SKIN CANCER DETECTION USING NEURAL NETWORK METHODS

By

MD. MUZAHIDUL ISLAM RAHI

16101111

FARHAN TANVIR KHAN

16101099

MOHAMMAD TANVIR MAHTAB

16101125

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

August 2019

Declaration

It is hereby declared that

1. The thesis submitted is my/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:

_____

MD. MUZAHIDUL ISLAM RAHI
16101111

Student's Full Name & Signature:

_____

FARHAN TANVIR KHAN
16101099

Student's Full Name & Signature:

_____

MOHAMMAD TANVIR MAHTAB
16101125

Approval

The thesis titled “Skin Cancer Detection using Neural Network Methods” submitted by

1. MD. MUZAHIDUL ISLAM RAHI (16101111)
2. FARHAN TANVIR KHAN (16101099)
3. MOHAMMAD TANVIR MAHTAB (16101125)

Of Summer, 2019 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on August 7, 2019.

Examining Committee:

Supervisor:
(Member)

_____
Dr. Md. Ashraful Alam
Assistant Professor, CSE
BRAC University

Co - Supervisor:
(Member)

_____
Dr. Md. Golam Rabiul Alam
Associate Professor, CSE
BRAC University

Departmental Head:
(Chair)

_____
Dr. Mahbubul Alam Majumdar
Professor and Chairperson, CSE
BRAC University

Abstract/ Executive Summary

This paper discusses about identifying 7 different types of skin cancers from images using different types of neural network. We have used HAM10000 as our dataset. This contains over 10000 labeled images of skin cancer. For our model, we have used CNN and by using Keras Sequential API, we have structured a new model. Later, for comparison and also accuracy we have pretrained data. These transfer learning model includes VGG11, VGG16, VGG19, RESNET50 and DENSENET121.

Keywords: neural network; machine learning; deep learning; transfer learning; pretrained data.

Acknowledgement

Firstly, all praise to the Allah for whom we could complete our thesis. Secondly our advisor and co advisor Dr. Md. Ashraful Alam and Dr. Md. Golam Rabiul Alam, who supported us throughout the thesis.

Table of Contents

CHAPTER 1

INTRODUCTION

In the world around us we see people with suffering from skin diseases but they seem to overlook thinking it might be just like any typical skin disease they might suffering from. But little do they know how dangerous this might get if they don't know what they are actually suffering from and take necessary precautions by going to the right doctor and get prescribed with proper medication. We know how people who live in the rural areas of underdeveloped countries or developing countries example Bangladesh, Sri Lanka, India etc. have so little or no access to doctors and let alone be doctors who are skin specialists. According to Oxford Journals, there are 13 to 15 lakh disease patients in Bangladesh, with around 2 lakh patients recently determined to have malignancy every year. In the coming years, the expansion in population and life span will bring about an increment in the quantity of malignancy patients in Bangladesh. At the point when joined with populace maturing, the expansion in malignant growth predominance is inescapable. Keeping pace with the interest will require a noteworthy government pledge to malignancy counteractive action and healing administrations in the coming years. Cancer is anticipated to be an undeniably significant reason for dismalness and mortality in Bangladesh in the following couple of decades. The evaluated rate of 12.7 million new malignant growth cases will ascend to 21.4 million by 2030. More than 66% of the absolute consumption on wellbeing is all through of-pocket installments. As indicated by the Bangladesh Bureau of Statistics, malignant growth is the 6th driving reason for death. Global Agency for Research on Cancer has assessed disease related passing rates in Bangladesh to be 7.5% in 2005 and 13% in 2030.

So when people like these have any skin disease, either they choose to avoid it or when they realize it's high time to consult a proper doctor, it's too late. Also doctors at those areas are not usually experts to understand skin diseases that can cause cancers. Furthermore, according to Skin Cancer Foundation, in USA alone, 9500 people are diagnosed with skin cancer every single day and more than two people die of the disease every hour. More people are diagnosed with skin cancer each year in the U.S than all other cancers combined. Also one in five Americans will develop skin cancer by the age of 70. This leads to an astounding annual cost of treating cancers in the U.S is estimated at around $8.1 billion. $4.8 billion are for nonmelanoma skin cancers and $3.3 billion for melanoma. So eve in a place like U.S where you have the best hospitals equipped with the best of the best doctors, still the rate of death because of cancer still very high. If more people could have gotten more precocious from an earlier time then they would have known that it's not any typical skin lesion, rather something life threatening. This is where our model of detecting skin cancer from mere images comes in. Our model will process the images and by the help of the algorithm we used, it will detect the probability of that skin disease of being cancerous.

At first we ran our own CNN model on the dataset (HAM10000) we got from ISIC archive. Later to get better and more accurate results, we used transfer learning.

Transfer learning is a machine learning method where it uses pre trained models given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide in related problems. It is an optimization that allows rapid progress or improved performance when modelling the second task.

In transfer learning we used VGG, ResNet and DenseNet Methods.

CHAPTER 2

Related Work

We have gone through several papers which aided us in doing our thesis work. They gave us the idea, motivation and guided us in difficult times. Below are some of the previous works of others that were related to our thesis topic. Our work focuses on detection of skin cancer using different neural networks so we looked into several methods which use neural networks and machine learning to try to solve the same problem.

One of the major works in this field was "Automatic Detection of Melanoma Skin Cancer using Texture Analysis" by Mariam A.Sheha Cairo University, Mai S.Mabrouk MUST University, Amr Sharawy Cairo University

This paper aims to detect Melanoma which is a skin related cancer and it is life-threatening. The paper talks about an automated method for melanoma diagnosis applied on a series of dermoscopy images. They have distinguished between melanocytic nevi and malignant melanoma using texture analysis in their system. The Artificial Neural Network they applied is MLP and represent the classes of the 2 diseases, whereas we used different neural networks such as ResNet, DenseNet and VGG models to classify 7 types of skin lesions. They have used 2 types of MLP: Automatic and Traditional. The authors have also addressed previous related work in achieving their goals. They have referred to various kinds of ways proposed that were used to improve accuracy of diagnosis. ELM (Epiluminescence Microscope) described in 1986 and it uses incident light, oil immersion and magnifier. In 2003,M. Wiltgen, A. Gerger and J. Smolle used TCA (Tissue Counter Analysis), which uses image partition into equal squares from which features are calculated. After this they have proposed their model which is given below:

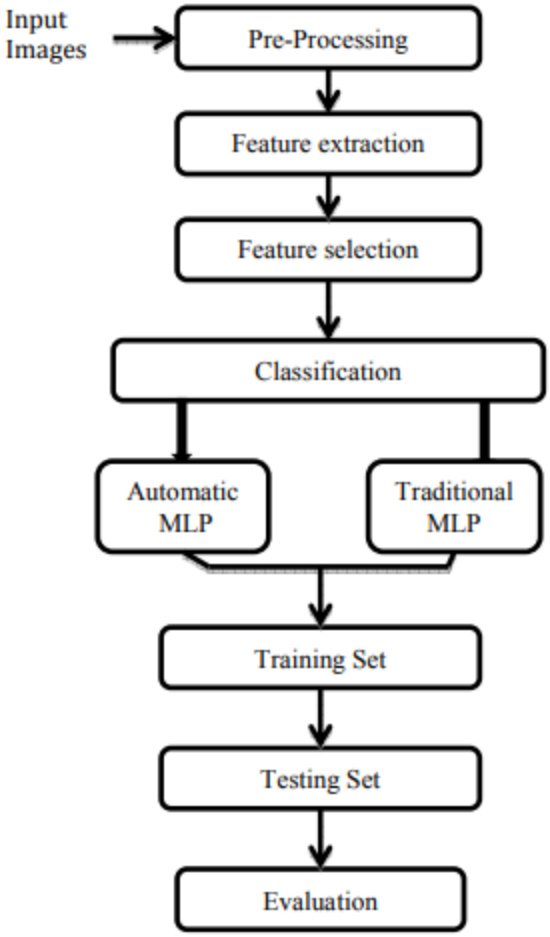

Figure: Automated diagnosis block diagram.

Classification is based on co-occurrence matrix texture features to distinguish between melanocytic nevi and melanoma. The images of these are given as inputs to a neural network classifier. MLP (Multilayer Perceptron) generates nonlinear boundaries and is a feed forward network. They have divided data into the following proportions: 60% to train, 20% to test and 20% for validation. In contrast to our own work which is 80% training and 20% testing and then a further 20% from the train data for validation purposes. The results of this paper is given below

Table 1: GLCM Features Value for Malignant and Nevi Class

Features	Malignant 1	Nevi1	Malignant 2	Nevi2	Malignant 3	Nevi3
Contrast	0.0956152	0.1592465	0.0875007	0.1920330	0.0985926	0.109134
Correlation	0.9785536	0.8042529	0.9861547	0.8904056	0.9826944	0.944586
Cluster Prominence	121.79231	23.69187	276.10653	72.473045	293.16681	33.63818
Dissimilarity	0.0942774	0.1107356	0.0794054	0.1447147	0.0831740	0.098493
Homogeneity	0.9529950	0.9492526	0.9610813	0.9320725	0.9599293	0.951783
Difference Variance	0.0956152	0.1592465	0.0875007	0.1920330	0.0985926	0.109134
Difference entropy	0.3146659	0.3507624	0.2806824	0.4269775	0.2929147	0.327785
Information measure of correlation 1	-0.7864709	-0.5797063	-0.8229712	-0.595457	-0.8117864	-0.739550
Information measure of correlation 2	0.9567472	0.7898776	0.9695077	0.8532012	0.9641905	0.929534
Inverse difference homogenous (INV)	0.9530765	0.9504054	0.9613402	0.9332840	0.9606344	0.952166
Inverse difference normalized (INN)	0.9895394	0.9881743	0.9912591	0.9843856	0.9909207	0.989165
Inverse difference moment normalized	0.9985310	0.9976929	0.9986747	0.9971896	0.9985152	0.998347

Table 2: Training Performance Measure

Classifier (MLP)	Sensitivity	Specificity	Accuracy
Automatic MLP	94.5%	92.3%	93.4%
Traditional MLP	100%	100%	100%

Table 3: Testing Performance Measure

Classifier (MLP)	Sensitivity	Specificity	Accuracy
Automatic MLP	70.5%	87.5%	76%
Traditional MLP	92.3%	91.6%	92%

Figure: Results from their model

Another work which we studied was “Neural Network Diagnosis of Malignant Melanoma from Color Images” by Fikret Ercal, Senior Member, IEEE Anurag Chawla, William V. Stoecker, Hsi-Chieh Lee, and Randy H. Moss, Senior Member, IEEE.

In this paper, the authors talked about using colored images of skin lesions combined with image processing techniques and an ANN(Artificial Neural Network) to separate melanoma from 3 other tumours: dysplastic nevi, intradermal nevi and seborrheic keratoses. To do this they used a feedforward artificial neural network with 14 inputs and 1 output stating whether or not the tumor is melanoma. This model was used and trained by backpropagation rule. They also said that they used NeuralWorks which is a commercial neural network software. Using these, they conducted 2 experiments: One with 14 input neurons, 7 hidden neurons and 1 output neuron and the other with 8 inputs, 4 hidden neurons and 1 output neuron.

There were a total of 240 images used in Experiment 1; 120 melanomas, 40 dysplastic nevi, 40 intradermal nevi and 40 seborrheic keratoses. For the 20/80 case, 20% of the images were used for training and the remaining 80% are used for testing, giving a total of 48 images for the training set and 192 images for the testing set. In the testing set, there would be (80% x 120) = 96 melanomas, and (192-96) = 96 nonmelanomas (always 50% of the testing set is melanomas and the remaining 50% is nonmelanomas). If 81 out of 96 melanomas were correctly diagnosed, then the percentage of true positives (the melanoma success rate) would be 81/96 = 84.4% and the number of false negatives (melanomas diagnosed as benign lesions) could be calculated as (96 - 84.4% x 96) = 15. If 74 out of 96 nonmelanomas were correctly diagnosed as nonmelanomas, then the percentage of true negatives (the nonmelanoma success rate) would be 74/96 = 77.1% and the number of false

positives (benign lesions diagnosed as melanomas) was calculated as (96 - 77.1% x 96) = 22. The overall success rate in this case turned out to be (81+74)/192 = 80.7%.

CHAPTER 3

Different types of Skin Cancer

The dataset we will on working on consists of 7 different types of skin diseases with over 10,000 images of skin diseases to train our model on. The 7 skin diseases we will be working on are:

1. Melanocytic nevi

2. Melanoma

3. Benign keratosis-like lesions

4. Basal cell carcinoma

5. Actinic keratoses

6. Vascular lesions

7. Dermatofibroma

Let's start talking about these diseases. What are they and why did we choose these 7 skin diseases to be the core of our work.

3.1 Melanocytic nevi

Melanocytic Nevi is a skin condition where an abnormally dark patch grows on the skin. Usually they are non-cancerous however people who have such type of skin condition since their birth have around 5-10 percent chance of developing a certain type of skin cancer known as melanoma.

Melanoma is a type of cancer, it usually develops from the thin coating of cells known as melanocytes. Usually diseases like melanomas occur in the skin but not in someone's mouth, eyes or intestines. Melanoma is very common amongst women and they are usually are seen growing on their legs. However for men, they occur on the back. It all starts from a mole and gradually with changes such as increase in size, changes in color of the mole, irregular edges and also itchiness. The main reason for such cancerous disease is Ultra Violet (UV) light from the sun.

Figure: Malignant melanoma

Melanocytic Nevi may cause individuals to be emotionally stressed because of the birth mark on their skin and this may affect their appearance and also their diet. This may affect children the most as they get affected emotionally and also tend to have behavioral problems.

The US National Cancer Institute predicted over 60,000 cases of melanoma and over 8,000 deaths would occur in the year of 2008 only in the States.

As melanoma has already been described before, it is quite debatable how the melanocytic nevi transforms into melanoma. But it appears that 10% of malignant melanomas have a precursor lesion, of which about 10% are melanocytic nevi.

Figure**:** Melanocytic Nevi

3.2 Melanoma

Melanoma is the most dangerous type of skin cancer. The growth of these cancerous cells occurs when the skin cells gets damaged by the unrepaired DNA. This mainly occurs because of UV rays that in turn triggers the mutation and leads the skin cells to grow abnormally. Melanomas often look like moles and most of the times they form from moles. Usually they look are black or brown in color but can also be skin, pink, red, purple, blue or white colored. The frequent exposure of UV rays leads to sunburn and also Melanoma and people who are inherited to this disease are at a greater risk. Melanoma can also form in the eyes but rarely does it occur in the internal organs such as the intestines.

Apart from UV radiation, tanning lamps also increases the risk of having this disease. So limiting oneself to the exposure to UV radiation can reduce the risk of melanoma. People who are under 40 have a higher risk of melanoma and most of them include women. If it can be detected early then melanoma can be treated.

Figure**:** Melanoma

Most common sign of someone having melanoma is when someone grows a new mole or there comes some changes in an existing mole of that person. Melanoma can occur anywhere on the body and in men they occur mainly in the back and in women they occur on the legs. As we mentioned above melanoma is affected by the sun, so areas of the body which are protected by the sun are where melanoma is very uncommon, for example the bottom and the scalp.

There are several other types of melanoma. Some of them are:

a. Superficial spreading melanoma

Almost 70% of the melanomas in the United Kingdom are superficial spreading melanomas. Melanomas like this is mostly found in people who have a pale skin and freckles but much less common in people who are dark skinned.

This melanomas usually grow forwards over downwards. However if they grow downwards, they can spread to other parts of the body.

Figure: Super spreading melanoma

b. Nodular Melanoma

These melanomas grow faster and downwards into the skin if not removed. They usually appear as lump that changes and the color might vary from black to red. Nodular Melanomas often grow on normal skin and mostly they grow on the head, neck, chest or back. Bleeding is a common sign.

Figure: Nodular Melanoma

c. Lentingo maligna melanoma

Any Melanoma has a 0.1 probability of being Lentingo Maligna Melanomas. These are commonly found in people who are old and spent time outdoors. It takes them time to develop, probably years but mostly in areas which are exposed to the sun.

They are flat and grow sideways in the surface of the skin. Usually they look like blotch but larger. Over time they get larger and their shape can change. They even might grow downwards into the deeper layers of skin.

Figure: Lentingo Maligna Melanoma

d. Acral Letiginous Melanoma

Acrak Ketiginous Melanomas are rare and can usually occur on the palms of the hands and the soles of the feet. More commonly they can be seen around thumbnail or big toenail. People with dark skin has greater risk of developing this disease but this still can occur in people of any skin color.

Figure: Acrak Ketiginous Melanoma

e. Amelanotic melanoma

Just like Acrak Ketinginous Melanomas, Amelanotic melanomas are also pretty rare and they account for 5 in 100 melanomas. Usually they don’t have color but sometimes they can be red or pink, or have light brown or grey edges.

Figure: Amelanotic melanoma

3.3 Benign keratosis-like lesions

Seborrheic keratosis is one of the most common skin growths in men which is non-cancerous. Usually it is brown, black or light tan that grows on the face, chest and shoulders or even at the back. The development has a waxy, layered, somewhat raised appearance. Seborrheic Keratosis doesn't become cancerous and have no risk if exposed to the sun yet they appear to look like skin cancer. Seborrheic keratosis are usually not painful and no treatment is required. Someone can also decide to get them removed if they become irritating. They are not painful but depending upon their size and location, they can be troublesome.

They grow singly or may grow in clusters and are round or oval shaped furthermore can be around 1 inch in size. The main reason for Seborrheic keratosis is unknown but are very common depending on someone's age. This is a kind of skin disease which can be inherited. But one has higher chance of developing this skin disease if that person is over 50.

Figure: Seborrheic Keratosis

3.4 Basal cell carcinoma

Basal cell carcinoma starts in the basal cells. This is a type of cell within the skin that makes new skin cells as the old one come off. They appear as a transparent bump on the skin and can also take other forms as well. It occurs on areas of the skin which are exposed to the sun, example- your head and neck. Most basal cell carcinomas are believed to be brought about by long haul introduction to bright (UV) radiation from daylight. Staying away from the sun and utilizing sunscreen may help ensure against basal cell carcinoma. This skin malignant growth seems less regularly on the storage compartment and legs, and basal cell carcinoma can — yet once in a while — happen on parts of your body typically shielded from the sun, for example, privates or ladies' breast.

Basal cell carcinoma shows up as an adjustment in the skin, for example, a development or a sore that won't mend. These adjustments in the skin, or injuries, more often than not have one of the accompanying attributes:

1. A silvery white, skin-hued or pink knock that is translucent, which means you can see somewhat through the surface. Small veins are frequently unmistakable. In individuals with darker skin tones, the sore would be darker yet at the same time to some degree translucent. The most well-known kind of basal cell carcinoma, this injury frequently shows up on the face, ears or neck. The sore may crack, drain and scab over.
2. A darker, dark or blue injury — or a sore with dull spots — with a marginally raised, translucent outskirt.
3. A level, layered, ruddy fix with a raised edge is increasingly normal on the back or chest. After some time, these patches can become very enormous.
4. A white, waxy, scar-like sore without an unmistakably characterized fringe, called morpheaform basal cell carcinoma, is the least normal. This sore is not entirely obvious, yet it might be an indication of an especially intrusive and distorting disease.

Figure: Basal cell carcinoma

3.5 Actinic keratoses

An actinic keratosis is an unpleasant, layered fix on your skin that creates from long stretches of presentation to the sun. It is most generally found all over, lips, ears, back of your hands, lower arms, scalp or neck.

Otherwise called a sun oriented keratosis, an actinic keratosis amplifies gradually and ordinarily causes no signs or indications other than a fix or little spot on your skin. These patches take a very long time to grow, generally first showing up in individuals more than 40.

The signs and manifestations of an actinic keratosis include:

1. Unpleasant, dry or layered fix of skin, typically under 1 inch (2.5 centimeters) in distance across
2. Level to marginally raised fix or knock on the top layer of skin

3. Now and again, a hard, wartlike surface
4. Shading as differed as pink, red or dark colored
5. Tingling or consuming in the influenced region

Actinic keratoses are found fundamentally on regions presented to the sun, for example, your face, lips, ears, hands, lower arms, scalp and neck

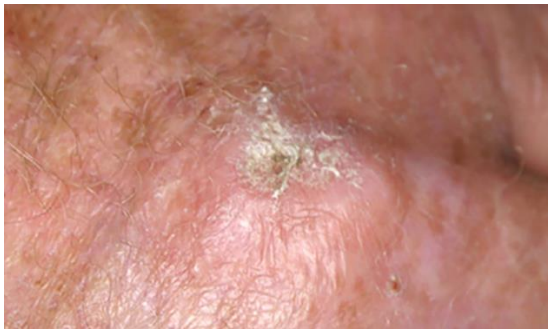

Figure: Actinic Keratosis

3.6 Vascular lesions

Vascular tumors incorporate neoplasms of the vascular framework, of which juvenile hemangiomas (IHs) are the most well-known. Vascular abnormalities, then again, comprise of lesions due to anomalous development of the vascular system, including the capillary, venous, arterial, and lymphatic systems.

This type of lesions are moderately basic variations from the norm of the skin and hidden tissues, all the more usually known as pigmentations. There are three noteworthy classes of vascular

injuries: Hemangiomas, Vascular Malformations, and Pyogenic Granulomas. While these skin colorations can seem to be comparable now and again, they each change as far as starting point and fundamental treatment.

Figure: Vascular lesions

3.7 Dermatofibroma

Dermatofibroma (stringy histiocytoma, histiocytoma cutis, sclerosing hemangioma) is a little, firm, amiable papule restricted to the skin. This sore typically happens in early adulthood and is generally situated on the lower furthest points of ladies. Dermatofibroma isn't basic in small kids and is never found in earliest stages. Most injuries are 3 to 4 mm in measurement, yet monster variations do exist. Horizontal weight on a dermatofibroma prompts dimpling of its surface, known as the dimple sign. The reason for dermatofibroma has not been unmistakably established, 8 but

rather it is accepted to be a considerate, receptive expansion of skin that happens because of minor injury, for example, a creepy crawly chomp. In grown-ups, the nearness of various dermatofibromas has been related with connective tissue maladies, for example, lupus erythematosus. Treatment for dermatofibroma is commonly not demonstrated except if the sore is symptomatic. On the off chance that the determination is being referred to, an excisional biopsy ought to be performed.

Figure: Dermatofibroma

CHAPTER 4

Dataset Description

The dataset we used is HAM10000 (“Human against Machine with 10000 training images”) dataset. The main reason behind using this dataset was because training the neural network fir automated diagnosis of pigmented skin lesions is effected by the small size of the images and the lack of different types of datasets with images of skin lesions. The images we collected were from diverse populations. By using this diverse images they applied different cleaning methods and got semi-automatic workflows utilizing specifically. In the final dataset they had 10015 images which are given for training set for academic reasons of machine learning and this training set is publically available through the ISIC achieve. This main dataset mainly can be used for comparisons and for machine learning. Over half of sores have been affirmed by pathology, while the ground truth for the remainder of the cases was either development, master accord, or affirmation by in-vivo confocal microscopy.

Design Type(s)	Database creation objective Data integration objective Image format conversion objective
Measurement Type(s)	Skin lesions
Technology Type(s)	Digital curation

Factor Type(s)	Diagnosis Diagnostic Procedure Age Biological sex Animal body part
Sample Characteristic(s)	Human Sapiens Skin of body

Figure: The summary of metadata

4.1 Background and Summary

Dermatoscopy is very famous identification method that identifies the infectious skin diseases from the skin and not so harmful skin diseases and it does it better than the naked human eye. Dermatoscopic pictures are additionally an appropriate source to prepare large neural systems to analyze pigmented skin sores consequently. Binder et al used these images successfully in 1994 to train and an artificial neural network to find the difference in different types of melanoma which is considered. Earlier same work was done but the problem was there was small size sample and lack of skin diseases affected by skin lesions. No matter how complex the neural network is because of the advancement in graphics card and machine learning techniques, new guide is set. And the automated identification system will soon be available everywhere and can be used to identify all kinds of lesions such as the ones which are brown, black or blue in color without the presence of any human.

For the training of the neural network, the algorithm for the identification require a large number of labeled images but the number of high quality images of the skin diseases can only be classified to a few different types of disease. In 2013 Mendonça et al. made 200 dermatoscopic pictures accessible as the PH2 dataset including 160 nevi and 40 melanomas5. Pathology was the ground truth for melanomas yet not accessible for generally nevi. Since the set is freely accessible (http://www.fc.up.pt/addi/) and incorporates far reaching metadata it filled in as a benchmark dataset for investigations of the PC determination of melanoma as of not long ago.

According to the book Iterative Atlas of Democracy, a cd-rom is available with 1044 images with 167 non-melanocytic lesions, and almost 20 images of diseases which are not even covered in the dataset. Despite the fact that this is one of the most various accessible datasets with respect to secured analyze, its utilization is likely restricted in light of its obliged openness.

The ISIC document (https://isic-archive.com/) is a gathering of different databases and as of now incorporates 13786 dermatoscopic pictures (as of February twelfth 2018). As a result of tolerant authorizing (CC-0), well-organized accessibility, and huge size it is as of now the standard hotspot for dermatoscopic picture examination look into. It is, in any case, one-sided towards melanocytic sores (12893 of 13786 pictures are nevi or melanomas). Since this entryway is the most exhaustive, in fact progressed, and open asset for computerized dermatoscopy, we will give our dataset through the ISIC file.

In view of the constraints of accessible datasets, past research concentrated on melanocytic sores (i.e the separation among melanoma and nevus) and ignored non-melanocytic pigmented injuries in spite of the fact that they are basic by and by. The confuse between the little decent variety of

accessible preparing information and the assortment of genuine information brought about a moderate execution of robotized analytic frameworks in the clinical setting in spite of magnificent execution in test settings. Building a classifier for different illnesses is more testing than parallel arrangement. As of now, solid multi-class forecasts are accessible for clinical pictures of skin illnesses however not for dermatoscopic pictures.

To help the exploration on mechanized analysis of dermatoscopic pictures we discharged the HAM10000 ("Human against Machine with 10000 preparing pictures") dataset. The dataset will be given to the members of the ISIC 2018 arrangement challenge facilitated by the yearly MICCAI meeting in Granada, Spain, yet will likewise be accessible to research bunches who don't take an interest in the test. Since we will likewise utilize this dataset to gather and give data on the presentation of human master analysis, it could fill in as benchmark set for the examinations of people and machines later on. So as to give more data to AI research bunches who expect to utilize the HAM10000 preparing set for research we portray the development and the points of interest of the dataset (Figure: Summary of Metadata) in detail.

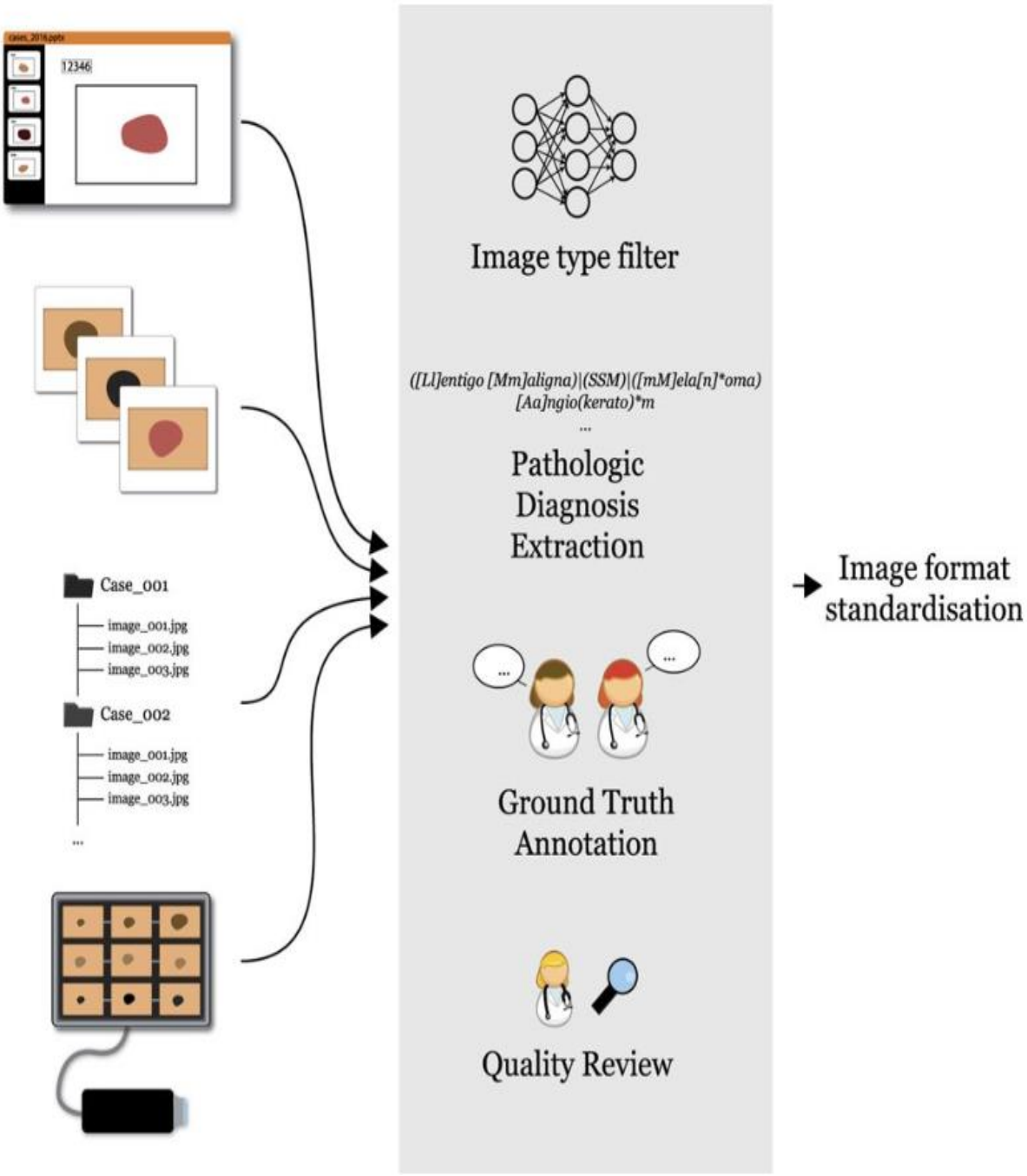

Figure: Image and data content from different sources were entered into a pipeline to organize and clear data. With final images being standardized and stored in a common format.

4.2 Methods

The 10015 dermatoscopic pictures of the HAM10000 preparing set were gathered over a time of 20 years from two unique destinations, the Department of Dermatology at the Medical University of Vienna, Austria, and the skin malignancy routine with regards to Cliff Rosendahl in Queensland, Australia. The Australian site put away pictures and meta-information in PowerPoint records and Excel databases. The Austrian site began to gather pictures before the time of computerized cameras and put away pictures and metadata in various configurations during various timeframes.

4.3 Extraction of images and meta-data from Power Point files

Each PowerPoint record contained successive clinical and dermatoscopic pictures of one schedule month of clinical workup, where each slide contained a solitary picture and a content field with a remarkable sore identifier. In light of the huge measure of information we connected a mechanized way to deal with concentrate and sort those pictures. We utilized the Python bundle python-pptx to get to the PowerPoint records and to get the substance. We iterated through each slide and consequently extricated and put away the source picture, the relating identifier, and the time of documentation, which was a piece of the record name.

4.4 Digitization of diapositivies

Prior to the presentation of computerized cameras, dermatoscopic pictures at the Department of Dermatology in Vienna, Austria were put away as diapositives. We digitized the diapositives with a Nikon Coolscan 5000 ED scanner with a two-overlap examine with Digital ICE and put away

records as JPEG Images (8-bit shading profundity) in most noteworthy quality (300DPI; 15×10 cm). We physically trimmed the filtered pictures with the injury focused to 800x600px at 72DPI, and connected manual histogram redresses to improve visual differentiation and shading multiplication (Fig. 2).

Figure: Original scanned image with remaining black border on the lower left, lesion off center, yellow hue and reduced luminance.

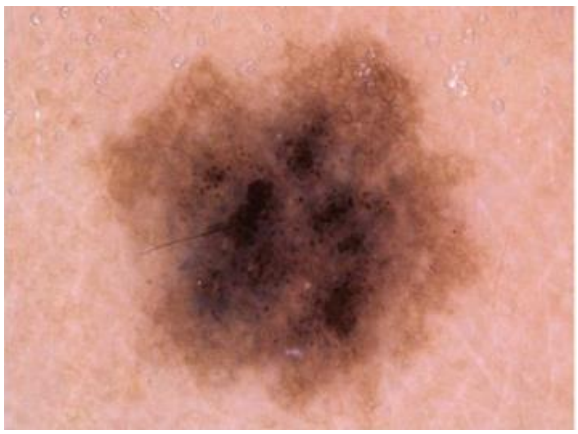

Figure: The final image after manual quality review

4.5 Extraction of data from a digital dernatoscopy system

The Department of Dermatology at the University of Vienna is furnished with the computerized dermatoscopy framework MoleMax HD (Derma Medical Systems, Vienna, Austria). We extricated cases from this framework by sifting SQL-tables with an exclusive apparatus given by the producer. We chose just non-melanocytic injuries with an agreement favorable analysis, nevi with >1.5 long stretches of computerized dermatoscopic development, and extracted sores with a histopathologic report. Histopathologic reports were coordinated physically to explicit injuries. From a progression of various consecutive pictures of a similar nevus we separated just the latest one. A few melanomas of this set were likewise captured with a DermLiteTM FOTO (3GenTM) camera. These extra pictures turned out to be additionally part of the ViDIR picture arrangement, where various pictures of a similar injury were marked with a typical identifier string. Unique pictures of the MoleMax HD framework had a goals of 1872x1053px (MoleMax HD) with non-quadratic pixels. We physically edited all MoleMax HD pictures to 800x600px (72DPI), focused the sore if vital, and returned the configuration to quadratic pixels

4.6 Filtering of dermatoscopic images

The source picture accumulations of the two destinations contained dermatoscopic pictures as well as clinical close-ups and outlines. Since there was no dependable comment of the imaging type, we needed to isolate the dermatoscopic pictures from the others. To manage the huge measure of information proficiently we built up a robotized technique to screen and order >30000 pictures, like Han et al: We hand-named 1501 picture records of the Australian accumulation into the classes "outlines", "close-ups" and "dermatoscopy". Utilizing the hand-named pictures as a preparation

set, we adjusted an InceptionV3-architecture13 (loads pre-prepared on ImageNet4 information) to group the pictures as indicated by picture type. In the wake of preparing for 20 ages with Stochastic Gradient Descent, with a learning rate instated at 0.0003, advance down (Gamma 0.1) at ages 7 and 13, and a cluster size of 64, we got a best 1 precision of 98.68% on our hold-out test set. This exactness was adequate to quicken the determination procedure of dermatoscopic pictures. The few remaining misclassified close-ups and reviews were expelled by turn in a subsequent modification.

4.7 Unifying pathologic diagnoses

Histopathologic conclusions demonstrated high fluctuation inside and between locales including grammatical mistakes, distinctive dermatopathologic phrasings, and various determinations per injury or unsure analyses. Cases with unsure judgments and crashes were rejected aside from melanomas in relationship with a nevus.

We brought together the analyses and shaped seven conventional classes, and explicitly stayed away from vague groupings. The histopathologic articulation "shallow spreading melanoma in situ, emerging in a previous dermal nevus", for instance, should just be designated to the "melanoma" class and not to the nevus class. The seven nonexclusive classes were picked for effortlessness and in see of the planned use as a benchmark dataset for the determination of pigmented injuries by people and machines. The seven classes secured over 95% of all pigmented sores inspected in every day clinical routine with regards to the two investigation locales. An increasingly itemized portrayal of the ailment classes is given in the utilization notes underneath.

4.8 Manual quality review

To discard cases with some attributes given below, a final validation was performed on all the images of the dataset.

- **Type -** Close-up and outline pictures that were not evacuated with programmed sifting
- **Identifiability -** Images in which things are identifiable such as tattoos, garment etc.
- **Quality -** Pictures that were out of center or had aggravating relics like blocking gel bubbles. We explicitly endured the nearness of terminal hairs
- **Content -** Totally non-pigmented sores and visual, subungual or mucosal sores

Remaining cases were assessed for fitting shading multiplication and luminance and, if important, redressed through manual histogram remedy.

4.9 Data Records

All the information about the HAM10000 dataset are stored in the Harvard Dataverse. The table also consists some other databases as well and the images of the HAM10000 dataset are compared to the existing datasets and the summary is given in the table below.

Dataset	License	Total images	Pathologic verification (%)	akiec	bcc	bkl	df	mel	nv	vasc
PH2	Research&Education[a]	200	20.5%	-	-	-	-	40	160	-
Atlas	No license	1024	unknown	5	42	70	20	275	582	30
ISIC 2017[b]	CC-0	13786	26.3%	2	33	575	7	1019	11861	15
Rosendahl	CC BY-NC 4.0	2259	100%	295	296	490	30	342	803	3
ViDIR Legacy	CC BY-NC 4.0	439	100%	0	5	10	4	67	350	3
ViDIR Current	CC BY-NC 4.0	3363	77.1%	32	211	475	51	680	1832	82
ViDIR MoleMax	CC BY-NC 4.0	3954	1.2%	0	2	124	30	24	3720	54
HAM10000	CC BY-NC 4.0	**10015**	53.3%	327	514	1099	115	1113	6705	142

Figure: Summary of the publicly available dermatoscopic image datasets in comparison

The following are the other datasets available at the Harvard Dataverse.

1. **Rosendahi image set (Australia)**

Sores of this piece of the HAM10000 dataset start from the workplace of the skin malignant growth routine with regards to Cliff Rosendahl (CR, School of Medicine, University of Queensland). We removed information from his database after institutional morals board endorsement (University of Queensland, Protocol-No. 2017001223). Pictures were exclusively taken by creator CR with either a DermLite Fluid (non-energized) or DermLite DL3 (3Gen, San Juan Capistrano, CA, USA) with submersion liquid (either 70% ethanol hand-wash gel, or ultrasound gel). Advanced pictures were put away inside PowerPointTM introductions. Each slide contained a dermatoscopic picture and a content field with a back to back remarkable sore ID connecting the picture to a different Excel database with clinical metadata and histopathologic analyze. Pictures were reported thusly sequentially beginning 2008 until May 2017. The Rosendahl arrangement comprises of 34.2GB of advanced slides (122 PowerPointTM-documents) from which we separated 36802 pictures with a coordinating histopathologic report as depicted previously. After evacuation of non-pigmented injuries, reviews, close-ups without dermatoscopy, and cases without or with an unseemly determination that did not can be categorized as one of the predefined conventional classes we touched base at the last dataset depicted in Table of the above figure.

2. **ViDIR image set (Austria)**

From the ViDIR Group (Department of Dermatology at the Medical University of Vienna, Austria) information sources from various occasions were accessible and prepared after morals council endorsement at the Medical University of Vienna (Protocol-No. 1804/2017).

3. Legacy Diapositives

The most seasoned asset of pictures of the ViDIR gathering goes back to the period before the availablility of advanced cameras when dermatoscopic pictures were taken with simple cameras and chronicled as diapositives. They were initially shot with the Heine Dermaphot framework utilizing drenching liquid, and delivered for instructive and authentic purposes with the E-6 method.

4. Current Database

Since 2005 we archived computerized dermatoscopic pictures with a DermLiteTM FOTO (3GenTM) framework (single cases additionally with Heine Delta 20) and put away pictures and meta-information in a focal database. This dataset incorporates triplets from a similar sore taken with various amplifications to empower perception of nearby highlights and general examples.

5. MoleMax Series

The Department of Dermatology in Vienna offers advanced dermatoscopic follow-up to high hazard patients to build the quantity of extractions of early melanoma while diminishing pointless extractions of nevi15. The time interim between follow-up visits more often than not runs from 6 a year. Rinner et al.16 as of late distributed a point by point depiction of this subsequent program. Since 2015 the MoleMax HD System (Derma Medical Systems, Vienna, Austria) is utilized for securing and capacity of dermatoscopic images.Most patients in the subsequent program have various nevi. Dermatologists of the Department of Dermatology in Vienna as a rule screen

"atypical" nevi yet additionally few arbitrarily chose subtle nevi, which are usually underrepresented in datasets utilized for AI.

4.10 Technical Validation

On the off chance that vital, a specialist dermatologist performed manual histogram remedy to modify pictures for visual review by human pursuers. We limited redresses to underexposed pictures and pictures with an unmistakable yellow or green shade and made no different changes. To delineate changes, we directed shading illuminant estimations after redresses as per the dim world assumption17 as appeared in Fig. 3. As wanted, rectifications moved yellow and green illuminants towards blue and red. We characterized four distinct sorts of ground truth:

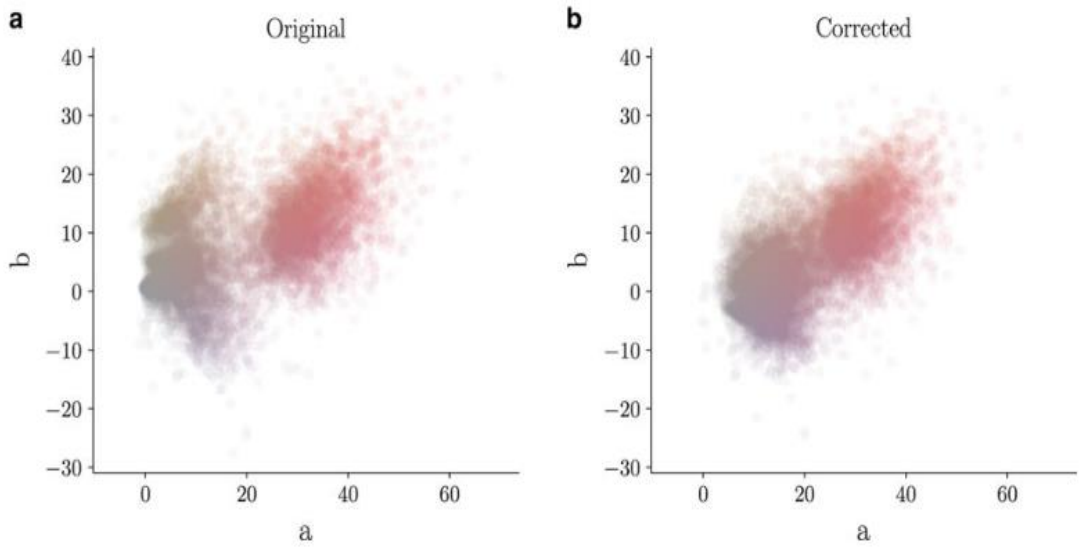

Figure: Values estimated with a grey-world assumption of all training-set images in Lab-color space before (left) and after (right) manual histogram changes.

1. **Histopathology**

Histopathologic determinations of extracted injuries have been performed by particular dermatopathologists. We filtered all accessible histopathologic slides of the current ViDIR picture set for later audit. We physically evaluated all pictures with the relating histopathologic determination and checked for believability. On the off chance that the histopathologic finding was doubtful we checked for test bungle and surveyed the report and reevaluated the slide if important. We rejected cases with uncertain histopathologic analyze (for instance: "support nevus however can't preclude developing melanoma in situ").

2. **Confocal**

Reflectance confocal microscopy is an in-vivo imaging procedure with a goals at close cell level18, and some facial kindhearted keratoses were confirmed by this technique. Most cases were incorporated into an imminent confocal study directed at the Department of Dermatology at the Medical University of Vienna that additionally included follow-up for one year.

3. **Follow-up**

In the event that nevi checked by computerized dermatoscopy did not show any progressions during 3 follow-up visits or 1.5 years we acknowledged this as proof of biologic kindheartedness. Just nevi, however no other benevolent analyses were marked with this kind of ground-truth since dermatologists for the most part don't screen dermatofibromas, seborrheic keratoses, or vascular

injuries. The nearness of progress was surveyed by creator HK who has over 20 years of involvement in advanced dermatoscopic development.

4. **Consensus**

For common kindhearted cases without histopathology or follow-up we give a specialist agreement rating of creators PT and HK. We connected the agreement mark just if the two creators freely gave the equivalent unequivocal considerate analysis. Injuries with this sort of ground-truth were generally captured for instructive reasons and did not need further development or biopsy for affirmation.

4.11 Diagnostic Categories

The following description of the diagnostic categories is meant for computer scientists who are not familiar with the dermatology literature:

1. **Akiec**

Actinic Keratoses (Solar Keratoses) and Intraepithelial Carcinoma (Bowen's disease) are common non-invasive, variants of squamous cell carcinoma that can be treated locally without surgery. Some authors regard them as precursors of squamous cell carcinomas and not as actual carcinomas. There is, however, agreement that these lesions may progress to invasive squamous cell carcinoma – which is usually not pigmented. Both neoplasms commonly show surface

scaling and commonly are devoid of pigment. Actinic keratoses are more common on the face and Bowen's disease is more common on other body sites. Because both types are induced by UV-light the surrounding skin is usually typified by severe sun damaged except in cases of Bowen's disease that are caused by human papilloma virus infection and not by UV. Pigmented variants exist for Bowen's disease20 and for actinic keratoses21, and both are included in this set. The dermatoscopic criteria of pigmented actinic keratoses and Bowen's disease are described in detail by Zalaudek and by Cameron.

2. **Bcc**

Basal cell carcinoma is a common variant of epithelial skin cancer that rarely metastasizes but grows destructively if untreated. It appears in different morphologic variants (flat, nodular, pigmented, cystic), which are described in more detail by Lallas.

3. **Bkl**

Benign keratosis" is a generic class that includes seborrheic keratoses ("senile wart"), solar lentigo - which can be regarded a flat variant of seborrheic keratosis - and lichen-planus like keratoses (LPLK), which corresponds to a seborrheic keratosis or a solar lentigo with inflammation and regression. The three subgroups may look different dermatoscopically, but we grouped them together because they are similar biologically and often reported under the same generic term histopathologically. From a dermatoscopic view, lichen planus-like keratoses are especially challenging because they can show morphologic features mimicking melanoma and are often biopsied or excised.

4. **Df**

Dermatofibroma is a benign skin lesion regarded as either a benign proliferation or an inflammatory reaction to minimal trauma. The most common dermatoscopic presentation is reticular lines at the periphery with a central white patch denoting fibrosis.

5. **Nv**

Melanocytic nevi are benign neoplasms of melanocytes and appear in a myriad of variants, which all are included in our series. The variants may differ significantly from a dermatoscopic point of view. In contrast to melanoma they are usually symmetric with regard to the distribution of color and structure.

6. **Mel**

Melanoma is a malignant neoplasm derived from melanocytes that may appear in different variants. If excised in an early stage it can be cured by simple surgical excision. Melanomas can be invasive or non-invasive (in situ). We included all variants of melanoma including melanoma in situ, but did exclude non-pigmented, subungual, ocular or mucosal melanoma.

7. **Vasc**

Vascular skin lesions in the dataset range from cherry angiomas to angiokeratomas and pyogenic granulomas. Hemorrhage is also included in this category. Angiomas are

dermatoscopically characterized by red or purple color and solid, well circumscribed structures known as red clods or lacunes.

The number of images in the datasets does not correspond to the number of unique lesions, because we also provide images of the same lesion taken at different magnifications or angles (Fig below), or with different cameras. This should serve as a natural data-augmentation as it shows random transformations and visualizes both general and local features.

Figure: Large source image were taken from different magnifications and angles

CHAPTER 5

Data Preprocessing

5.1 The libraries required:

Python Libraries: os, cv2, itertools, matplotlib.pyplot, numpy, panda, seaborn, glob, PIL

Pytorch Libraries: torch, torch.autograd, torch.utils.data, torchvision

Sklearn (Scikit-Learn) Libraries: sklearn.metrics, sklearn.model_selection, sklearn.preprocessing

Keras Libraries: keras, keras.utils.np_utils, keras.utils.np_utils, keras.layers.normalization, keras.optimizers, keras.preprocessing.image, keras.callbacks,keras.models, keras.layers

5.2 Making Dictionary of images and labels

Here, we made the image path dictionary by joining the folder path from base directory base_skin_dir and merging the images from both the HAM10000 images part1.zip and HAM10000 images part2.zip folders in jpg format. This is so that we can label these images from 0 to 6 according to the disease. This dictionary is also useful for displaying more human-friendly labels that we have showed later on.

5.3 Reading and Processing Data

We have perused the csv by joining the path of image folder which is the base folder where every one of the pictures are put named base_skin_dir. After that we made some new sections which is effectively comprehended for later reference, for example, we have made segment way which contains the image_id, cell_type which contains the short name of disease type and finally we have

made the segment cell_type_idx in which we arrange the cancer kind in to index from 0 to 6. We also created a new column that is a copy of the lesion_id column. If the same lesion_id is repeated then it is a duplicate, otherwise it is not.

	lesion_id	image_id	dx	dx_type	age	sex	localization	path	cell_type	cell_type_idx	duplicates
0	HAM_0000118	ISIC_0027419	bkl	histo	80.0	male	scalp	../input/ham10000_images_part_1/ISIC_0027419.jpg	Benign keratosis-like lesions	2	duplicated
1	HAM_0000118	ISIC_0025030	bkl	histo	80.0	male	scalp	../input/ham10000_images_part_1/ISIC_0025030.jpg	Benign keratosis-like lesions	2	duplicated
2	HAM_0002730	ISIC_0026769	bkl	histo	80.0	male	scalp	../input/ham10000_images_part_1/ISIC_0026769.jpg	Benign keratosis-like lesions	2	duplicated
3	HAM_0002730	ISIC_0025661	bkl	histo	80.0	male	scalp	../input/ham10000_images_part_1/ISIC_0025661.jpg	Benign keratosis-like lesions	2	duplicated
4	HAM_0001466	ISIC_0031633	bkl	histo	75.0	male	ear	../input/ham10000_images_part_2/ISIC_0031633.jpg	Benign keratosis-like lesions	2	duplicated

Figure: The addition of new columns in the dataset

5.4 Data Cleaning

In this step we have checked for Missing values (null values) and datatype of each field and we filled the null values by their mean and then checked the presence of null values until there is any. As from our database, the only age has null values which is 57 so we will fill the null values by their mean.

We identified lesion_id's that have duplicate images and those that have only one image and we filter out images that don't have duplicates. From our database, 5514 unduplicated and 4501 duplicated images. Then we filtered out images that don't have duplicates.

```
unduplicated    5514
duplicated      4501
Name: duplicates, dtype: int64
```

Figure: Number of duplicate images

5.5 Exploratory data analysis (EDA)

Here, we will look into different aspects of this dataset, their distributions and actual counts of this dataset.

I. Number of people affected with each type of cancer

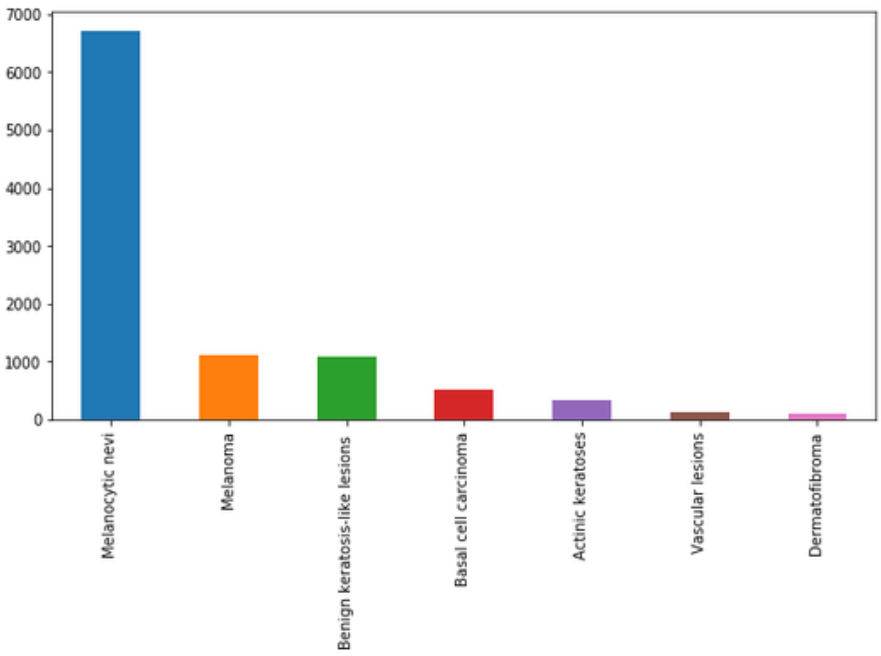

Figure: Number of people affected with each type of cancer

It appears from the above plot that in this dataset cell type Melanecytic nevi has extremely huge numbers of occasions in contrast with other cell types.

II. Plotting of Technical field (ground truth) which is dx_type to see the distribution of its 4 categories of acquired results.

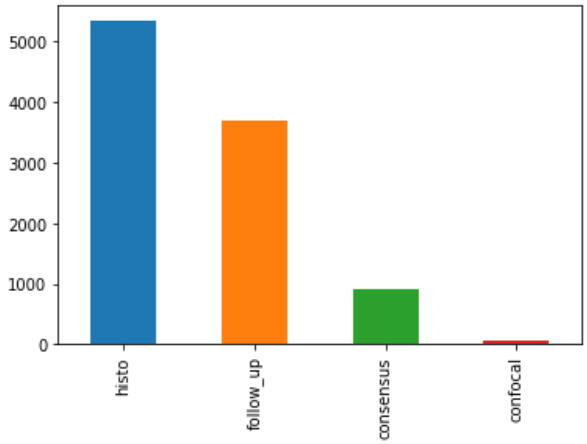

Figure: Distribution of its 4 categories of acquired results

III. Plotting the distribution of localization field in skin area

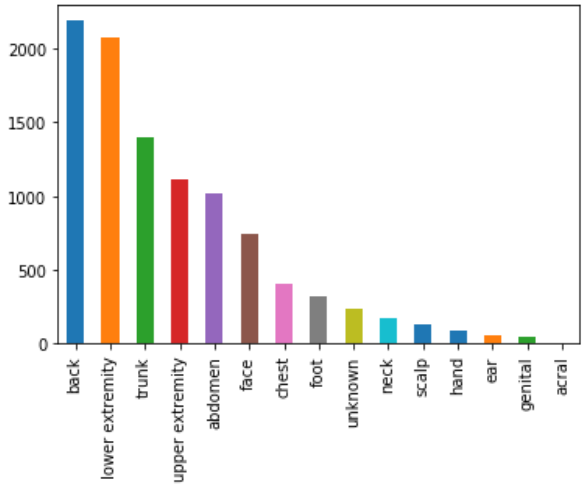

Figure: Distribution of localization field in skin area

It seems back, lower extremity, trunk and upper extremity are heavily compromised regions of skin cancer compared with other body parts.

IV. Distribution of Age

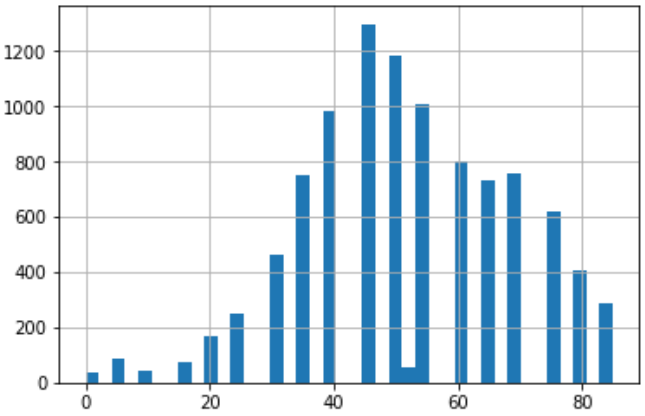

Figure: Distribution of Age

It appears that there are bigger instances of patients having age from 30 to 60.

V. Distribution of Age with distribution of skin cancer types

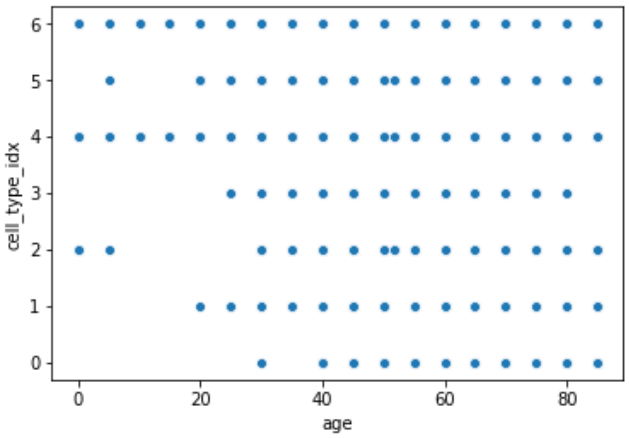

Figure: Distribution of Age with distribution of skin cancer types

It seems that skin cancer types 0, 1, 3 and 5 which are Melanocytic nevi, Melanoma, Basal cell carcinoma and Vascular lesions are not much prevalent below the age of 20 years. But we 4 and 6 which are Actinic keratosis and Dermatofibroma, being active throughout all ages.

VI. Distribution with Gender

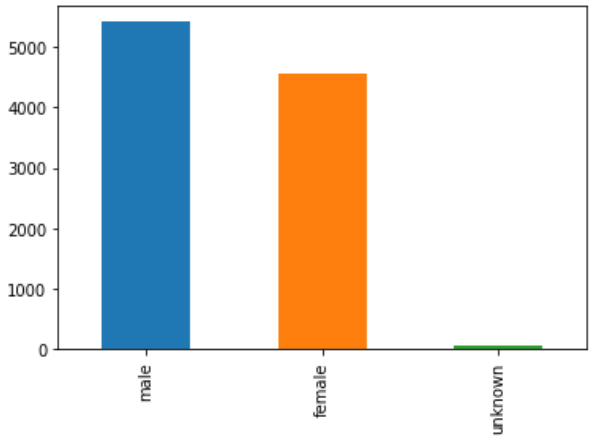

Figure: Distribution with Gender

We see more males are being affected than female. Below the factor plot shows similar results.

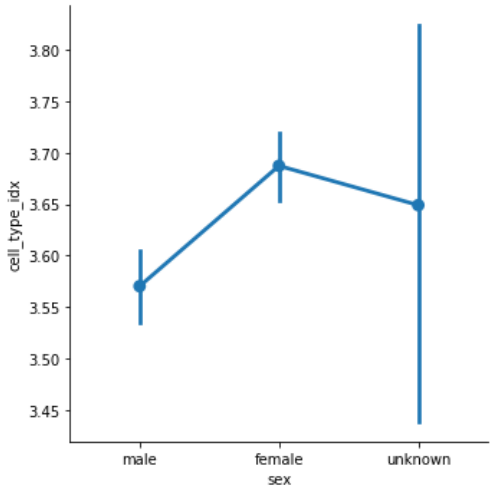

Figure: Factor plot with Sex

CHAPTER 6

Model Training

6.1 Loading and resizing of images

In this step images will be loaded into the column named image from the image path from the image folder. We also resize the images as the original dimension of images are 450 x 600 x3 which TensorFlow can't handle, so that's why we resize it into 100 x 75 x 3. We have also used many transfer learning methods. For VGG, RESNET and DENSENET we need images of 224 x 224 x 3, so we reshaped them into that for that method.

6.2 Test-Train Split

We have split the dataset into training and testing set of 80:20 ratio. This is a common practice in data science.

The training dataset is the underlying dataset used to prepare a calculation to see how to apply advancements, for example, neural systems, to learn and create complex outcomes. It incorporates both info information and the relating anticipated yield. The purpose behind the training dataset is to give our calculation "ground truth" information.

The test dataset, be that as it may, is utilized to survey how well our calculation was prepared with the preparation dataset. We can't just reuse the training dataset in the testing stage in light of the fact that the calculation will definitely "know" the normal yield, which invalidates the point of testing the calculation.

6.3 Normalization

We normalized train set and test set by subtracting from their mean values and then dividing by their standard deviation. This is normalize all values from 0-255 to 0-1.

Normalization is a strategy frequently connected as a major aspect of information planning for AI or machine learning. The objective of normalization is to change the estimations of numeric sections in the dataset to a typical scale, without contorting contrasts in the scopes of qualities. It is required just when features have various ranges like in our situation.

6.4 Label Encoding

Labels are 7 different classes of skin cancer types from 0 to 6. The indexing are as follows:

a. Melanocytic nevi – 0
b. Melanoma – 1
c. Benign keratosis-like lesions - 2
d. Basal cell carcinoma - 3
e. Actinic keratoses - 4
f. Vascular lesions - 5
g. Dermatofibroma – 6

6.5 Splitting training and validation split

We split the train set in two parts: a small fraction (10%) became the validation set which the model is evaluated and the rest (90%) is used to train the model.

Validation dataset is used to minimize over-fitting. We are not adjusting the weights of the network with this data set, we are just verifying that any increase in accuracy over the training dataset actually yields an increase in accuracy over a data set that has not been shown to the network before, or at least the network hasn't trained on it (i.e. validation data set). If the accuracy over the training data set increases, but the accuracy over the validation data set stays the same or decreases, then we are over-fitting your neural network and you should stop training.

6.6 Data Augmentation

From the above statistics of each category, we can see that there is a serious class imbalance in the training data, because almost half of the data is of Melanocytic nevi. To solve this problem, we think we can augment our data.

Also as to abstain from over fitting issue, we have to extend artificially our HAM 10000 dataset. We can make your current dataset considerably bigger. The thought is to change the preparation information with little changes to recreate the varieties.

Methodologies that adjust the preparation information in manners that change the cluster portrayal while keeping the mark the equivalent are known as information increase strategies. Some mainstream augmentations used are grayscales, horizontal flips, vertical flips, random crops, color jitters, translations, rotations, and much more.

By applying only several these changes to our training data, we can without much of a stretch twofold or triple the quantity of preparing models and make an exceptionally robust model.

For our model we randomly rotate some training images by 10 degrees randomly zoom by 10%, some training images randomly shift images horizontally by 10% of the width, randomly shift images vertically by 10% of the height. Once our model is ready, we fit the training dataset.

Chapter 7

Model Building and Evaluation by CNN Model using Keras Sequential API

7.1 What is CNN?

Convolutional Neural Networks are built to detect patterns directly from pixel images with minimum amount of preprocessing. Convolutional Neural Networks (CNN) is aimed towards data processing which is made up of a grid like structure. So a digital image is optimal with this class of neural networks since it is essentially a binary representation of data which is made up of a series of pixels in a grid-like structure and each of the cells holds a pixel value which shows the color and brightness of each pixel.

When we humans see an image, our brain processes this large chunk of information by the use of neurons. They are connected to each other in a network and works in a way such that the entire field of vision is covered. Each neuron responds to stimuli in the receptive field only and following the same analogy, each neuron in a CNN process data only in its receptive field. Layers are arranged so that they break the image into finer, simpler patterns such as lines, curves, etc. and then into more complex shapes such as objects and facial features. In a way, CNN can be used to provide vision to computers.

Convolutional Neural Networks usually have 3 layers - Convolution, Pooling & Fully Connected. Convolution layer is where most of the calculations are done. Here, 2 matrices are taken - one of them is the kernel and the other is the restricted portion of the receptive field. In terms of size, the kernel matrix is smaller than an image, but it has more details. So, if the image is made up of 3

RGB channels, the kernel will have a smaller height and width in spatial terms but the depth extends up to all the channels. In this layer, a dot product between the 2 matrices is done.

The kernel slides across the height and width of the image according to the stride (size of the slide). This produces a 2D representation of that image and this is called the activation map. The activation map gives the response of the kernel at each position of the image.

If input of size is W x W x D and Dout number of kernels with a spatial size of F with stride S and amount of padding P, then the size of output volume can be determined by the following formula:

$$W_{out} = \frac{W - F + 2P}{S} + 1$$

Figure: Formula for Output Weight during Convulation

This will yield an output volume of size W_{out} x W_{out} x D_{out}.

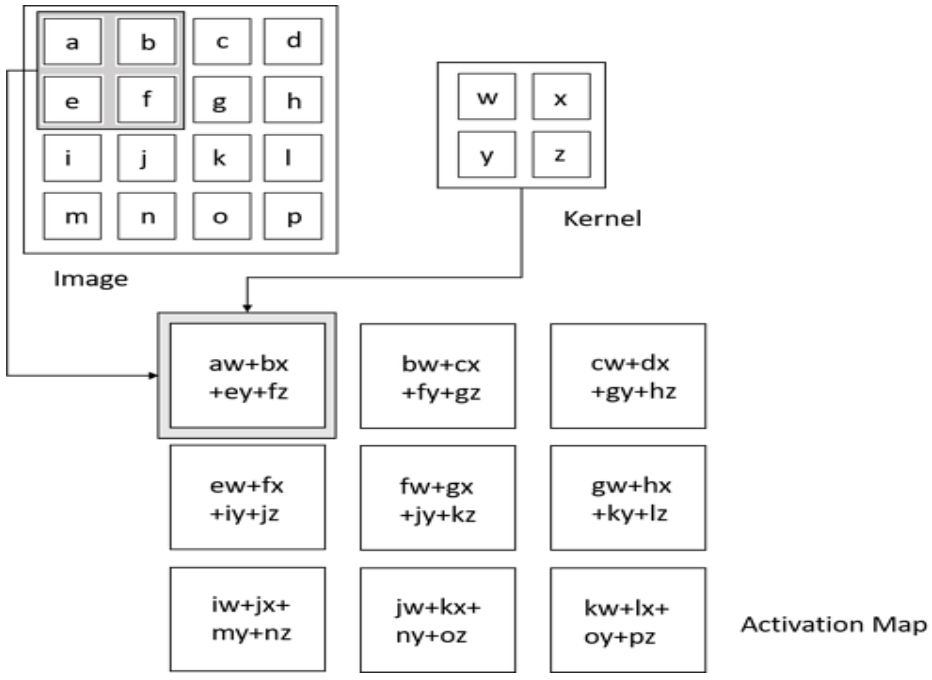

Figure: Convolution Operation (Source: Deep Learning by I. Goodfellow, Y. Bengio)

Next, the pooling layer finds the maximum output at specific locations by summarizing nearby outputs. This reduces the size of the representation and thus decreases computational time and weights. Max pooling is the most popular among several pooling functions, and this is what we have used in our thesis. Max pooling calculates the maximum output from a neighborhood.

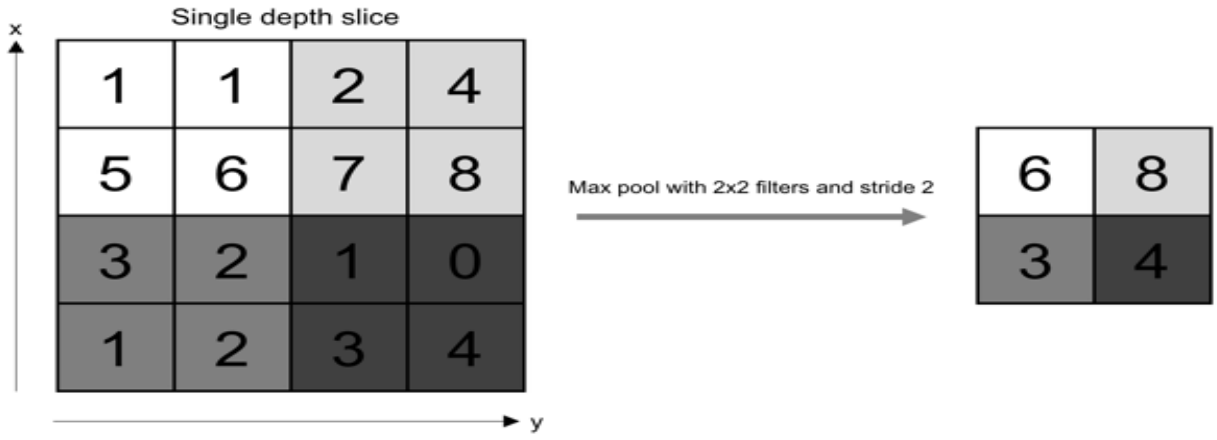

Figure: Pooling Operation (Source: O'Reilly Media)

If we have an activation map of size W x W x D, a pooling kernel of spatial size F, and stride S, then the size of output volume can be determined by the following formula:

$$W_{out} = \frac{W - F}{S} + 1$$

Figure: Output Weight during MaxPooling

This will yield an output volume of size W_{out} x W_{out} x D.

Pooling is essential for the computer to recognize an object which might appear anywhere on the image.

In fully connected layer, the neurons are all connected with the previous and the next layer. It's not possible to calculate this by simple dot multiplication. This layer maps the representation between the input and output.

There is a problem however. Convolution is a linear operation but images not. Immediately after convolutional layer, non-linearity layers are placed so that the activation map is not linear. There are several types of non-linear operations, the most used are described below:

a) Sigmoid: This has the mathematical formula $\sigma(\kappa) = 1/(1+e^{-}\kappa)$. It takes real valued number and converts it into range 0-1. The disadvantage with this non-linearity is that when the activation is at either ends, the gradient nearly nullifies. If local gradient is very close to 0, the backpropagation will almost certainly be 0. Furthermore, if the input data is always positive, then the output will be either all positive or all negative. This produces a zig-zag pattern of gradient for weight.
b) Tanh: The real-valued numbers are converted in range [-1,1]. The activation saturates but the output is zero centered.
c) ReLU: Abbreviation of "The Rectified Linear Unit" calculates the function $f(\kappa)=\max(0,\kappa)$. The activation is threshold at zero.

ReLU is more reliable and the convergence is much faster (by six fold). And this is what we have used in our thesis. A problem might occur during the training stage if a large gradient flows such

that it updates it such that the neuron never gets further updated. We have worked around this by setting a suitable learning rate.

Researchers established that it is typically better to go deeper when it comes to convolutional neural networks. However, it is observed that after traversing some depth, the performance actually degrades. There are better convolutional networks today with a huge database from ImageNet. ImageNet is a dataset of over 15 million labeled high-resolution images with around 22,000 categories. So we have used this to our advantage in increasing our accuracy obtained from the results.

7.2 Our Model

We used the Keras Sequential API, where you have just to add one layer at a time, starting from the input.

The first is the convolutional (Conv2D) layer. It is like a set of learnable filters. We chose to set 32 filters for the two firsts conv2D layers and 64 filters for the two last ones. Each filter transforms a part of the image (defined by the kernel size) using the kernel filter. The kernel filter matrix is applied on the whole image. Filters can be seen as a transformation of the image.

The CNN can isolate features that are useful everywhere from these transformed images (feature maps).

The second important layer in CNN is the pooling (MaxPool2D) layer. This layer simply acts as a down sampling filter. It looks at the 2 neighboring pixels and picks the maximal value. These are

used to reduce computational cost, and to some extent also reduce over fitting. We have to choose the pooling size (i.e. the area size pooled each time) more the pooling dimension is high, more the down sampling is important.

Combining convolutional and pooling layers, CNN are able to combine local features and learn more global features of the image.

Dropout is a regularization method, where a proportion of nodes in the layer are randomly ignored (setting their weights to zero) for each training sample. This drops randomly a proportion of the network and forces the network to learn features in a distributed way. This technique also improves generalization and reduces the over fitting.

'Relu' is the rectifier (activation function max(0,x). The rectifier activation function is used to add non linearity to the network.

The Flatten layer is use to convert the final feature maps into a one single 1D vector. This flattening step is needed so that you can make use of fully connected layers after some convolutional/ maxpool layers. It combines all the found local features of the previous convolutional layers.

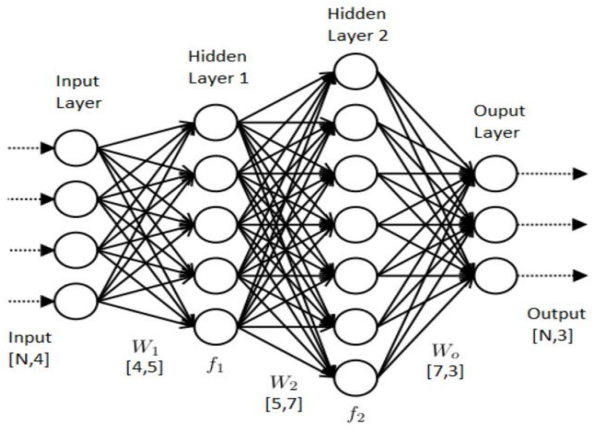

Figure: A model describing forward & backward propagation (Source: Data Science Central)

In the end we used the features in two fully-connected (Dense) layers which is just artificial neural networks (ANN) classifier. In the last layer (Dense(10,activation="softmax")) the net outputs distribution of probability of each class. The diagram above shows 1D vector as the input layer.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 75, 100, 128)      3584
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 75, 100, 128)      147584
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 37, 50, 128)       0
_________________________________________________________________
dropout_1 (Dropout)          (None, 37, 50, 128)       0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 37, 50, 64)        73792
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 37, 50, 64)        36928
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 18, 25, 64)        0
_________________________________________________________________
dropout_2 (Dropout)          (None, 18, 25, 64)        0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 18, 25, 64)        36928
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 18, 25, 64)        36928
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 9, 12, 64)         0
_________________________________________________________________
dropout_3 (Dropout)          (None, 9, 12, 64)         0
_________________________________________________________________
flatten_1 (Flatten)          (None, 6912)              0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               884864
_________________________________________________________________
dropout_4 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 7)                 903
=================================================================
Total params: 1,221,511
Trainable params: 1,221,511
Non-trainable params: 0
_________________________________________________________________
```

Figure: Architecture of our Model

7.3 Setting Optimizer

We have used ADAM Optimizer for both the models. The Adam optimization algorithm is an extension to stochastic gradient descent that has recently seen broader adoption for deep learning applications in computer vision and natural language processing. This function will iteratively improve parameters (filters kernel values, weights and bias of neurons) in order to minimize the loss percentage.

Once our layers are added to the model, we need to set up a score function, a loss function and an optimization algorithm. We define the loss function to measure how poorly our model performs on images with known labels. It is the error rate between the observed labels and the predicted ones. We use a specific form for categorical classifications (>2 classes) called the "categorical_crossentropy". The most important function is the optimizer. This function will iteratively improve parameters (filters kernel values, weights and bias of neurons ...) in order to minimize the loss. We chose Adam optimizer because it combines the advantages of two other extensions of stochastic gradient descent.

Specifically:

1. Adaptive Gradient Algorithm (AdaGrad) that maintains a per-parameter learning rate that improves performance on problems with sparse gradients (e.g. natural language and computer vision problems).

2. Root Mean Square Propagation (RMSProp) that also maintains per-parameter learning rates that are adapted based on the average of recent magnitudes of the gradients for the

weight (e.g. how quickly it is changing). This means the algorithm does well on online and non-stationary problems (e.g. noisy).

Adam realizes the benefits of both AdaGrad and RMSProp. Adam is a popular algorithm in the field of deep learning because it achieves good results fast.

The metric function "accuracy" is used is to evaluate the performance our model. This metric function is similar to the loss function, except that the results from the metric evaluation are not used when training the model (only for evaluation).

7.4 Annealing

In order to make the optimizer converge faster and closest to the global minimum of the loss function, we used an annealing method of the learning rate (LR).

The LR is the step by which the optimizer walks through the 'loss landscape'. The higher LR, the bigger are the steps and the quicker is the convergence. However the sampling is very poor with a high LR and the optimizer could probably fall into a local minima.

It’s better to have a decreasing learning rate during the training to reach efficiently the global minimum of the loss function.

To keep the advantage of the fast computation time with a high LR, we decreased the LR dynamically every X steps (epochs) depending if it is necessary (when accuracy is not improved).

With the ReduceLROnPlateau function from Keras.callbacks, we chose to reduce the LR by half if the accuracy is not improved after 3 epochs.

7.5 Fitting the Model

We fit the model into x_train, y_train. In this step we have chosen batch size of 10 and 50 epochs as small as your batch size will be more efficiently your model will train and we have chosen 50 epochs to give the model sufficient epochs to train.

```
Epoch 46/50
721/721 [==============================] - 16s 22ms/step - loss: 0.5560 - acc: 0.7868 - val_loss: 0.6286 - v
al_acc: 0.7893
Epoch 47/50
721/721 [==============================] - 16s 23ms/step - loss: 0.5555 - acc: 0.7939 - val_loss: 0.6255 - v
al_acc: 0.7805

Epoch 00047: ReduceLROnPlateau reducing learning rate to 1e-05.
Epoch 48/50
721/721 [==============================] - 16s 23ms/step - loss: 0.5516 - acc: 0.7861 - val_loss: 0.6266 - v
al_acc: 0.7855
Epoch 49/50
721/721 [==============================] - 16s 22ms/step - loss: 0.5656 - acc: 0.7881 - val_loss: 0.6278 - v
al_acc: 0.7830
Epoch 50/50
721/721 [==============================] - 16s 23ms/step - loss: 0.5457 - acc: 0.7902 - val_loss: 0.6294 - v
al_acc: 0.7893
```

Figure: Validation Accuracy (the last 5 epochs) for CNN Model

7.6 Model Evaluation using CNN

```
2003/2003 [==============================] - 1s 499us/step
802/802 [==============================] - 0s 395us/step
Validation: accuracy = 0.789277  ;  loss_v = 0.629361
Test: accuracy = 0.757364  ;  loss = 0.632409
```

Figure: Testing and validation accuracy for CNN Model

Now we will check the testing accuracy and validation accuracy of our model, plot confusion matrix and also check the misclassified images count of each type.

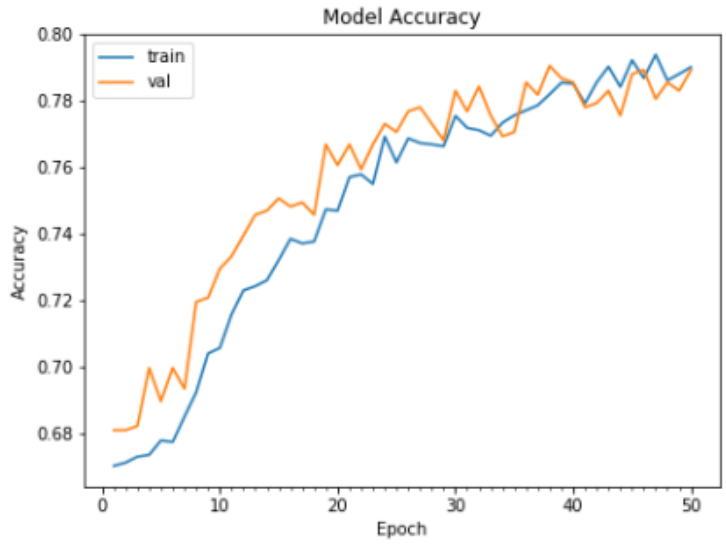

Figure: Model Accuracy for CNN Model

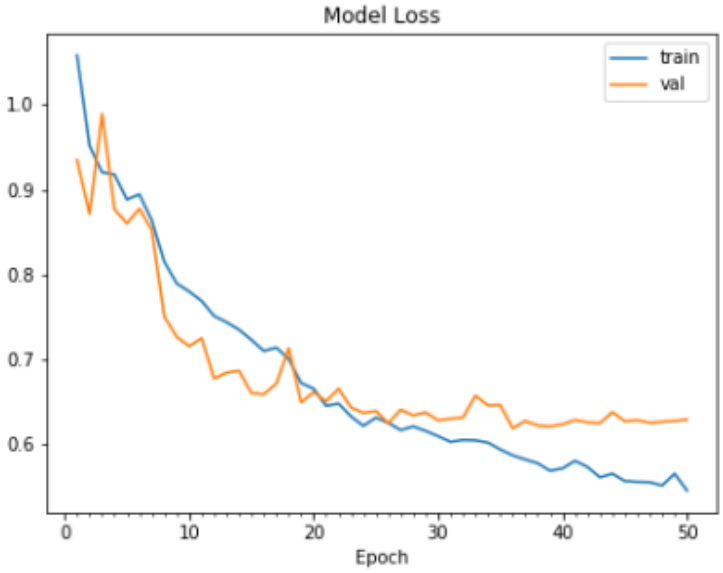

Figure: Model Loss for CNN Model

These graphs show our model accuracy and model loss for both of training and validation set.

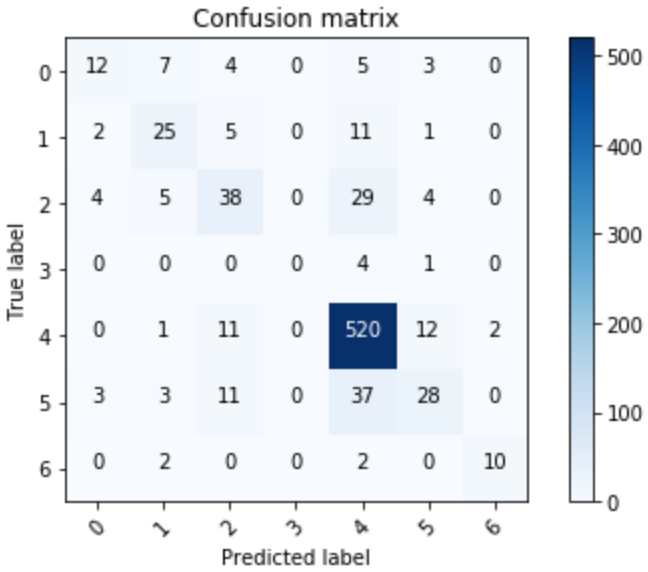

Figure: Confusion matrix for CNN

In the field of machine learning and specifically the problem of statistical classification, a confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one (in unsupervised learning it is usually called a matching matrix). Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class (or vice versa).The name stems from the fact that it makes it easy to see if the system is confusing two classes (i.e. commonly mislabeling one as another).

It is a special kind of contingency table, with two dimensions ("actual" and "predicted"), and identical sets of "classes" in both dimensions (each combination of dimension and class is a variable in the contingency table).

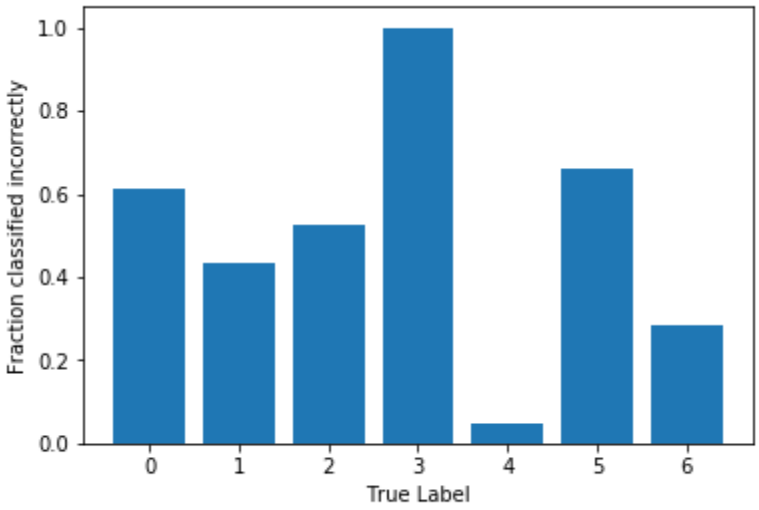

Figure: Fraction classified incorrectly for CNN Model.

This graph gives us an idea about which has the most incorrect predictions. It looks like Basal cell carcinoma (3) has the most incorrect and Actinic keratoses (4) showed most correct results.

CHAPTER 8

Model Building and Evaluation using RESNET50

8.1 What is Resnet (Residual Neural Network)?

A residual neural system (ResNet) is a artificial neural system (ANN), that expands on develops known from pyramidal cells in the cerebral cortex. Residual neural systems do this by using skip associations, or alternate ways to hop over certain layers. Usually, ResNet models are actualized with twofold or triple-layer skips that contain nonlinearities (ReLu) and bunch standardization in the middle. An extra weight matrix might be utilized to get familiar with the skip weights which are known as HighwayNets. In the context of residual neural networks, a non-residual network may be described as a plain network.

One reason for skipping layers is to keep away from the issue of vanishing slopes, by reusing activation from a past layer until the adjoining layer learns its weights. During preparing, the loads adjust to quiet the upstream layer, and intensify the already skipped layer. In the least complex case, just the loads for the nearby layer's association are adjusted, with no unequivocal loads for the upstream layer. This works best when a solitary non-direct layer is ventured over, or when the middle of the road layers are on the whole straight. In the event that not, at that point an unequivocal weight network ought to be educated for the skipped association.

Skipping adequately rearranges the system, utilizing less layers in the underlying preparing stages. This velocities learning by diminishing the effect of disappearing inclinations, as there are less layers to proliferate through. The system at that point continuously reestablishes the skipped layers as it learns the element space. Towards the finish of preparing, when all layers are extended, it

remains nearer to the complex and in this manner adapts quicker. A neural system without remaining parts investigates a greater amount of the element space. This makes it progressively powerless against irritations that reason it to leave the complex, and requires additional preparation information to recoup.

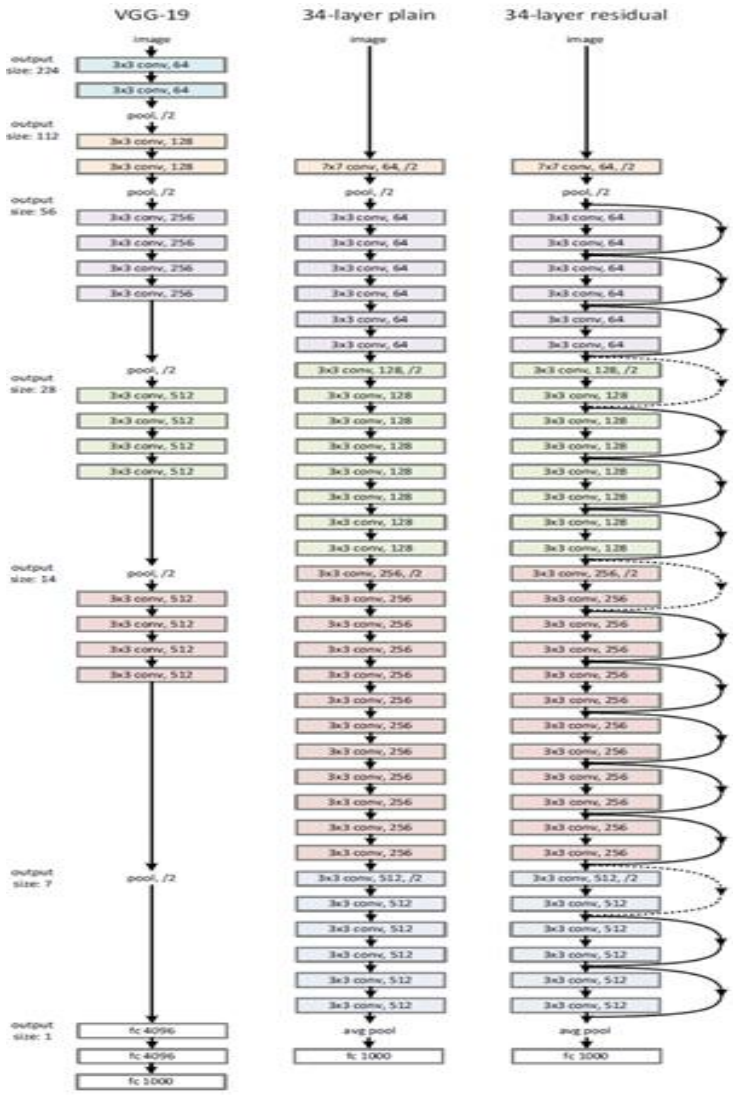

Figure: Difference between Plain and Residual Networks (Source: medium.com)

8.2 ResNet50

Trained on over 1 million pictures from ImageNet, this network consists of 50 layers. Images can be grouped into 1000 object categories. This network is capable of working on a wide range of images due to its rich network. Whereas plain networks are stacked one over the other, residual networks utilize skip-connections as shortcuts to jump over some layers.

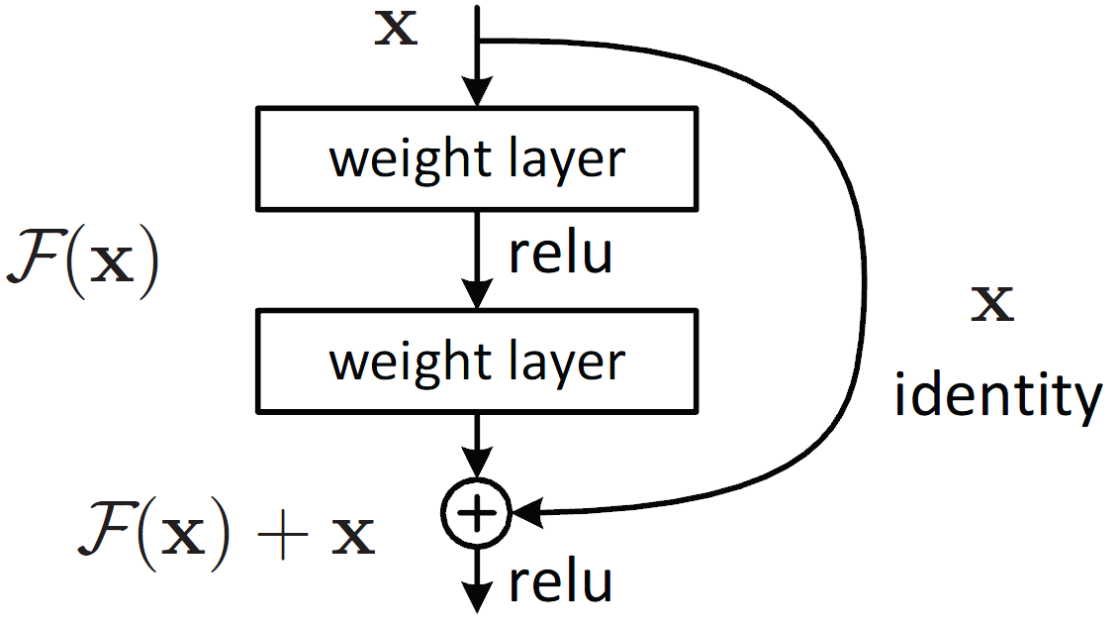

Figure: ResNet50 Algorithm (Source: neurohive.io/en/popular-networks/resnet/)

ResNet models are integrated with double or even triple layer skips that contain non-linarites (ReLu) and in between there are batch normalizations. Skipping over layers are done so that vanishing gradient problem can be avoided. This is done by reusing the activations from a previous layer until the adjacent layers knows its weight. Jumping over layers therefore simplifies the network and uses less number of layers during the training stage. The network has an image input size of 224-by-224.

8.3 Model Evaluation using ResNet50

```
[epoch 10], [iter 100 / 1124], [train loss 0.25444], [train acc 0.90031]
[epoch 10], [iter 200 / 1124], [train loss 0.25240], [train acc 0.90422]
[epoch 10], [iter 300 / 1124], [train loss 0.26108], [train acc 0.89979]
[epoch 10], [iter 400 / 1124], [train loss 0.25361], [train acc 0.90312]
[epoch 10], [iter 500 / 1124], [train loss 0.25032], [train acc 0.90500]
[epoch 10], [iter 600 / 1124], [train loss 0.25092], [train acc 0.90484]
[epoch 10], [iter 700 / 1124], [train loss 0.25462], [train acc 0.90237]
[epoch 10], [iter 800 / 1124], [train loss 0.25331], [train acc 0.90285]
[epoch 10], [iter 900 / 1124], [train loss 0.25478], [train acc 0.90201]
[epoch 10], [iter 1000 / 1124], [train loss 0.25308], [train acc 0.90303]
[epoch 10], [iter 1100 / 1124], [train loss 0.25251], [train acc 0.90347]
------------------------------------------------------------
[epoch 10], [val loss 0.33129], [val acc 0.89720]
------------------------------------------------------------
**************************************************
best record: [epoch 10], [val loss 0.33129], [val acc 0.89720]
**************************************************
```

Figure: Validation Accuracy for the last epoch for ResNet50

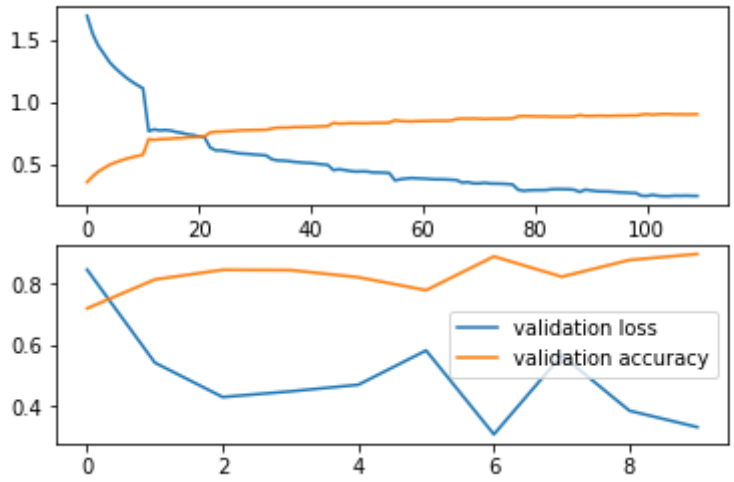

Figure: Validation loss and Validation Accuracy for ResNet50

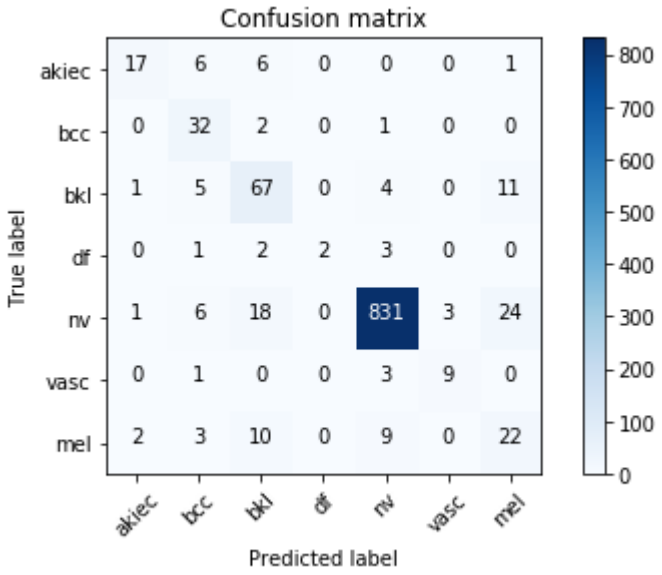

Figure: Confusion Matrix for ResNet50

	precision	recall	f1-score	support
akiec	0.81	0.57	0.67	30
bcc	0.59	0.91	0.72	35
bkl	0.64	0.76	0.69	88
df	1.00	0.25	0.40	8
nv	0.98	0.94	0.96	883
vasc	0.75	0.69	0.72	13
mel	0.38	0.48	0.42	46
accuracy			0.89	1103
macro avg	0.74	0.66	0.65	1103
weighted avg	0.91	0.89	0.89	1103

Figure: Precision, Recall, F1 score and support for ResNet50

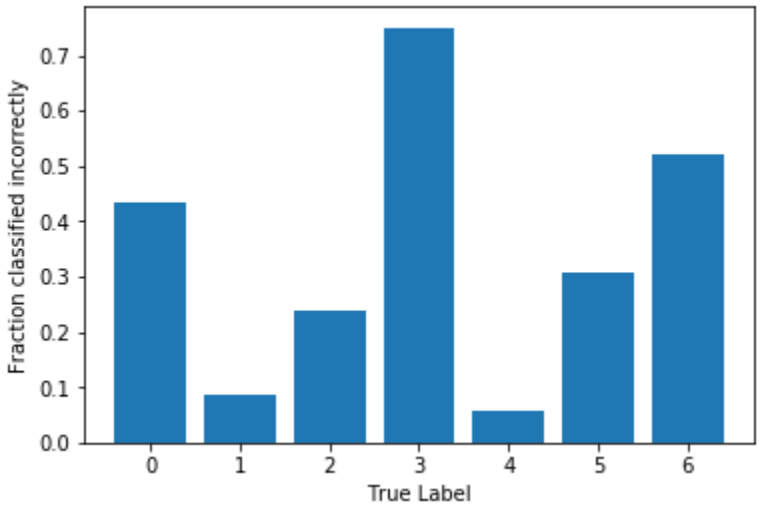

Figure: Fraction classified incorrectly for ResNet50

CHAPTER 9

Model Building and Evaluation using DENSENET121

9.1 DENSENET121

Due to the path from input to output layer being so huge, sometimes the information flowing might be lost before they reach the other end. DenseNets simplify connectivity pattern between layers in the other architectures such as Residual Networks.

DenseNets take advantage of potential of the network by reusing feature. Every layer is connected with each other. Due to this connection, DenseNets need less parameters than traditional CNN because the need to learn redundant feature maps is eliminated. It is proven that in some types of ResNets, not all layers are useful and they can simply be dropped. ResNets are huge with every layer having its own learned weights. To counter this, DenseNets use narrow layers with the addition of smaller sets of new feature-maps. Additionally, DenseNets have every layer with access to the gradients from the loss function and the original input image which helps solve the problem during training stage with information flow and gradients.

DenseNets avoid summing the output feature maps of the layer with incoming feature maps. Instead they concatenate. DenseNets are divided into DenseBlocks where the feature maps remain the same in size and dimensions but the amount of filters applied changes in between. The in-between layers are called Transition Layers and handle downsampling by batch normalization, one 1x1 convolution and 2x2 pooling layers.

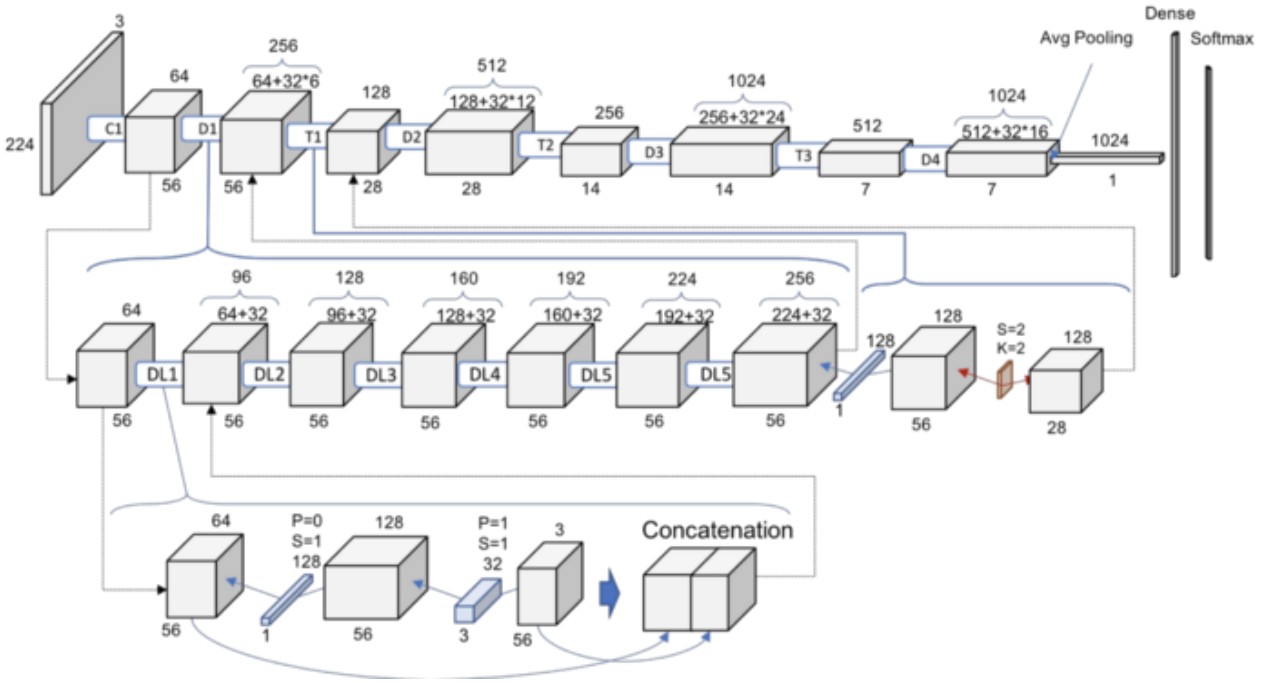

Figure: Full schematic representation of ResNet-121 (Source: towardssdatascience.com)

In the figure above, every layer is adding to preceding volume 32 feature maps. For this reason, it goes from 64 to 256 after 6 whole layers. Transition Block works as 1x1 convolution with 128 filters then a 2x2 pooling with 2 strides. This results in volume and feature maps to be halved after every Transition block. After this we can perform a more expensive 3x3 convolution with the chosen 32 feature maps of growth rate. Then, the input volume and output of the 2 operations are concatenated. This adds new information to the shared knowledge of the network.

9.2 Model Evaluation using DenseNet121

```
[epoch 10], [iter 100 / 1124], [train loss 0.17042], [train acc 0.93719]
[epoch 10], [iter 200 / 1124], [train loss 0.17969], [train acc 0.93437]
[epoch 10], [iter 300 / 1124], [train loss 0.17854], [train acc 0.93469]
[epoch 10], [iter 400 / 1124], [train loss 0.18066], [train acc 0.93391]
[epoch 10], [iter 500 / 1124], [train loss 0.17936], [train acc 0.93469]
[epoch 10], [iter 600 / 1124], [train loss 0.18095], [train acc 0.93458]
[epoch 10], [iter 700 / 1124], [train loss 0.18087], [train acc 0.93384]
[epoch 10], [iter 800 / 1124], [train loss 0.18313], [train acc 0.93305]
[epoch 10], [iter 900 / 1124], [train loss 0.18356], [train acc 0.93292]
[epoch 10], [iter 1000 / 1124], [train loss 0.18233], [train acc 0.93281]
[epoch 10], [iter 1100 / 1124], [train loss 0.18084], [train acc 0.93349]
------------------------------------------------------------
[epoch 10], [val loss 0.34539], [val acc 0.89976]
------------------------------------------------------------
************************************************
best record: [epoch 10], [val loss 0.34539], [val acc 0.89976]
************************************************
```

Figure: Validation Accuracy for the last epoch for DenseNet121

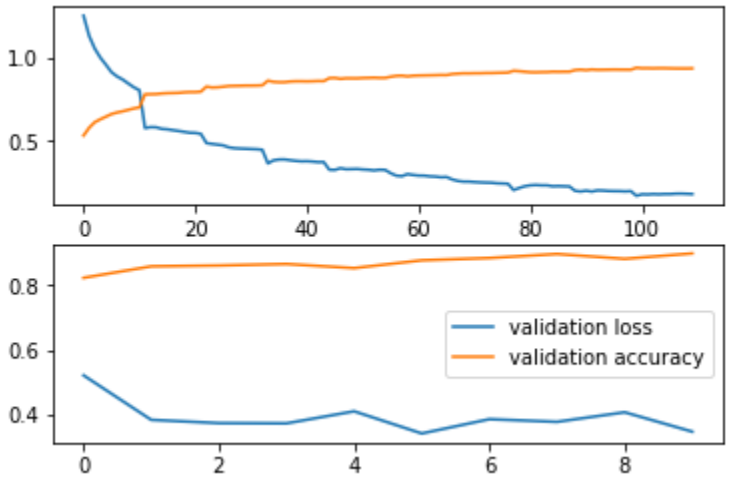

Figure: Validation loss and Validation Accuracy for DenseNet121

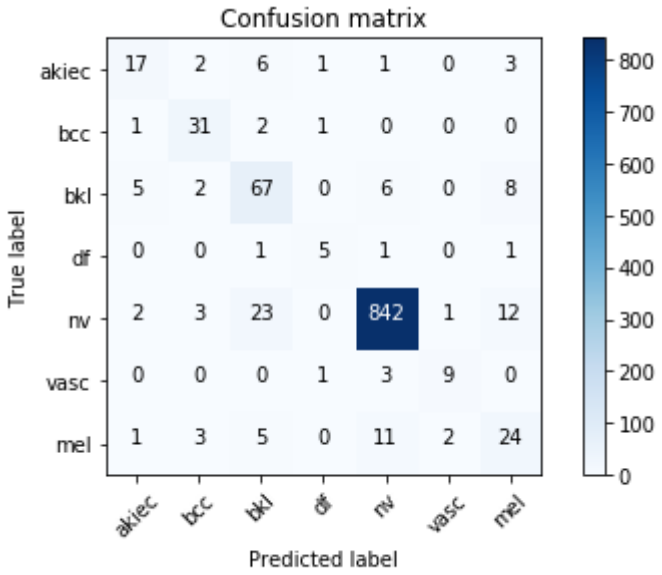

Figure: Confusion Matrix for DenseNet121

	precision	recall	f1-score	support
akiec	0.81	0.57	0.67	30
bcc	0.59	0.91	0.72	35
bkl	0.64	0.76	0.69	88
df	1.00	0.25	0.40	8
nv	0.98	0.94	0.96	883
vasc	0.75	0.69	0.72	13
mel	0.38	0.48	0.42	46
accuracy			0.89	1103
macro avg	0.74	0.66	0.65	1103
weighted avg	0.91	0.89	0.89	1103

Figure: Precision, Recall, F1 score and support for DenseNet121

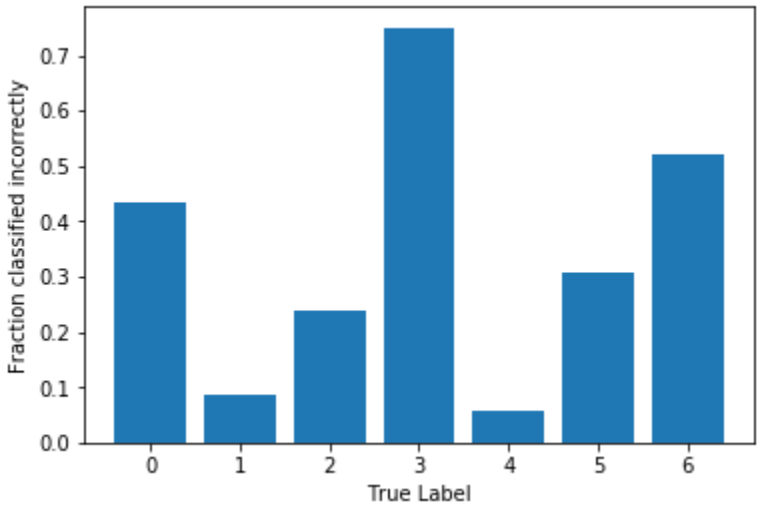

Figure: Fraction classified incorrectly for DenseNet121

CHAPTER 10

Model Building and Evaluation using VGG Models

10.1 VGG16

VGG architecture comes from VGG group, Oxford. They consist of multiple 3 x 3 kernel sized filters appearing sequentially. It is better to have several smaller kernels than one large one because they increase the depth of network and this in turn helps to learn more complex features. VGG-16 D CNN has 144 million parameters with 16 convolutional layers, 5 max pooling layers followed by 3 fully-connected layers and finally a soft-max layer. ReLu is applied to all the hidden layers. The architecture of VGG-16 model is given below:

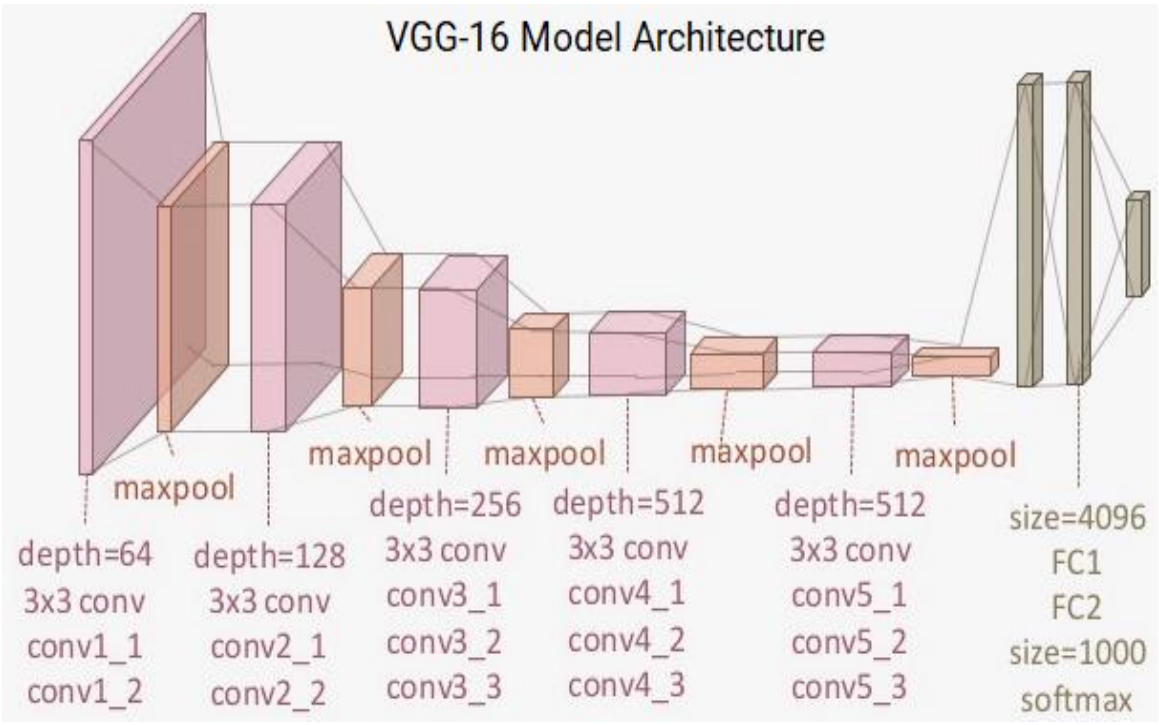

Figure: VGG-16 Model Architecture (Source: towardsdatascience.com)

10.2 VGG19

Just like VGG16, VGG-19 is a convolutional neural system that is prepared on in excess of a million pictures from the ImageNet database. The system is 19 layers profound and can arrange pictures into 1000 item classifications, for example, console, mouse, pencil, and numerous creatures. Therefore, the system has learned rich element portrayals for a wide scope of pictures. This system also has an image input size of 224-by-224.

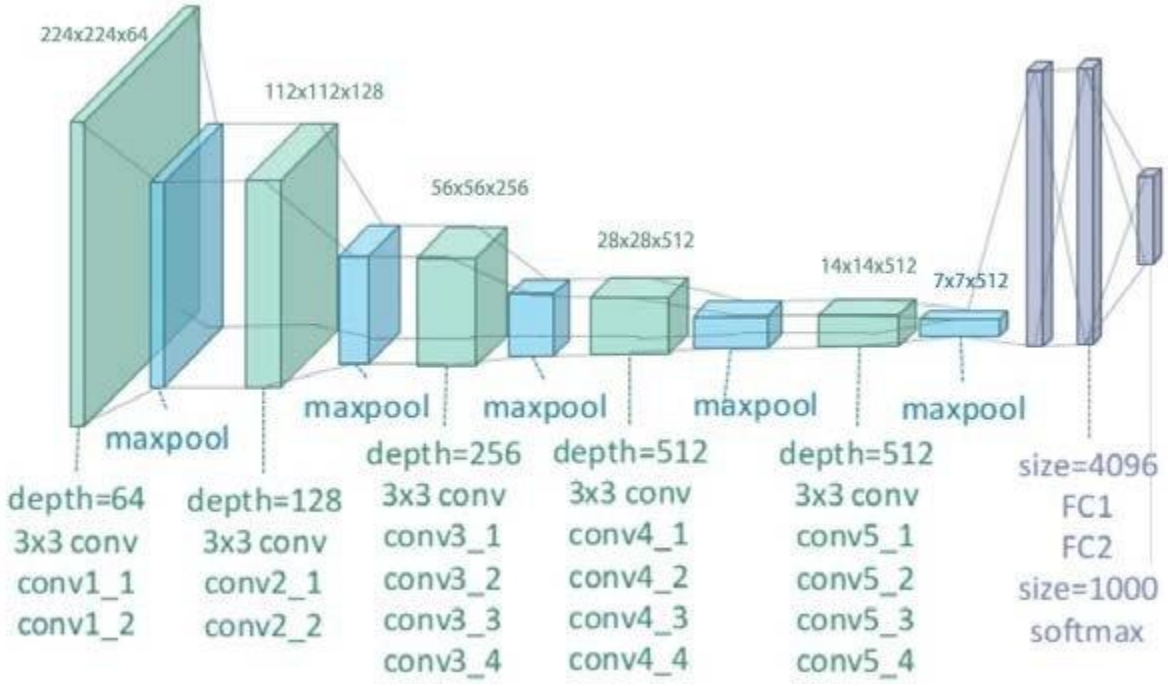

Figure: VGG-19 Model Architecture (Source: researchgate.net)

10.3 Model Evaluation using VGG16 and VGG19

Surprisingly, using VGG16 and VGG19, we failed to get a desired outcome with 10 epochs. Both the models showed similar outcomes, so we showed the last two epochs out of the ten.

```
[epoch 9], [iter 100 / 1124], [train loss 1.94475], [train acc 0.15313]
[epoch 9], [iter 200 / 1124], [train loss 1.94549], [train acc 0.15391]

[epoch 9], [iter 300 / 1124], [train loss 1.94531], [train acc 0.15687]
[epoch 9], [iter 400 / 1124], [train loss 1.94527], [train acc 0.15508]
[epoch 9], [iter 500 / 1124], [train loss 1.94486], [train acc 0.15669]
[epoch 9], [iter 600 / 1124], [train loss 1.94456], [train acc 0.15812]
[epoch 9], [iter 700 / 1124], [train loss 1.94441], [train acc 0.15848]
[epoch 9], [iter 800 / 1124], [train loss 1.94427], [train acc 0.15809]
[epoch 9], [iter 900 / 1124], [train loss 1.94450], [train acc 0.15840]
[epoch 9], [iter 1000 / 1124], [train loss 1.94404], [train acc 0.15969]
[epoch 9], [iter 1100 / 1124], [train loss 1.94414], [train acc 0.15963]
------------------------------------------------------------
[epoch 9], [val loss 1.84751], [val acc 0.80256]
```

Figure: Epoch 9 of VGG16 showing training accuracy and validation accuracy

```
[epoch 10], [iter 100 / 1124], [train loss 1.94283], [train acc 0.16812]
[epoch 10], [iter 200 / 1124], [train loss 1.94390], [train acc 0.16391]
[epoch 10], [iter 300 / 1124], [train loss 1.94400], [train acc 0.16135]
[epoch 10], [iter 400 / 1124], [train loss 1.94392], [train acc 0.15844]
[epoch 10], [iter 500 / 1124], [train loss 1.94372], [train acc 0.15950]
[epoch 10], [iter 600 / 1124], [train loss 1.94438], [train acc 0.15833]
[epoch 10], [iter 700 / 1124], [train loss 1.94407], [train acc 0.15978]
[epoch 10], [iter 800 / 1124], [train loss 1.94411], [train acc 0.15941]
[epoch 10], [iter 900 / 1124], [train loss 1.94400], [train acc 0.16056]
[epoch 10], [iter 1000 / 1124], [train loss 1.94400], [train acc 0.16044]
[epoch 10], [iter 1100 / 1124], [train loss 1.94396], [train acc 0.16037]
------------------------------------------------------------
[epoch 10], [val loss 1.91417], [val acc 0.00714]
------------------------------------------------------------
```

Figure: Epoch 10 of VGG16 showing training accuracy and validation accuracy

Both the epoch shows high irregularity in validation accuracies and the failure in having reasonable training accuracy leads to under-fitting. As a solution to this, we chose to apply our model using VGG11 BN (Batch Normalization).

10.4 VGG11 BN (Batch Normalization)

VGG-11 already obtains 10.4% error rate. VGG uses pre trained data to increase accuracy. VGG11 has 133 million parameters whereas VGG16 has 138 million and VGG19 has 144 million parameters. VGG11 is the predecessor to VGG16 and VGG19.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure: Architecture of VGG Models

The figure above shows the various models of VGG models. VGG11 BN is the first model in the above diagram and also sometimes known as Model A of VGG.

10.5 Model Evaluation using VGG11 BN

```
------------------------------------------------------------
[epoch 9], [val loss 0.45042], [val acc 0.83815]
------------------------------------------------------------
[epoch 10], [iter 100 / 1124], [train loss 0.36983], [train acc 0.86031]
[epoch 10], [iter 200 / 1124], [train loss 0.37107], [train acc 0.85938]
[epoch 10], [iter 300 / 1124], [train loss 0.36981], [train acc 0.86135]
[epoch 10], [iter 400 / 1124], [train loss 0.37699], [train acc 0.85852]
[epoch 10], [iter 500 / 1124], [train loss 0.37434], [train acc 0.85856]
[epoch 10], [iter 600 / 1124], [train loss 0.37031], [train acc 0.86021]
[epoch 10], [iter 700 / 1124], [train loss 0.37153], [train acc 0.85848]
[epoch 10], [iter 800 / 1124], [train loss 0.36963], [train acc 0.85906]
[epoch 10], [iter 900 / 1124], [train loss 0.36968], [train acc 0.86007]
[epoch 10], [iter 1000 / 1124], [train loss 0.37009], [train acc 0.85966]
[epoch 10], [iter 1100 / 1124], [train loss 0.37495], [train acc 0.85847]
------------------------------------------------------------
[epoch 10], [val loss 0.41358], [val acc 0.84810]
------------------------------------------------------------
```

Figure: Validation Accuracy for the last epoch for VGG11

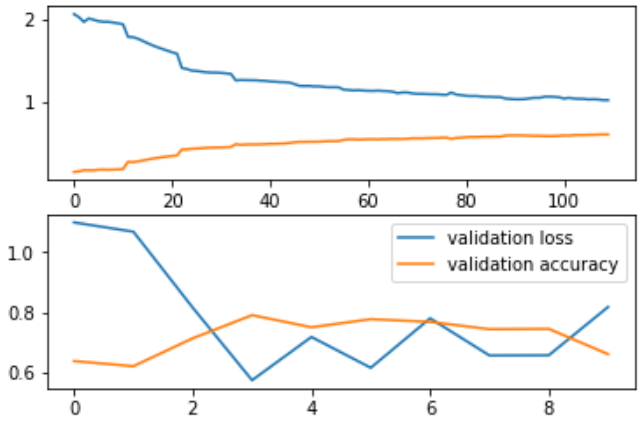

Figure: Validation loss and Validation Accuracy for VGG11

	precision	recall	f1-score	support
akiec	0.55	0.57	0.56	30
bcc	0.63	0.94	0.76	35
bkl	0.48	0.72	0.58	88
df	0.33	0.38	0.35	8
nv	0.98	0.89	0.93	883
vasc	0.91	0.77	0.83	13
mel	0.40	0.63	0.49	46
accuracy			0.85	1103
macro avg	0.61	0.70	0.64	1103
weighted avg	0.89	0.85	0.87	1103

Figure: Precision, Recall, F1 score and support for VGG11

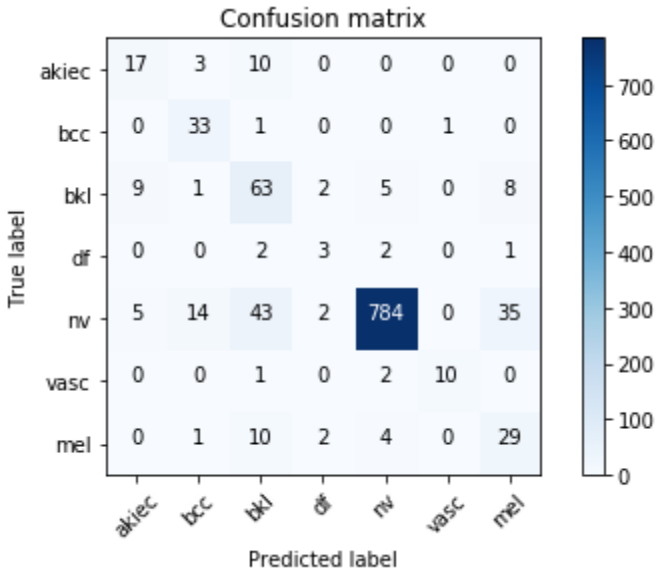

Figure: Confusion Matrix for VGG11

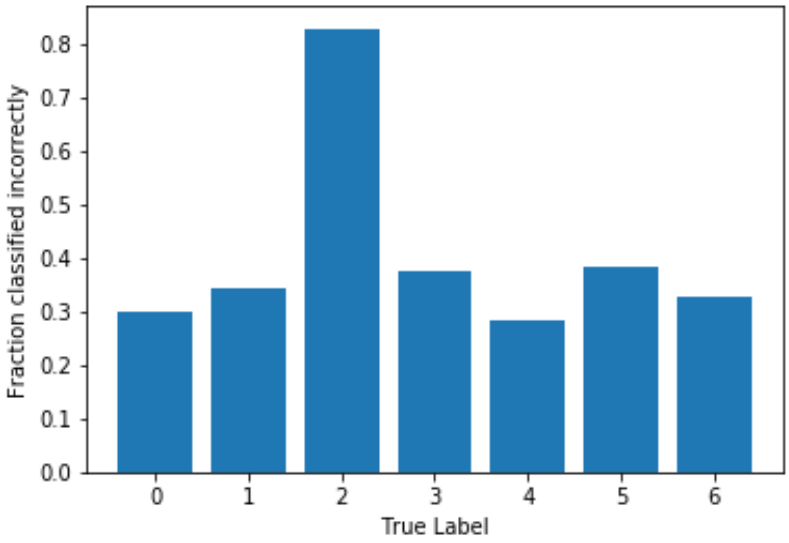

Figure: Fraction classified incorrectly for VGG11

CHAPTER 11

Conclusion

This paper discusses about the concept about identifying different types of skin cancer using a CNN model by Keras Sequential API. We have set our kernel size for our convolution and filter size for our max-pooling. We finally flattened these into 1D single vector and then into the ANN from Keras. Our database consisted of 10015 images and we split it into 80:20 for test and train. We further split the train set into 10% validation. Using our model, we were getting around 79% accuracy in training with the validation set and around 76% using the test set where we ran our model for 50 epochs. We wanted to increase our accuracy and thus we used some known algorithms, such as VGG11, ResNet50 and DenseNet121 which used pre-trained data from ImageNet. These 'transfer learning' algorithm increases the magnitude of our dataset and thus increases our efficiency of the model. Using these we were able to reach 90% accuracy in training and minimal amount of loss. For all of these models, we used 10 epochs and for evaluation we drew graphs of loss and accuracy, confusion matrix, incorrect predictions and so on. We also calculated precision, support, F1 score and recall. We believe that a system which is able to detect 90% accurately just with the image can work in areas where this model is efficient in comparison to detection with human eyes.

References

1. Binder M, Steiner A, Schwarz M, Knollmayer S, Wolff K, PehambergerH: “Application of an artificial neural network in epiluminescencemicroscopy pattern analysis of pigmented skin lesions:a pilot study.” Br J Dermatol1994, 130:460-465.
2. G. Larsson, M. Maire and G. Shakhnarovich, “FractalNet: Ultra-Deep Neural Networks without Residuals,” 2017.
3. G. Huang, Z. Liu and L. van der Maaten, “Densely Connected Convolutional Networks,” 2018.
4. https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/
5. https://keras.io/getting-started/sequential-model-guide/
6. https://machinelearningmastery.com/use-pre-trained-vgg-model-classify-objects-photographs/
7. https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5
8. https://tensorflow.rstudio.com/keras/reference/application_vgg.html
9. https://www.nhs.uk/conditions/melanoma-skin-cancer/
10. K. He, X. Zhang, S. Ren and Jian Sun, “Deep Residual Learning for Image Recognition,” 2015.
11. Margarita Sordo“Introduction to Neural NetworksinHealthcare”October 2002.
12. M. L. Astion and P. Wilding, “Application of neural networks to the interpretation of laboratory data in cancer diagnosis,” Clinical Chemistry, vol. 38, no. 1, pp. 34-38, 1991
13. R. Kumar, K. Gredd and J. Schmidhuber, “Highway Networks,” 2015.

14. Piccolo D, Ferrari A, Peris K, Diadone R, Ruggeri B, Chimenti S “Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study”. Br JDermatol2002, 147:481-486
15. S. E. Umbaugh, “Computer vision in medicine: Color metrics and image segmentation methods for skin cancer diagnosis,” Ph.D. dissertation, Elect. Eng. Dep., Univ. Missouri-Rolla, Rolla, MO: UMI Dissertation Services, Ann Arbor, MI, 1990.
16. W. V. Stoecker, C. S. Chiang and R. H. Moss, “Texture in skin images: Comparison of three methods to determine smoothness,” Computerized Med. /mag. Graphics, vol. 16(3), pp. 179-190, 1992.
17. www.cancer.net/patient/Cancer+Types/Melanoma?sectio nTitle=Statistics