>	<b>10</b>
3.1	Introduction . . . . .	10
3.2	Data Prepration . . . . .	10
3.3	Data Description . . . . .	11
3.4	Data Visualisation: . . . . .	12
<b>4</b>	<b>Data Preprocessing</b>	<b>16</b>
4.1	Rescaling . . . . .	16
4.2	Data Augmentation . . . . .	16
4.3	Removed unwanted images from data Augmentation . . . . .	18
4.4	Normalize data . . . . .	18
<b>5</b>	<b>Binary Classification</b>	<b>19</b>
5.1	Introduction . . . . .	19
5.2	Loss function for Binary Classification . . . . .	19
5.3	Optimizing Function . . . . .	19
5.4	Activation Function . . . . .	20
5.5	Metrics for this Binary Classification . . . . .	20



---

<b>6</b>	<b>Anemia Detection using Resnet50</b>	<b>21</b>
6.1	Motivation . . . . .	21
6.2	Preliminaries . . . . .	21
6.2.1	Architecture . . . . .	21
6.2.2	Implementations . . . . .	22
6.2.3	Result . . . . .	23
<b>7</b>	<b>Anemia Detection using Mobile net v2</b>	<b>24</b>
7.1	Motivation . . . . .	24
7.2	Preliminaries . . . . .	24
7.2.1	Architecture . . . . .	24
7.2.2	Implementations . . . . .	25
7.2.3	Result . . . . .	25
<b>8</b>	<b>Anemia Detection using Densenet121</b>	<b>27</b>
8.1	Motivation . . . . .	27
8.2	Preliminaries . . . . .	27
8.2.1	Architecture . . . . .	27
8.2.2	Implementations . . . . .	29
8.2.3	Result . . . . .	29
<b>9</b>	<b>Future Work and Conclusion</b>	<b>30</b>
9.1	Conclusion . . . . .	30
9.2	Future Work . . . . .	30

# Chapter 1

## Introduction

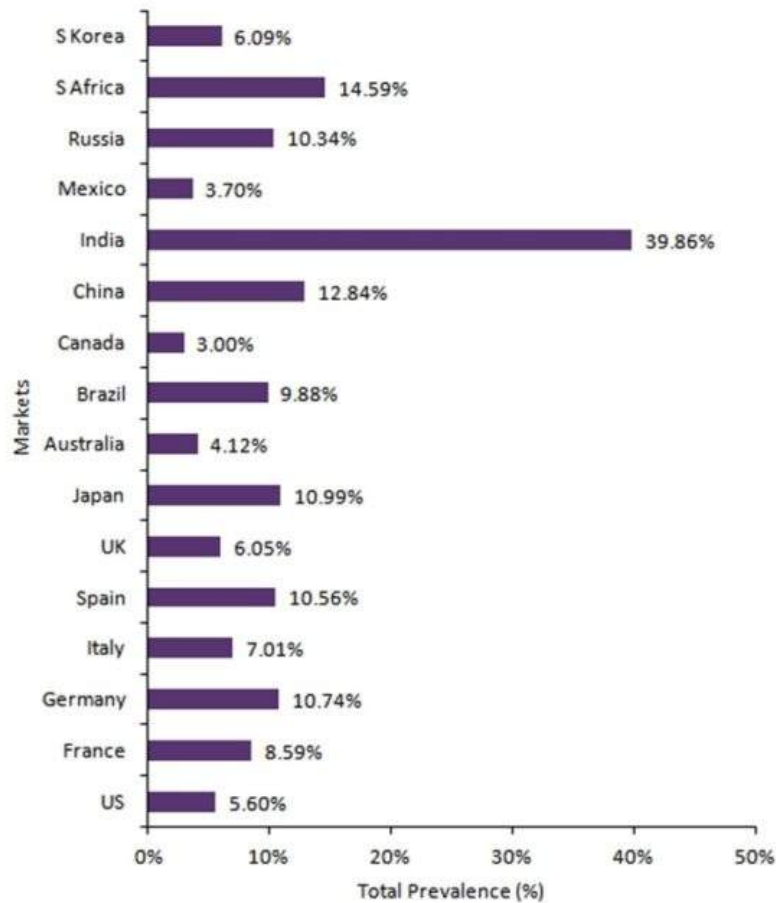
### 1.1 What is Anemia?

Anemia is defined as condition in which the blood doesn't have enough healthy red blood cells.[18] It is one of the most common nutritional deficiency diseases observed globally affecting both developing and developed countries. This result in the reduction in oxygen flow in the body. Anemia can be temporary or long term, and it can range from mild to severe. This varied range of severity is because of the various cases person suffers from it. It may be passed on from genes, iron deficiency, due to blood loss, due to other diseases etc. This causes the varies things like shortness of breath, dizziness, headache, tiredness and weakness etc. In a routine blood test, anemia is reported as a low hemoglobin or hematocrit. According to the National Family Health Survey in 2019 68.4 per cent children, 21.7 per cent men and 53.2 per cent women surveyed suffered from anaemia. In order to give the idea about the prevalence Global Data epidemiologists analyzed 16 major pharmaceutical markets(US, France, Germany, Italy, Spain, UK, Japan, Australia, Brazil, Canada, China, India, Mexico, Russia, South Africa, and South Korea). Where Total prevalence is defined as including both diagnosed and undiagnosed cases. Here India is highest with the total of 40 percent as shown in the figure.[10]

### 1.2 Fundus images

The fundus is the back of the eye and includes the retina, optic nerve, and retinal blood vessels. In the fundus images we try to capture all of this via the special cameras through a dilated pupil, providing a color picture of the back of the eye. Procedure to do so is easy, it hardly takes one-two minutes and it is completely non invasive.[2] After the imaging is done it's able to provide a picture of the retina, the retinal vasculature (blood vessels), and the optic nerve head, where retinal blood vessels

Figure 1: Total Prevalence of Anemia in 16MM



Source: GlobalData

Figure 1.1: Total prevalence of anemia in 16 major countries[18]

enter the eye. It is 2 dimensional(2-D) image and for a single person their could be multiple images based on which eye and position where the camera is focusing.

## 1.3 Deep Learning

Machine Learning and artificial intelligence is the new cornerstones of new revolution in computing. This things help in identifying the patterns, make relevant predictions about the future. In here, Deep learning can be considered as the subset of the machine learning.[11] It is a field that is based on learning and improving on its own by examining computer algorithms. While machine learning uses simpler concepts, deep learning works with artificial neural networks, which are designed to imitate how

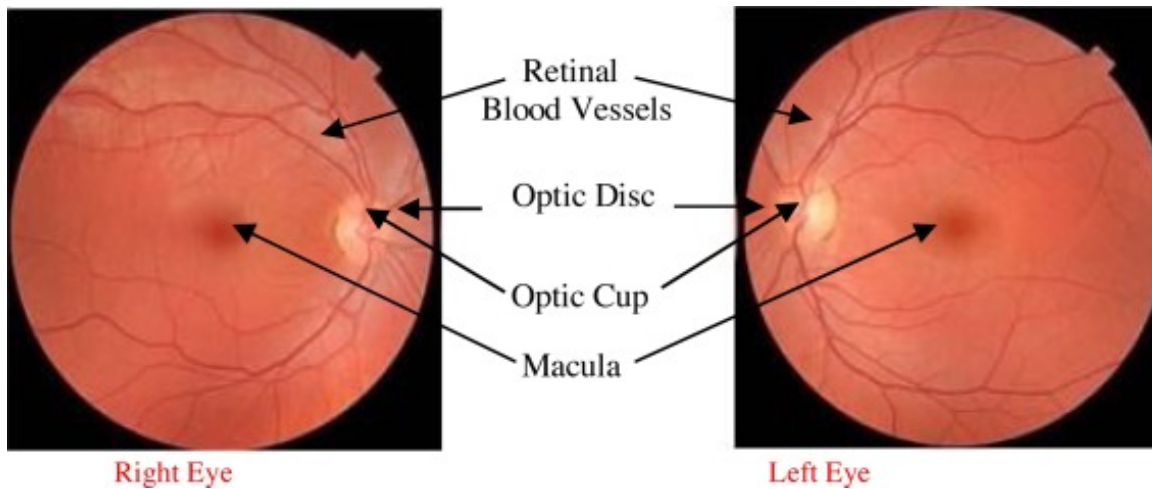


Figure 1.2: Sample fundus image

humans think and learn. Amazing part about this is deep learning has aided image classification, language translation, speech recognition. It can be used to solve any pattern recognition problem and without human intervention. Deep Neural Networks (DNNs) are such types of networks where each layer can perform complex operations such as representation and abstraction that make sense of images, sound, and text. Neural networks are layers of nodes, much like the human brain is made up of neurons. Nodes within individual layers are connected to adjacent layers. The network is said to be deeper based on the number of layers it has. Deep learning systems require large amounts of data to return accurate results; accordingly, information is fed as huge data sets.

Here, in our project we are trying to be focusing on the image data and Hence we are going to be using the convolution neural network , commonly called as CNN.In CNN's their is convolution layers and Each convolutional layer contains a series of filters known as convolutional kernels with the integer values.Each pixel is multiplied by the corresponding value in the kernel, then the result is summed up for a single value for simplicity representing a grid cell, like a pixel, in the output. Their the other important layer is pooing which is basically summarises the features present in a region of the feature map generated by a convolution layer. Dense layer is just a regular deeply connected neural network layer and finally to produce the output there is output layer.

## 1.4 Problem Statement

In this problem we have to build an classification model that can detect the anemic and non-anemic patients from the fundus retinal images. We are using image dataset

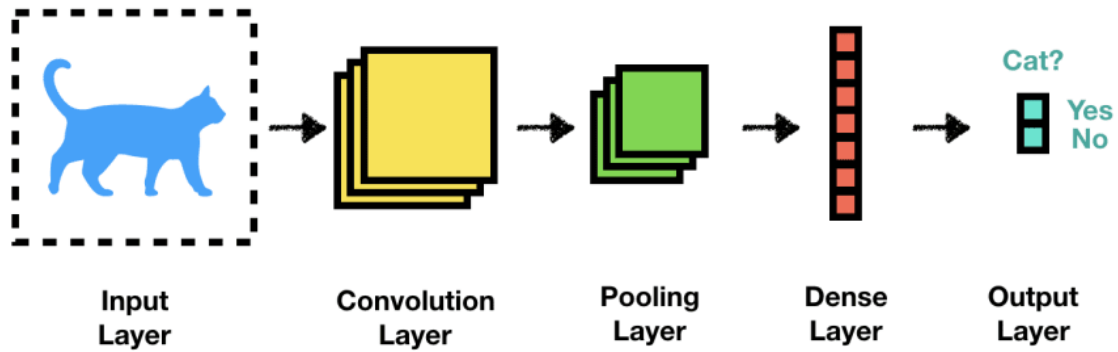


Figure 1.3: Convolution Neural Network

of retinal fundus images provided by the Sankara Nethralaya. This is a binary classification problem since one image can belong either to an anemic or a non-anemic patient. Here, from the output we try to see what features of the images are important and can we say that the optic disc of the retinal fundus image is an important feature of the image to consider in this model or not?

# Chapter 2

## Related Work

In 2020 There is an research paper by Akinori Mitani and its team [4] named Detection of anaemia from retinal fundus images via deep learning which was published by nature biomedical engineering. In this research paper what they try to do is to use the retinal fundus images macula centered from the large dataset provided by the UK bio-bank where the number of participants is 57,163 and total number of images among those are 114,326 images because each patient with two images of left and right eye. Here they try to build the model from those images to predict whether the patient have the anemia or not . As their are WHO standard range of hemoglobin to see a person has anemia or not based on it's gender, here they used 3 of those category WHO range standard anemia, moderate anemia for detecting serious cases and approximate anemia for the non gender based common setup . Based on this they try to three types of model fundus-image based model , metadata only model(Here age, sex ) for baseline purpose and combined model of both metadata and fundus image model.They used AUC to the measure InceptionV4 for the modelling.After this they try to see how explainable the results are actually .

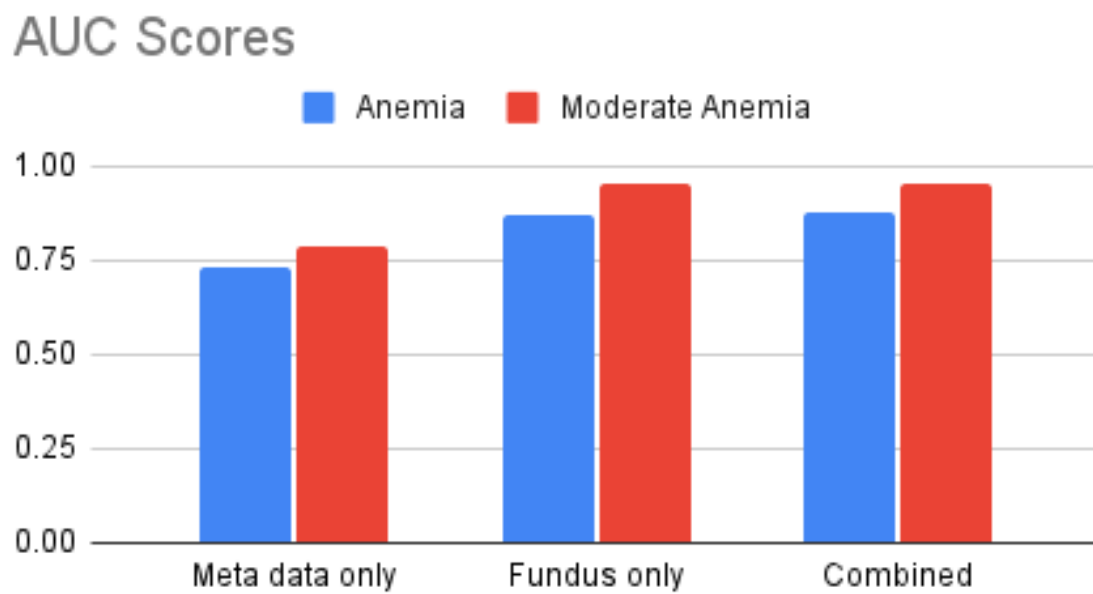


Figure 2.1: Convolution Neural Network

# Chapter 3

## Data Prepration & Visualisation

### 3.1 Introduction

The dataset that we have used is provided by the Sankara Nethralaya, Chennai which is Eye hospital and research center. Here, we got the raw dataset of 400 patients fundus images with each patients have multiple images of both left and right eye and metadata which could be used in this as it originally has 669 columns of different values of all the 400 patients and from which we select the meaningful one which is mentioned in matani's paper as well which is sex,age, blood pressure both diastolic and systolic with the hemoglobin which we will try to predict and based on only this we prepare data for anemic and non anemic classes.

### 3.2 Data Prepration

As the data we have is raw which we got directly from the hospital we have to remove the unnecessary data which we are not going to be used and curate in the way we want.

1. Here, first we merged the data of different folders for each patient to a combined one but keeping the image ID intact. Then in the original data every patient have two image of eye taken from normal camera to capture how the eye looks like externally, we will remove that.
2. We separate out the images with no optic disc in the image as per our problem statement we are looking in that region and trying to see if there is any importance for that region.
3. Based on CSV file we have we make two folders out of it. anemic and non anemic which will be based on the hemoglobin contains. As WHO specified the



hemoglobin contains for Anemic and non-anemic patients for male 13g/dl and female 12 g/dl .Here we see image id of the image and from CSV it corresponding hemoglobin and put in whether it is anemic or non anemic.

4. Now we split the dataset in test , training and validation data. Here we are going to be using 80% training , 10% validation and 10% testing.

### 3.3 Data Description

The information regarding the data in image will look like [Image-Id][patient-image-no.][male/female][hemoglobin] as follows (as shown in Fig3.1)

1. Id - An image Id. Each Id corresponds to a unique patient, and will contain an underscore .
2. patient-image-no. As a patient can have the multiple images this is to indicate the unique image for a patient.
3. male/female: This is to show weather a patient is male or the female .So, if later case if we want to identify the anemic patient we can directly use it where 1 is male and 2 is female.
4. hemoglobin: this is an actual result of hemoglobin in g dl which is basically we are mainly focusing on and decide whether it is anemic or not.
5. Label - This is class which we create with different sub-directories for whether anemic or non anemic .

There are total 6263 images .You can see the all images having anemic or non anemic labels.This data do have the imbalanced classes which we can see in Figure 3.2 and one example image so see how it looks like with the original name in the description.

There are three folders of images training, validation and testing folder with each have anemic and non anemic in them. There are total 720 anemic and 5370 non anemic training images,72 anemic and 546 non anemic validating images and 72 anemic and 547 non anemic testing images. Each image is single 2D slice having resolution among 2240\*1488,2588\* 1958,1504\*1000 pixel.This is due to the different cameras used for the images.we will convert it into the same resolution n the data preprocessing part.We can see that this data is highly imbalanced.

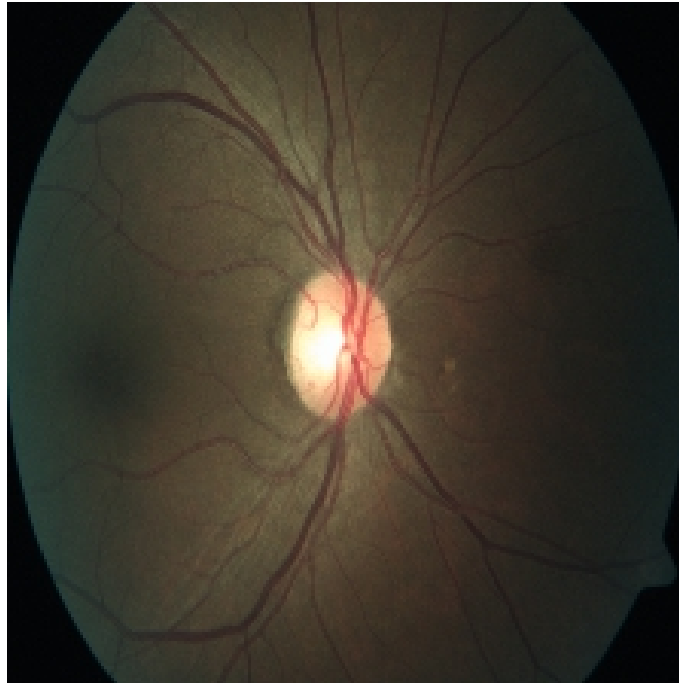


Figure 3.1: Examples of sample image named 142\_00\_002\_15.2.jpg

### 3.4 Data Visualisation:

let us have look at different kind of Anemic and Non Anemic images have some idea from those images. There is no practical generalization which can be seen from the fundus images and say this is anemic and non anemic but through deep learning we will try to do that later. Here are sample of image data of both the classes.

Apart from it if we see the distribution of data among the Age we can look for Age histogram as in figure , similarly sex histogram for the distribution of data among sex is in Figure.

Apart from this we have Hemoglobin which is the basically determine whether patient has anemia or not and since it mainly affects based on sex we have plot it separately for men and women. Here remember that threshold as per the WHO for anemic patient is 13 g/dl for men and 12 g/dl for women.

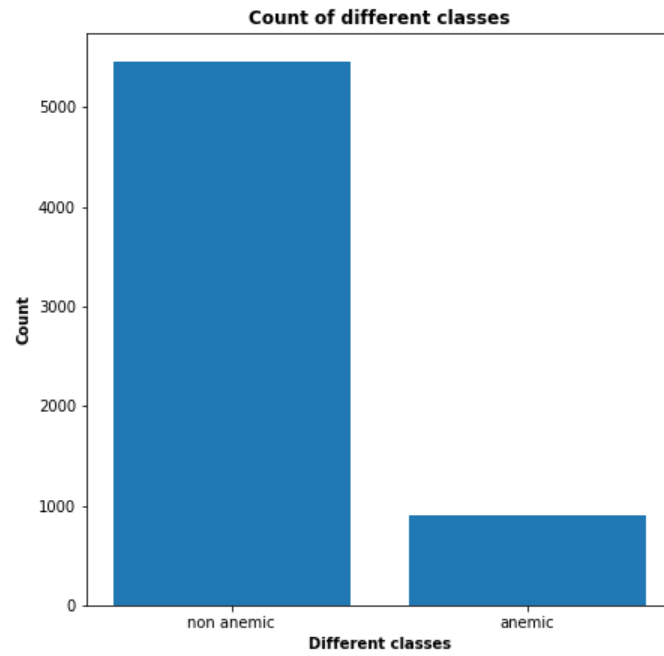


Figure 3.2: Count plot of label images

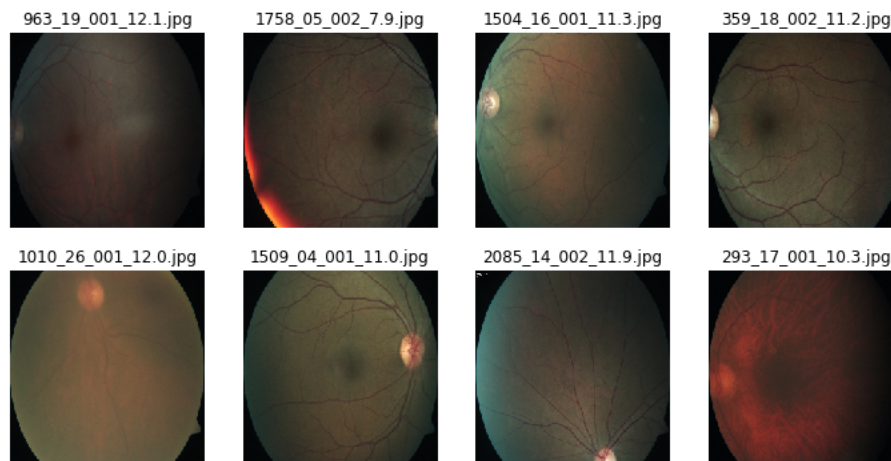


Figure 3.3: Anemic patients sample images

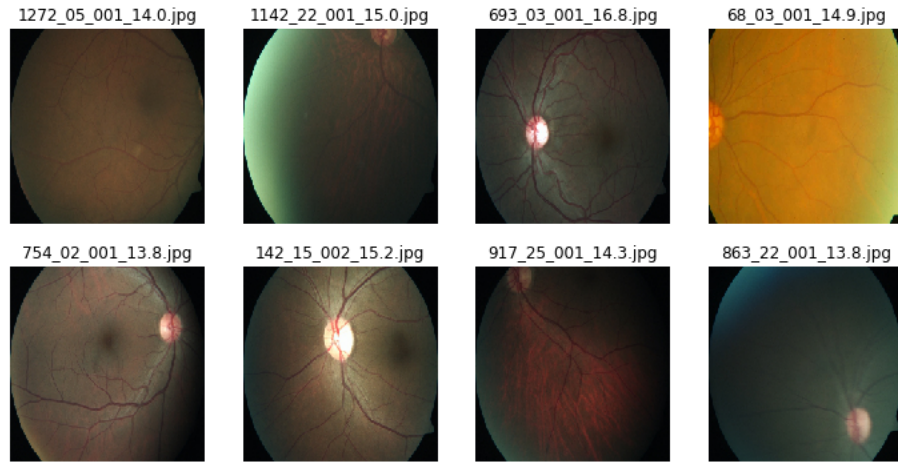


Figure 3.4: Non\_anemic sample images

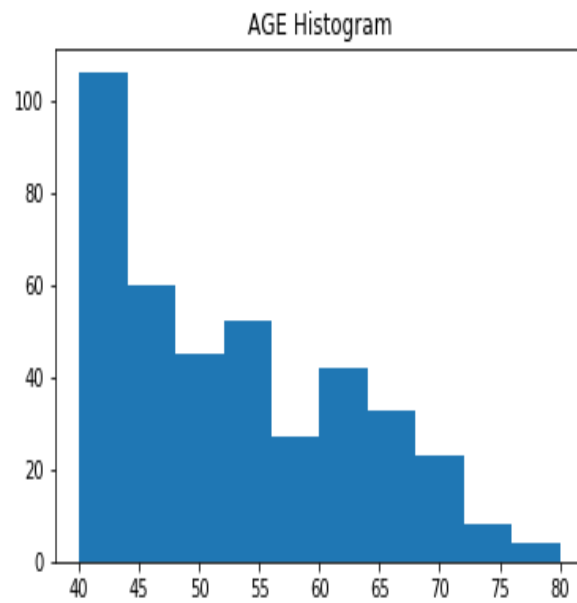


Figure 3.5: Age Histogram among Data

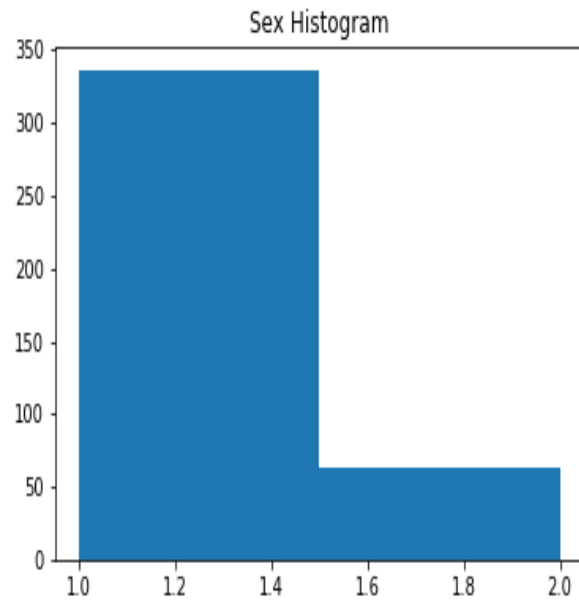


Figure 3.6: Sex histogram Among Data

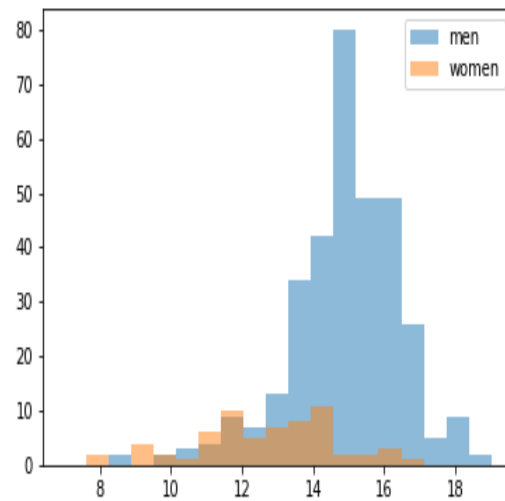


Figure 3.7: hemoglobin histogram with different sex Among Data

# Chapter 4

## Data Preprocessing

In any Machine Learning process, Data Preprocessing is the major step in which the data gets transformed, or Encoded in such state that now the machine learning model can easily use it. In other words, the features of the data can now be easily interpreted by the algorithm. We have taken the reference from this website. [7]

### 4.1 Rescaling

As already mentioned before we have images of different resolutions among 2240\*1488, 2588\*1958, 1504\*1000 pixel and to make it work in deep learning model we need it to be squared image and since 224\*224 resolution is commonly used in standard models and we have the resource limitation to use bigger networks as parameters is going to get increase we will be working with each image rescaled to resolution 224\*224.

### 4.2 Data Augmentation

This is very important part as we have the imbalanced data the way we are going to be handling it here is by using the data augmentation in the training samples. In that way it will be more generalized and can look in validation and test dataset which are the fundus images without the augmentation and say model is robust enough or not. Here we have used [14] this paper which is Development of a Fundus Image-Based Deep Learning Diagnostic Tool for Various Retinal Diseases and it shows the data augmentation that can be used. Which is width shift 0.4 , height shift 0.2, rotation range 90 degree, zoom range 0.1, horizontal flip to be true and shear range to be 30. This is also been presented in the figure 4.1.

Here, we can see from the Figure 4.2 that how the before and after augmentation images of this 4 samples looks like. In case of anemic we have augmented 12 from a single image and for non anemic 2 from a single image. Later we will remove the

Argument	Value
Width_shift_range	0.4
Height_shift_range	0.2
Rotation_range	90
Zoom_range	0.1
Horizontal_flip	True
Vertical_flip	True
Shear_range	30

Figure 4.1: Data Augmentation used in the image

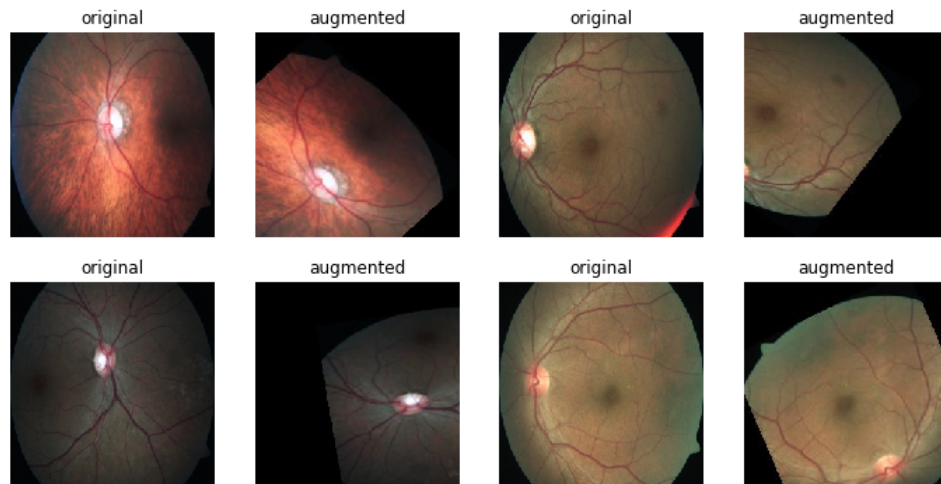


Figure 4.2: Data Augmentation sample of image before and after augmentation

unwanted data and expect it to be around balanced data. After data augmentation we have 5769 anemic and 5833 non\_anemic images.

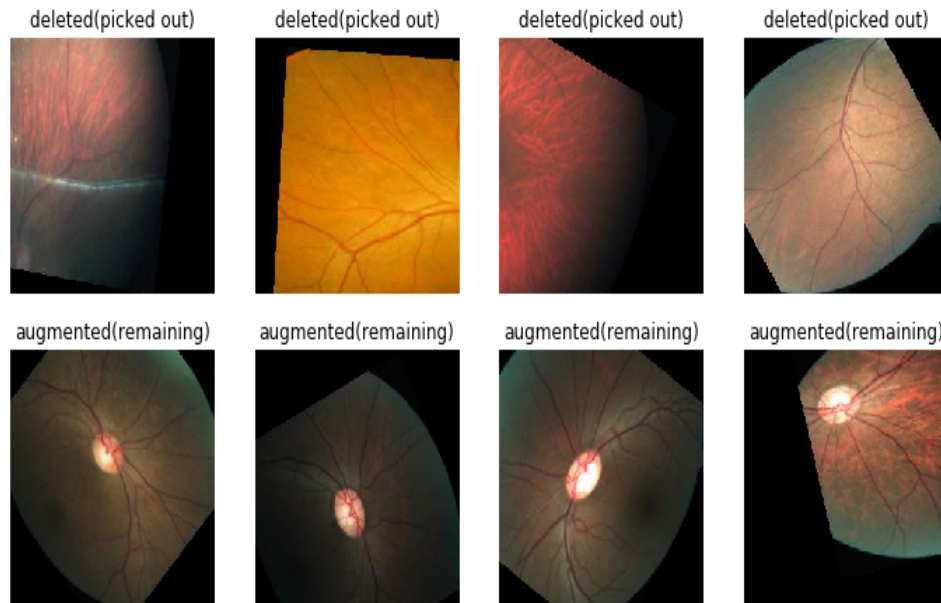


Figure 4.3: Picked out and remaining sample of images after this is done

### 4.3 Removed unwanted images from data Augmentation

As we use the data augmentation optic disc in images will shift to right or left or top or bottom either way data augmentation may lead to loss optic disc in the image and our aim here was to train model with optic disc and see it's importance. Hence it becomes necessity to pick that data again with optic disc on it. After this is done we are left with 4430 anemic images and 4342 non anemic images. So, in this process we have actually deleted 1339 anemic and 1491 non anemic augmented images. Sample of deleted and taken images in figure 4.3

### 4.4 Normalize data

Here, we will try to do the pixel normalization which is to scale pixel values to the range 0-1. In general RGB image have 3 channels and in each channel intensity of each pixel is between 0 to 255. The main reason to do this is Neural networks process inputs using small weight values, and inputs with large integer values can disrupt or slow down the learning process.



# Chapter 5

## Binary Classification

### 5.1 Introduction

Binary classification is one of the most common and frequently tackled problems in the machine learning domain [1]. Binary classification as the name suggests deals with two labels that is it assigned to every instance in a dataset. In our problem statement we have to find a model that inputs the patient's eye image of a fundus and it will output a value which will denote the prediction whether based on that fundus image that patient is anemic or not.

### 5.2 Loss function for Binary Classification

Binary crossentropy is a loss function that is used in binary classification tasks. These are tasks that answer a question with only two choices (yes or no, A or B, 0 or 1, left or right). Mathematically, it is the preferred loss function under the inference framework of maximum likelihood. It is the loss function to be evaluated first and only changed if you have a good reason. Cross-entropy will calculate a score that summarizes the average difference between the actual and predicted probability distributions for predicting class 1. The score is minimized and a perfect cross-entropy value is 0. We are using this in keras and for that loss='binary\_crossentropy' is used in the function.

### 5.3 Optimizing Function

Adam extension to stochastic gradient descent that has seen broader adoption for deep learning applications in computer vision and natural language processing. It has the adaptive learning rate. It makes use of the average of the second moments of the gradients (the uncentered variance). Popular in deep learning because it achieves good

results fast.[8]

## 5.4 Activation Function

As the model which we are going to be using is CNN which doesn't have activation function till it gets flatten . After that we have used we add 256 layer dense with Relu(Rectified Linear Unit ) Activation [3] and after that in the final output layer with output function as softmax. ReLU for short is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. It has become the default activation function for many types of neural networks because a model that uses it is easier to train and often achieves better performance. Softmax is useful because it uses weighted sum values into probabilities

## 5.5 Metrics for this Binary Classification

### Precision

It calculates the proportion of positive predictions, those are actually correct.It calculates the accuracy for the minority class basically.In our problem we are taking anemic and non anemic classes.Precision = TruePositives / (TruePositives + FalsePositives)[16]

### AUC

Area Under the Curve (AUC)[16] is the measure of the ability of a classifier to distinguish between classes and is used as a summary of the ROC curve.It is a probability curve that plots the TPR against FPR at various threshold values and essentially separates the 'signal' from the 'noise'.This curve is called ROC(Receiver Operator Characteristic).Here, higher the AUC, the better the performance of the model at distinguishing between the positive and negative classes

### Weighted log loss\_metric

Logarithmic loss[16] (related to crossentropy) measures the performance of a classification model where the prediction input is a probability value between 0 and 1. The goal of our machine learning models is to minimize this value. A perfect model would have a log loss of 0. Log loss increases as the predicted probability diverges from the actual label.When we give weight to each of its class then its weight log loss.

# Chapter 6

## Anemia Detection using Resnet50

### 6.1 Motivation

We generally perceive that "the deeper the model better it is" when it comes to convolutional neural network(CNN's). This makes sense, since the models should be more capable (their flexibility to adapt to any space increases) because they have a bigger parameter space to explore. However, it has been noticed that after increase in some depth, the performance degrades. This was one of the bottlenecks of earlier networks before the 2015. ResNet gives us the residual learning framework to ease the training of networks that are substantially deeper and still performance of it doesn't degrades. It has won the 1st place on the ILSVRC 2015 for classification & localization task. [12] The depth of representation is very important for many visual recognition tasks and we have used this deep representation to generate class activation maps to indicate the discriminating image regions used by the CNN to identify that category .

### 6.2 Preliminaries

#### 6.2.1 Architecture

Let us consider  $H(x)$  as an underlying mapping to be fit by a few stacked layers (not necessarily the entire net), with  $x$  denoting the inputs to the first of these layers. One hypothesizes that multiple nonlinear layers can asymptotically approximate complicated functions , then it is equivalent to hypothesize that they can asymptotically approximate the residual functions, i.e.,  $H(x) - x$  (assuming that the input and output are of the same dimensions). So rather than expect stacked layers to approximate  $H(x)$ , we explicitly let these layers approximate a residual function  $F(x) := H(x) - x$ . The original function thus becomes  $F(x) + x$ . Although both forms should be able

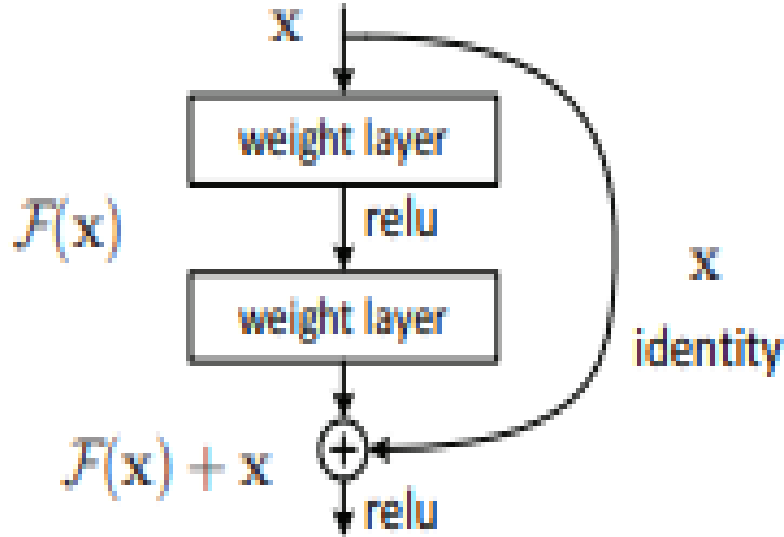


Figure 6.1: Residual learning: a building block. [6]

to asymptotically approximate the desired functions (as hypothesized), the ease of learning might be different. From this above observation taking a little bit further on that. Based on the plain network fig we insert shortcut connections which turn the network into its counterpart residual version. The identity shortcuts can be directly used when the input and output are of the same dimensions. In fig we can see that the ResNet consists on one convolution and pooling step( on orange) followed by 4 layers of similar behavior. Each of the layers follow the same pattern. They perform 3 x 3 convolution with fixed feature map dimensions (F) [64, 128, 256, 512] respectively, by passing the input every 2 convolutions. Furthermore, the width (W) and Height (H) dimensions remain constant during the entire lay.

### 6.2.2 Implementations

We have used the fundus images of Eye in this. As we have discussed in data preprocessing step we have rescaled, then data augment, then picked out for optical centers and then actually pixel normalized. This data now is been passed to the standard Resnet50 model taken from the Keras where from the last layer one dense layer of 256 of relu and then an output softmax layer of 2 output is been build. Here the optimization function used is Adam and used Binary crossEntropy as loss function. Metric which we are concerned with is AUC and precision. We have taken the early stopping [5] and ReduceLROnPlateau [6] for this. The batchsize is 64 in training per

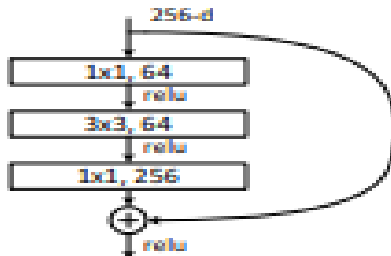


Figure 6.2: Building block for ResNet-50 [6]

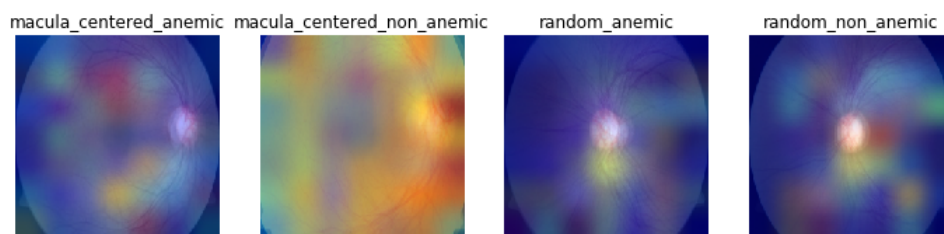


Figure 6.3: mean Gradcam results of 6 images of each 4 sample classes-resnet50

epoch. This training takes more than 3 hours of training. This training is done on the google colab completely.

### 6.2.3 Result

After the training is done with multiple hyperparameters and number of epoch we picked out the best among those in terms of AUC. Here, on the test data we have AUC of 0.6223, precision 0.59 and loss of 0.74. In terms of limitation of resources and data this is the quite initial result which can be improved further.

For the evaluation of this model to actually see whether we can conclude anything from it or not we have used the Gradcam [17] which is the visualization technique to see the result and which area affects on original image. For this we have taken the 6 images of the sample and 4 groups (macula\_centered\_anemic, macula\_centered\_non\_anemic, random\_anemic, random\_non\_anemic) in which images chosen such that they are similar looking in the class group. After this we have taken all 6 grad result and combine and mean it in one image and superimpose it in 6 original similar looking images.

Here, from this image fig 6.3 of gradcam we can see that for the non anemic cases gradcam is focusing much more on the optical center region than macula. But for the anemic case there is focus on optical disc but not the strong one (Otherwise the color would be more in terms of red, yellow here it is greenish or blueish).

# Chapter 7

## Anemia Detection using Mobile net v2

### 7.1 Motivation

A new mobile architecture, MobileNetV2 [15], that improves the state of the art performance of mobile models on multiple tasks and benchmarks as well as across a spectrum of different model sizes. It supports the task like support classification, detection and and semantic segmentation very well. It is build on the idea of using depthwise separable convolution as efficient building blocks. as with traditional residual connections, shortcuts enable faster training and better accuracy. As we have the limited resource on google collaboratory and it offers the quite low number of parameter as compare to other networks.

### 7.2 Preliminaries

#### 7.2.1 Architecture

Here it introduces two new features to the architecture: 1) linear bottlenecks between the layers, and 2) shortcut connections between the bottlenecks. Residual blocks connect the beginning and end of a convolutional block with a skip connection. By adding these two states the network has the opportunity of accessing earlier activation's that weren't modified in the convolutional block. This approach turned out to be essential in order to build networks of great depth. This idea as an inverted residual block because skip connections exist between narrow parts of the network which is opposite of how an original residual connection works. When you run the two snippets above you will notice that the inverted block has far fewer parameters. In figure bottleneck architecture is there in this t: expansion factor, c: number of output channels, n:

Input	Operator	$t$	$c$	$n$	$s$
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	-

Figure 7.1: Bottleneck Architecture

repeating number,  $s$ : stride.

### 7.2.2 Implementations

We have used the fundus images of Eye in this. As we have discussed in data preprocessing step we have rescaled, then data augment, then picked out for optical centers and then actually pixel normalized. This data now is been passed to the standard Mobilenetv2 model taken from the Keras where from the last layer one dense layer of 256 of relu and then an output softmax layer of 2 output is been build. Here the optimization function used is Adam and used Binary crossEntropy as loss function. Metric which we are concerned with is AUC and precision. We have taken the early stopping [5] and ReduceLROnPlateau [6] for this. The batchsize is 64 in training per epoch. This training takes more then 2.5 hours of training. This training is done on the google colaboratory completely.

### 7.2.3 Result

After the training is done with multiple hyperparameters and number of epoch we picked out the best among those in terms of AUC. Here, on the test data we have AUC of 0.671887 , precision 0.67 and loss of 0.584. In terms of limitation of resources and data this is the quite initial result which can be improved further.

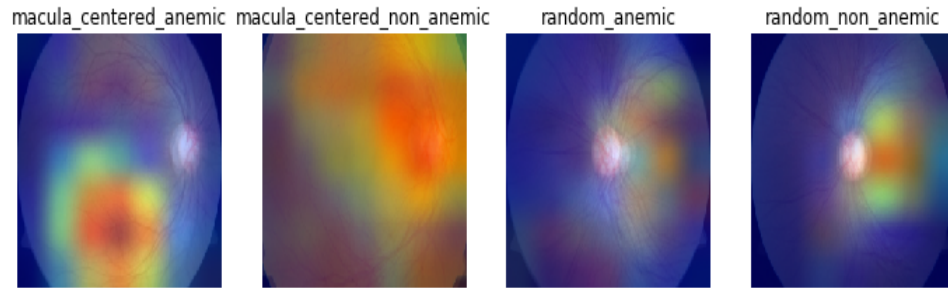


Figure 7.2: Mean Gradcam results of 6 images of each 4 sample classes-mobilenetv2

For the evaluation of this model to actually see whether we can conclude anything from it or not we have used the Gradcam [17] which is the visualization technique to see the result and which area affects on original image. For this we have taken the 6 images of the sample and 4 groups (macula\_centered\_anemic, macula\_centered\_non\_anemic, random\_anemic, random\_non\_anemic) in which images chosen such that they are similar looking in the class group. After this we have taken all 6 grad result and combine and mean it in one image and superimpose it in 6 original similar looking images. Here, from this image fig ?? of gradcam we can see that for the non anemic cases gradcam is focusing much more on the optical center region than macula. But for the anemic case there is not at all focus on optical disc or quite minor one.



# Chapter 8

## Anemia Detection using Densenet121

### 8.1 Motivation

DenseNets [9], are the next step on the way to keep increasing the depth of deep convolutional networks After the resnet's [12] and VGG 19 [13].As the problem when layers get deeper that they get vanished before getting to the other side (i.e. output layer).Author solve that problem by ensuring maximum information (and gradient) flow.Here the main idea is instead of drawing representational power from extremely deep or wide architectures, DenseNets exploit the potential of the network through feature reuse.

### 8.2 Preliminaries

#### 8.2.1 Architecture

Here,no need to learn redundant feature maps. some variations of ResNets have proven that many layers are barely contributing and can be dropped. In fact, the number of parameters of ResNets are big because every layer has its weights to learn. Instead, DenseNets layers are very narrow (e.g. 12 filters), and they just add a small set of new feature-maps.Here also each layer has direct access to the gradients from the loss function and the original input image.

DenseNets make the first difference with ResNets right here. DenseNets do not sum the output feature maps of the layer with the incoming feature maps but concatenate them. DenseNets are divided into DenseBlocks, where the dimensions of the feature maps remains constant within a block, but the number of filters changes between them.Here, is what our model looks like for the image of  $224*224*3$ .

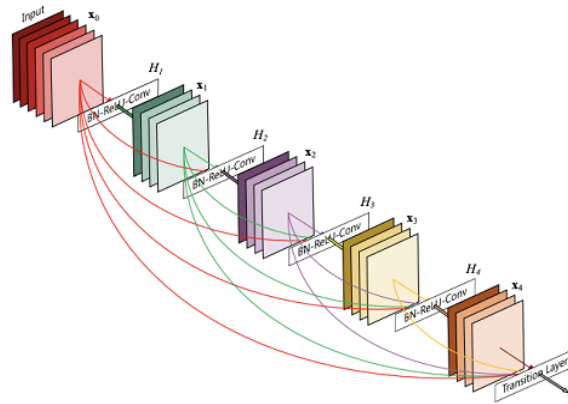


Figure 8.1: Densenet basic flow

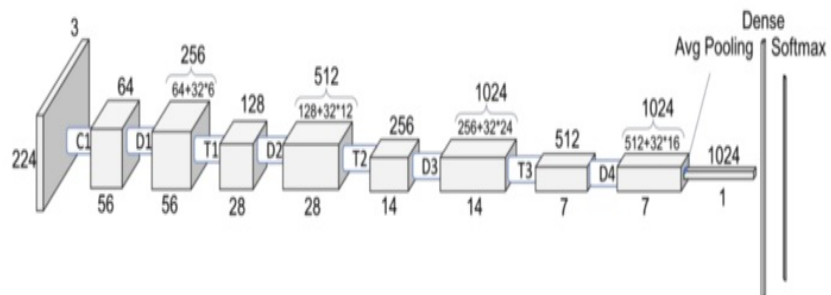


Figure 8.2: Densenet model

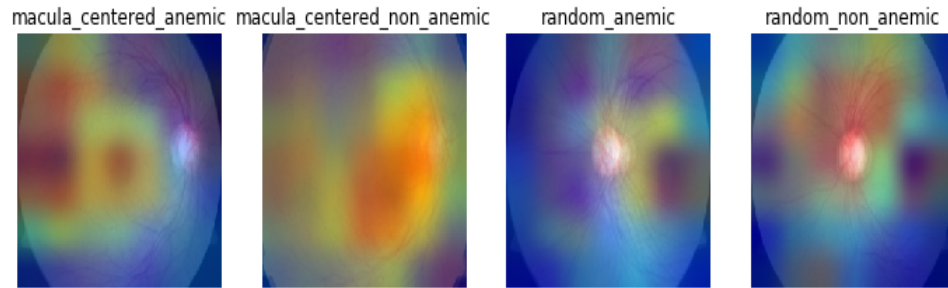


Figure 8.3: Mean Gradcam results of 6 images of each 4 sample classes-densenet121

## 8.2.2 Implementations

We have used the fundus images of Eye in this. As we have discussed in data preprocessing step we have rescaled, then data augment, then picked out for optical centers and then actually pixel normalized. This data now is been passed to the standard Densenet121 model taken from the Keras where from the last layer one dense layer of 256 of relu and then an output softmax layer of 2 output is been build. Here the optimization function used is Adam and used Binary crossEntropy as loss function. Metric which we are concerned with is AUC and precision. We have taken the early stopping [5] and ReduceLROnPlateau [6] for this. The batchsize is 64 in training per epoch. This training takes more then 2.5 hours of training. This training is done on the google colaboratory completely.

## 8.2.3 Result

After the training is done with multiple hyperparameters and number of epoch we picked out the best among those in terms of AUC. Here, on the test data we have AUC of 0.6649 , precision 0.587 and loss of 0.687. In terms of limitation of resources and data this is the quite initial result which can be improved further.

For the evaluation of this model to actually see whether we can conclude anything from it or not we have used the Gradcam [17] which is the visualization technique to see the result and which area affects on original image. For this we have taken the 6 images of the sample and 4 groups (macula\_centered\_anemic, macula\_centered\_non\_anemic, random\_anemic, random\_non\_anemic) in which images chosen such that they are similar looking in the class group. After this we have taken all 6 grad result and combine and mean it in one image and superimpose it in 6 original similar looking images. Here, from this image fig 8.3 of gradcam we can see that for the non anemic cases gradcam is focusing much more on the optical center region then macula. But for the anemic case there is not at all focus on optical disc or quite minor one.

# Chapter 9

## Future Work and Conclusion

### 9.1 Conclusion

In this project we have Achieved an AUC 0.671 with precision of 0.67 in the mobilenetV2 ,AUC 0.6223 with precision of 0.59 in the Resnet50 and AUC 0.6649 with precision of 0.587 in the densenet121.From all of this model we used gradcam mean images to actually which part this models are concentrating and observed that in case of the non anemic patient most of the focus was around the optic disc area but that's not the same case for the anemic patient.

### 9.2 Future Work

More data based study with model not limited upto the resource. Here the major limitation was the limited GPU because of which we didn't able to try even further with bigger network models.

# Bibliography

- [1] Binary classification <https://medium.com/deep-learning-with-keras/which-activation-loss-functions-part-a-e16f5ad6d82a>
- [2] fundus image <https://www.opsweb.org/page/fundusimaging>
- [3] Agarap, A.F.M.: Deep learning using rectified linear units (relu) (2018)
- [4] Akinori Mitani, Abigail Huang, Subhashini Venugopalan, Greg S. Corrado, Lily Peng, Dale R. Webster, Naama Hammel, Yun Liu, Avinash V. Varadarajan: Detection of anaemia from retinal fundus images via deep learning (2020)
- [5] Brownlee, J.: A gentle introduction to early stopping to avoid overtraining neural networks <https://machinelearningmastery.com/early-stopping-to-avoid-overtraining-neural-network-models/>
- [6] Brownlee, J.: Understand the impact of learning rate on neural network performance <https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/>
- [7] dayanand, T.: data preprocessing in deep learning <https://towardsdatascience.com/data-preprocessing-and-network-building-in-cnn-15624ef3a28>
- [8] Diederik P. Kingma, J.B.: Adam: A method for stochastic optimization (2015)
- [9] Gao Huang, Zhuang Liu, L.v.d.M.K.Q.W.: Densely connected convolutional networks (2016)
- [10] of Hospital Management is part of GlobalData Plc., T.: anemia prevalence <https://www.hospitalmanagement.net/comment/anemia-prevalence-nears-40-india/>
- [11] Jiang, J.: An introduction to deep learning <https://towardsdatascience.com/an-introduction-to-deep-learning-af63448c122c>
- [12] Kaiming He, Xiangyu Zhang, S.R.J.S.: Deep residual learning for image recognition (2015)

- 
- [13] Karen Simonyan, A.Z.: Very deep convolutional networks for large-scale image recognition (2014)
- [14] Kyoung Min Kim, Tae-Young Heo, A.K.J.K.K.J.H.J.Y., Min, J.K.: Development of a fundus image-based deep learning diagnostic tool for various retinal diseases, <https://www.mdpi.com/2075-4426/11/5/321/htm>
- [15] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh : Mobilenetv2: Inverted residuals and linear bottlenecks (2018)
- [16] Mohammed Sunasra: Performance metrics for classification problems in machine learning <https://medium.com/@MohammedS/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d>
- [17] Ramprasaath R. Selvaraju, Michael Cogswell, A.D.R.V.D.P.D.B.: Mobilenetv2: Inverted residuals and linear bottlenecks (2016)
- [18] Sabrina Felson, M.: What is anemia? <https://www.webmd.com/a-to-z-guides/understanding-anemia-basics>