Integration of Handcrafted and Deep Neural Features for Melanoma Classification

by

Mohammad Saminoor Rahman
17201136
Md. Jubayer Hossain
17301177
Siful islam
16201050
Md.Nafiul Kabir
17101256
Md.Kamrul Hasan Sujon
16201070

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

Department of Computer Science and Engineering
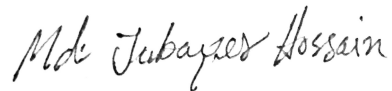BRAC University
September 2021

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at BRAC University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Mohammad Saminoor Rahman
17201136

Md. Jubayer Hossain
17301177

Md.Kamrul Hasan Sujon
16201070

Md.Nafiul Kabir
17101256

Siful islam
16201050

Approval

The thesis titled "Integration of Handcrafted and Deep Neural Features for Melanoma Classification" submitted by

1. Mohammad Saminoor Rahman (17201136)

2. Md. Jubayer Hossain (17301177)

3. Md.Kamrul Hasan Sujon (16201070)

4. Md.Nafiul Kabir (17101256)

5. Siful islam (16201050)

Of Summer, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on September 26, 2021.

Examining Committee:

Supervisor:
(Member)

_____
Md. Ashraful Alam, PhD
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Co-Supervisor:
(Member)

_____
Md. Tanzim Reza
Lecturer
Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

_____
Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
BRAC University

Abstract

Deep neural networks (DNNs) are widely utilized to automate medical image interpretation in many forms of cancer diagnosis and to support medical specialists with fast data processing. Although man-made characteristics have been used to diagnose since the 1990s, DNN is fairly new in this field and has shown extremely promising results. The fundamental goal of this study is to detect melanoma cancer in its early stages by obtaining a remarkable outcome with greater accuracy. Our purpose is to address the problem of an increase in skin cancer patients throughout the world, as well as an exponential increase in the danger of mortality from not commencing the diagnosis at an early stage, as a result of late detection. We propose that the research works on handcrafted features and merges the result with deep learning approaches with the initial help with a huge dataset of raw images. The DNN model used in this research has multiple layers with various effective filtering processes called batch normalization and dropout also with added layers named flatten and dense. In this process, images are classified to predict melanoma cancer at an early stage with Mean Shift, SIFT, and Gabor separately then the output was ensembled with later added Raw images results to give better accuracy. With an early integration model for separate featured databases and with a late and full integration model for ensemble with various results from the early integrated model we got our results. As a result, this neural network has provided an accuracy of 90% in early models and in late and full integration 86% and 84% respectfully, which is higher than other conventional approaches.

Keywords: Skin cancer, DNN, Handcrafted feature, Melanoma, Ensemble, Image segmentation, Confusion Matrix.

Dedication

This thesis team wants to dedicate the work to the development of the melanoma cancer detection process. This work might help the current researcher get better results. We tried to find a new way to get better results for detecting melanoma cancer.

This research is dedicated to our father, who showed us that the greatest kind of knowledge is gained for the purpose of learning.

It is also dedicated to our mother, who told us that if you do things one step at a time, even the most difficult work can be completed.

Acknowledgement

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

ACS American Cancer Society

AUC Area Under the Curve

BCC Basal Cell Carcinoma

CNN Convolutional Neural Networks

$DCNN$ Deep Convolutional Neural Networks

DL Deep Learning

DLS Deep Learning and Secuirity

DNN Deep Neural Network

$ECOC$ Error Correcting Output Codes

FC Fully Connected

FN False Negative

FNA Fine Needle Aspiration

FP False Positive

GUI Graphical User Interface

$ILSVRC$ ImageNet Large Scale Visual Recognition Challenge

$ISBI$ International Symposium for Biomedical Imaging

$ISIC$ International Skin Imaging Collaboration

KNN The k-nearest Neighbors

$MBConV$ Mobile Invented Bottleneck Convolution

MCC Merkel Cell Carcinoma

ML Machine Learning

MR Runs scored by Home team

MRI Magnetic Resonance Imaging

PCA Primary Component Analysis

$ReLU$ Rectified Linear Unit

RGB Red Green Blue

ROI Reigon Of Interest

SCC Scam Cell Carcinoma

SRC Sparse Representation-based Classification

SVM Support Vector Machines

TN True Negative

TP True Position

Chapter 1

Introduction

Skin cancer is an out-of-control proliferation in the external skin layer of aberrant cells in the epidermis produced by unrepeated DNA damage that causes mutations. These alterations lead to a fast multiplication of the skin cells and malignant tumors. Basal cell carcinoma (BCC), scam cell carcinoma (SCC), and melanoma and Merkel cell carcinoma (MCC) are major kinds of skin cancer. But of all skin tumors with a poor survival rate, melanoma is deadly.

Melanoma is a kind of cancer that arises from melanocytes, which are skin cells that generate the pigment melanin, which gives skin its color. Melanomas look a lot like moles and may occasionally develop from them. They may be found on any part of the body, even in places where the sun isn't usually present. Melanoma is often initiated by sunburn, which is caused by severe, intermittent sun exposure. The usage of tanning beds also raises the risk of melanoma. Melanoma may be detected early, which may enhance the patient's chances of survival. Melanoma kills one person in the United States per hour [36]. More than 207,390 new cases of melanoma are predicted to be diagnosed in the United States in 2021, with around 106,110 of them invasive. The most serious of the three most frequent types of skin cancer is melanoma. Melanomas may be treated and cured if they can be identified and treated early enough. Melanoma will kill roughly 7,180 people in 2021, according to estimates. Skin malignancies, such as melanoma, are caused by broken DNA (mutations) in skin cells, which leads to uncontrolled cell development. UV radiation from the sun or tanning beds causes DNA damage in our skin cells. Some of the damage is repaired by our immune system, but not all of it. The DNA damage that remains over time may lead to mutations that cause skin cancer. Melanoma is caused by a combination of variables, including genetics (family history), skin type or color, hair color, freckling, and the number of moles on the body.

Melanoma is linked to pale skin tones, sunburns, and past genetic factors, all of which are linked to a weakened immune system and excessive UV exposure [29]. If melanoma is not detected early on, it may form and spread through the first layer of the skin epidermis, eventually causing the lymph vessel and blood to come into contact. Form, color, and size are significant elements in identifying skin cancer, and irregular moles with forms, color changes, and diameters more than 6mm will show up. Various non-invasive approaches have been recommended for analyzing cancer detection and classifying it as benign or melanoma.

Due to a dearth of data and a concentrate on specific tasks like dermoscopy and classification of histological images, earlier work on dermatologic classification lacked

the capacity for medical professionals to generalize. [9] Dermoscopic images are created using a specific instrument, whereas historical images are created via invasive biopsy and microscope, both of which provide highly uniform images. Photographic images (for example, smartphone shots) have a wide range of magnification, perspective, and light, making categorization considerably more challenging. [7] Many previous systems required extensive preprocessing, lesion segmentation, and domain visual feature extraction before classification. Our method, on the other hand, does not need any bespoke features. It uses a single network for both photographic and dermoscopic images, and it is completely trained using picture labels and raw pixels. The current work makes use of small data sets, with less than a thousand photographs of skin lesions that do not create new images well. [3][5][11]

Recently, Deep Convolutional Neural Networks (DCNN) have achieved tremendous success in many machine learning applications and have shown an outstanding performance in various computer-assisted diagnosis applications. In terms of accuracy, DCNN outperformed highly qualified dermatologists. DCNNs classify the lesions based on high-level features rather than the conventional method incorporating the low-level dermoscopic visual information that requires a segmentation step beforehand to extract those features. [13]

1.1   Motivation

Melanoma skin cancer is the deadliest cancer among all the cancers and if we take a look more than 132,000 cases have been diagnosed every year [31]. Several Computer-Aided Identification Systems have been designed to help medical practitioners in the examination of skin lesions and the diagnostic of melanomas, including CT scans and MRIs. Other systems like Deep learning augment lesions in real-time and enhance melanoma diagnosis accuracy.

To improve the odds of survival, if the melanoma is found early enough, minimally invasive surgery may be recommended. Dermatoscopy is among the most often utilized imaging techniques by dermatologists. It causes the skin lesion to have a greater surface area, making it easier for the dermatologist to detect. This strategy, albeit effective, can only be used effectively by qualified doctors. Since we are having challenges, we must try to find new ways to diagnose and treat melanoma. Patients with melanoma use a computer-aided diagnosis (CAD) system to detect the disease. The CAD tool offers an easy-to-use environment for new dermatologists. CAD diagnostic software may be utilized as a kind of further analysis in the process of cancer diagnosis for melanoma.

The purpose of our research is to determine the relationship between the existence of symmetry and the identification of melanoma. Our objective is to conduct research to determine which characteristics are the most discriminating. Additionally, we considered a variety of shape and symmetry features to determine which ones had the most influence. This is the first study of its kind to our knowledge, since there has been little research on this kind of comparison in the field of deep learning.

1.2   Problem Statement

Melanoma is a kind of skin cancer that is caused by an abnormal proliferation of melanocytes (the cells that give the skin its tan or brown color). Melanoma is rare cancer that often affects people with fair skin. However, melanoma metastasizes far more often than other tumors. When the survival rate falls below 14 percent, the situation becomes considerably more serious. Skin cancer has a survival rate of around 97 percent when discovered early. In light of these results, researchers are examining ways for diagnosing and treating melanoma skin cancer at an early stage. The issue with therapy is that skin cancer must be detected early.

There has been a myriad of research and studies that show various ways to detect skin cancer with the help of image processing and deep neural networks. Though most of the researchers in their paper show how to process the images and about their models, most of the time their approaches are old and outdated. Furthermore, there are also new data-sets with more data and new improved hardware are implemented in the research nowadays.

Thus, we have two objectives in this paper: to enhance existing processes and DNN models. Additionally, our objective is to investigate which diagnostic criteria are most effective in detecting sickness early. In this study, we also intend to improve the accuracy level to a better position. Additionally, we wish to classify skin cancers, whether they are melanoma or another kind of skin cancer. This is accomplished via the use of a machine learning technique known as deep learning, which is meant to predict data from a large number of inputs, such as images or other inputs.

1.3   Research Objective

(a) To concentrated on cutting-edge research on deep learning algorithms for melanoma diagnosis.

(b) To ascertain current research trends, obstacles, and possibilities in the area of melanoma diagnosis.

(c) To conduct an evaluation of the behaviors and outcomes associated with skin cancer prevention.

(d) Finally, the desired system is a less difficult deep learning framework in which accuracy may be raised further while maintaining the architecture's simplicity.

Chapter 2

Literature Review

According to Jinen Daghrir[34] the hair detection preprocessing approach provided in [6] is used in their paper and a few small adjustments were made. There are multiple steps to the suggested hair removal technology. The 2-D derivatives of Gaussian (DOG) of the blue component of the pictures are used to detect hairlines from dermoscopic images. The DOG can detect hairlines in four different directions with pinpoint accuracy. The hair lines are then separated from the background using thresholding and they used the Otsu method. Finally, morphological dilation is used to increase and improve hairlines.

According to the author, Chan et al[4], They used a KNN classifier, which detects an examined lesion's K nearest neighbors and classifies it as malignant or benign based on the features of the majority of neighbors. The features are mapped by a fully connected layer with a SoftMax activation function after the convolution layers have extracted them and the pooling layers have down-sampled them. They were also the least accurate when only looking at the five nearest neighbors. The SVM classifier, on the other hand, surpasses the KNN classifier in terms of efficiency and adaptability. Their suggested approach uses a majority vote to combine the predictions of three distinct techniques. Most systems depend entirely on two physician-defined rules: the ABCD and the Blue-Black rule, both of which have been found to have limits and have proven unsuccessful in some circumstances.

In another paper, according to researcher Kadampur [35] in their paper mentioned Deep learning is used to provide augmented dermatological aid. The average accuracy of diagnosis with this model is 98.89 percent. The machine-assisted diagnosis given here solves the problem of dermatologists in public health being late, inaccurate, and scarcity and authors also said CNNs are the greatest state-of-the-art approaches for picture classification since they essentially mirror the human visual cognition system.

Haenssle et al. [24] used a deep convolutional neural network to categorize a binary diagnostic category of dermatoscopy melanocytic pictures, with 86.6 percent sensitivity and specificity. Dorj et al. created a multiclass classification utilizing ECOC SVM and deep learning CNN in Ref. [23]. The method was to classify multiclass data using ECOC SVM and pre-trained AlexNet Deep Learning CNN. This study claims an average accuracy of 95.1 percent. Their research introduced a platform

that allows someone without a programming background to create complicated deep learning models. It hinted at general techniques and looping patterns in the building of deep learning models, allowing for more flexibility in creating deep learning classifiers. The paper described the capabilities of the DLS tool and explained how to use it to build a Deep Learning Model. The procedure for preparing data utilizing skin cell pictures and testing it in the DLS model to detect cancer cells was described in the paper. The DLS models achieved an AUC of 99.77 percent in detecting cancer cells from the images of cancer cells.

In another paper, Esteva et al.[13] said that automated image categorization of skin lesions is a difficult job due to the fine-grained diversity of skin lesions appearance. Deep CNNs [10] [2] provide prospective for broad and highly varied tasks across multiple fine-graining categories of objects. Here they showed skin lesions classified by employing a single CNN, directly trained picture end-to-end utilizing inputs from pixels and disease labels alone. Their dataset is based on an open-access dermatological repository combination. The technique is meant to use taxonomy to create training classes that are clinically and aesthetically comparable to particular illnesses. During training, a factor of 720 increases the visuals. All network layers are optimized to the same global learning rate of 0.001. CNN outperforms dermatologists using photographic and dermoscopic pictures in the classification of skin cancer.

In another paper, researcher Acosta et al.[37] offered an automated classification approach for cutaneous lesions in digital dermatoscopic pictures to detect the presence of melanoma. This approach consists of two main stages: In the first stage, Mask $R_C NN$ makes a selection of candidates, i.e., it filters regions within the image that may contain a skin lesion. In Stage 2, they crop a bounding box of the image around the region of interest (ROI) they then train, validate and test the model in this stage using the dataset from the ISIC 2017. In the second stage, ResNet152 needs to be adjusted in a process known as training. The performance of the classifier is evaluated using specificity, sensitivity, accuracy, and balance accuracy metrics, in which the reference class is the " malignant" class. They have shown the best performing values for each training scenario. The proposed model achieves an increase in accuracy and balanced accuracy of 3.66 percent and 9.96 percent on the test data set. Unlike previous models, the specificity and sensitivity achieve a high score, which indicates that the model is good for accurate discrimination.

In another reviewed paper we found, NAIR wt al.[38] and his colleagues created a deep learning architecture to handle skin cancer screening which only focuses on it. At 70 percent accuracy, the model was able to categorize the melanoma into a benign or malignant class. Their theory on skin cancer would be that it occurs in two phases: when the illness appears, and when it progresses. In the first phase, data is collected and prepared, followed by the training of the constructed Deep CNN model. Phase 2 is about implementing the idea in real-time and then bringing the display of the results to the GUI. A dense layer is a linear procedure in which the input is coupled to every output by weight. The feature classification function which is retrieved by the convolution layer and downsampled by the pooling layer is responsible for categorizing the features. In every layer of the test picture, a probable illness characteristic is sought out. In the CNN model's updated CNN version,

illness affection may be spotted in photos. In phase 2 they have applied a GUI brings the model to life in a more convincing way.

Chanki Yu [28] utilized coevolutionary neural networks for early detection to do a dermoscopy of acral melanoma, incorporating healthy nevi pictures. Double cross-validation was performed to check system accuracy; half the dataset was utilized for training and a half was used for testing. By using coevolutionary neural networks across the whole data set of the images, their system accuracy was about 83.51 percent and 80.23 percent.

2.1 Importance of Study

The goal of our research study is to improve the quality of existing diagnostic systems by approaching handcrafted features, Gabor and SIFT, and classification methods which are basically deep learning-based features. These methods are aimed at the early detection of melanoma skin cancer. Our approached process provides a better tool for screening and detecting lesions that are considered to be suspicious allowing early treatment and improved survival rate. This study focuses on the proposed methods that are a combination of segmentation methods for skin cancer image to detect the lesion border (edges), followed by handcrafted features like Mean Shift, Gabor and SIFT and other classification methods. To our knowledge, this type of comparison research hasn't been done much in the field of deep learning.

Chapter 3

Background Study

3.1 Basic Idea

Cancer develops when the cells of the body grow abnormally. Cancerous cells may form in almost any area of the body and then spread to other areas.

Skin cancer is an out-of-control proliferation in the external skin layer of aberrant cells in the epidermis produced by unrepeated DNA damage that causes mutations. These alterations lead to a fast multiplication of the skin cells and malignant tumors. Basal cell carcinoma (BCC), scam cell carcinoma (SCC), and melanoma and Merkel cell carcinoma are major kinds of skin cancer (MCC). But of all skin tumors with a poor survival rate, melanoma is deadly.

Melanoma is a kind of cancer that arises from melanocytes, which are skin cells that generate the pigment melanin, which gives skin its tan or brown color. The incidence of melanoma is substantially lower than the incidence of several other skin malignancies. However, because melanoma has a considerably higher risk of spreading, being diagnosed and treated early is critical.

3.1.1 BCC and SCC

The basal cell layer is located at the bottom of the epidermis, known as the basal cell layer. Cells upon on the surface of the skin are constantly replaced with new cells so that the previous cells can be worn away.

The cells convert from epidermal cells to squamous cells as they go higher in the epidermis. The epidermis's upper (outer) layer squamous cells are constantly being shed and replaced.

The epidermal is divided from the deeper layers of the skin by the basal lamina. Once every skin cancer has progressed to surrounding tissue, it often develops to further deeper layers.

3.1.2 Statistics

Studies show that Melanoma kills one person in the United States per hour [36]. In the study of American Cancer Society (ACS) found that more than 207,390 new cases of melanoma are predicted to be diagnosed in the United States in 2021, with around 106,110 of them invasive. The most serious of the three most frequent types of skin cancer is melanoma. Melanomas may be treated and cured if identified and

treated early enough. Melanoma will kill roughly 7,180 people in 2021, according to estimates.

While melanoma is approximately twenty times more prevalent in whites than African Americans, the ratio of African American and white cases is around 20:1. Whites have a lifetime risk of melanoma of 2.6 percent, Blacks have a risk of 0.1 percent, and Hispanics have a risk of 0.6 percent. It is common for a multitude of risk variables to influence each individual's risk. Although melanoma is seen more commonly in males overall, the chance of developing the disease increases with age.

3.1.3   Risk Factors

A cancer risk seems to enhance an individual's chances of developing a disease such as cancer. Cancer risk factors vary according to their kind. Even if we take precautionary precautions against the sun, lifestyle adjustments may help minimize our risk of sunburn as well as other skin damage. We will always have the qualities we were born with, regardless of our upbringing. Despite the existence of multiple risk factors, melanoma may develop.

Some Risk factor are given below:

(a) UV rays are a primary risk factor for melanoma in the majority of cases. UV rays are derived from sunlight. UV rays can also be found in tanning beds and sun lamps. It has an adverse effect on the DNA found in the skin cells, which is involved in cell growth.

(b) A benign pigmented tumor, referred to as a mole, may not cause concern; however, individuals with a high number of moles are much more likely to develop melanoma.

(c) There are some other moles (Atypical moles, Congenital melanocytic moles) which may develop melanoma if not taken proper steps.

(d) The risk is quite higher for fair skin people with light hair (red or blonde), blue or green eyes or skin that burns easily are at higher risk

(e) If first degree relatives (parents, siblings, children) has had melanoma then the chance of getting melanoma is close to 10

(f) Having weak immune system is another reason for caching diseases easily.

(g) Melanoma is more common in older people who are older than 60, but it is also found in people younger than 30 (especially younger women).

(h) In USA, men have a higher rate of melanoma than women.

3.1.4   Detection Process

It is quite difficult to completely eliminate a mole once it has been established, since it generally remains consistent in color, size, and form for many years. While most moles vanish, those that are larger than six millimeters or appear in an odd location, or show signs of melanoma, or have some kind of skin cancer, is a possible symptom of skin cancer or melanoma in progress. The helpful indication is a certain region on your skin that appears completely distinct from all the other areas.

The ABCDE Rule is another guidance to usually detect signs of melanoma.

(a) A represent Asymmetry where one half of a mole is different than other half

(b) B represent Border where the side line is ragged, irregular, blurred or notched.

(c) Color is another factor. If the color of the mole is not same for all the place or have different patches with lighter color (Pink, Red, White or blue)

(d) If the Diameter is more than 6 millimeters than that is another reason though in some cases the size may vary

(e) If the mole starts to Evolve to different size shape or color.

Though these are not the only ones. If a sore doesn't heal up properly, swelling in the sidelines, even change in sensation are also important signs.
It is a fact that a small portion of melanomas starts places other than skins, such as fingernails or toenails, inside the mouth, in the iris part of the eye. So checking with a doctor for slight changes are a must to detect melanoma in the early stages.)
There are 4 stages of melanoma cancer but a medical professional can detect at early stages of melanoma with medical tests like Skin biopsy, Punch biopsy, Optical biopsy, etc. If the cancer cell is already spread through the body, then Fine needle aspiration (FNA) biopsy, Surgical lymph node biopsy, Sentinel lymph node biopsies are needed to get the lab tests.
Using image tests like Chest x-ray, Ultrasound, CT scan, MRI scan, and PET scan melanoma cancer also be detected.

3.2   Initial Problem to Acquire Data

Diagnosis and control of skin lesions have depended primarily on dermatology on direct observation and other non-invasive assessments, avoiding intrusive treatments since they may obliterate lesions and prevent the development of these diseases from being clinically monitored [32]. Computer-supported diagnostic approaches for highly chosen patient groups have shown great sensitivity and may be valuable for professionals in order to decrease the danger of lack of melanomas [26]. One of the non-invasive ways used for diagnosis and surveillance is the use of confocal microscopes for clinical application, which produce clear imagery since they only catch light from the plane of the sample in focus. The confocal microscopy of in vivo reflection produces high-quality pictures in real-time and is used to evaluate different dermatological conditions [32]. Confocal reflectance microscopic may minimize needless excisions without missing instances of melanoma [22]. Research findings in another paper of the same author [20] reveal that dermoscopy is more accurate than visual examination alone; the findings are given in [21] reveal that visual examination is not used to identify melanoma.

3.3 Use of DNN in Medical Melanoma

The development of robust skin lesion categorization systems is based on machine learning (ML) and deep learning (DL) methods. When the other fails, each strategy is better. Their performance is directly linked to the number of research data produced. Machine learning tactics are therefore less potent than those based on deep learning when dealing with huge data sets, but when dealing with small datasets ML is more efficient. Preprocessing, segmentation, extraction, and classification are used in machine learning. It is challenging to extract and choose relevant attributes from a dataset in machine learning. Building an algorithm based on the fusion of DL and ML hopefully will give a better result.

From a technical point of view, several dermatogenic detectors based on Convolutional Neural Networks (CNN), such as the GoogLeNet, AlexNet, ResNet, VGGNet assemblies [25], the R CNN [33], the Mask and Region-based Convolutional Neural Network, the DeeplabV3+ method ensemble [30], and many other CNN structures, making this approach work as one of the most powerful effective feature extraction and classification [17].

Chapter 4

Methodology

4.1 Structure of The Proposed Model

The proposed work started with collecting the raw data and all the data were selected randomly. Later early data processing were applied to make the images more informational for the neural network model. Processed raw images were then filtered
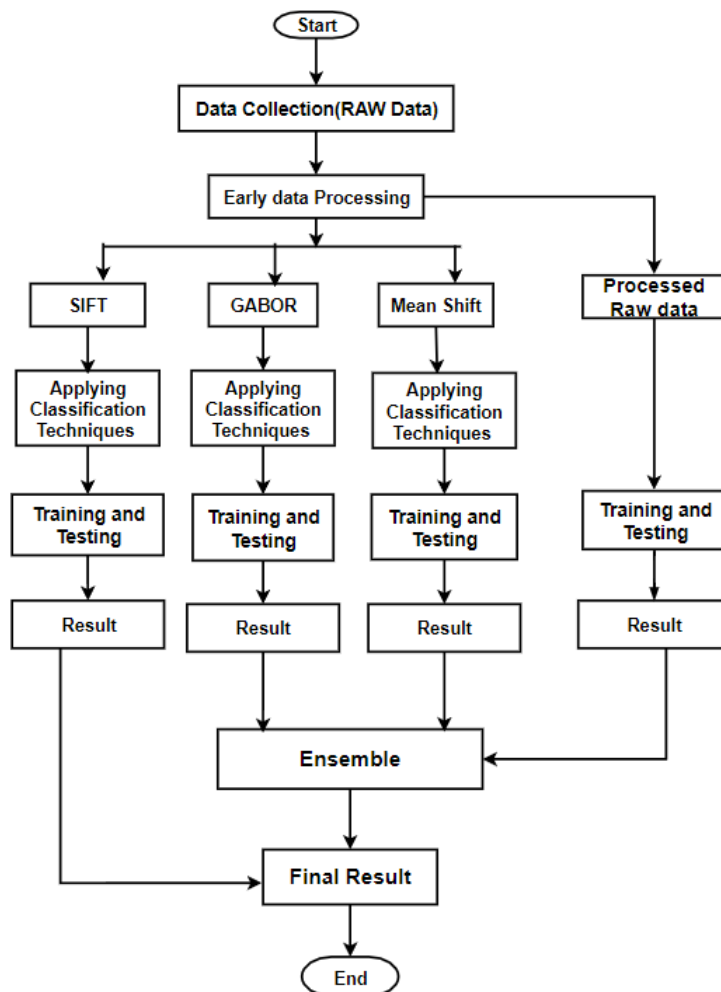

Figure 4.1: Architecture of The Proposed Model

by SIFT, Gabor and Mean Shift feature. After all the data were saved to separate
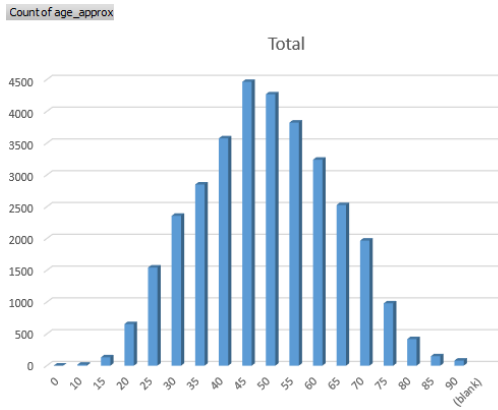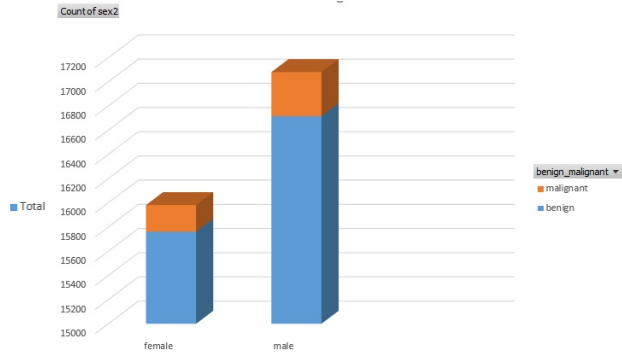
Figure 4.2: Age Graph

Figure 4.3: Sex Graph

dataset, each training and validation process started. With the confirmation of validation, final dataset were uploaded into the neural network and by ensembling the separate results from each feature network, final result was ready.

4.2 Data Collection

The used data are open source data from prestige data-set created by renowned medical hospitals and labs. Two different data-sets were used for this research purposes and those are

(a) ISIC 2020 data-set and

(b) PH2 data-set

4.2.1 ISIC2020 Data-set

The ISIC (Intenational Skin Image Collaboration) 2020 data-set contains more data than ISIC 2017 data-set which is already an advantage because most of the previous work are done with the ISIC 2017 data-set. International Skin Imaging Collaboration (ISIC) created the data-set, which includes images from the Barcelona hospital, the Medical University of Vienna, the Memorial Sloan Kettering Cancer Center, the Melanoma Institute Australia, the University of Queensland, and the University of Athens' medical school. ISIC 2020 data-set contains more than 33 thousand image in which 584 melanoma skin cancer. These images are RGB (red, green, blue) color system and the resolution of the pictures are 1024*1024 pixels. The age graph figure 4.2 shows that most cases are aged between age 45-50. Moreover, we see in figure 4.3 that mostly men are coming with the melanoma cases. Also in figure 4.5 that mostly men are affected by the disease.

In figure 4.4 we see the tumor/mole placement in a human body (which places are more vulnerable) and the number of affected parts in a human body from the data-set.
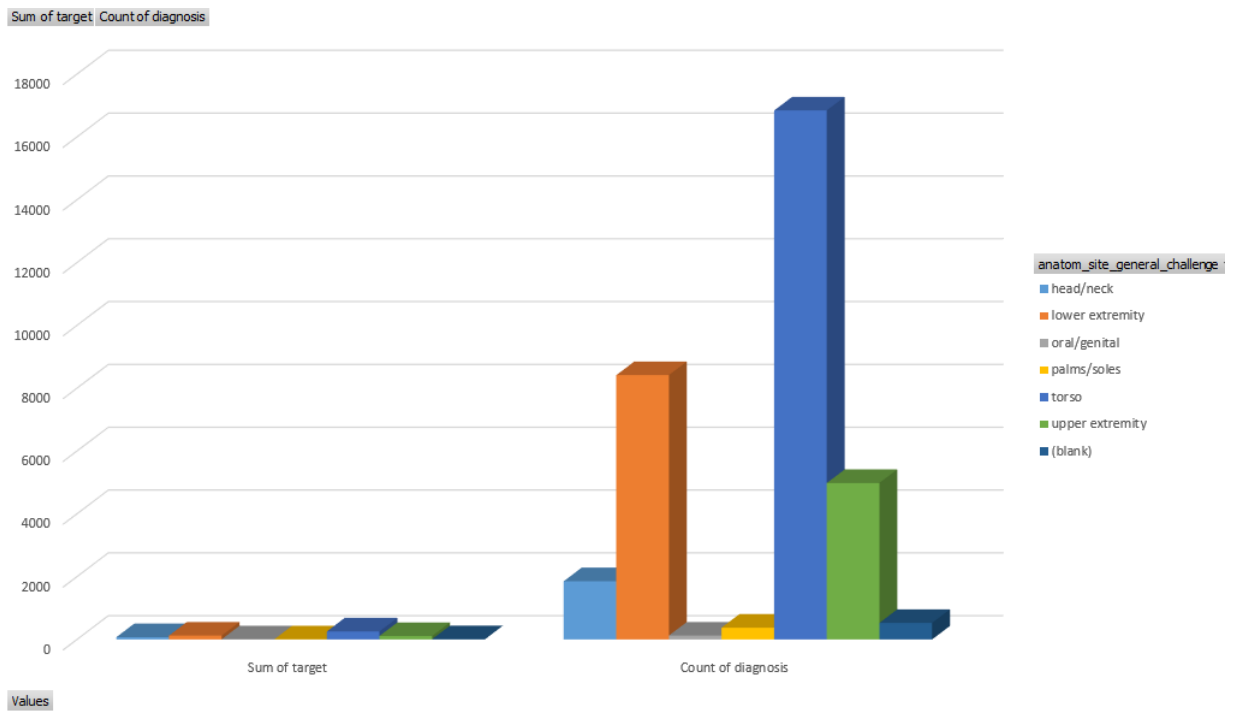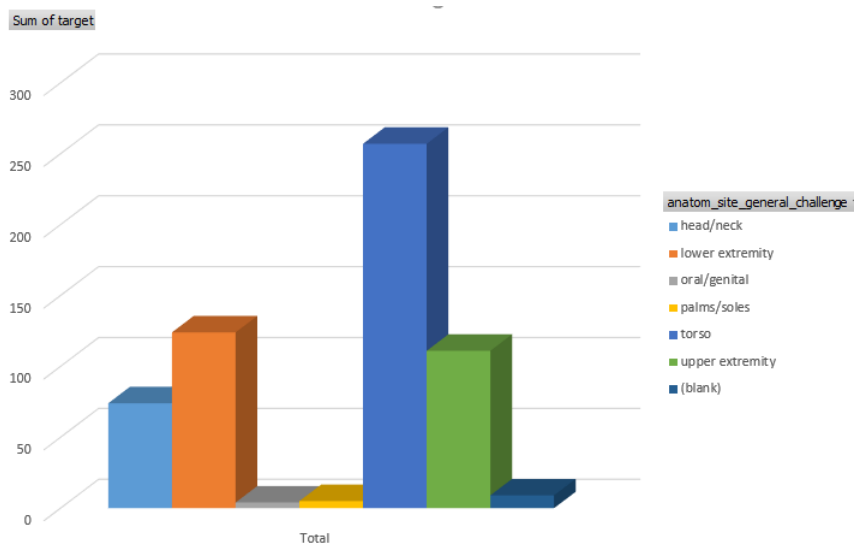
Figure 4.4: All Data with Melanoma and non-Melanoma

Figure 4.5: Different Melanoma Cases

4.2.2   PH2 Data-set

The PH2 data-set contains close to 200 images in which 20 percent are malignant melanoma skin cancer and 80 percent cases are non-melanoma cases (40 percent common nevi, 40 percent atypical nevi). These images are also within RGB (red, green, blue) color system. The picture resolution are 764*575 pixels. The example of the database is given in figure 4.6.
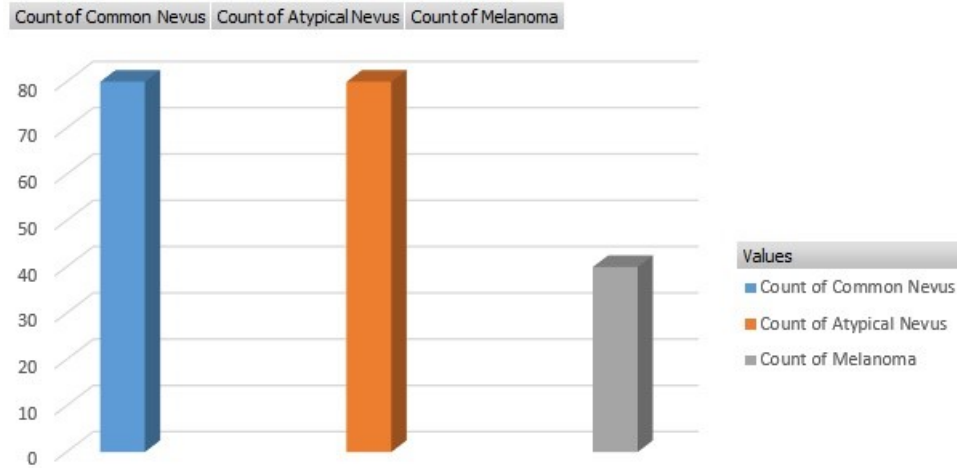
Figure 4.6: PH2 Data-set

4.3   Proposed Features

Deep neural networks (DNNs) have been widely used to automate the evaluate medical pictures for illness diagnosis and to assist human experts by processing large amounts of data rapidly. For optical coherence tomography image-based skin illness categorization, three alternative integration frameworks will be used to mix-handmade and deep features. To integrate the handcrafted feature at the input and fully connected layers using existing networks and to train and test our technique,we will use a publicly available data-set which we shared earlier.

4.3.1   Handcrafted Features

Handcrafted features are descriptions of qualities that are determined via the use of a range of algorithms that take the visual content of the image into account. The form, skeleton, texture, and color of the lesion are taken from the image. To begin, a brief pre-processing step is performed to remove any possible artifacts in the lesion; this aids in segmentation and feature extraction. To begin, divide the original image into two images: an object image and a texture image. The object will be used to segment the lesion and to extract shape, skeleton, and color features.

SIFT Feature

SIFT (Scale-Invariant Feature Transform) is a technique for invariantly defining local parts of a picture, including tiny features. SIFT is often defined as a two-step process

that entails identifying important characteristics (using techniques such as the DoG methodology) and computing SIFT descriptors. The extraction steps are shown in Figure 1. While other technologies such as a Harris detector or a simple regular grid of significant locations may be used in place of the first step, it is crucial to verify the ultimate result is correct regardless. Local gradient histograms at important places were sampled on a 4x4 grid and assigned a unique identification, scale, and prevailing gradient direction within the zone. SIFT features obtained from a rotated and scaled image are invariant in terms of scale and rotation, implying that the results of a rotated and scaled image should roughly mirror SIFT features computed on the original image. It enables more dynamic modifications in classification concerns. Healthy leaves will have a far lower number of key points than symptomatic leaves, since key points are primarily triggered by leaf venation and edges.
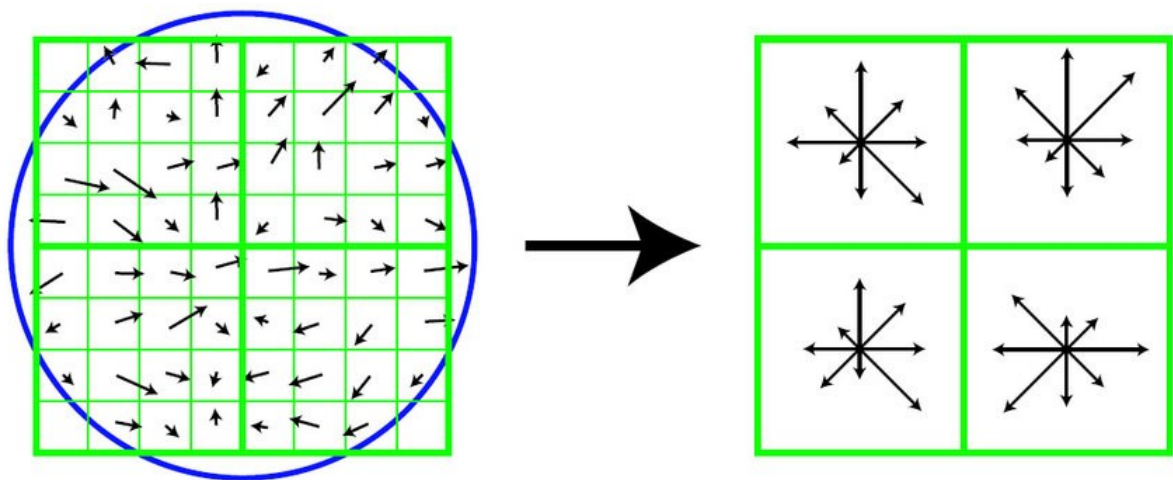
Figure 4.7: SIFT feature extraction

Gabor Feature

In Gabor's filter, an analysis is made to see whether a specified frequency content exists in an image in specified directions around a given point of examination. Many individuals choose to utilize Gabor filters in order to simulate the receptive fields of basic visual cortex cells. Extracted directly from gray scale character pictures, Gabor features are generated via Gabor filters, which were developed specifically to simulate character structures based on statistical information. Deep models used Gabor wavelets for initialization as the input layer. To provide better results on low-quality pictures, an adaptive sigmoid function is used to the outputs of Gabor filters. Feature extraction is a crucial component of classical image recognition techniques in Gabor. To summarize, we conclude that using Gabor features will lead to the incorporation of Gabor filters, a design approach distinct from the one we used before. By manipulating the essence of CNNs, we may apply the Gabor effects on each convolutional layer. Due to this, the DCNNs will also be able to better resist size and orientation changes in feature representations by virtue of their steerable properties.

Figure 4.8: Gabor feature extraction

4.3.2 Mean Shift Feature

The basic principle of the MeanShfit mean shift algorithm, which is a general segmentation algorithm, is: a certain amount for a given samp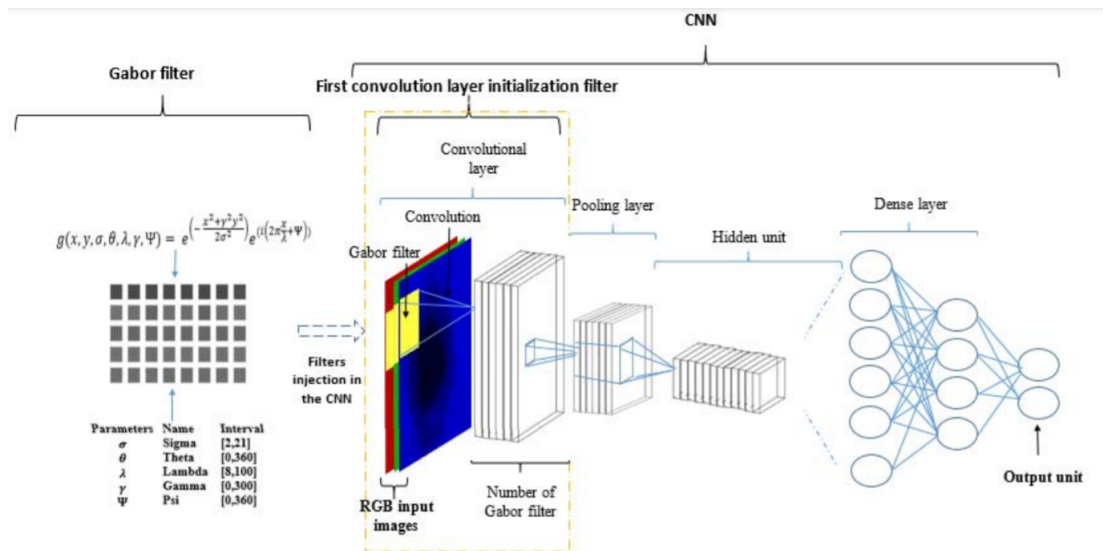le, where the optionally a sample to the sample is designated a center point of the circular area, the circle is obtained samples within the shaped area centroid, i.e., the maximum point density, then the above-described iterative process continues until the final converges. The advantage of the mean shift approach is that it maintains the borders and other discontinuities. By smoothing them with a weighted mean of their neighbors in both space and color, pixels which spread across the discontinuity are systematically eliminated. A two-dimensional grid is frequently used to describe an image using p-dimensional vectors. For each pixel of the image and the set of neighboring pixels within each parameter, the spatial radius and color distance, h_s and h_r is computed. For this neighboring pixel collection, the new spatial center and color mean values are calculated. These updated values will serve as the new focus for the next iteration. This will be repeated until the space and color media settle or the maximum number of iterations is achieved. There are two stages to the mean shift procedure:

(a) Building a probability density in a feature space,

(b) Mapping each point to the maximum (mode) of the density that is closest to it.

4.4 Neural Network Model

There are plenty of DNN models which works effectively. DNN is usually composed of various convolution layers and max or average pooling layers. There are plenty of pre-trained models which are good, well known, and prepared with billions of pictures. In Fig. 6, all the accuracy of the pre-trained model can be seen. Here we can see all the recent models and how well they perform. From all these pre-trained

model we will use EfficientNet model, Specifically the EfficientNet-B4. All efficient net models have high accuracy and using the B7 version might give higher accuracy but we will use B4 as the desired accuracy can be obtainable. In future work, for better comparison B7 model can be used. There are some other ground-breaking architectures are AlexNet, VGG-16 and ResNet, etc. EfficientNet-B4 can achieve
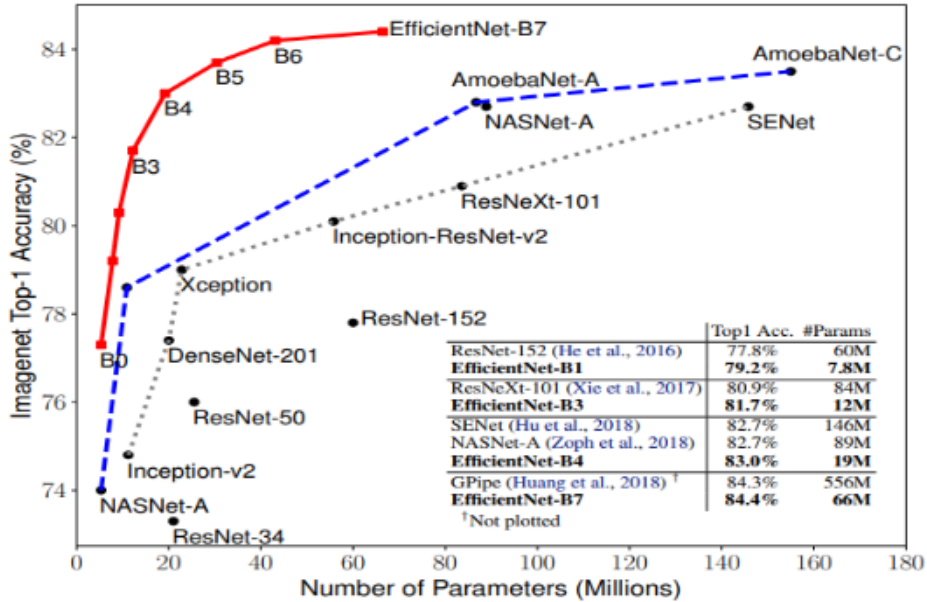
Figure 4.9: Some Popular Pre-trained Model

better ImageNet accuracy and it is also very smaller compared to other models. Its main building block is mobile inverted bottleneck MBConv. For this the model can use shortcut directly between the bottlenecks which works with very little number of channels. In the below table we can see the structured layers of EfficientNet-B4. We also have dropout layer and predicted layer. The dropout rate of the used layer is 0.3. The model also have 11,310,636 trainable parameters, other 87,303 parameters are non-trainable.

Table 4.1: Layers and Parameters of EfficientNet-B4

Layer(type)	Output Shape	Parameters
Input Layer	[(None,300,300,3)]	0
Efficientnet-B4	(None,10,10,1536)	10783535
Flatten	(None, 153600)	0
Dropout	[(None,300,300,3)]	0
Dense	(None, 4)	614404
Total Parameter		11,387,939
Trainable Parameter		11,310,636
Non-trainable Parameter		87,303

The modules and architecture the EfficientNet-B4 are given below:

Figure 4.10: Active module of EfficientNet-B4

Figure 4.11: Architecture of EfficientNet-B4

4.4.1 EfficientNet-B4 Layer Details

There are few layers which was used in our neural network such as: Conv2D, Depthwise Conv2D, Pooling, Dropout, Flatten and Dense.

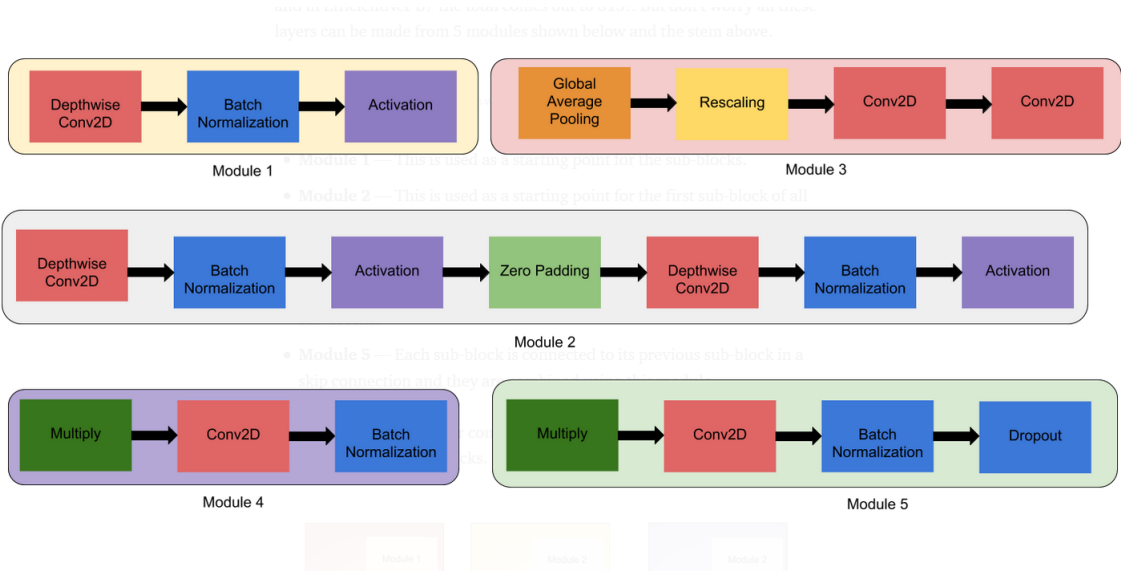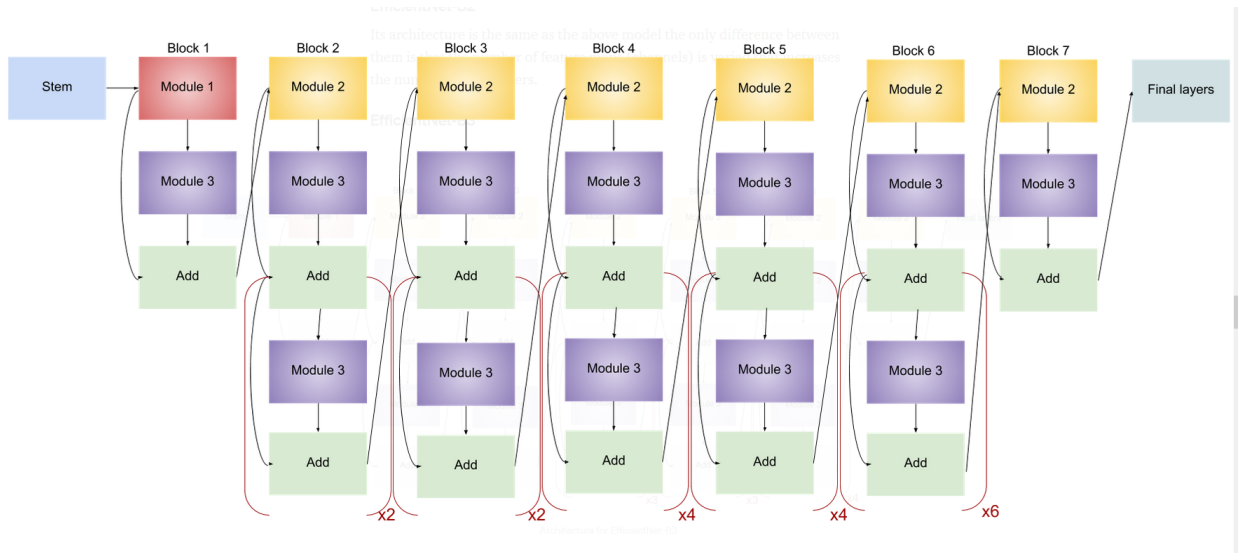(a) *Conv2D:* It was needed in the beginning. The option Conv2D defines the number of filters to learn from the convolutionary layer. Layers are used in the early stages of a network architecture. The layer closest to the actual input picture, for example, it is less used to learn revolutionary filters and the layers further in the network, i.e., closer to the output predictions. The intermediate layers of Conv2D are more filters than the early levels of Conv2D, but less filters than the close layers.
The basic idea of the Conv2D layers is to convert the input image into an abstract image. This format may be used to classify highly linked layers. The multidimensional feature map generated by the last Conv2D layer must be converted to a one-dimensional format since dense layers may take only one-dimensional entrance. Because of flatten layer this is feasible.

(b) *Dropout:* Dropout means that a set of neurons chosen at random are ignored during training (i.e., neurons). Dropout is a method to tackle these two problems. It avoids over fitting and permits the use of many alternative designs in the neural networks. In addition, "Dropout" refers to portions that are no longer active in the neural network (both hidden and public). Deleting a unit means deleting it from the network, including its incoming and exit connections, for a specified amount of time. The units to be dropped are randomly selected. Individual units with a constant probability p, which may be determined by a test set or set to 0.5, which appears almost ideal for a range of networks and activities, are maintained. The method used to extract a "thinned" neural network is known as dropout from a wider network. All units that have avoided the drop out of the thinning network (Figure 4.12). A collection of neural networks of 2n thinned may be termed a neural network of n-units.
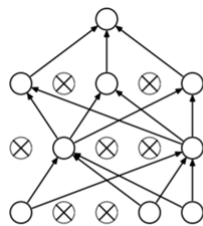
Figure 4.12: After Applying Dropout

(c) *Pooling:* This block lowers the amount of data, the number of parameters and the calculation required by over-fitting. 2D max pooling may be compared in order to reduce image size. The pooling block 2D Max constitutes the best method of pooling. A tensor that reduces the number of parameters and calculations that are required to avoid over-fitting and produces a smaller tensor than its input. The 2D Max panels locate the max in every window by sliding the supplied data across a rectangle

(window). The size of the window depends on the horizontal and vertical pooling factors, and how many steps the window takes on the horizontal and vertical steps. Max body blocks are then added to a large percentage of the original picture after one or more convolutionary blocks to support the collection of inner convolutionary blocks. Before pooling, a 4x4 portion of the tensor affects a 2x2 filter. Once convolution blocks are considered detectors of that feature, Max pooling only retains the highest value for a feature within the pooling rights. There is a distinct treatment of each channel and therefore each feature.

(d) *Flatten:* A tensor flatten operation is a common operation inside convolutional neural network. This is because convolutional layer outputs that are passed to fully connected layers must be flatted out before the fully connected layer will accept the input. To flatten a tensor, we need to have at least two axes. This makes it so that we are starting with something that is not already flat. FTS contains similar properties from both ReLU and Swish, and a threshold T parameter is attached, which could improve the classification accuracy. Mathematically, ReLU is defined as [8]

$$ReLU(x) = \{x, x \geq 0 \& 0, x < 0 \qquad (4.1)$$

ReLU has introduced to tackle the issues such as gradient vanishing/exploding and squashing problems by the Sigmoid activation functions in deep neural networks [18][27]. Formally, Sigmoid activation function can be defined as [1]

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \qquad (4.2)$$

To construct the FTS activation function, this study first amends the original ReLU function by multiplying its linear identity part (when x $\geq$ 0) with Sigmoid activation function. Where the idea can be simply expressed by FTS(x) = ReLU(x) * Sigmoid(x). With this amendment, this study has noticed that the FTS at x $\geq$ 0 has a similar property to a recent activation function introduced by Google Brain called "Swish" [19]. However, the derivative of Swish has a large portion of the no sparse property thus probably trigger higher computational complexity. Meanwhile, FTS retains the hard zero property at the other side as in ReLU which eventually deactivated most of the neurons when during both forward and backward propagation.

(e) *Dense:* The dense layer is a highly integrated neural network layer, where each neuron receives input from all preceding layers. Modeling methods indicate that the thick layer is the most often utilized one. Matrix-vector multiplication occurs within the backgrounds as a result of the thick layer. A back propagation algorithm will train and update the matrix's variables, which are really parameters. The dense layer serves mainly to increase the dimensionality of the vector. It also applied in vector operations like rotation, scaling, and translation to the underlying matrices. It uses sequential keras to build a model when dense is enabled. This time, EfficientNet-B4 is an output.

4.5 Feature Integration Network

This research has three separate integration architectures to merge handcrafted characteristics with profound learning. The structures of the suggested networks include,

    (a) Early integration network,

    (b) Late integration network and

    (c) Full integration network.

The handmade original function map was first added as an early integration feature image. The late integration technique integrated the FC layers' output maps and used Softmax to complete the final classification. The complete approach to integration was distinct from the other two and may be characterized as a synthesis of the two.

4.5.1 Early Integration Network

The first stage of integration is the extraction of relevant attributes from both perspectives. The integration feature is the process through which extracted features are combined into a single set of features. The storage of the input feature vectors in a single vector is one of the simplest ways for feature integration. A modeling approach is then used to the combined feature vector to provide the final option or output. Early integration allows for the recognition of cross-modal linkages at the feature level, and particular modeling methods are required to reduce costs and complexity in comparison to other melting approaches that need more modeling units [33, 34]. Primary component analysis (PCA)and efficient net B4 is used in first stage. Additionally, since various modalities have varying sensing and processing speeds, they may be asynchronous. Due to the requirement that the functional vectors be from the same time period, this issue should be approached in stages [33]. While feature integration is the most often utilized technique for early integration, one modality may be utilized to perform a specific startup or preparation while the balance of the work is completed in the other modality. During the first stages of
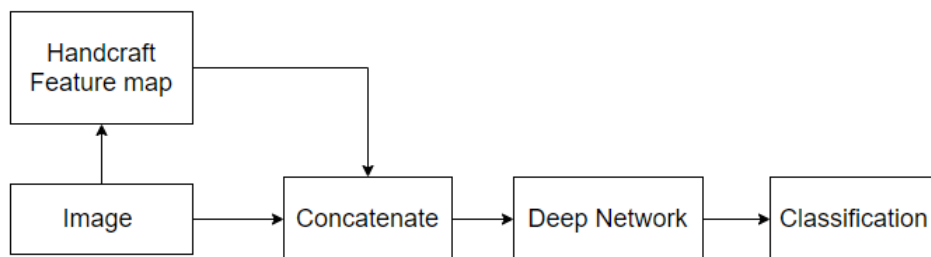
Figure 4.13: Early Integration Model

integration, both authentic and generated feature photographs were used. The original image was combined with those from SIFT/Gabor and utilized as a network input. In this post, we attempted to combine both SIFT image characteristics and the original image at an angle of around 90 or 180 degrees, to create an RGB color image and utilize it as a deep network input. Gabor use the same processing as

21

SIFT to combine real and fictitious elements and generate input with the original image. In the early integration phases, classic DNNs like VGG-16, efficient net B4 have been deployed.

4.5.2 Late Integration Network

This technique employs a distinct modeling approach for each modality, which takes the attributes of one modality as an input and creates an output selection. The decision integration unit then takes all of these and merges them to get the final output. The most fundamental processes used at this step are weighting and summarization [33]. Information from the original and feature photos was integrated in a late integration network at the conclusion of the deep network. Two distinct deep networks were utilized to extract features from the original and feature photos. The outputs of both networks were concatenated 2D vectors which were then entirely connected to the layer of categorization. Comparing the late integration network with the early integration network requires about two network characteristics for late integration. The outputs of the modeling procedures have the same form, which makes it simpler to combine them than to combine feature vectors. Compared to early integration techniques, asynchronous techniques are easy at decision level, and the system with the number of modalities may be scaled down. Another advantage of this technique is that it enables customized approach to be used for each technique.
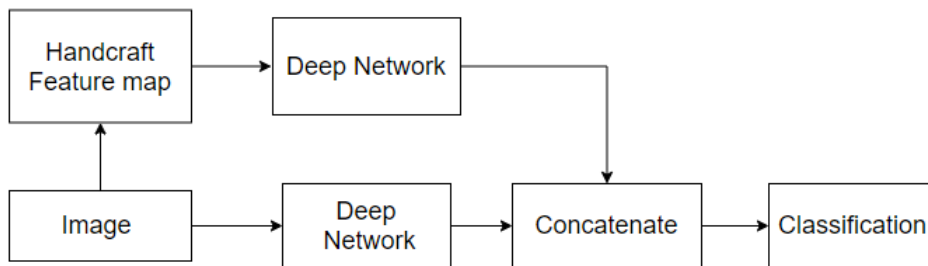
Figure 4.14: Late Integration Model

4.5.3 Full integration network

The original and produced images were changed after each convolutionary block of the communication based network. The highest results from all the processed network are combined in another neural network. The ensembled model provides final result with all the previous result combined and give the an average value as a result. There are chances to customize the model and the model is customized in a way to provide the average highest values as an output.
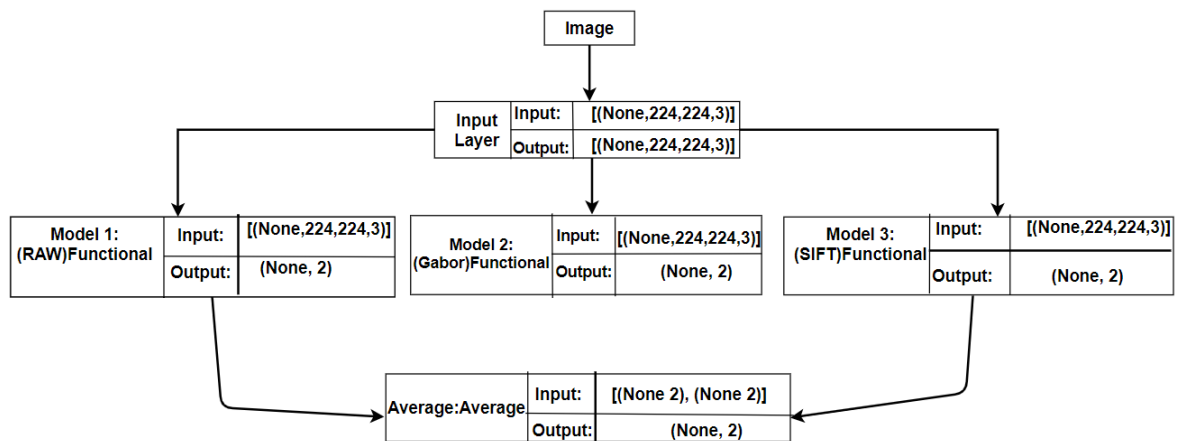
Figure 4.15: Full Integration Model

Chapter 5

Implementation

Pre-processing is a technique that has been used for better image quality and clarity. SIFT, Gabor also Mean Shift are used for late processing part. For early processing, to eliminate unnecessary defects and other characteristics of different images which is important for later processing part where main features are going to be used.

5.1 Early Processing Methods

Firstly 2994 raw images of malignant melanoma and non-malignant melanoma were selected both from described data sets. Data selection was done randomly so that the result would not become biased, also it was ensured that the selected data are efficient enough for the research work. These raw images were then pushed through various early processing phases. All these phases are described below,

(a) *Initializing:* Initial work was to make sure the resolution of every image is the same. Because two individual datasets are used in this research and both datasets have different image resolutions, optimized resolutions were selected 224x224 for every image. selecting this resolution made the trained model work more efficiently. Then the data were converted into their RGB values and saved in an array.

(b) *Creating different sets:* Among the selected 2994 raw data, 70 percent data (2094 images) was chosen as training data and 15 percent (450 images) for validation, and the remaining data (450 images) for testing purposes. More data was selected to train and validate because, after training, it can be a surety to have to check if the output results are giving biased results or not.

(c) *Normalization:* The image RGB values have been divided properly to provide values from 1 to 0.

(d) *Magnitude:* The brightness of a picture is a representation of how quickly it varies. It contains two distinct bits of information. The gradient's magnitude determines how efficiently the picture is shifting, whereas the gradient's direction indicates which direction the picture is moving the most swiftly. Consider a picture as a landscape where we are provided a depth rather than a brightness at each spot. The direction of the gradient would be upward to any point on the area. Whenever a brief step upward

is taken, the size of the gradient indicates that how rapidly the position grows. Since the depth has a magnitude and a direction, it would be only logical to demonstrate it as a vector. The magnitude of the gradient is determined by the length of this vector, whereas the gradient orientation is specified by its axis. finally, express the gradient with a separate vector at each picture position since the gradient may vary.

(e) *Enhanced brightness:* An image must have the proper brightness for easy viewing. Brightness refers to the overall lightness or darkness of the image. Brightness can be simply increased or decreased by simple addition or subtraction, to the image matrix. First, calculate the mean brightness. To do that min brightness was selected as 0 and max as 1 and if the brightness is close to 0 then just separated those and later enhanced those image's quality and brightness.
Since we already know, that each image has a matrix at its behind that contains the pixel values. At this point, simply add a value to each of the matrix values of the image until it gets perfect brightness. For a better training process, we simply calculated the RGB values, and then, all we did is just add a value to the trainee's picture matrix. After adding some value to the picture, an improved image was ready that works well for the proposed training process.

(f) *Image energy:* Image energy is basically showing the energy values within the images.

(g) *Energy density:* It's the rate of change in the brightness or magnitude of the pixels over local areas. This is especially true for edges of the things inside the image and because of the nature of compression, these areas are the hardest to compress and therefore it's a solid guess that these are more important, they are often edges or quick gradients. These are the different contexts but they refer to the same thing.

(h) *PCA:* Large data sets becoming more and more frequent, yet it may be difficult to comprehend them. The PCA technology enhances both accuracy and the risk of data loss to reduce the complexity of larger datasets. It accomplishes this by generating additional statistically independent variables which increase differences in a certain sequence. A variety of methods were proposed in order to accomplish this goal, but one of the oldest and most frequently utilized component analysis (PCA).
The PCA technique uses example pictures to generate an Image Component, an average picture, and finally an estimate of parameters for all of the input images. As many differences as feasible are highlighted in data, PCA showing the majority of the remaining variability as appropriate for each subsequent element. It is a technology for detecting and identifying data patterns in a way that maximizes their similarities and contrasts. PCA identifies and compresses significant patterns in data collection by decreasing features and therefore minimizing data loss.

- PCA Algorithm: PCA is an image collection assessment method. PCA is utilized for an evaluation of a single-colored image in this case. This picture consists of several columns. In the fundamental PCA

technique, each row is utilized as an input. If the data consistency is strong, the HOTELLING transform is carried out into the rows in an image. For the picture to be separated by the required number of rows of PCA methods, we must first identify the difference between rows using a methodology. The variance value is computed for the best outcome between the rows of the image matrix.

$$cov(a, b) = E[(a - (A \cup a))(b - (A \cup b))] \qquad (5.1)$$

For this which found out similarity between the rows of an image matrix:

$$Similarity(i) = cov((i)row; (i+1)row) here, i = 1, 2, 3..., n-1 \quad (5.2)$$

(a) Melanoma

(b) Melanoma

(c) Non Melanoma
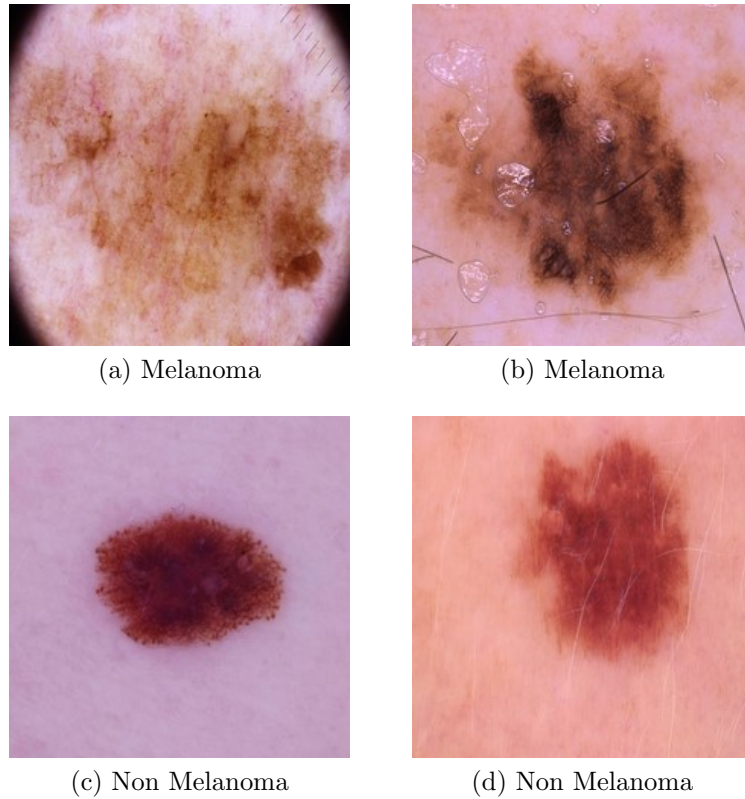
(d) Non Melanoma

Figure 5.1: Early Processed Original Images

(i) *Arranging variable:* Variable for melanoma and non-melanoma cells are converted into binary values for better understating where melanoma has been labelled as 1 and non-melanoma as 0 value.

(j) *Data augmentation:* To provide better information and understanding it was made sure all data are re-scaled properly. The main goal here to add more information to the data so that it helps the research, the information might exists already but by updating those elements to a certain extent which basically helps the neural network to train and work properly.

- The rotation range is set by 45 degrees (-45, 45).
- Width shift range is set by 0.2.

- Height shift range is set by 0.2
- Shear range is also set to 0.2
- Enhanced zoom range by 20 percent
- Horizontal and vertical flip is also done.
- fill mode set to nearest so that if a pixel is missing any value then it can fix it up from the nearest pixels.

After early processing is done, the images we got are in the figure 5.1,

5.2 Late Processing Methods (Applying filters)

After early pre-processing is finished handcrafted filters were applied which are Gabor, SIFT and Mean Shift and created separated image files.

5.2.1 Gabor filter

To begin, specifics data (pictures) were checked while getting Gabor effects using Gabor filters. First clarity was checked and then enhanced the color of the images. pixelated values of the images were transferred to the NumPy array. With the energy density from the previous early processing stage, bandwidth was calculated. frequency was also calculated and saved on another array. For theta, another NumPy array was used and got the result by using this value (0, np.pi, np.pi/6). With all

(a) Gabor Filtered Melanoma   (b) Gabor Filtered Melanoma
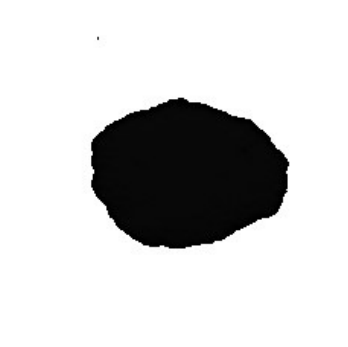
(c) Gabor Filtered Non Melanoma   (d) Gabor Filtered Non Melanoma

Figure 5.2: Gabor Filtered Images

the arrays (image, frequency, bandwidth, theta). the newly created magnitude for the images was saved. On the newfound magnitude data, Gaussian smoothing was applied. To do that, firstly sigma values were calculated which was chosen to be half of the applied frequency. After using Gaussian smoothing reshaping was applied to the images by applying PCA. After PCA dimensionally reduced images with ready with Gabor filter. Then Gabor filters were applied to all the images. The featured images are given in figure 5.2.

5.2.2   SIFT filter

After taking the original images we reshaped the images and enhanced the pixels. We applied the features to the images and concatenated the feature with the layers. Then we reshaped the feature array and saved it. We saved the label and features to the data frame and shuffled the data. Then we converted the features and the information of the data to an array and saved it. that is how we got the SIFT featured data ready with 224x224 pixels. The featured images are shown in figure 5.3.

(a) SIFT Filtered Melanoma   (b) SIFT Filtered Melanoma

(c)   SIFT   Filtered   Non   (d)   SIFT   Filtered   Non
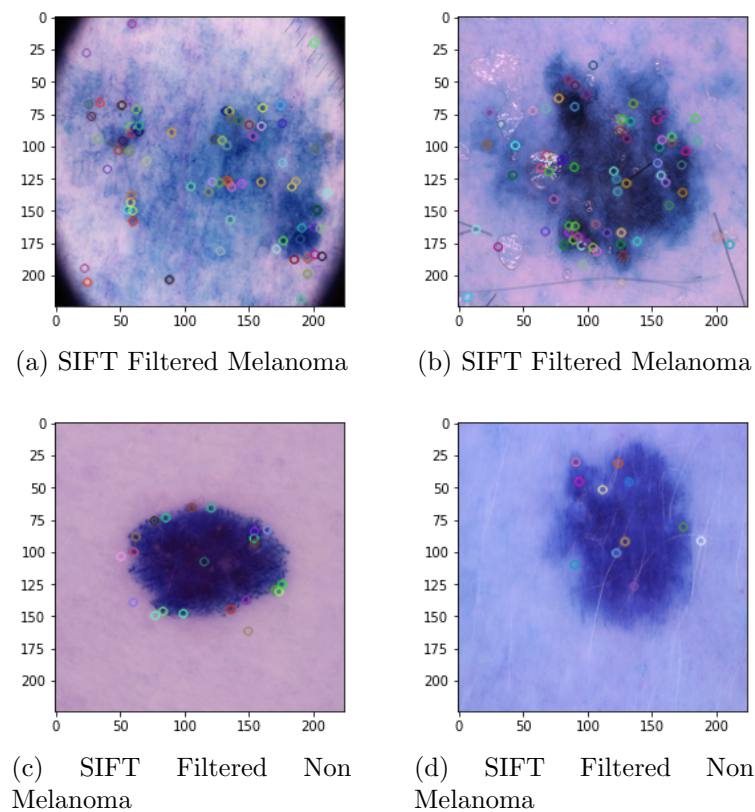Melanoma   Melanoma

Figure 5.3: SIFT Filtered Images

5.2.3   Mean Shift filter

Firstly, cv2.imshow was used to display an image in a window. Automatically adjusts the window to the image size. Then the images are stored in SRC using cv2.imread. Following that, a demo build of the shift demo function was created.

The images were submitted by using cv2.pyrMeanShiftFiltering to filter the image using an integrated Python approach. Color graph segmentation is a feature of the middle shift technique, while the related function is the pyrMeanShiftFiltering. This feature does not split the picture, but instead divides the color of the image smoothly, which may be divided into the same color, but also smooth color details, sections dissolved, and smaller colors. Instead of the splitting suffix "segment," the OpenCV Filter Suffix "Filter" is thus utilized. the utilization of the cv2.waitkey function was done then. This method would not function properly without cv2.imshow. The picture will be shown 6 seconds before it is shut off if the wait time is set to 6000. The image will be shown indeterminately until the escape key is pressed. Now the Mean shift filter is applied.
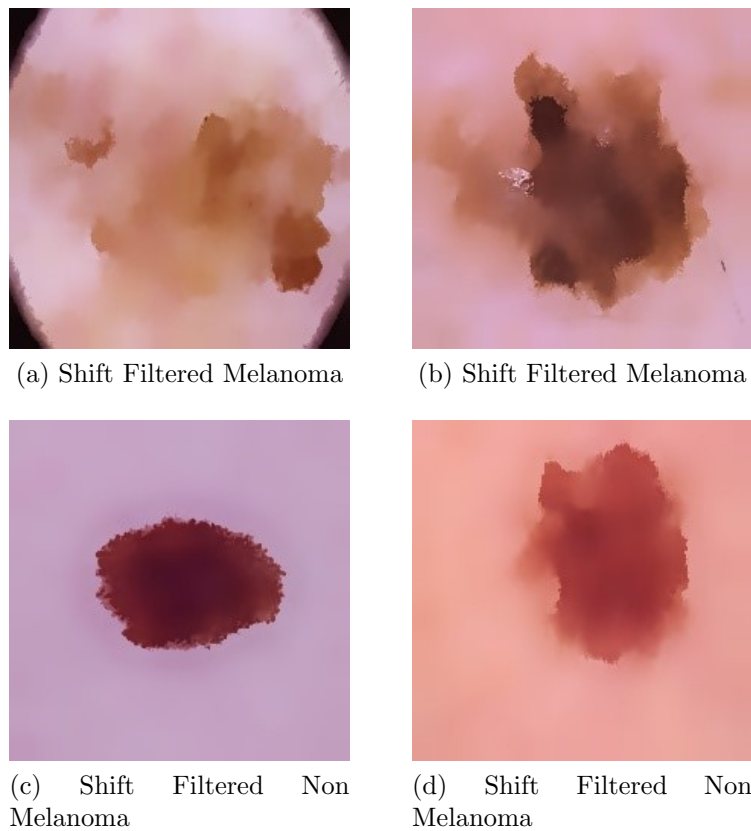

(a) Shift Filtered Melanoma

(b) Shift Filtered Melanoma

(c) Shift Filtered Non Melanoma

(d) Shift Filtered Non Melanoma

Figure 5.4: Mean Shift Filtered Images

5.3   Training Process

The pre-processed dataset has 2994 images and all these images are re-shaped and re-scaled to 224 x 224 pixels. We focus well enough on the early data processing and late data processing with various features to expand the variation in the robustness of the data for the research work. Data augmentation with a 45-degree rotation range flipped and zoomed for additional increased data was done. Form all the data around Seventy(70) percent of data (2094 images) have been selected randomly for training purposes. The neural network has been trained with a sequential technique. Some basic parameters rates are given below:

- Batch Size: The used batch size is 10. So, every chunk of the dataset has around 300 data. these ten 3x3 kernel filters have acted as a feature map with a pixel size of 224 x 224.

- Learning Rate: After each stage of optimization, the learning rate shows how large or little weight changes are. The weight of the neural network changes considerably if you select a high learning rate. The initial learning rate is 0.0001 which is set by EfficientNet-B4 which is very low and also effective. The value of the learning rate will increase after the training phase has started.

- Epoch: In the sample training dataset has the opportunity to modify the parameters of the internal model once each period. A period has one or more lots. The number of times is typically large, from centuries to thousands, allowing the learning process to continue to an acceptable level until the inaccuracy of the model has been minimized. An epoch and an iteration are sometimes used interchangeably. In the dataset, the value of epoch is 30 and the deep neural approach tried 30 iterations every time while training the data.

- Dropout Rate: Dropout rate is basically shows the data of the over-fitting. Dropout is a simple way to reduce over-fitting in the neural network.

$$DropoutRate = \frac{LearningRate}{Epoch} \tag{5.3}$$

- Early Stopping: Early Stopping is an extra feature to identify the validation loss. If the Epoch parameter starts to give a value that might reduce the accuracy rate then the early stopping parameter stops the value accuracy count and shows the best accuracy value till the end.

These Parameters are fixed for the training process of every filter that was used in this research, though all the parameters are the same, they are flexible for different filters such as Epoch and early stopping.

5.3.1 Raw Image Training Process

For the raw image original data were used in the dataset. The first selected 2994 data were gathered together and uploaded into the neural network. For clarity and re-scaling purposes the augmentation process was applied and when all data were ready training process started. 2094 data were selected for training the DNN model. 30 epoch was done every time but because of the early stopping parameter it can be seen that the accuracy stopped increasing after the 25th epoch step.

5.3.2 Gabor Filter Training Process

With the late processing methods, ready Gabor filtered images are collected. Those images were then uploaded into the DNN model. Firstly data augmentation method was applied and re-scaled was done to those images. The early stopping method stopped registering the value of accuracy from the 18th epoch.

5.3.3 SIFT filter Training Process

Using the dataset from SIFT filter, the data augmentation method was used after uploading to the network model. A custom SIFT filter was build and used here and in the training process 5 folds were created and the data were separated into these folds and training began. The model did not follow the early stopping method.

5.3.4 Mean Shift Filter Training Process

Using the same protocol, the dataset created from the Mean Shift filter is uploaded to the custom EfficientNet-B4 model. After re-augmenting the data the main training process was started. From the 26th epoch step, the accuracy value remains the same.

5.4 Validation

For validation purposes, previously selected 450 images which are more than 15 percent of the dataset. The selection process of these data was random. These randomly chosen validation data have been used in order to avoid the disruption of the noise or random fluctuation it learned to form the training phase. Furthermore, validation helps to dispense the direction and instruction to the network model. Its result also represents that the final training process is successful.
Because we have different filterers, these data were also faltered using each filtration method and used in separate models after the training process is done.

5.5 Performance Measure

To evaluate our research study, we need to measure the performance whether we are getting a better result or not, some measurement procedure has followed. Those procedures are recall or sensitivity, precision, specificity, accuracy, and F score.
The recall or sensitivity basically calculated the true positive values are correct or not. the equation is presented by equation (5.4) [12].

$$RecallorSensitivity = \frac{TP}{TP + FN} \tag{5.4}$$

The precision determines the value of the actual positive from the true positive value. It also represents how accurate our neural model's predictions are. the equation is represented by equation (5.5). [14].

$$Precision = \frac{TP}{TP + FP} \tag{5.5}$$

Accuracy equation (5.6) describe the accuracy value [16].

$$Accuracy = \frac{TP + TN)}{TP + TN + FP + FN} \tag{5.6}$$

F score (F1 value) shown in the equation (5.7) is basically the harmonic mean of both precision and recall [15]. It is used to evaluate binary classification systems of

positive and negative. It creates a single matrix that determines the best result for the used value of precision and recall.

$$F1Measure = \frac{2 * (Sensitivity * Precision)}{Sensitivity + Precision} \tag{5.7}$$

In these equations TP represents the number of true positive, TN represents the number of true negative and for FP and FN, the number of false-positive and false-negative respectively.

Chapter 6

Result And Evaluations

The neural model has been trained with 2094 images in the training process and 450 images in the validation process. The images we used to train the neural model and make it ready are three different handcrafted filter images and also raw images. Also, the variable parameters and layers of the neural network remain the same during all these testing and training phases. The efficiency of the neural network and the progression of how well it provides the results all are determined by the graph of the result and the confusion matrix of melanoma and non-melanoma's positive and negative value.

6.1 Early Integration Results

At the early integration process we have four different results to show, as the early integration includes raw data, Mean Shift featured data, Gabor featured data and SIFT featured data.

6.1.1 Raw Image

Before applying the test data into the trained model and training accuracy came as output with validation accuracy, training loss, and validation loss. If the training loss is reduced then the accuracy gets high and at a certain point, training loss remains the same.
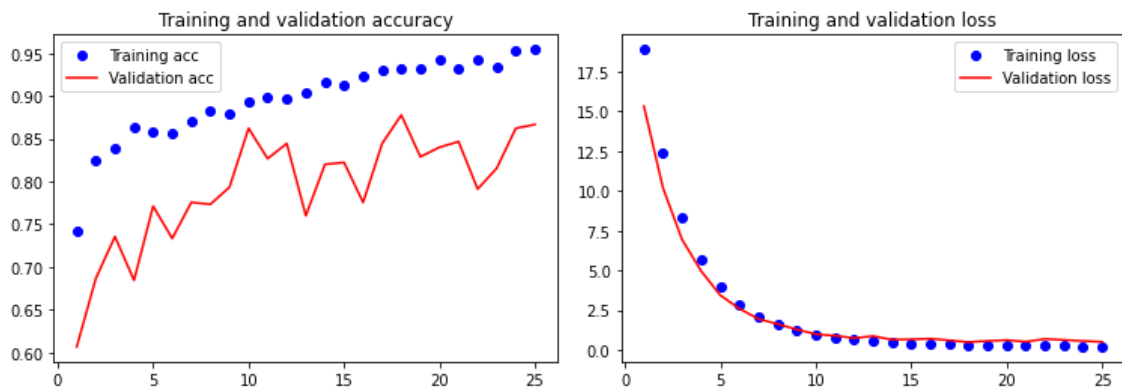

Figure 6.1: Training and Validation Accuracy (Raw images) Figure 6.2: Training and Validation Loss (Raw images)

From the graph, it is clear that the validation testing work is giving different results so the raw process is working perfectly. Then the neural network was ready to apply the testing data of the raw image and the result is given below in the table,

Table 6.1: Raw Image Result

Parameters	Result(percentage)
Accuracy	90
Precision	91.7
Recall	88
F-Measure	89.8

We also created a confusion matrix to see the true and false positive also true and false negative values. The top left of the confusion matrix gives true positive value and the top right part shows true negative, which means from melanoma data how many the model can accurately determine and how many it can not. Also in the bottom left it shows how much data the model can not determine accurately and on the bottom right how many it can determine accurately from the non-melanoma dataset.
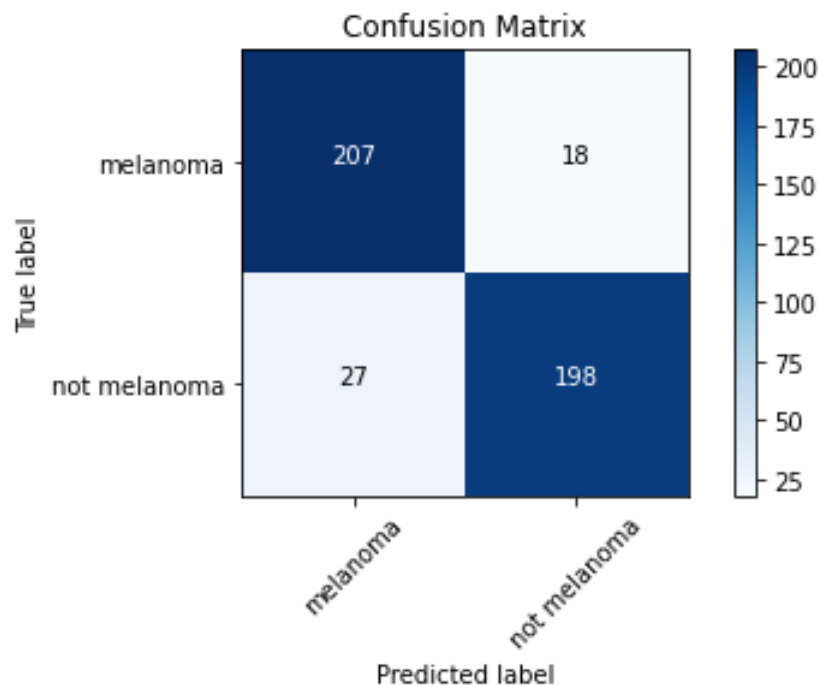

Figure 6.3: Confusion Matrix (Raw Images)

From this, we can see that the model can detect 207 melanoma and 198 non-melanoma cases accurately.
At last, for better recognizing where the cancer cell can situate LIME framework was used. This framework mainly shows green and red marks. Green resembles those parts where the model has predicted correctly and red shows the incorrect portion. If an image does not have any red part but has the green part that means that the image does not have that much incorrect feature. After lime feature is applied the picture shows green colored places, where the model found cancer cells.
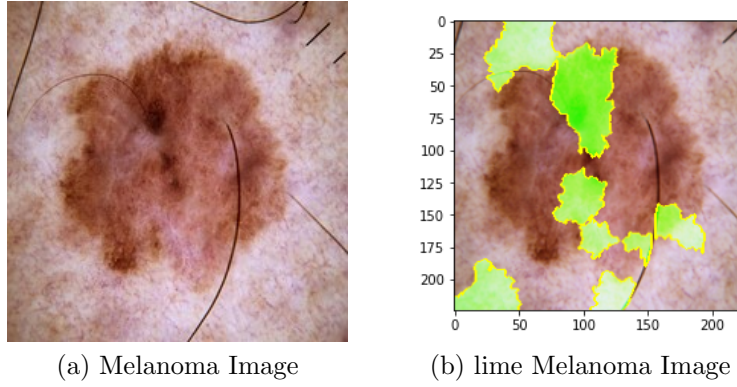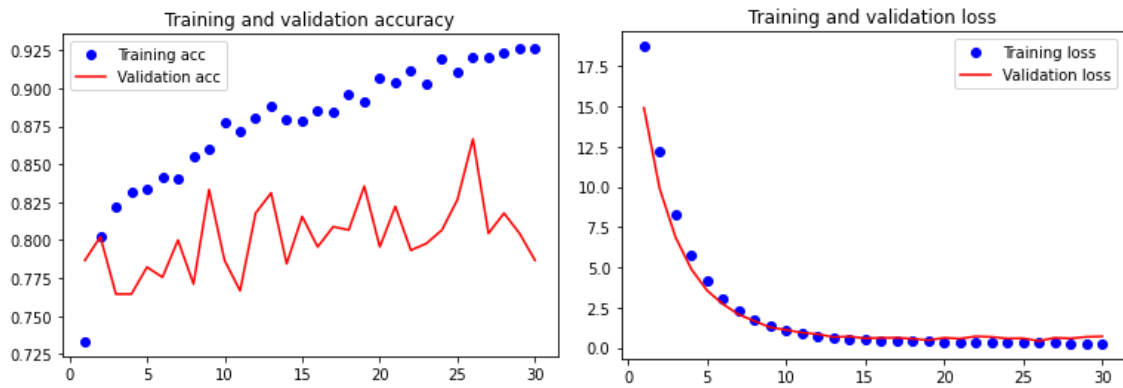
(a) Melanoma Image   (b) lime Melanoma Image

Figure 6.4: Lime Featured Melanoma images

Figure 6.5: Training and Validation Accu-Figure 6.6: Training and Validation Loss
racy (Mean Shift)   (Mean Shift)

6.1.2 Mean Shift Result

Form figure 6.5 and 6.6 we can see the mean shift feature's training and validation
accuracy and loss.

Now we apply the testing data of mean shift featured image and the results are
given below in the table,

Table 6.2: Mean Shift Feature Result

Parameters	Result(percentage)
Accuracy	87.8
Precision	86
Recall	90.2
F-Measure	88.1

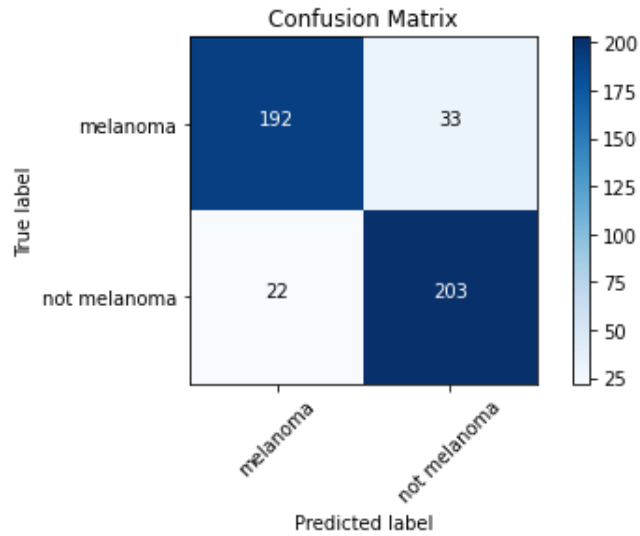The confusion matrix of mean shift is shown in figure 6.7,

35

Figure 6.7: Confusion Matrix (Mean Shift)

From this we can see that the model can detect 192 melanoma and 203 non melanoma cases accurately.

6.1.3 Gabor Feature Result

Form figure 6.8 and 6.9 we can see the gabor feature's training and validation accuracy and loss.

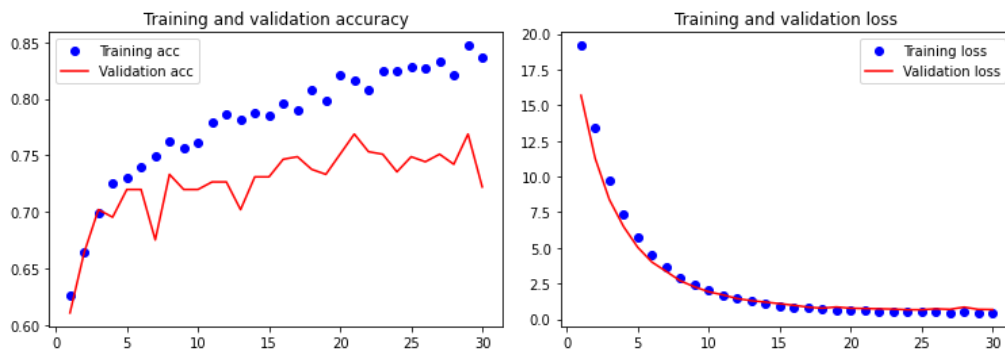Now we apply the testing data of gabor featured image and the results are given below in the table,

Figure 6.8: Training and Validation Accu-Figure 6.9: Training and Validation Loss
racy (Gabor)   (Gabor)

Table 6.3: Gabor Feature Result

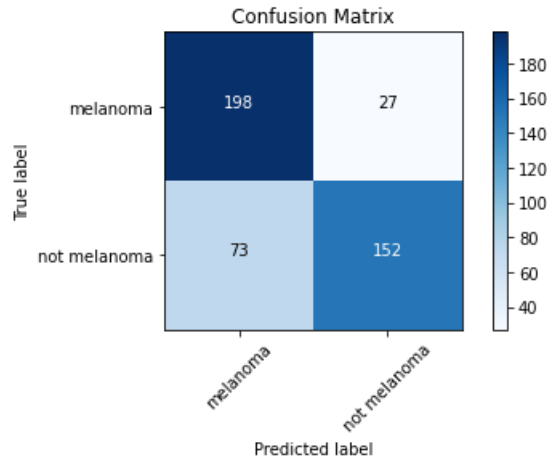| Parameters	Result(percentage)
Accuracy	77.8
Precision	84.9
Recall	67.6
F-Measure	75.2

The confusion matrix,

Figure 6.10: Confusion Matrix (Gabor)

From this, we can see that the model can detect 198 melanoma and 152 non-melanoma cases accurately. From the result, it is clear that non-melanoma case's accuracy is very low here. That is because when we apply the Gabor filter to the image, it becomes a black and white image. In some less severe cases when the mole in the human body is not melanoma but the mole is big enough in the skin image, the converted Gabor image shows a big proportion of that mole and in some cases, the model can not detect the proper depth of that black proportion. That is
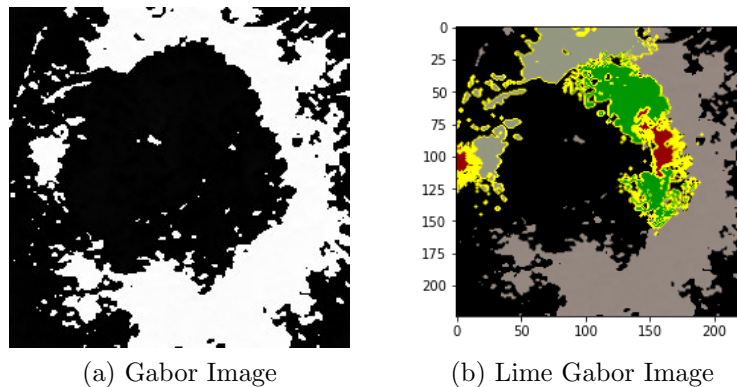

(a) Gabor Image   (b) Lime Gabor Image

Figure 6.11: Lime Featured Gabor images

why the Gabor feature accuracy is getting lower. If the model could properly detect non-melanoma cases then the accuracy might come higher. The promising part is with the gabor feature we can not only detect the cases but also recognize where

the cancer cell is situated. Though lime framework shows better recognizing feature in figure 6.11.

6.1.4 SIFT Feature Result

Form figure 6.12 we can see the SIFT feature's training and validation accuracy and loss. The validation accuracy came 98.56 percent which is very high.
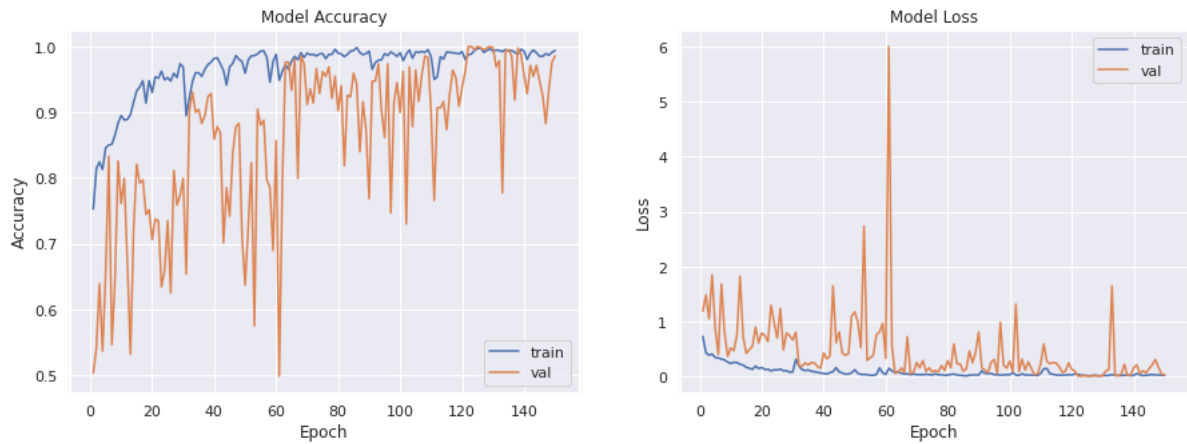

Figure 6.12: Accuracy and Loss Graph (SIFT)

Now we apply the testing data of SIFT featured image and the results are,

Table 6.4: SIFT Feature Result

Parameters	Result(percentage)
Accuracy	81.11
Precision	81
Recall	81
F-Measure	81

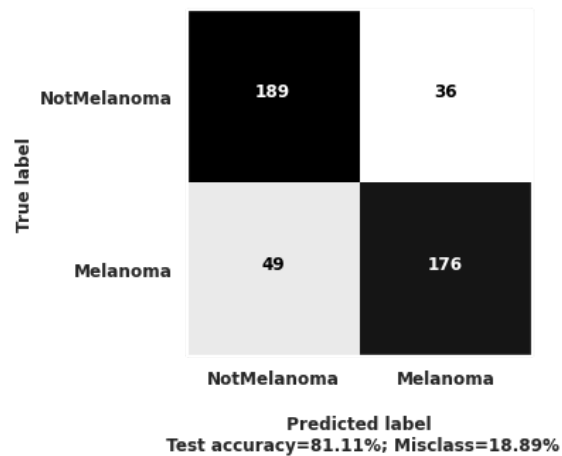The confusion matrix is shown in figure 6.13,

Figure 6.13: Confusion Matrix (SIFT)

From this, it can be seen that the model can detect 189 melanoma and 176 non-melanoma cases accurately. Though the validation accuracy was high, the testing accuracy was not high enough as validation accuracy. The testing accuracy we got is 81.11 percent but the validation accuracy is 98.56 percent. The lime featured images are,

(a) SIFT Image   (b) lime SIFT Image
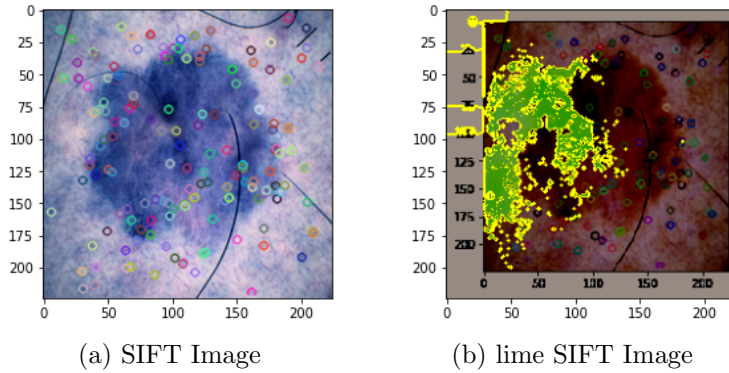
Figure 6.14: Lime Featured SIFT images

6.2 Late Integration Results

As we know this decision integration unit takes these results and merges them to get the final output result. After getting all the early integration model results we can see that the best results are mean shift result and raw image result. Now we are going to use these models' best accuracy results as an input to this neural model. After giving the input, the decision integration unit merged and later delivered the result,

Table 6.5: Late Integration Result (Raw and Shift)

Parameters	Result(percentage)
Accuracy	86
Precision	96.6
Recall	74.7
F-Measure	84.2

Also, the confusion matrix shows the result that the used model can detect melanoma cases perfectly, but the problem remains with the non-melanoma cases. We could also merge these two models because the base of these models are similar to another.
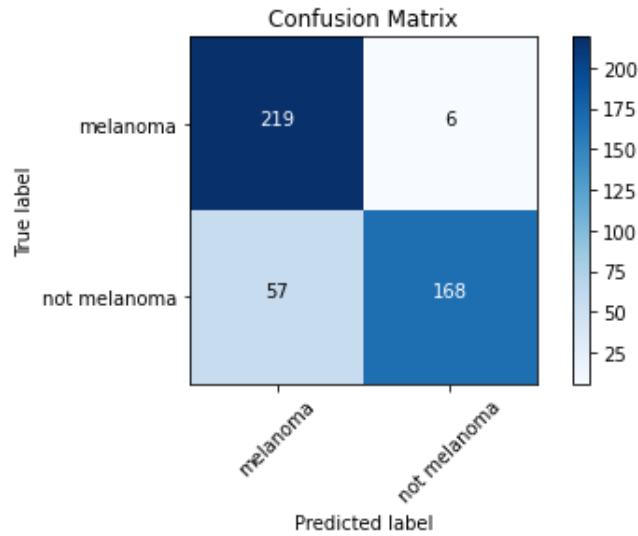
Figure 6.15: Confusion Matrix (Late Integration)

6.3 Full Integration Result

If we look closely the models we have been using in the early integration models are made similarly so that in the final part we can work on these models together. For raw, mean shift, and Gabor we used a similar neural model but for SIFT we had to change the model's base and create a very custom neural network, different from other models. So in the final part, we worked with all the features and original model's result but could not add SIFT model. The final result was also created using the previous values and results. For the full integration model, the accuracy, precision, recall, and f1 measure are,

Table 6.6: Full Integration Result

Parameters	Result(percentage)
Accuracy	84
Precision	94.2
Recall	72.4
F-Measure	82

Also the confusion matrix is,

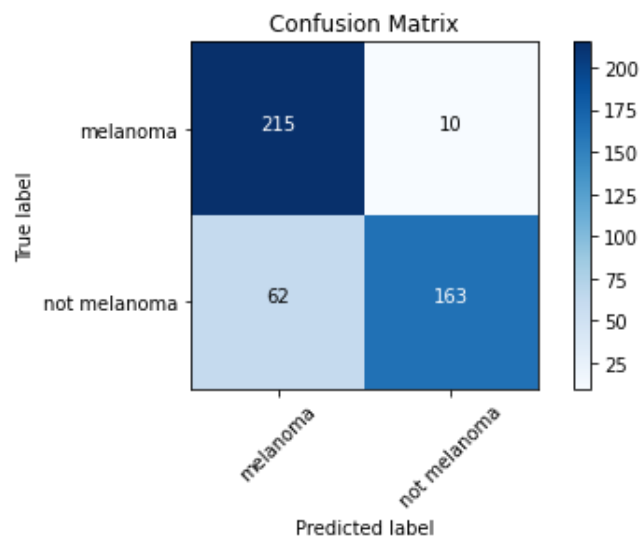Figure 6.16: Confusion Matrix (Full Integration)

Chapter 7

Conclusion And Future Works

7.1   Conclusion

To recapitulate, Deep neural networks (DNNs) have been widely used to automate the evaluate medical pictures for illness diagnosis and to assist human experts by processing large amounts of data rapidly. DNN can also provide extremely promising results. The fundamental goal of this study was to detect melanoma cancer in the early stages with greater accuracy which is already achieved. By applying different handcrafted features (Raw, Gabor, SIFT, and Mean Shift) and later with the neural networks (EfficientNet-B4), we got separate features result in which raw images and mean shift result came highest, 90% and 87.8% respectfully. We also used VGG-16 as another neural network but the accuracy results were lower than EfficientNet-B4. The EfficientNet-B4 gave a better result because it is a more structured and multi-layered model. Later with a late and full integration process where we ensembled different early integration results we got an accuracy of 86%(Late integration) and 84%(Full integration). This result automatically reduces the time to find melanoma skin cancer, as this system can provide the output accurately and instantly. The accuracy might improve using some findings of our which we described in Future works. This system has higher accuracy and is also cost-effective as well which can fill the need in medical testing to detect melanoma cancer in early stages.

7.2   Future Works

There are quite some chances to develop our work in the future. While working on this project we came up with some shortcomings that we have and developing those criteria can be the future work for others. Some of the ideas are described below,

- We used EfficientNet-B4 and VGG-16 while doing the research and EfficeintNet-B4 gave better results, EfficientNet-B7 which is the latest one can be used for this work that might provide more accuracy and better results because in EffiecientNet-B7 there is more layer to work with. The result may vary for different models, even with fewer layers too.

- Our Gabor feature result was low because of non-melanoma cases. A new processing method could have been used to do Gabor feature work that might boost the accuracy close to 90 percent.

- There is a scope to use a fully custom DNN model which might increase the accuracy of each feature and also the final results.

- Better hardware which is focused on research purposes can be used because we only used our personal, daily used devices to complete this research.

- Grad cam can be a solution to have 3D depth information about the images and where the cancer cell is situated.

- Adding hair removal techniques might provide more accurate results. This method can be added next in the future.

Bibliography

[1]   J. Han and C. Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning," in *International workshop on artificial neural networks*, Springer, 1995, pp. 195–201.

[2]   Y. B. Yann Lecun, "Convolutional networks for images, speech, and time-series," *The handbook of brain theory and neural networks*, 1995.

[3]   M. Binder, H. Kittler, A. Seeber, A. Steiner, H. Pehamberger, and K. Wolff, "Epiluminescence microscopy-based classification of pigmented skin lesions using computerized image analysis and an artificial neural network.," *Melanoma research*, vol. 8, no. 3, pp. 261–266, 1998.

[4]   T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.

[5]   H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy," *The lancet oncology*, vol. 3, no. 3, pp. 159–165, 2002.

[6]   Q. Abbas, M. E. Celebi, and I. F. Garcıa, "Hair removal methods: A comparative study for dermoscopy images," *Biomedical Signal Processing and Control*, vol. 6, no. 4, pp. 395–404, 2011.

[7]   Y. S. Kiran Ramlakhan, "A mobile automated skin lesion classification system," *2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*, 2011. DOI: 10.1109/ICTAI.2011.29.

[8]   A. L. Maas, A. Y. Hannun, A. Y. Ng, *et al.*, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. icml*, Citeseer, vol. 30, 2013, p. 3.

[9]   A. Masood and A. Ali Al-Jumaily, "Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms," *International journal of biomedical imaging*, vol. 2013, 2013.

[10]   Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[11]   D. Gutman, N. C. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, and A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic)," *arXiv preprint arXiv:1605.01397*, 2016.

[12]   I. G. Diaz, "Incorporating the knowledge of dermatologists to convolutional neural networks for the diagnosis of skin lesions," *arXiv preprint arXiv:1703.01976*, 2017.

[13] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[14] H. Fan, F. Xie, Y. Li, Z. Jiang, and J. Liu, "Automatic segmentation of dermoscopy images using saliency combined with otsu threshold," *Computers in biology and medicine*, vol. 85, pp. 75–85, 2017.

[15] Y. Filali, A. Ennouni, M. A. Sabri, and A. Aarab, "Multiscale approach for skin lesion analysis and classification," in *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, IEEE, 2017, pp. 1–6.

[16] Y. Filali, M. A. Sabri, and A. Aarab, "An improved approach for skin lesion analysis based on multiscale decomposition," in *2017 International Conference on Electrical and Information Technologies (ICEIT)*, IEEE, 2017, pp. 1–6.

[17] M. H. Jafari, E. Nasr-Esfahani, N. Karimi, S. R. Soroushmehr, S. Samavi, and K. Najarian, "Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma," *International journal of computer assisted radiology and surgery*, vol. 12, no. 6, pp. 1021–1030, 2017.

[18] M. Masrob, M. Rahman, and G. George, "Design of a neural network based power system stabilizer in reduced order power system," in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, IEEE, 2017, pp. 1–6.

[19] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," *arXiv preprint arXiv:1710.05941*, 2017.

[20] J. Dinnes, J. Deeks, N. Chuchu, D. Saleh, S. Bayliss, Y. Takwoingi, C. Davenport, L. Patel, R. Matin, C. O'Sullivan, *et al.*, "Cochrane skin cancer diagnostic test accuracy group. reflectance confocal microscopy for diagnosing keratinocyte skin cancers in adults," *Cochrane Database Syst Rev*, vol. 12, 2018.

[21] J. Dinnes, J. J. Deeks, M. J. Grainge, N. Chuchu, L. F. di Ruffano, R. N. Matin, D. R. Thomson, K. Y. Wong, R. B. Aldridge, R. Abbott, *et al.*, "Visual inspection for diagnosing cutaneous melanoma in adults," *Cochrane Database of Systematic Reviews*, no. 12, 2018.

[22] J. Dinnes, J. J. Deeks, D. Saleh, N. Chuchu, S. E. Bayliss, L. Patel, C. Davenport, Y. Takwoingi, K. Godfrey, R. N. Matin, *et al.*, "Reflectance confocal microscopy for diagnosing cutaneous melanoma in adults," *Cochrane Database of Systematic Reviews*, no. 12, 2018.

[23] U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9909–9924, 2018.

[24] H. A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A. B. H. Hassen, L. Thomas, A. Enk, *et al.*, "Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Annals of oncology*, vol. 29, no. 8, pp. 1836–1842, 2018.

[25] B. Harangi, "Skin lesion classification with ensembles of deep convolutional neural networks," *Journal of biomedical informatics*, vol. 86, pp. 25–32, 2018.

[26] L. F. di Ruffano, Y. Takwoingi, J. Dinnes, N. Chuchu, S. E. Bayliss, C. Davenport, R. N. Matin, K. Godfrey, C. O'Sullivan, A. Gulati, *et al.*, "Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults," *Cochrane Database of Systematic Reviews*, no. 12, 2018.

[27] J. Xiao, Z. Liu, P. Zhao, Y. Li, and J. Huo, "Deep learning image reconstruction simulation for electromagnetic tomography," *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3290–3298, 2018.

[28] C. Yu, S. Yang, W. Kim, J. Jung, K.-Y. Chung, S. W. Lee, and B. Oh, "Correction: Acral melanoma detection using a convolutional neural network for dermoscopy images," *PloS one*, vol. 13, no. 4, e0196621, 2018.

[29] M. Gaana, S. Gupta, and N. S. Ramaiah, "Diagnosis of skin cancer melanoma using machine learning," *Available at SSRN 3358134*, 2019.

[30] M. Goyal, A. Oakley, P. Bansal, D. Dancey, and M. H. Yap, "Automatic lesion boundary segmentation in dermoscopic images with ensemble deep learning methods," *arXiv preprint arXiv:1902.00809*, 2019.

[31] S. Hosseinnezhad, M. Zadshir, X. Yu, H. Yin, B. K. Sharma, and E. Fini, "Differential effects of ultraviolet radiation and oxidative aging on bio-modified binders," *Fuel*, vol. 251, pp. 45–56, 2019.

[32] M. Maarouf, C. Costello, S. Gonzalez, I. Angulo, C. Curiel-Lewandrowski, and V. Shi, "In vivo reflectance confocal microscopy: Emerging role in non-invasive diagnosis and monitoring of eczematous dermatoses," *Actas Dermo-Sifiliográficas (English Edition)*, vol. 110, no. 8, pp. 626–636, 2019.

[33] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering," *International journal of medical informatics*, vol. 124, pp. 37–48, 2019.

[34] J. Daghrir, L. Tlig, M. Bouchouicha, and M. Sayadi, "Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach," in *2020 5th International Conference on Advanced Technologies for Signal and Image Processing ATSIP)*, IEEE, 2020, pp. 1–5.

[35] M. A. Kadampur and S. Al Riyaee, "Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images," *Informatics in Medicine Unlocked*, vol. 18, p. 100 282, 2020.

[36] A. K. Sakshi Niraj Supaha Yashswita Sanjay Jirange, "Melanoma skin cancer detection using image processing," *International Journal of Future Generation Communication and Networking*, vol. 13, no. 2s, 423–428-423–428, Jul. 2020.

[37] M. F. J. Acosta, L. Y. C. Tovar, M. B. Garcia-Zapirain, and W. S. Percybrooks, "Melanoma diagnosis using deep learning techniques on dermatoscopic images," *BMC Medical Imaging*, vol. 21, no. 1, pp. 1–11, 2021.

[38] R. Nair, S. Vishwakarma, M. Soni, T. Patel, and S. Joshi, "Detection of covid-19 cases through x-ray images using hybrid deep neural network," *World Journal of Engineering*, 2021.