Detection of skin cancer using Convolutional Neural Network

by

Abu Sa-adat Mohamed Moon-Im Al Ahsan
12301023
Shadman Monsur Alif
15101012
Junaid Bin Kibria
15101032
Prince Elvis Gomes
15101037

A thesis submitted to the Department of CSE
in partial fulfillment of the requirements for the degree of
B.Sc. Engineering in CSE

Department of Computer Science and Engineering
Brac University
October 2019

Declaration

We, hereby declare that this thesis is based on the results found by ourselves. Materials of work found by other researcher are mentioned by reference. This Thesis, neither in whole or in part has been previously submitted for any degree.

Student's Full Name & Signature:

ABU SA-ADAT MOHAMEDMOON-IM AL AHSAN
12301023

SHADMAN MONSUR ALIF
15101012

JUNAID BIN KIBRIA
15101032

PRINCE ELVIS GOMES
15101037

Approval

The thesis/project titled "Detection of skin cancer using Convolutional Neural Network" submitted by

1. ABU SA-ADAT MOHAMED MOON-IM AL AHSAN (12301023)
2. SHADMAN MONSUR ALIF (15101012)
3. JUNAID BIN KIBRIA (15101032)
4. PRINCE ELVIS GOMES (15101037)

Of Summer, 2019 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on September 26, 2019.

Examining Committee:

Supervisor:
(Member)

Dr.Jia Uddin
Associate Professor
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Name of Program Coordinator
Designation
Department
Brac University

Head of Department:
(Chair)

Mahbub Alam Majumdar
Professor
Department of Computer Science and Engineering
Brac University

Abstract

One of the most common and fatal cancer in the universe is skin cancer which arise from skin of epidermis, the topmost layer of the skin, it can happen anywhere in the body. We can find out the cancer by early detection. Skin cancer detection is a time consuming process and very critical. So in clinical applications, the machine learning analysis of skin cancer is failed to give correct images for a model. In our paper we followed three pre-processing steps which are: a) removing the shadows from the image which is illumination correction processing, b) to find the border of the skin lesion in the segmentation part, c) feature extraction by doing the ABCD framework. Our thesis makes an attempt to implement the method of Convolutional Neural Network. Using this classification, we find out the best result in inception v3 which was trained on skin lesions and we got the accuracy of 82.4%. So, our primary focus of this thesis is to differentiate between cancerous and non-cancerous image. Then our goal is to reduce importance of one of the painful process in cancer detection which is known as biopsy. Biopsy is removing tissue from a body and later it goes to many laboratory tests.

Keywords: Skin Cancer; Detection; Machine Learning; Convolutional Neural Network; Prediction; Cross Validation

Dedication

Firstly We want to dedicate our work to our Parents. Because without them we can not reach this far. We also want to thank Almighty for successful completion of our work.

Acknowledgement

Firstly we would like to thank the almighty for allowing us to initiate our thesis to put our best endeavor and effectively finish it.

Secondly, we offer our earnest and cordial gratitude to our respected Supervisor Dr. Jia Uddin for his contribution, enthusiasm, guidance and immense support in conducting the thesis and preparation for the report. His guidance, incorporation, involvement has incited us throughout our research and has been a motivator throughout our research.

Last but not least, we are thankful to our faculties, seniors, friends and our family who have been motivated inspired us in our journey indirectly with our research. We would also like to concede the assistance we got from numerous resources over the internet, especially from the fellow researcher's work.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

IPL Indian Premier League

LBW Leg before Wicket

MR Runs scored by Home team

MRN Home Team Run Rate

ODI One day International

OR Runs scored by the opponent team

ORN Opponent Team Run Rate

$T20$ Twenty Twenty

Chapter 1

Introduction

1.1 Motivation

Globally, the rate of skin disease are increasing and this diseases led people to skin cancers. It can be detected and treated early if possible by looking at the skin images. Thus we can predict that it is deadly or not. But sometimes people are not recognize that and later it will be so late for him. Also if people get the treatment from the healthcare system sometimes that cannot be possible in and also treatment can be costly or begin the process much later it could be starting. In diagnosis part it can be determined by biopsy, it is a process that doctors remove a small amount of tissue from the body to examine. Like X-rays detects the digital image of the internal image of something, especially the part of the body. But, for skin cancer, the only way to get out of it by biopsy process which is painful and also after the procedure there can be bleeding or infection. Biopsy process needs many medical professional and also it includes so much time and sometime accuracy is not correct. Because the technicalities, the diagnosing patient may die. So, due to this enormous complexity, we need more expert doctors to cure the people. Sadly, the expertise doctors are less than the affected people. So, doctors are now looking for other options which is machine learning algorithms.

This paper motivated us to think about the mechanism to recognize the disease or cancer and using technology to speed up the process for the doctors because image processing can save a tremendous amount of time and also increase the proficiency of research. So we remove the biopsy process and represent our three-way step to proceed with our program. Then we applied Convolution neural network algorithm on the images. CNN gives us the more accuracy than the current method.

1.2 Objectives

Our aim is to remove the painful medical process biopsy, so we prepare a model which gives us the accuracy of the skin cancer. Furthermore, if it can also tell us what kind of skin disease it has. So, in our paper our dataset will be preprocessed by the three steps which are a) illumination correction processing, b) segmentation of a lesion, c) feature extraction. Later for better accuracy we applied Convolutional Neural Network which is used in our paper to classify the skin cancer. We get better accuracy by taking more images and also by the iteration process. Thus, accuracy

can be improved. We proposed our model in Tensorflow in backend. We are applying Tensorflow because it is developed by Google which is a machine learning system.

1.3 Skin Cancer

As stated by the World Health Organization, cancer is the second most dangerous reason of death and about 9.6 million people are dead in 2018.**ferlay2012globocan** Cancer is a growing part of the body. It can affect any part of the body and it is mainly happening by the large group of diseases. Furthermore, it is the growth of the abnormal cell in the body. Among the cancers the most common cancers are: a) Lung cancer, b) Skin cancer, c) Breast cancer, d) Prostate cancer, e) Stomach cancer. Among all the cancers we will be working on Skin cancer as this is growing day by day. According to the World Health Organization, 132,000 melanoma skin cancer happens each year in the world. As we are working on skin cancer, we found two types of skin cancer, a) Melanoma, and b) Non-Melanoma. So, under non-melanoma there is Basal cell carcinoma and Squamous cell carcinoma. Basal cell carcinoma starts in the basal cells and develops in the rooted layer of the skin. When basal cell carcinoma occurs there is a bump in the skin which is transparent. On the other hand, the second most common cancer is squamous cell cancer which occurs middle and upper layer of the skin. The treatment of these two is painful and sometimes it cannot be figured.

These happen because

- Non-melanoma happens on the outer skin or the parts which are not covered where people get the ultraviolet light from the sun and the body part is ears, face, neck, forearms.

On the other hand, malignant melanoma is the major reason of death and it happens mainly because of the generic and personal characteristics. The strongest risk of this malignancy is one kind of mole in the body and mostly occurs on the fair-skinned people. Secondly, this malignancy occurs on the people who have a pale complexion, blue eyes, and red or fair hair. Thirdly, sunburn at an early age also causes of this malignancy. If this malignancy treated early then it can be curable but if not then it can be fatal and people might be dead. An estimation of 7320 people dead because of melanoma.

Recent times, the detection of skin cancer is done by visually. There will be a screening, a biopsy process, and hispathological examination.

1.4 Thesis Orientation

The rest of the thesis is organized as follows

- Chapter 2 discusses the fundamental of image processing and algorithms.
- Chapter 3 presents the working procedure of the research.
- Chapter 4 demonstrates the results found in our research.
- Chapter 5 concludes the thesis and states the future research direction.

Chapter 2

Background Study

2.1 Literature Review

Chandrahasa M , Varun Vadigeri and Dixit Salecha in their paper talks about the early detection of skin cancer via smartphone application by analyzing properties of the cancer, Asymmetry, Border, Color variation, Diameter and Expansion(ABCDE).These attributes are found by analyzing the datas using different methods of image processing like Grey scale conversion, Segmentation, border extraction and analysis of histogram**chandrahasa2016detection**

In [2013] J Abdul Jaleel: put forth the idea of skin disease detection based on Entropy of maximum threshold, Gray Level Co-occurrence Matrix (GLCM) for the extraction of features, and Artificial neural network(ANN) to classify the data into groups. Back-propagation neural (BPN) Network is also used for the purpose of classifying the data **jaleel2013computer**

Uzma Bano Ansari [2017] discussed how the identification of skin cancer in an early stage can help a patient both from pain and fast recovery of the disease. The different methods that were involved in bringing out the results are SVM(Support Vector Machine), Thresholding and GLCM(Gray Level Co Occurrence Matrix) **ansari2017skin**

In [2012] Mariam A.Sheha, Mai S.Mabrouk, Amr Sharawy : The paper discusses a particular way for the diagnosis of skin cancer named melanoma applying it on a group of cancer affected digital images. GLCM has been used to extract features from the images. Further MLP (multilayer perception) was used as a classifier to generate a difference between the positive and negative cancer images**sheha2012automatic**

The idea behind the work of X.Zang, J. Sun and S. Ren is to one day detect a malignant lesion just by taking a picture with the smartphone camera. For this to be successful they used fully convolutional residual neural network (FCRN) as a tool for segmenting the images. And for classification they used Deep Residual Neural Network(DRNN). Their work is inspired from the work of L. Yu, H. Chen, Q. Dou, J. Qin, and P. A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks"[5]. Their work is based on improving the techniques that the paper 'Automated Melanoma recognition using deep residual neural network' used **he2016deep**

In their paper Aya Abu Ali and Marzouqi used Convolutional neural network(LightNet architechture) to detect melanoma skin cancer. They used a started with 170 images but then using image augmentation along with rotation and cropping they increased their dataset to 6120 images. They created a total of 17 layers in the hidden section. For the final output they used an activation function named 'Softmax'. Their final output produced a specificity of 98%.

Advantage: accuracy is close 100%

Disadvantage: Computational load is very high and detects only melanoma cancerous cells

Angel Alfonso Cruz-Roa, John Edison Arevalo Ovalle, Anant Madabhushi, and Fabio Augusto Gonz alez Osorio used a Deep Learning Architecture for Image Representation, Visual Interpretability and Automated Basal-Cell Carcinoma Cancer Detection. Experimental evaluation was performed on set of 1,417 images from 308 regions of interest of skin histopathology slides, where the presence of absence of basal cell carcinoma needs to be determined. Their accuracy was 91.4%

Advantage: Their results are achieved comparatively faster with a decent amount of dataset

Disadvantage: They are only able to detect basal cell cancerous cells

We used 600 images to feed it into the input layer. After applying the CNN we achieved an accuracy of 82.4%. The advantage of our implementation is that the computational time is very low and we can detect both Melanoma and Basal cell carcinoma at the same time unlike the above two papers.

2.2 Purpose of Image processing

The important five reasons as to why image processing is necessary. They are:

1. Visualization – Finding out the elements in an image that are not visible without processing.
2. Image restoration – In order to produce a clearer picture.
3. Retrieval of Image – Sorting of only parts of the image that are actually important and leave out the unimportant ones.
4. Pattern Recognition – Divides the image into several parts and analyses the individual parts separately.
5. Recognition of image – figure out the meaning of the individual parts that were analyzed.

2.3 Fundamental steps of digital image processing

1. Image Enhancement
2. Compression
3. Segmentation

2.3.1 Image Enhancement

Enhancement of the image is basically turning the image into a clearer version of it. In this case the image is neither compressed nor the resolution is increased. What is does is that it highlights the raw image and brings in a more visible picture. It could also change the hue or saturation to make the image more vibrant. There are tons of apps now that can bring details via these methods.[Figure 4] shows a change in light condition between two images to bring more details. **ShrevyasReport**

Figure 2.1: Image Enhancement

2.3.2 Compression

Compression is a technique that reduces the size of the picture by reducing the pixels from it. The image will still be understandable but the quality will take a hit. Even though the quality will degrade it is necessary to compress the data for a storage that is limited. Lower image resolution would mean greater number of images can be transferred through a particular bandwidth. This also saves a huge load of time during processing as the read time will be much lower. In the following figure the picture lost some of the resolution in order to accommodate with the size available. This resulted in loss of detail [figure 2.2]**ShrevyasReport**

Figure 2.2: Compression

2.3.3 Segmentation

Segmentation divides the image into several constituents or objects. If we are automatically generate a segmentation of image then it would be quite hard to achieve in digital processing of image. After an in-depth segmentation of the image, the image is a lot closer to being successfully solved from imaging problems. This way individual objects in an image can be determined. In this figure each of the fruit are separately determined to distinctively differentiate between objects [Figure 2.3]. **pal1993review**

Figure 2.3: Segmentation of Image

2.4 Algorithms

2.4.1 Support Vector Machine (SVM)

Support vector machine is differentiation of classes which finds out a hyperplane in N-dimensional place and also a supervised method of classification and regression**hyperplane** So, SVM is mainly finding a line or separating two things with a line. If we want to draw a line between the dots and squares it would be easy, just draw a line between the dots and square.**svm**

Figure 2.4: Finding a line to separate

Drawing a line we get

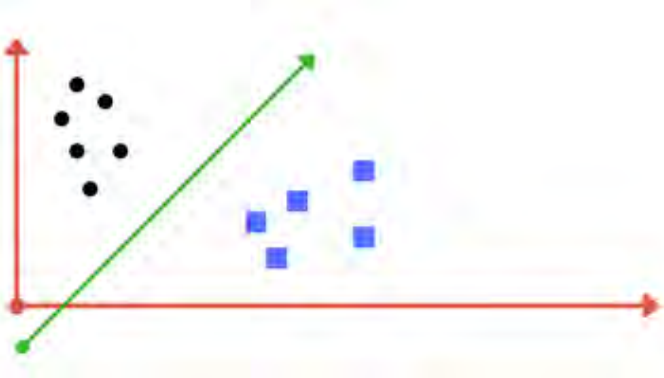

Figure 2.5: Green line separates these classes

But there are so many hyperplanes we can give in this diagram to detached data point. Earlier we said that, the given data portrayed by N- dimensional spaces and that means illustrates the number of data we use in our algorithm, which also represent as a particular coordinate. So, SVM takes input which is our data and giving us the output by giving a line which separates the data in particular numbers. So, this line is actually hyperplane which differentiates the two classes**hyperplane**

In left diagram there are many lines and red squares and blue circles. So we have to differentiate an optimal hyperplane between those two things. To find an optimal hyperplane, the line should be far from those red squares and blue circles because it will not give us correct answer and it will be noisy. So, the more the distance the more we get the outcomes from all data points**hyperplane** Also, there are data points which influence the position and orientation of the line and nearer to the hyperplane known as support vectors. It is used to maximize the margin of the data we are training.

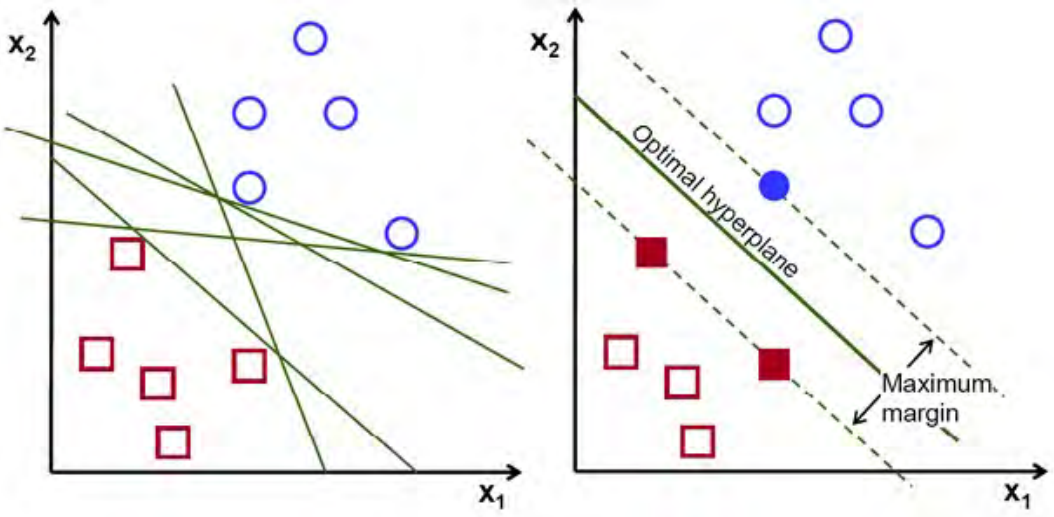

Figure 2.6: Optimal Hyperplanes

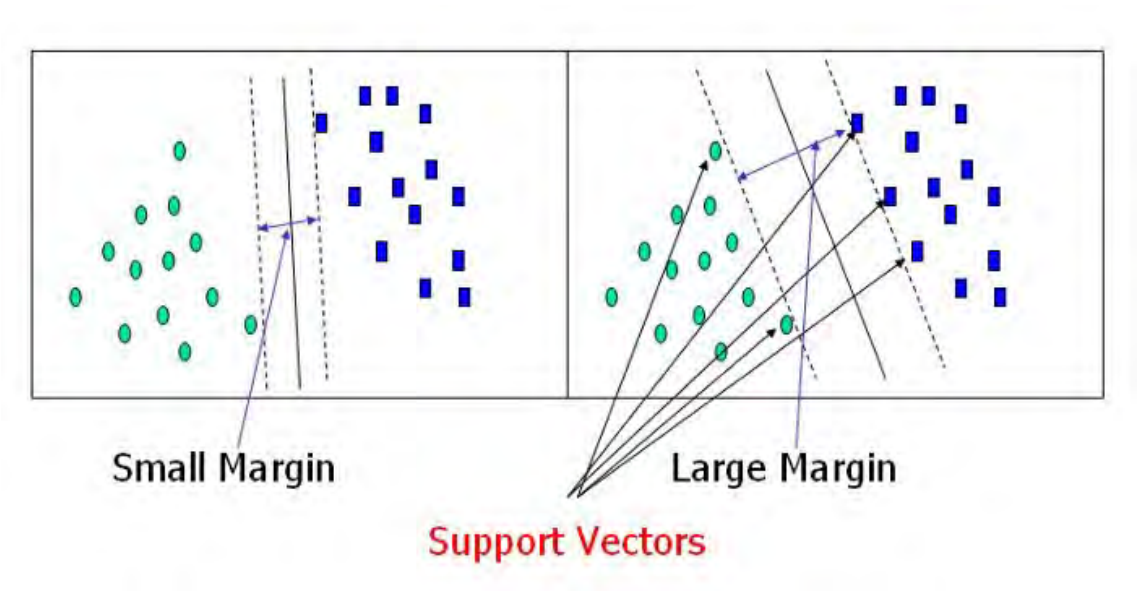

Figure 2.7: Support Vectors

Hyperplanes helps to categorize data points and works as a decision boundaries. There can be many possible hyperplanes to separate the data points also it depends on the number of the features. Suppose, the number of feature is 2, it gives us a line. On the other hand, if the number of feature is 3,**hyperplane** it gives us a two dimensional place as shown in the Fig 2.10

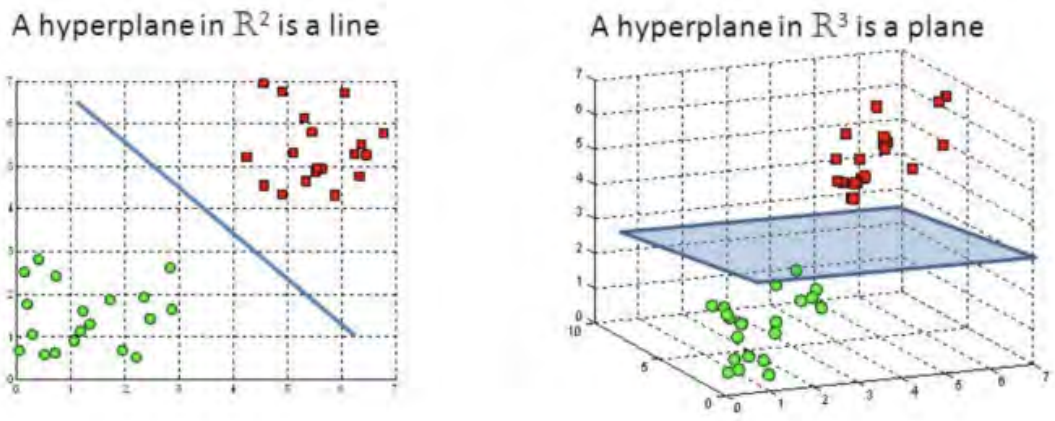

Figure 2.8: 2D and 3D feature Hyperplane

These are the data that can be linearly separable but there are also data which are not linearly separable. To do this, we have to add an axis known as z- axis because we can see the translucent separation of the lines which are more visible.**svm** So, the points on z plane is

$$z = x^2 + y^2 \tag{2.1}$$

Here z- origin is manipulating the data as a distance of point shown in the Fig below

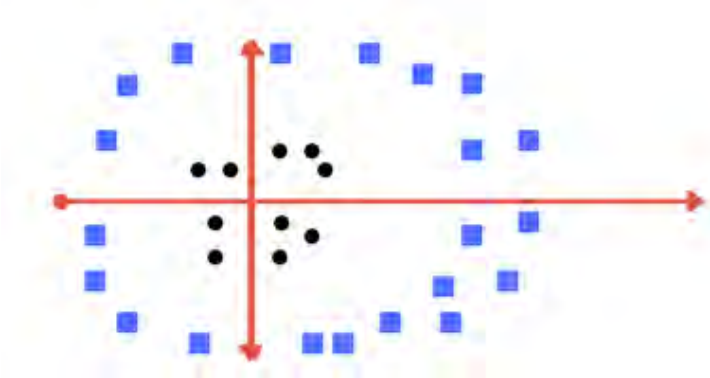

Figure 2.9: Finding a line

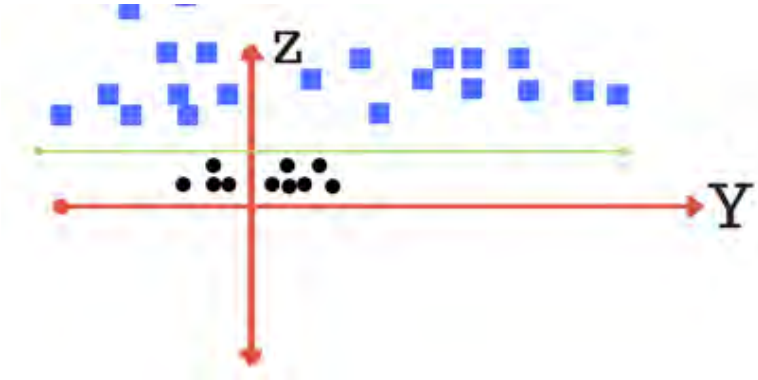

Figure 2.10: Line separation

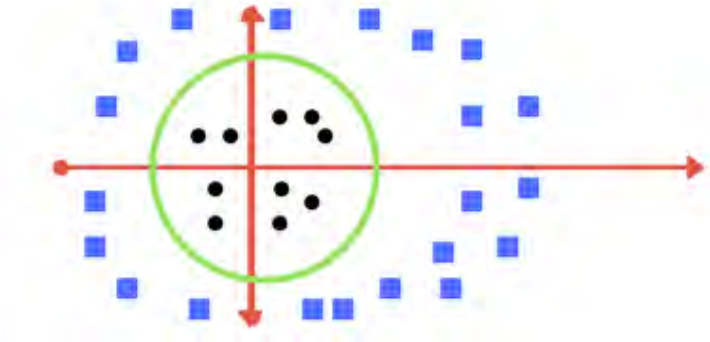

Figure 2.11: Line changing into circle

Now, we get a line which is separating those things and also is mapping a circular boundary which known as Kernels or the name of the techniques is Kernel trick. It converts the low dimensional input spaces into higher dimensional input spaces. In non-linear separation we used this kernel trick. In microarray data this algorithm gives us better performance and it is a powerful classifier**svm**

Thus we can get maximum margin to find a plane and which is our moto. It can also give us more confidence to classified data point which we can use in future.

2.4.2 K-nearest Neighbor (KNN)

One of the most supervised learning technique which is broadly used in machine learning and data mining is K- Nearest Neighbor. It is used in regression and classification K-NN is simple, easy to understand, useful and efficient. This algorithm is mainly instance based, non- parametric and lazy learning. In instance- based, it record the training data. This training instances is used for knowledge in prediction phase. In non- parametric, we can find the data in real world and it doesn't give us any assumption of the underlying data. Though it is call lazy algorithm but it is not, generalization technique is not used in here for the training data. K-NN keep repose all the training data because of the lack of generalization. Thus, the progress of this algorithm is very fast. Hence, all the training data we data we are using is need in the training phase.. When there is no knowledge about distribution how to use the data we use K-NN.

Suppose, we have a data about capsicums. a) Red capsicums and b) Green capsicums c) And the other capsicum which we don't know the color. We have to find out the color of the capsicum. Assume, R is red capsicum, G is green capsicum. So we have a dataset row belongs to capsicum and column belong to its characteristics. So, we have out data and its classifications. We can find the color of the capsicum by N ways which means we have to see data's characteristics. If it is similar to the red capsicum, then it is red and if it is similar to the green capsicum then it is green. This thing is done by mathematical calculations to measure the distance between the classifications.

This is called actually feature similarity. To find out nearest neighbor we have to follow some steps, which are:

- Take the unclassified data,
- Find out the distance between known and unknown data by doing some methods (Euclidian method, Manhattan method, Minkowski or weighted method)
- Give a parameter K to find smaller distance
- Check the shortest distance list and count that
- Receive the most times correct data
- Lastly, categorize the new data with last step.

Here, we see there are three types of colors, which is red (belongs to red capsicum) and green (belongs to green capsicum) and the blue one which we don't what the color of this. Now, if we calculate the measurement of the new data we can get which one is the smallest distance. As, we can see three capsicums belongs to smaller circle which is the closest data (k=3) and inside the data there are two green capsicums and one red capsicum. So, the blue one which not classified now we know the color of it which is green. This happens because of the feature similarity and now it is classified as green.

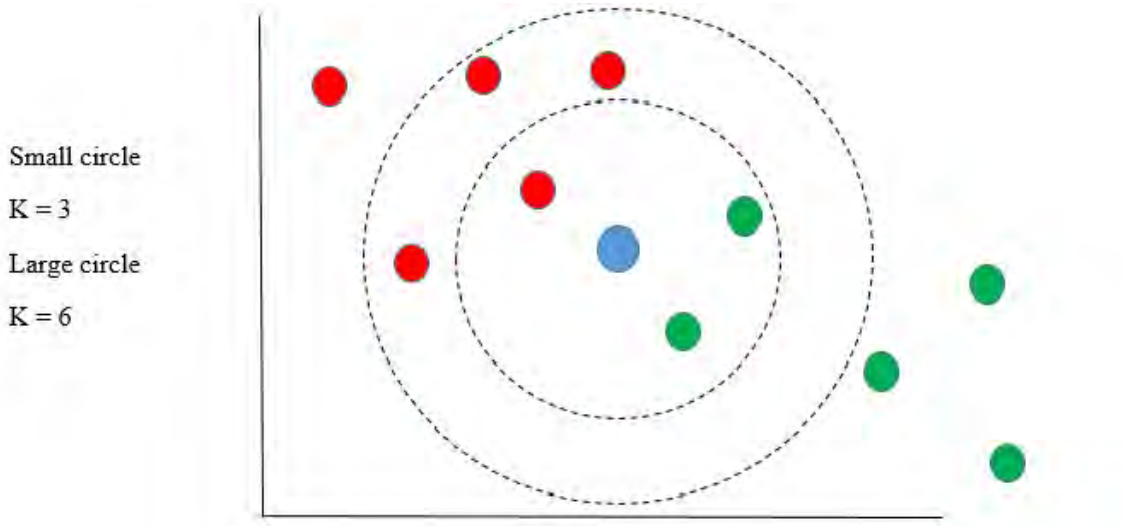

Figure 2.12: KNN Example

So, to find out K for KNN algorithm there is a process called "Parameter Selection" to find out the better accuracy for choosing the right value. There are some steps to choose the right value. They are:

- We have to find out the best value of "K" because there is no mathematical way to determine. "Unknown" is the assuming part of the training data.
- If we take small values of K, it will be noisy and also detached from the main body of system.
- On the other hand, larger values give us the perfect, smoother decision boundaries.
- Cross validation is the other way to find out "K" from the training dataset. We have to take a validation dataset which is the small section of the training dataset. In this way we predict the label, using (k=1), (k=2), or (k=3) and so on... by this way we can get the best value for the data.
- There is a formula to find out K which is k = sqrt(N); N= no. of samples we are using our training dataset.
- In other way, we can give the odd value of K because there can be a confusion between the classes and we have to avoid it.

As we said earlier, we use this algorithm for classification and regression problem. To find the object, it classified by the smallest distance and similarities of the nearest neighbor. So, we are applying this algorithm for classification in our paper. So, we have to find an appropriate value for its classification because without this value we cannot be successful.

2.5 Convolutional Neural Network

2.5.1 Convolutional Neural Network Architectures

LetNet 5 (1998)

In 1998, LeCun et al first introduced this 7 convolutional network which can classify digits. Later on, many banks used this to recognize hand written checks digitized in 32x32 pixel greyscale. For processing images with higher resolution, it needs more convolutional layers which are also larger in size. There are 3 main ideas behind building this network. They are local receptive fields, shared weights and special subsampling.

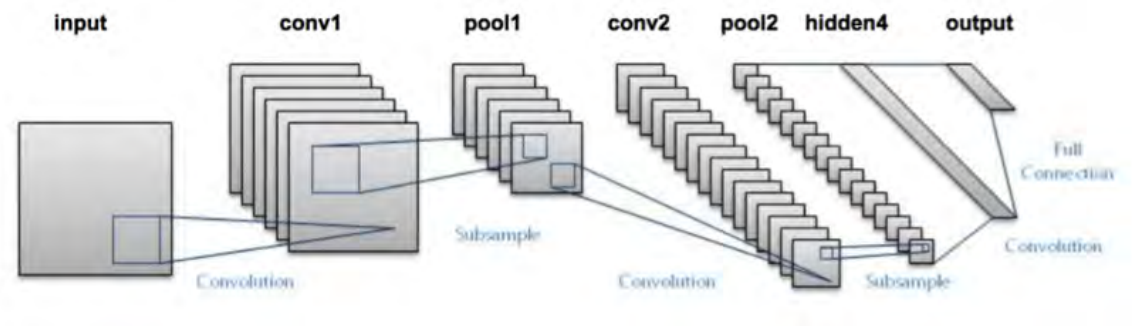

Figure 2.13: LeNet 5

AlexNet

AlexNet was developed by Alex Krizhevsky, Geoffrey Hinton and Liya Sutskever in 2012. This network is very similar framework to LetNet 5 but it was deeper because it has more filters per layer along with stacked convolutional layers.It has more data and it is a much bigger model than LetNet7 hidden layers, 650K units, 60M params). In 2012 , AlexNet got massive success by outperforming all the prior competitors as they won the challenge by reducing the top 5 error from 26 percent to 15.3 percent.

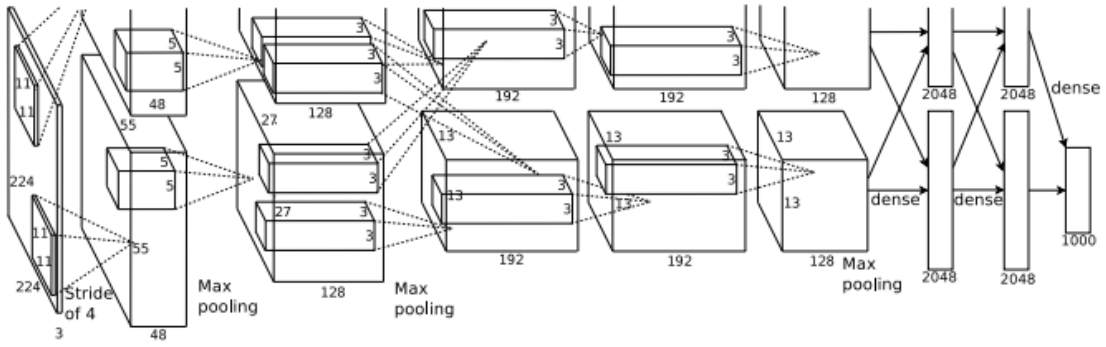

Figure 2.14: AlexNet

ZFNet

ZFNet was designed by Dr Fergus and his PhD student at the moment, Dr. Mathew D. Zeiler in NewYork University. Hence, this network is called ZFNet, based on surname, Zeiler and fergus. The image classification error rate was improved massively in ZFNet at time of comparing with Alexnet. This network maintained almost same structure and framework like AlexNet and got great achievement by tweaking the hyper parameters. This network successfully won the ILSVRC 2013.

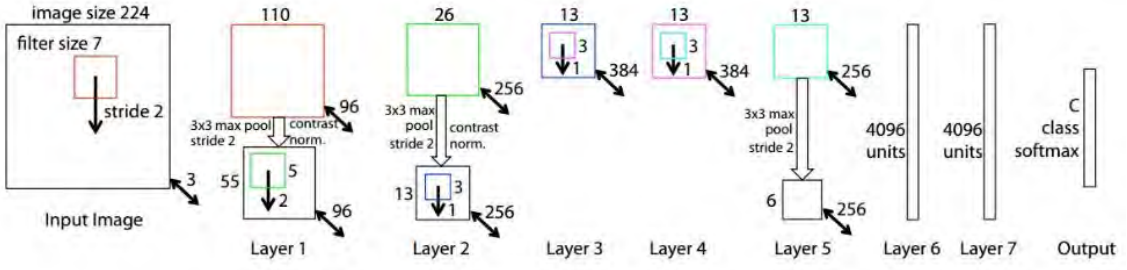

Figure 2.15: ZFNet

GoogLeNet

From the word "GoogLeNet" we can realize that this is from Google. Here, "LeNet" used for paying tribute to Prof. Yan LeCun's Letnet. GoogLeNet was the winner of ILSVRC (ImageNet Large Scale Visual Recognition Competition) 2014. This was a big improvement over ZFNet and AlexNet and it had a very lower rate of error than the first runner-up VGGNet. This was also known as Inception V1 but later on, V2, V3 and V4 are also designed by them. In GoogleNet the top 5 error rate was only 6.67 percent which was a great achievement. GoogLeNet architecture consisted of 22 layers CNN and reduced number of parameters significantly from 60 million (AlexNet) to 4 million.

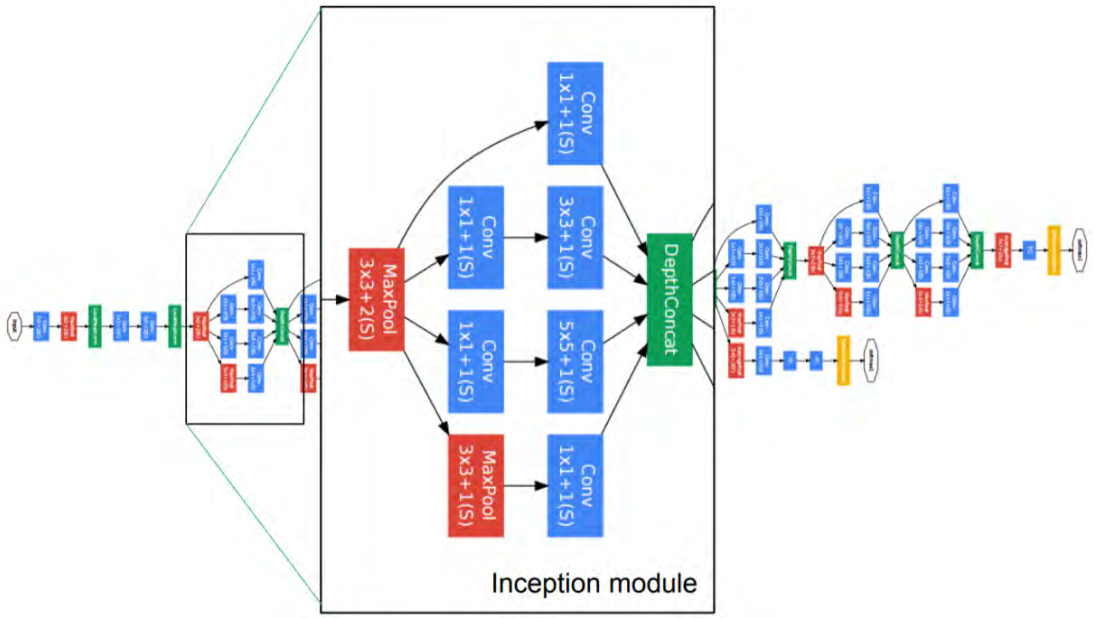

Figure 2.16: GoogleNet

VGGNet

VGGNet was designed by Simonyan and Zisserman in 2014. This network became runners-up at the ILSRC 2014 competition. Like AlexNet, this network has only 3x3 convolutions but it has a lot of filters. VGGNet ha 16 convolutional layers.VGGNet has around 138million parameters which is a bit difficult to handle since GoogLeNet which was also invented on the same year, had only 4million parameters.

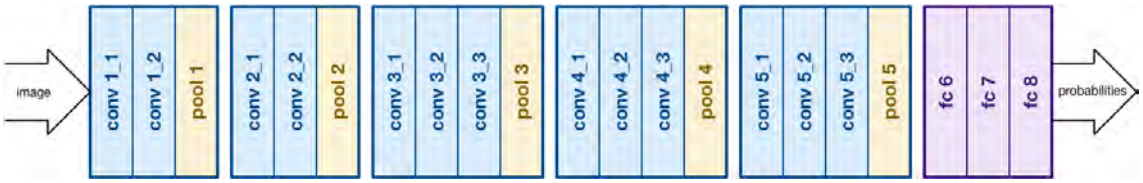

Figure 2.17: VGGNet

ResNet

ResNet was developed by kaiming He and was first introduced in ILSVRC 2015. ResNet is also called Residual Neural Network. They used around 152 layers while still having lower complexity tan VGGNet. It won ILSVRC 2015 with a top 5 error rate of 3.57percent which beats human-level performance on dataset. While GoogLeNet used inception modules, ResNet used residual connections.

Figure 2.18: ResNet

Table 2.1: From the light of above discussion, the summary table is shown below

Summary Table			
Year	CNN	Top5 error rate	No of parameters
1998	LeNet(5)	0.96%	60 thousand
2012	AlexNet	15.3%	60million
2013	ZFNet	14.8%	
2014	GooglLeNet	6.67%	4million
2014	VGGNet	7.3%	138million
2015	ResNet	3.8%	

Chapter 3

Working Procedure

3.1 Dataset

To begin with, after collecting the data our first task was to prepare them properly. After checking the CSV file we found that images were in different formats. Then we needed to convert all of them into JPG format. Moreover, although in the CSV file all the image of same disease were in a sequence, we found that, in the "image folder" there was not such sequence. Thus, we needed to separate them by the help of CSV file. Later on, we checked the CSV file and found that images were scaled in 600x425 format. Then, we resized all the images into 150x150 format because previously the images were too large and it would take too much time at time of processing. Nextly, we used TensorFlow in backend and Keras for running the machine learning and the library for image processing. We used the default dataset of keras to train our program and later on we got accuracy after the analysis. The reason behind getting so low accuracy was that in the default library there were basic images like the images of flowers, animals etc. After that, for getting better accuracy we added three pre-processing steps and intended to apply neural network. The three stages are illumination correction preprocessing, lesion segmentation and feature extraction.

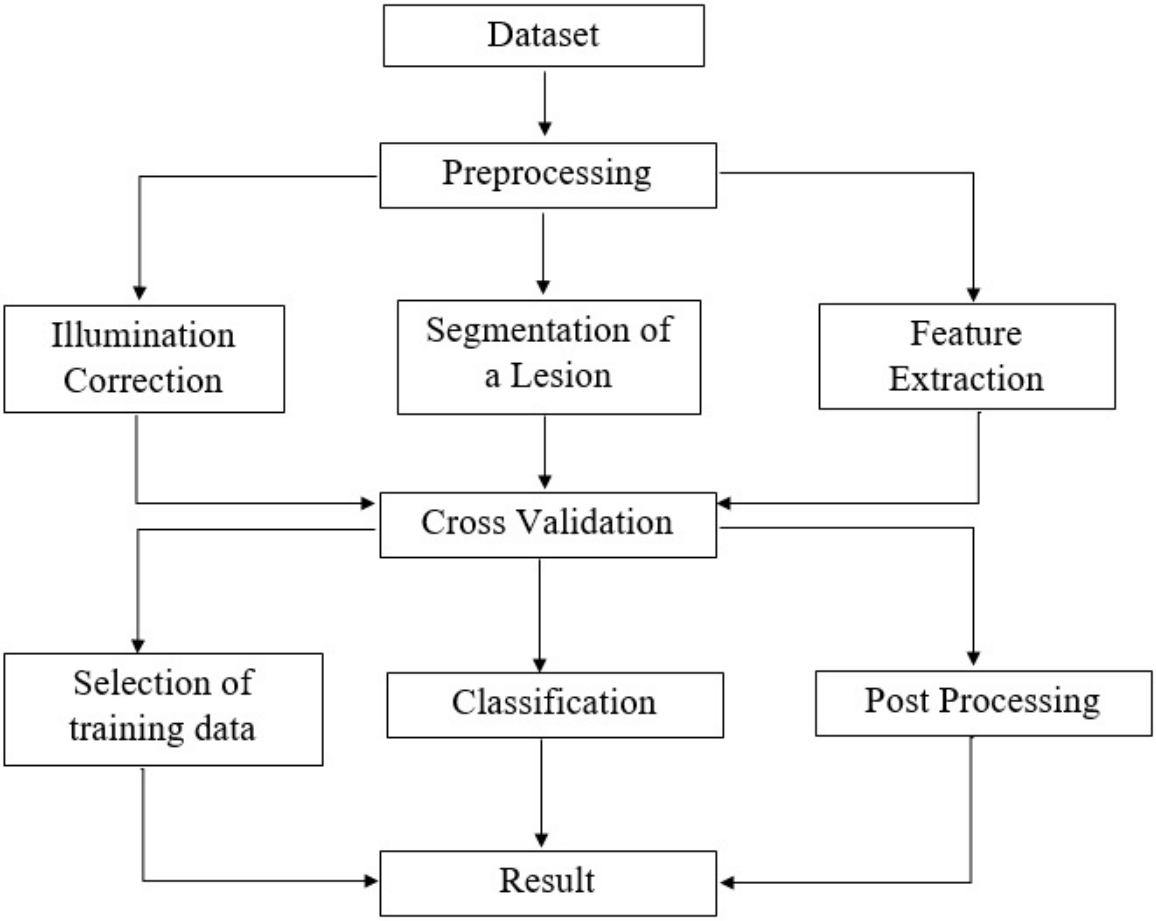

Figure 3.1: Workflow

3.2 Preprocessing

3.2.1 Illumination Correction Processing

In the pre-processing steps illumination or radiance correction processing is one the main step in our paper. We used standard camera to take correct images of the wounded skin. In our images, sometime we shadows or the bright areas. So, first work is to remove those extra shadows and brightness. This work is done by illumination correction processing. We applied here Multistage Illumination Correction Modeling (MSIM), it actually remove the reflectance component. So, we have to first estimate the radiance component and calculate the proportion of light using this. Finally, our classication model is introduced for identifying our input fundus image related to which type of skin disease or cancer among the 8 diseases.

We use this process because it has some features.

- This process works on complicated texture.
- After carry away the shadow, it sustains the consistency of the skin lesion.

- It give us corrected pictures after the process.

After the classification, the original RGB image is altered into HSV color to get the illumination map.

Then, it is send into segmentation segment.

3.2.2 Segmentation of a lesion

As our work is to find out the corrected image so we have to apply this pre-processing steps. Second step is segmentation which works done by finding border of the skin lesion. It approaches texture distinctiveness based on the lesion border. In our paper, it gives us the separation of the images containing into each pixels. As our images are medical images it can give us the perfect outline or contour of the image.

3.2.3 Feature Extraction

To classify our image we used some features which are extracted after the segmentation part. In this part image is represented by N- dimensional image space. We used the most important parameter here which is known as ABCD parameter.

- A stands for Asymmetry
- B stands for border anomaly
- C stands for color
- D stands for diameter.

From many classification algorithms we used KNN, SVM, CNN. We used KNN because it records the training data which is used for regression and classification. Moreover, we used Support Vector Machine (SVM) is used for segmented the lesion as cancerous or non- cancerous disease. It gives us the perfect classification results which makes hyper planes to classify set of data.

3.3 Cross Validation

In machine learning, sometimes we have to split our data into training and testing data. This method is called cross validation. It gives us the measurement of our training dataset based on our model's performance. So, in our paper, we divide our dataset following some steps.

- Split the data into training set and testing set
- In training set we train our data which we take 70% from our dataset
- To evaluate the model on testing set we take 20% from our dataset.

- For different data points we repeat the steps.

This way we trained our model and for validation it repeat multiple times. For unbiased evaluation we use validation dataset. For validation, we used 10% of the total data. Sometimes if we trained our model too well it would be fit in our training set and give us the error. If the error rate is too high then our model will be overfitting model. On the other hand if the data is not fitted, it cannot be generalized to the new data.
To overcome the drawbacks, we use K- Fold Cross Validation in our paper. In K-Fold Cross Validation, we divide K into number of sections where each of the section work as a testing set at some point where the first fold is used for testing and the other K- 1 folds are used for training. In second step the second fold is used for testing and the rest are used for training. It continues until the last fold comes. For validate the model each time different fold is used. One of the best thing of K- Fold cross validation that it tested each fold accurately and also take time for the testing k-1 times. Then, we average the result of each set and gain our accuracy.
As we see, each and every section took part in testing and training set, so it give use perfect accuracy. We can increase or decrease the value of K. Increasing the value of k, the model will be become biased. On the other hand, if we decrease the value of k, the model will be less biased. In our paper we used the value of k = 27.

we also used VGG16 in our thesis and it gives us the accuracy which we needed. On the other hand for better accuracy we applied inception v3. This performs very well in low computational power.

3.3.1 Selection of training data

We use 10015 images in our thesis. In our model it took 600 images by default.

3.3.2 CNN Implementation

The initial stage of Convolutional Neural Network is to import the sequential model from Keras that will help set the layers of nodes sequentially from input to the output layer. Following this task we will set the image size that we are going to feed in to this network. Here the images will be 150 by 150 and they will RGB colors.

```
model = models.Sequential()

model.add(layers.Conv2D(128, (5, 5), activation='relu',input_shape=(150, 150, 3)))
```

Here the 2 dimensional convolution is used. The parameter (5,5) specifies the size of the filter in width and height.

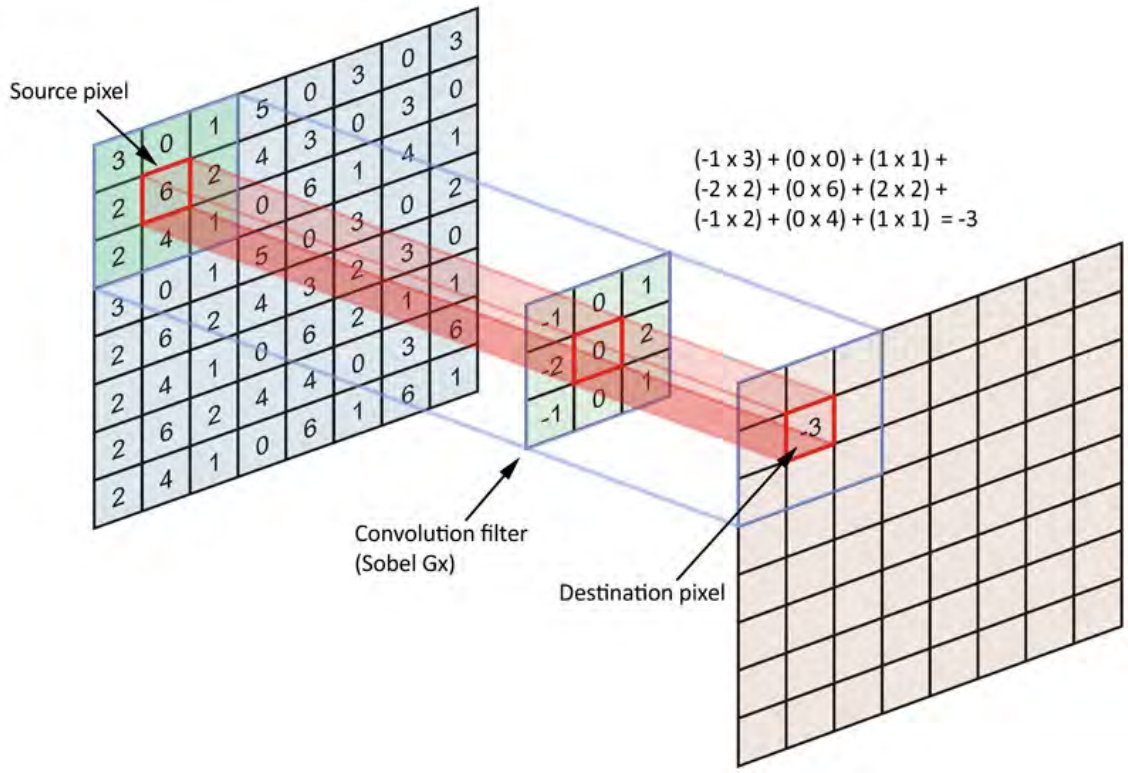

Figure 3.2: Convolution

This example perfectly shows how the source pixel is converted to destination pixel using a 3 by 3 filter. For our case the source image is 150 by 150. The 5 by 5 filter will matrix multiply with the top left part of the image according to the size of the filter. After the first value of the destination pixel is calculated the frame moves by a defined pixel known as Stride. If the Stride is 1 then the frame in the source image will move one pixel while it will move 2 pixels for 2 and follow the same procedure until it fully creates the destination pixels.

Width of Destination Pixel= (Width of source image-spatial extent+2*Zero Padding)/Stride+1

Height of Destination Pixel= (Height of source image-spatial extent+2*Zero Padding)/Stride+1

The digit 128 determines the number of filters to be used. This way there will be 128 different sets of output image from convolution. This is called a feature map.

Next up comes the activation function. We used Rectified Linear Unit(ReLU) for the computation of the values in the nodes of hidden layer.

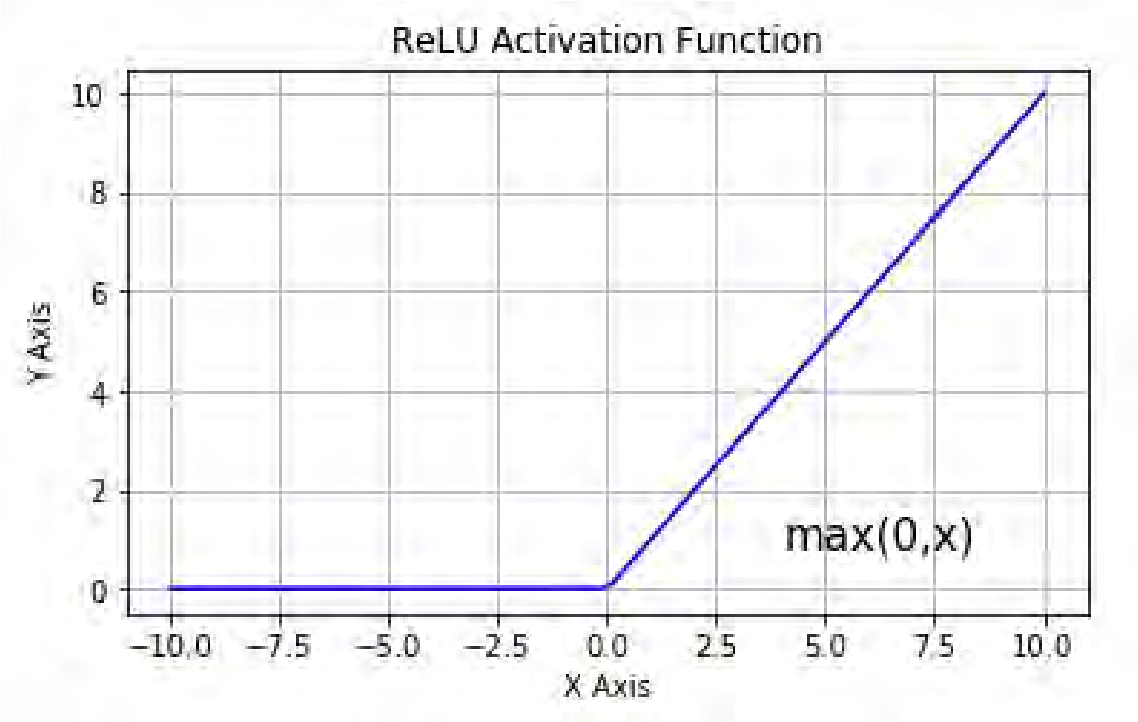

Figure 3.3: Rectifier Linear Unit

If the value is negative, then it is set as 0. But if the value comes as any positive number then it is set as the number itself. This way the value stays within the range 0 and 1.

```
model.add(layers.MaxPooling2D((5, 5)))
```

Then comes MaxPooling of the convoluted image. The purpose of this is to reduce the size of the image so that the number of parameters are vastly reduced. This in turn lowers the computational load and thus reduces cost.

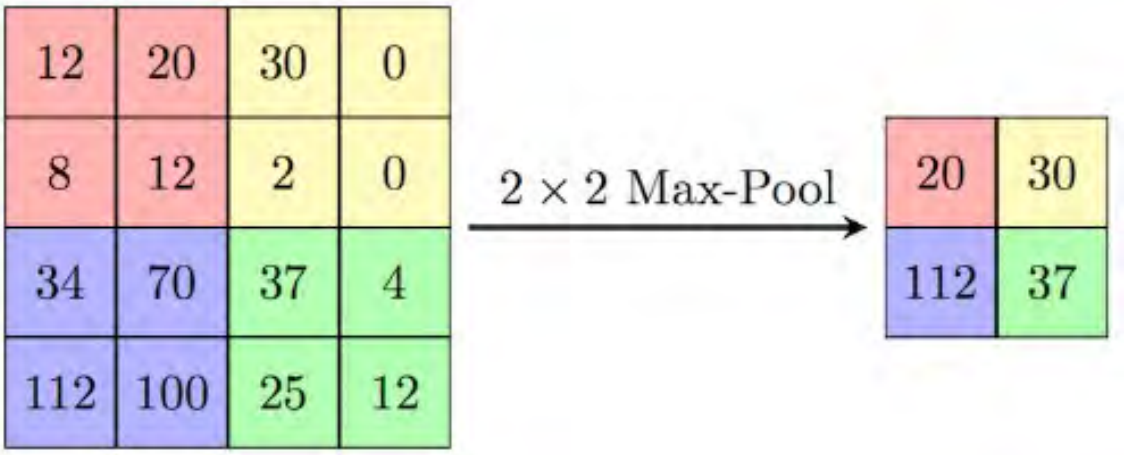

Figure 3.4: Maxpool

This example shows that from each segment the MaxPool brings out the largest value from the convoluted image using a 2 by 2 MaxPool. In our case we used 5 by 5 MaxPooling.

```
model.add(layers.Conv2D(64, (4, 4), activation='relu'))
model.add(layers.MaxPooling2D((4, 4)))
model.add(layers.Dropout(0.2))
model.add(layers.Conv2D(32, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((3, 3)))
model.add(layers.Dropout(0.25))
model.add(layers.Flatten())
model.add(layers.Dropout(0.4))  #Dropout for regularization
```

Then we carried out the same procedure several times to make hidden layers upon hidden layers but with different number and sizes of filter. The dropout layer helps the training set from being overfit. Dropout as the name suggests simply drops some of the nodes within the hidden layer for the output to have a better validation and accuracy.

Until now we had a feature map containing several output images in 2D space. Now we need to convert it to 1D space. And for this we used the function flattened.

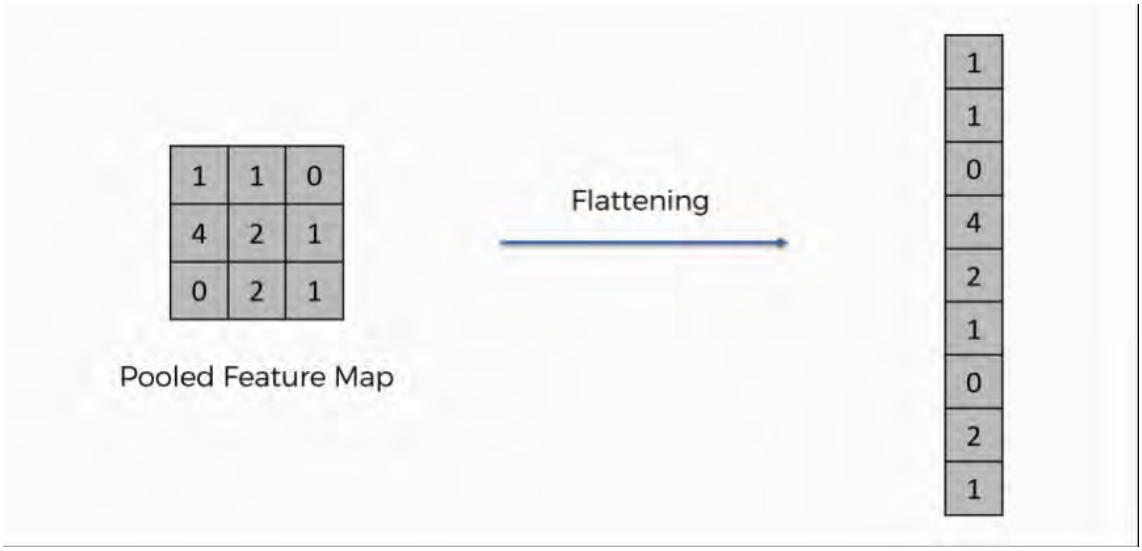

Figure 3.5: Flattening

```
model.add(layers.Dense(512, activation='relu'))
model.add(layers.Dense(256, activation='relu'))
model.add(layers.Dense(128, activation='relu'))
model.add(layers.Dense(64, activation='relu'))
model.add(layers.Dense(32, activation='relu'))
model.add(layers.Dense(1, activation='sigmoid'))
```

Then the number of nodes in a particular layer is defined until it reaches the result. For this we used the Dense function. From the image we can see the nodes kept on decreasing until it gets to 1 in the result layer. Up till the result layer ReLU function has been used as an activation function. Then finally to get the result the sigmoid function is used.

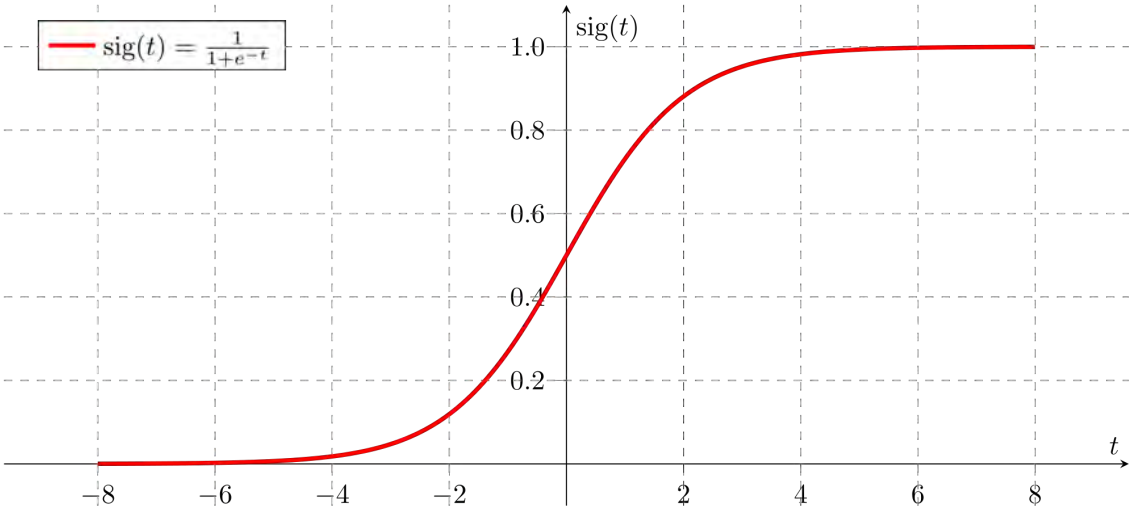

Figure 3.6: Sigmoid function

When the value gets too negative it reaches close to 0. As the value of t increases the value nears 1 but never goes beyond 1. The range similar to the ReLU function, this also has a range from 0 to 1.

Chapter 4

Experimental Results and Analysis

4.1 Dataset

We used dataset from International Skin Imaging Collaboration (ISIC)**isic** This is a combined academia and industry effort for improving skin cancer detection. It is mainly focused on melanoma more to reduce melanoma mortality using digital skin imaging technologies. This dataset is sponsored by International Society for Digital Imaging of the Skin (ISDIS). They are working on developing standards to be used in skin imaging. They have given special attention to privacy and interoperability. They also published some articles for skin lesion imaging. ISIC archive contains

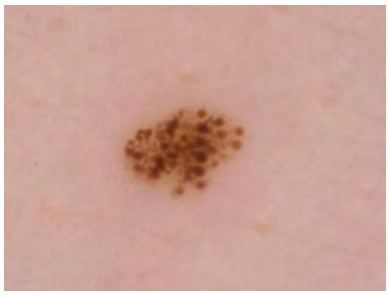

Figure 4.1: (a) Melanoma

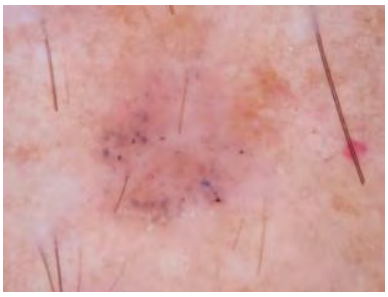

Figure 4.2: (b) Basal Cell

the largest available collection of quality controlled dermoscopic images. It contains over 23000. All this images are collected from clinical and diagnostic centers all around the globe. All this images are gone through few steps before being published to ensure quality control and privacy. This dataset is specially organized for using it in machine learning and neural networks. They also organize machine learning challenges every year starting from 2016. During 2016 challenge, they found average sensitivity of specificity of pigmented lesions was 82 percent and 59 percent respectively. In 2017 challenge, they able to specificity up to 74.7 percent. Main goal of isic melanoma project is to reduce melanoma related deaths and unnecessary biopsies. We have used this dataset not only for melanoma but also basal cell and also other related skin diseases. We used dataset of 10015 people of different age and gender.

We have used this dataset for identifying actinic keratosis, basal cell, dermatofibroma, melanoma, nevus, pigmented benign keratosis and vascular lesion.

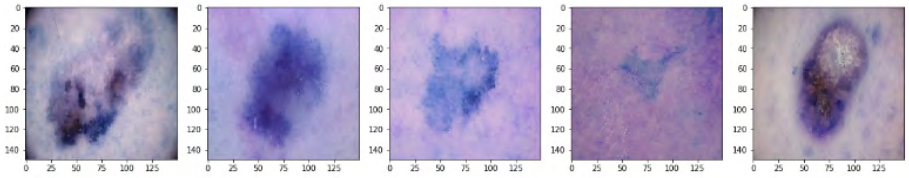

Figure 4.3: Randomly selected images from dataset

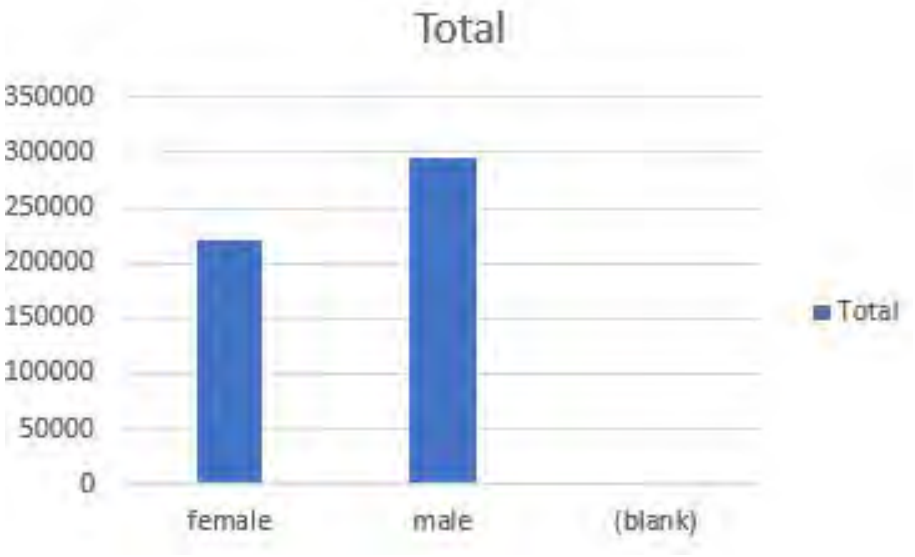

Figure 4.4: Age vs Gender

In this diagram we can see that total sum of age of female is 2,30,000 and of male is 2,90,000.

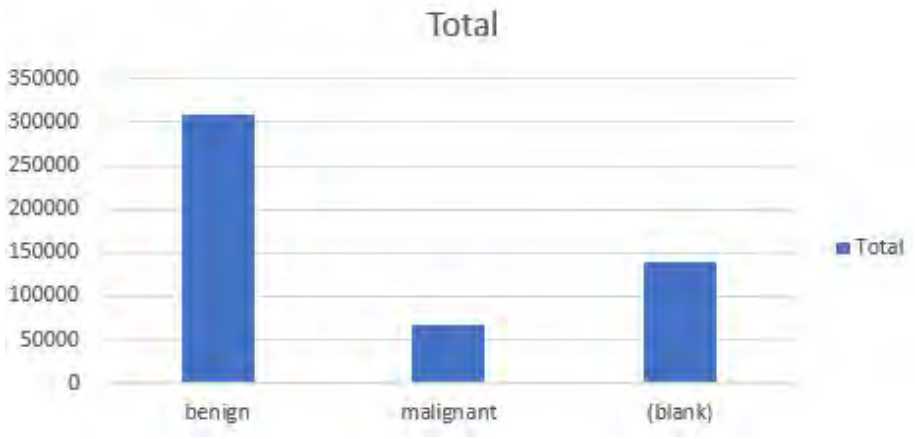

Figure 4.5: Age vs Melignent

From the above diagram it represent total sum of age of benign and malignant diagnosed people.

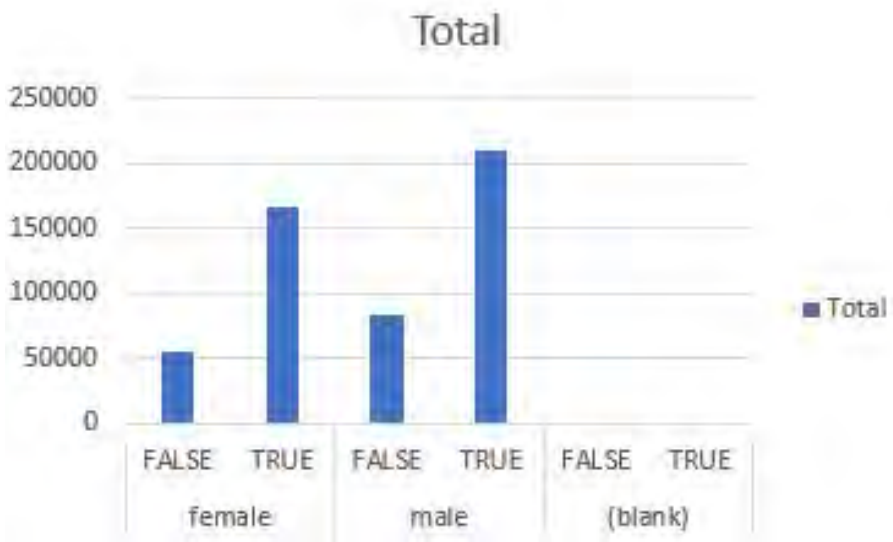

Figure 4.6: Age vs Gender and Melanocytic

From the above diagram we can illustrate sum of total age based on gender and melanocytic data. Here total sum of age of non melanocytic is 51,000 and of melanocytic is 1,65,000 of female. Again total sum of age of non melanocytic is 80,000 and of melanocytic is 2,10,000 of male.

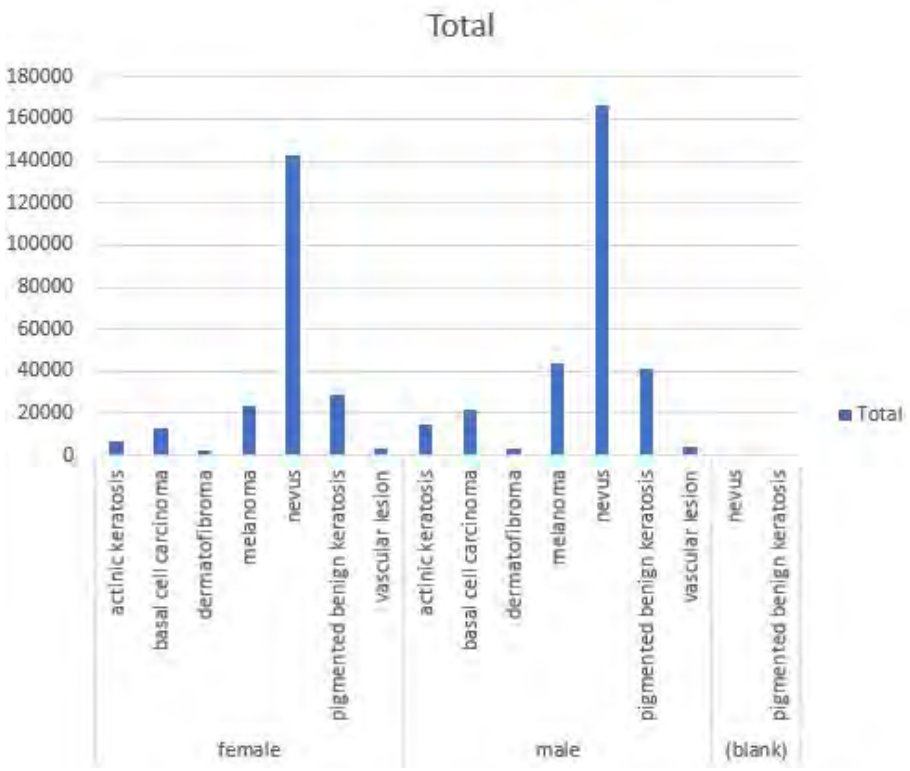

Figure 4.7: Age vs Gender and Diagnosed Disease

On the above diagram we can see total age based on gender and diagnosed disease.

4.2 Training and Results

At first we used this dataset on a pretrained model. At this stage we found very low accuracy of 15.5%. Because pretrained models are trained with variety of image as dataset.

Later, We used VGG-16 a convolutional neural network model to train our dataset.We took 2400 image of melignent and 2400 of non-melignent image at this stage.We ietarate for 32 times to train. After validation, we got 64% accuracy. In this stage accuracy has increased at a large scale. Because the model we used here is well defined and mature model then then our previously approached models.

Again we used another model named Inception V3 to train our dataset. At this stage we increased total number of image a bit. We used 5000 image for both side to train and iterate for 45 times. Here we were able to achieve accuracy of 72%.

Lastly we developed a proposed model. We used to train our dataset on this model. In this stage we used 6000 image and iterate for 70 times. Thus we got 82.4% accuracy.

For accuracy we used the default equation for accuracy that use below mentioned equation.

$$A\gamma = \frac{TP + TN}{TP + FP + FN + TN} \tag{4.1}$$

From the images below we see graph generated based on training and validation based on epoches. To avoid overfitting we used cross validation. In both of the images we can observe that intersection between training and validation is same. So we can can say the model we proposed is well balenced and working perfactly.

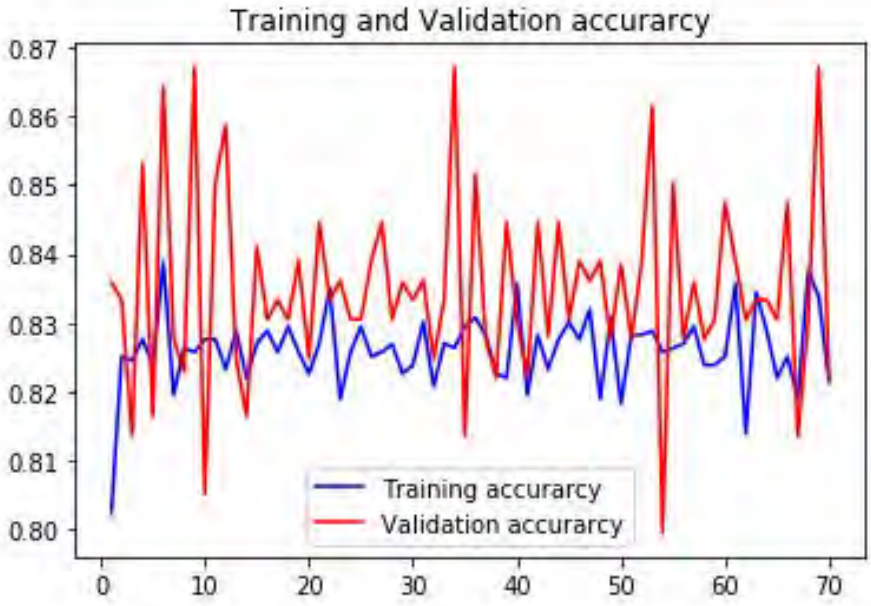

Figure 4.8: Training and validation accuracy

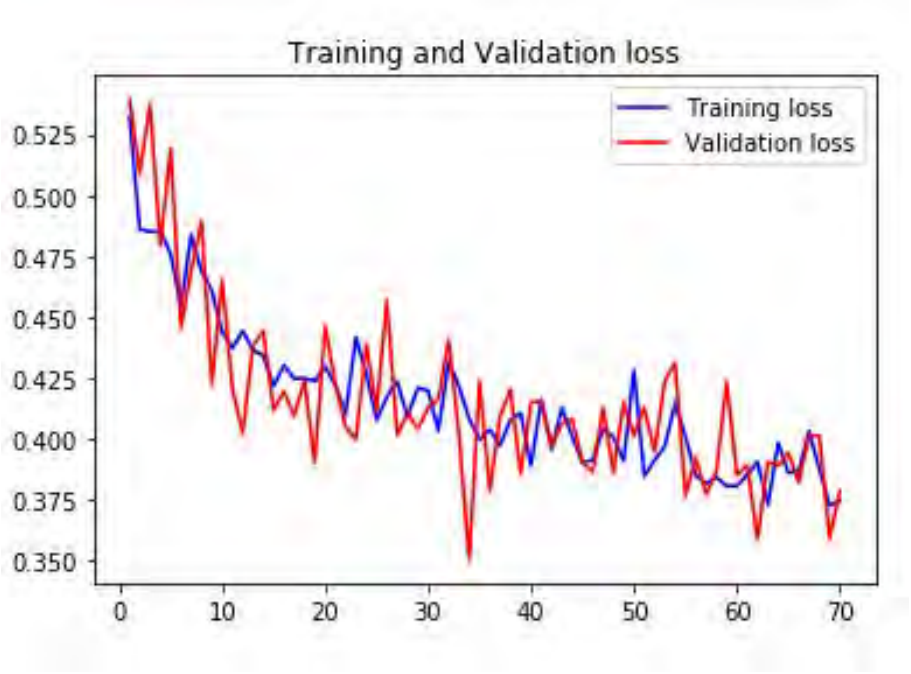

Figure 4.9: Training and validation loss

Chapter 5

Conclusion and Future Work

5.1 Conclusion

Since we got good accuracy rate in our program, this will function quite properly in real life application as our dataset was much enriched. We hope that our program will help medical team to detect diseases more efficiently. Main goal of this thesis was to reduce mortality rate from skin cancer and also reducing unnecessary biopsy since biopsy is very time consuming and at the same time painful process. We hope that outcome from our work will reduce this hectic process. Thus our hard work will be paid off. We found 82.83% accuracy with our model.

5.2 Future Work

In this dataset we worked on around 10015 images and we worked on around 8 type of skin cancers and diseases. In near future we have planned to increase quantity of images in dataset and apply new algorithms to improve the accuracy. Moreover, we have plan to add more images of other type of skin diseases and train our program in a certain way so that it can be capable of detecting more diseases. Moreover, using the algorithm we have a plan to work on other type of medical image processing. For instance, apart from skin cancer there are various types of other cancers. We have planned to work on them.