DERMOSCOPIC IMAGE CLASSIFICATION OF SKIN LESIONS USING DENSENET

A PROJECT REPORT

Submitted by

ABIRAMI R (811519104005)

ANGELINE JOY ALEX (811519104009)

DURGA S (811519104028)

GOWTHAMY R M (811519104037)

in partial fulfillment of the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

K. RAMAKRISHNAN COLLEGE OF ENGINEERING (AUTONOMOUS),

TRICHY – 621 112.

APRIL 2023

K.RAMAKRISHNAN COLLEGE OF ENGINEERING (AUTONOMOUS)

BONAFIDE CERTIFICATE

Certified that this project report **"DERMOSCOPIC IMAGE CLASSIFICATION OF SKIN LESIONS USING DENSENET"** is the bonafide work of **"ABIRAMI R, ANGELINE JOY ALEX, DURGA S** and **GOWTHAMY R M"** who carried out the project work under my supervision.

SIGNATURE   SIGNATURE

Dr. T.M. NITHYA, M.E, Ph.D.,   Mrs. M. RUBA, M.E, (Ph.D).,

Associate Professor   Assistant Professor

HEAD OF THE DEPARTMENT   **SUPERVISOR**

Computer Science & Engineering   Computer Science & Engineering

K.Ramakrishnan College of   K.Ramakrishnan College of

Engineering (Autonoumous)   Engineering (Autonomous)

Samayapuram,   Samayapuram,

Trichy – 621 112   Trichy – 621 112

Submitted for the Project Viva-Voce Examination held on ………………….

INTERNAL EXAMINER   **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

DECLARATION

I hereby declare that the work entitled **"DERMOSCOPIC IMAGE CLASSIFICATION OF SKIN LESIONS USING DENSENET"** is submitted in partial fulfilment of the requirement for the reward of the degree in B.E., Anna University, Chennai, is a record of our own work carried out by me during the academic year 2022-2023 under the supervision and guidance of **Mrs.M.RUBA, M.E, (Ph.D).**, Assistant professor, **Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering (Autonomous).** The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not been submitted for the award of any degree or diploma, either in this or any other university.

ABIRAMI R

(811519104005)

I certify that the declaration made by above candidate is true.

Mrs. M. RUBA, M.E., (Ph. D).,

Assistant Professor/CSE

DECLARATION

I hereby declare that the work entitled **"DERMOSCOPIC IMAGE CLASSIFICATION OF SKIN LESIONS USING DENSENET"** is submitted in partial fulfillment of the requirement for the reward of the degree in B.E., Anna University, Chennai, is a record of our own work carried out by me during the academic year 2022-2023 under the supervision and guidance of **Mrs.M.RUBA, M.E,(Ph.D).,** Assistant professor, **Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering (Autonomous).** The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not been submitted for the award of any degree or diploma, either in this or any other University.

ANGELINE JOY ALEX

(811519104009)

I certify that the declaration made by above candidate is true.

Mrs. M. RUBA, M.E., Ph. D.,

Assistant Professor/CSE

DECLARATION

I hereby declare that the work entitled **"DERMOSCOPIC IMAGE CLASSIFICATION OF SKIN LESIONS USING DENSENET"** is submitted in partial fulfilment of the requirement for the reward of the degree in B.E., Anna University, Chennai, is a record of our own work carried out by me during the academic year 2022-2023 under the supervision and guidance of **Mrs. M.RUBA, M.E,(Ph.D).,**Assistant professor, **Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering (Autonomous).** The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not been submitted for the award of any degree or diploma, either in this or any other University.

DURGA S

(811519104028)

I certify that the declaration made by above candidate is true.

Mrs. M. RUBA, M.E., (Ph.D).,

Assistant Professor/CSE

DECLARATION

I hereby declare that the work entitled **"DERMOSCOPIC IMAGE CLASSIFICATION OF SKIN LESIONS USING DENSENET"** is submitted in partial fulfilment of the requirement for the reward of the degree in B.E., Anna University, Chennai, is a record of our own work carried out by me during the academic year 2022-2023 under the supervision and guidance of **Mrs. M.RUBA, M.E,(Ph.D).,**Assistant professor, **Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering (Autonomous).** The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not been submitted for the award of any degree or diploma, either in this or any other University.

GOWTHAMY R M

(811519104037)

I certify that the declaration made by above candidate is true.

Mrs. M. RUBA, M.E., (Ph.D).,

Assistant Professor/CSE

ABSTRACT

Globally skin cancer is one of the main causes of death in humans. Early diagnosis plays a major role in increasing the prevention of death rate caused due to any kind of cancer. Conventional diagnosis of skin cancer is a tedious and time-consuming process. To overcome this an automated skin lesion classification must develop. Automated skin lesion classification is a challenging task due to the fine-grained variability in the visibility of skin lesions. In this work dermoscopic images are obtained from the International Skin Image Collaboration Archive 2016. In the proposed method the analysis and classification of skin lesions is done with the help of a Convolution Neural Network (CNN) along with the hand-crafted features of dermoscopic image using Scattered Wavelet Transform as additional input to the fully connected layer of CNN, which leads to an improvement in the accuracy for identifying Melanoma and different skin lesion classification when compared to the other state of the art methods. When raw dermoscopic image is given as an input to the CNN and feature values of segmented dermoscopic image as input to the fully connected layer as an additional information, the proposed method gives a high classification accuracy .

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Today, skin cancer is a public health and economic issue, that for long Year shave been approached with the same methodology by the dermatology field [1]. This is troublesome when we analyze that for the last 30 years the numbers of cases diagnosed with skin cancer have increased significantly [2]. It is more troublesome when money comes in the equation, seeing that millions of dollars are being spent in the public sector [3]. A major part of this is spent in the individual analysis of the patient. Where the doctor analyzes the lesion and acts on the pieces of evidence seen. If any of these steps were to be optimized, it could mean a decrease in expenditure for the whole dermatology sector. Dermatology is one of the most important fields of medicine, with the cases of skin diseases outpacing hypertension, obesity and cancer summed together. That is accounted because skin diseases are one of the most common human illness, affecting every age, gender and pervading many cultures, summing up to between 30% and 70% of people in the United States.

This means that in any given time at least 1 person, out of 3, will have a skin disease [4]. Therefore, skin diseases are an issue on a global scale, positioning on 18th in a global rank of health burden worldwide [5]. Furthermore, medical imaging can show itself as a resource of high value as dermatology has an extensive list of illness that it has to treat. In addition, the field has developed its own vocabulary to describe the lesions. However,verbal descriptions have their limitations and a good picture can replace successfully many sentences of description and is not susceptible to the bias of the message carrier. Inaddition, recommended way to detect early skin diseases is to be aware of new or

changing skin growths [6]. Analysis with the naked eye is still the first resource used by specialists along with techniques such as ABCDE, that consists of scanning the skin area of interest for asymmetry, border irregularity, uniform colors, large diameters and evolving patches of skin over time [7]. In this way, the analysis from medical images is analogous to the analysis with the naked eye and thus can be applied the same techniques and implications. This supports the idea that skin cancer often is detectable through naked eye and medical photography. Worldwide the most common case of cancer is skin cancer, been melanoma, basal and squamous cell carcinoma (BCC and SCC) the most frequent types of the disease [8].This type of the disease is most frequent in countries with the population with predominant white skin or in countries like Australia or New Zealand [9]. In Brazil, it is estimated that for the biennium of 2018-2019, there will be 165,580 new cases of non- melanoma skin cancer (BCC and SCC mostly) [10]. Moreover, it is visible that the incidence of these types of skin cancer Risen for many years. This increase can be due to the combination of various factors, such as longer longevity of the population, more people being exposed to the sun and better cancer detection [8]. In the United States, the numbers add up to 9,730 deaths estimated for 2017 [6]. Skin cancer accounts for more than 1,688,780 cases (not including carcinoma in situ, nor nonmelanoma cancers) in the US alone in the year of 2017 [6]. Despite skin cancer being the most common type of cancer in society, it does not represent a great death rate in its first stages, since the patient has asurvival rate of 97%. However, if the patients are diagnosed in the later stages the 5-year survival rate decreases to 15%. In Brazil, were expected to occur 114,000 new cases of nonmelanoma skin cancer in 2010. From that, it was expected that 95% were diagnosed in early stages. However, even with early diagnosis this amount of cases means around R$37 million (Reais) to the health

The first is composed mostly of skin cancers and the latter being composed of any lesion that does not pose a major threat. One counter example of this division is the actinic keratosis, that presents itself as a potential SCC, as it has the potential to develop into it. Thus, actinic keratosis is classified as a precancerous lesion [11]. Furthermore, this work analyzed and chose 12 lesions in total, 4 malignant and 8 benign (being 1 precancerous), as seen on Figure 1. The lesions where chosen mainly on the public data available online, to be described in subsection III-A. Seeing the problems involved in diagnosing skin lesions, this work envisions to create a learning model to classify skin lesions in one of 12 conditions of interest. With this purpose, the classifier aims to correct distinguish lesions analyzing clinical images with the condition. Furthermore, this can prove to be a useful tool to aid patients and doctors on a daily basis operation. Furthermore, this work is done with the vision of being the stepping stone for newer approaches that democratize and distribute access to health care. A good lesion classification model may be the motor that will accelerate the construction of tools that puts the possibility of early diagnosis and alert on patient's hands, even far isolated patients, where few doctors can reach. These tools may save many lives and reduce several costs with the treatment of late- stage diseases. The related work in this field proved that there are many algorithms capable of tackling this problem, but there is an astonishing difference between shallow and deep methods in machine learning. With that in view, this work will guide its efforts in using deep neural networks To achieve its main objective. For this to happen, the gathering of good practices and techniques used to approach the classification of clinical images is needed.

1.1.1 SCOPE OF DEEP LERNING

Deep learning algorithms seek to explore and analyze the unknown structure in the input distribution to discover useful representations with the help of feature learning, continually evolving and improvising with the input of additional data.The scope of deep knowledge for its self-adaptive feature is boundless in today's time. The deep learning discipline isn't simple, and its complete capabilities have not yet been explored. But it is a potent branch of machine learning and has a lot of scopes soon. It will adopt a core set of standards and tooling frameworks.It is destined to be endorsed and demanded by several brands and businesses. Fast coding through simplified programming frameworks.It will be increasingly used in the field of designing and building generative designing tools.Deep learning has been one of the most dynamic and versatile branches of data science that is ever- transforming and has immense potential. It stands to be one of the most promising career paths with a diverse range of scopes and opportunities.

1.1.2 OVERVIEW OF DEEP LEARNING

Deep learning models are based on deep neural networks with large set of labeled data and back propagation and forward propagation techniques. These are capable of Supervised and Unsupervised Learning. Create and Train Deep Learning Models Photographic style transfer image colorization generating textures and stylized images Visual and textual question answering visual recognition and description and visual art processing object detection document processing character motion synthesis and editing Person identification face recognition and verification action recognition in videos Human action recognition action recognition classifying and visualizing motion capture sequences Handwriting generation and prediction automate and machine translation Named entity recognition mobile vision and advertising conversational agents calling genetic variants and Bioinformatics Cancer detection X-ray CT reconstruction Epileptic Seizure Prediction hardware acceleration Robotics, speech and audio processing Information retrieval, object recognition and computer vision Financial fraud detection, Medica

1.1.3 CONCEPTS OF DEEP LEARNING

In machine learning we (1) take some data, (2) train a model on that data, and (3) use the trained model to make predictions on new data. The process of training a model can be seen as a learning process where the model is exposed to new, unfamiliar data step by step. At each step, the model makes predictions and gets feedback about how accurate its generated predictions were. This feedback, which is provided in terms of an error according to some measure (for example distance from the correct solution), is used to correct the errors made in prediction. The learning process is often a game of back- andforth in the parameter space: If you tweak a parameter of the model to get a prediction right, the model may have in such that it gets a previously correct prediction wrong. It may take many iterations to train a model with good predictive performance. This iterative predict-and-adjust process continues until the predictions of the model no longer improve

1.1.4 METHODOLGY OF DEEP LEARNING

DL also represents learning methods from data where the computation is done through multi-layer neural networks and processing. The term "Deep" in the deep learning methodology refers to the concept of multiple levels or stages through which data is processed for building a data-driven model

DL can be considered as one of the core technology of AI, a frontier for artificial intelligence, which can be used for building intelligent systems and automation. More importantly, it pushes AI to a new level, termed "Smarter AI".

As DL are capable of learning from data, there is a strong relation of deep learningDL technology is capable to change the current world, particularly, in terms of a powerful computational engine and contribute to technology-driven automation, smart and intelligent systems accordingly, and meets the goal of Industry 4.0.

CHAPTER 2

LITERATURE SURVEY

A literature survey is a text written by someone to consider the critical points of current knowledge including substantive findings, as well as theoretical sections.

2.1 MSIM: MULTISTAGE ILLUMINATION MODELING OF DERMATOLOGICAL PHOTOGRAPHS ILLUMINATION CORRECTED SKIN LESION ANALYSIS

As the time of melanoma detection increases, the risk of spreading of melanoma to other organs through the lymph increases thereby increasing the mortality rate. This risk can be reduced with the help of CAD techniques. Classification of melanoma using CAD systems include machine learning or deep learning based techniques which differ in the steps involved for the task. Machine learning based melanoma classification involve three major stages: categorization, feature extraction, and lesion segmentation. Before segmentation, the dermatoscopy images need to be pre-processed. Basic pre- processing procedures include eliminating varying lighting effects

2.2 FAST DENSITY-BASED LESION DETECTION IN DERMOSCOPY IMAGES

A suitable combination of pre-processing procedures is necessary for accurate lesion segmentation. Prior identification of type of noise or artifacts present in the image is necessary for proper selection of the pre – processing technique. But a generalization of pre – processing techniques is not possible due to variation of artifacts in different dermoscopic images.After pre – processing only segmentation can be performed. Decomposition is the process of dividing the afflicted part of the epidermis from the lesion. Thresholding, region-based, and edge-based methods are three basic categories of segmentation techniques. Clustering

2.3 DEEP LEARNING ENSEMBLES For MELANOMA RECOGNITION IN DERMOSCOPY IMAGES

The advantage with these classifiers is that they can be trained with lesser amount of data. But their robustness to classification is low. Handcrafted feature based diagnostic performance is found to be still unsatisfactory due to high intra - class and low inter - class variations in melanoma. Application of CNN in medical images include segmentation of a desired region or classification of diseases. CNN is used for segmentation of medical images such as x- ray, MRI, ventricle segmentation. CNN is used for diagnosis tasks such as tumour identification, tuberculosis diagnosis, lung cancer screening. There are several works that focus on end to end deep learning based melanoma classification such as deep learning ensembles .

2.4 FAST DENSITY-Based LESION DETECTION IN DERMOSCOPY IMAGES

Accurate lesion segmentation requires a proper combination of pre - processing steps. Prior identification of type of noise or artifacts present in the image is necessary for proper selection of the pre – processing technique. But a generalisation of pre – processing techniques is not possible due to variation of artifacts in different dermoscopic images. After pre – processing only segmentation can be performed. Segmentation means separating the lesion (affected region) from the normal skin region. Segmentation methods can be broadly classified as thresholding, edge based and region based methods Thresholding based segmentation involves clustering.

2.5 AUTOMATIC LESION BOUNDARY DETECTION IN DERMOSCOPY IMAGES USING GRADIENT VECTOR FLOW

Segmentation is done based on the zero-crossings of the laplacian of gaussian they perform poorly when there is a smooth transition between skin and lesion and also without well-defined boundaries resulting in leakage of contour through gaps in the edges. Region based methods include the multi scale region growing.

CHAPTER 3

SYSTEM ANALYSIS

The process of analyzing the system that existed and alterations that are made in the proposed system is stated in system analysis.

3.1 EXISTING SYSTEM

Extensive reviews related to medical imaging using machine learning techniques using SVM have been publishes SVM based, automated diagnostic model skin disease tend to use handcrafted features due to their inability to extract adaptive features. The functional connectivity (FC) patterns representing disease region correlation are popular features of existing SVM based diagnosis model. Despite its popularity SVM has been criticized for its poor performance on raw data and for requiring the expert use of design techniques to extract informative features.

3.1.1 LIMITATIONS

- Existing methods takes a long time to process.
- Accuracy of the system is low
- Segmentation process is very complex
- Impossible to classify between the beningn and maligant state.

3.2 PROPOSED SYSTEM

To classify dermoscopic images into benign and malignant a recently developed deep learning algorithm called DenseNet is used. Classification of dermoscopic images are done automatically without applying lesion segmentation or complex image pre - processing. The proposed work involves analysing the performance of this architecture on selecting different optimizers, for selecting best optimizer for the network.

3.2.1 ADVANTAGES

- Quick calculation time
- It is very versatile
- The system has a very high accuracy
- No assumptions about the data- no need to make additional assumptions, tune several parameters, or build a model, this makes curcial in non linear data case

CHAPTER 4

SYSTEM REQUIREMENTS

Every project needs certain hardware components or other software resources to be present. These prerequisites are known as system requirements and are often used as a guideline as opposed to an absolute rule.

4.1 SOFTWARE REQUIREMENTS

Operating System : Windows 7 / 8/ 10

Language : Python

IDE : Anaconda, Notebook

4.2 HARDWARE REQUIREMENTS

Processor : Intel core i3

Ram : 8 GB

Hard Disk : 120 GB

CHAPTER 5

MODULE DESCRIPTION

Module description describes about the various modules that are used in the database to perform the task.

5.1 DATASET COLLECTION

Appropriate datasets are required at all stages of object recognition research, starting from training phase to evaluating the performance of recognition algorithms. All the images collected for the dataset were downloaded from the Internet, searched by name on various sources in different languages.

5.2 IMAGE PROCESSING AND LABELLLING

Images downloaded from the Internet were in various formats along with different resolutions and quality. In order to get better feature extraction, final images intended to be used as dataset for deep neural network classifier were preprocessed in order to gain consistency. Furthermore, procedure of image preprocessing involved cropping of all the images manually,in order to highlight the region of interest.

5.3 AUGUMENTATION PROCESS

The main purpose of applying augmentation is to increase the dataset and introduce slight distortion to the images which helps in reducing over fitting during the training stage. Image data augmentation is a technique that can be used to artificially expand the size of a training dataset by creating modified versions of images in the dataset. Training deep learning neural network models on more data can result in more skillful models, and the augmentation techniques can create variations of the images that can improve the ability of the fit models to generalize what they have learned to new images.

5.4 NEURAL NETWORK TRAINING

The main goal of training the network is for neural network to learn the features that distinguish one class from the others. Therefore, when using more augmented images, the chance for the network to learn the appropriate features has been increased.

5.5 TESTING TRAINED MODEL WITH VALUATION DATA

Finally the trained network is used to detect the disease by processing the input images in valuation dataset and results are processed.

CHAPTER 6

SYSTEM DESIGN

System architecture is the process of defining the architecture, modules, interfaces and data for a system to satisfy specified requirements.

6.1 ARCHITECTURE DIAGRAM

An architecture diagram is a graphical representation of a set of concepts that are part of architecture, including their principles, elements and components.

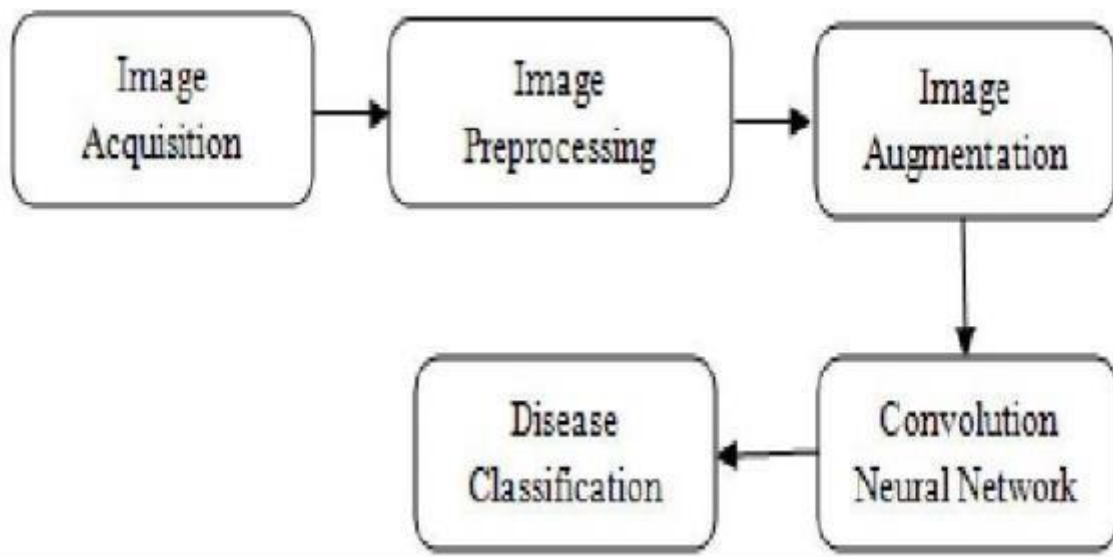**Figure 6.1 Architecture Diagram**

6.2 DATA FLOW DIAGRAM

LEVEL 0:

In level 0 the data are collected and preprocessing data and the collected data are saved into CNN and train by the image classification.
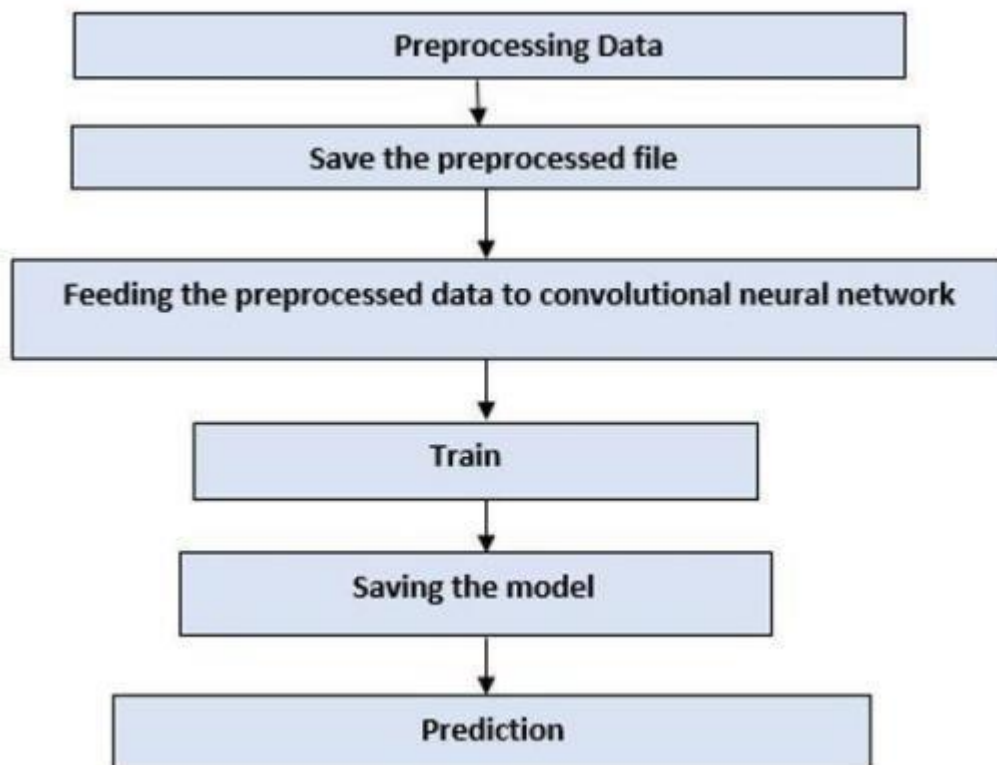

Figure 6.2 DFD Level 0 Diagram

6.3 USE CASE DIAGRAM

In software and systems engineering, a use case is a list of actions or event steps, typically defining the interactions between a role (known in the Unified Modelling Language as an actor) and a system, to achieve a goal. The actor can be a human or other external system. In systems engineering, use cases are used at a higher level than within software engineering, often representing stakeholder goals. The detailed requirements may then be captured in the System Modelling Language (SysML).Use Case analysis is an important and valuable requirement analysis technique that has been widely used in modern software engineering. Use case driven development is a key characteristic of many process models and frameworks such as ICONIX, the Unified Process (UP), the IBM Rational Unified process (RUP), and the Oracle Unified Method (OUM).
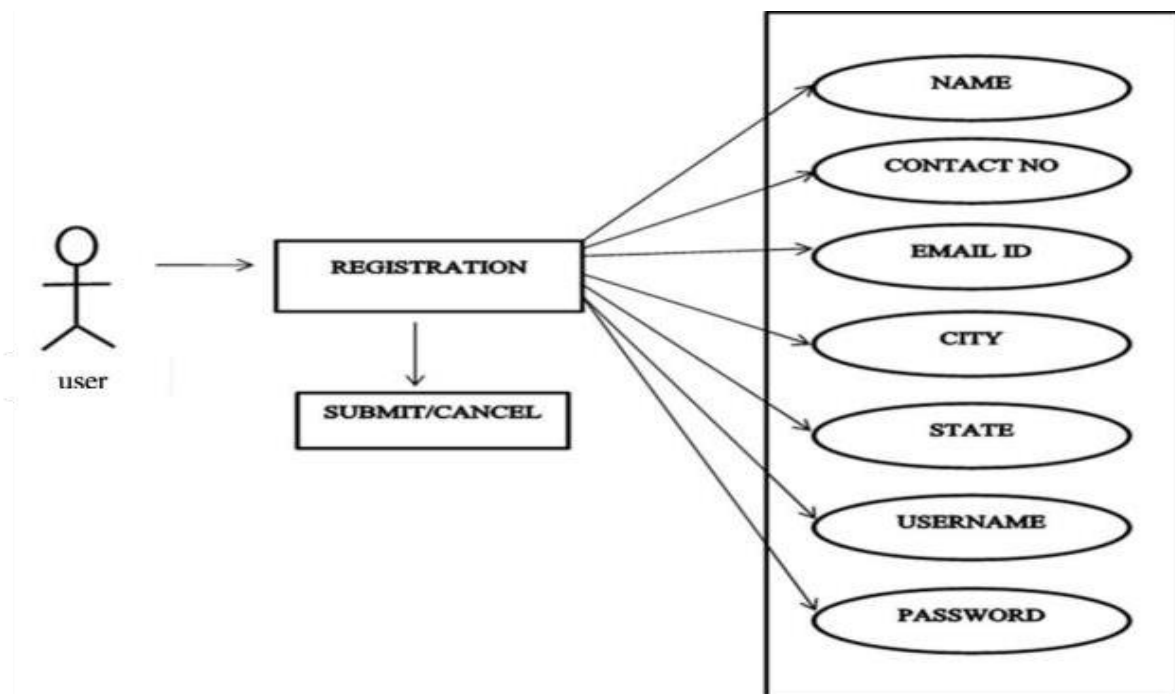
Figure 6.3 usecsase digram

6.4 ER DIAGRAM

An entity-relationship model (or ER model) describes interrelated things of interest in a specific domain of knowledge. A basic ER model is composed by entity types (which classify the things of interest) and specifies relationships that can exist between entities. In software engineering, an ER model is commonly formed to represent things a business needs to remember in order to perform business processes.

Figure 6.4 ER Diagram

6.5 ACTIVITY DIAGRAM

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modelling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.
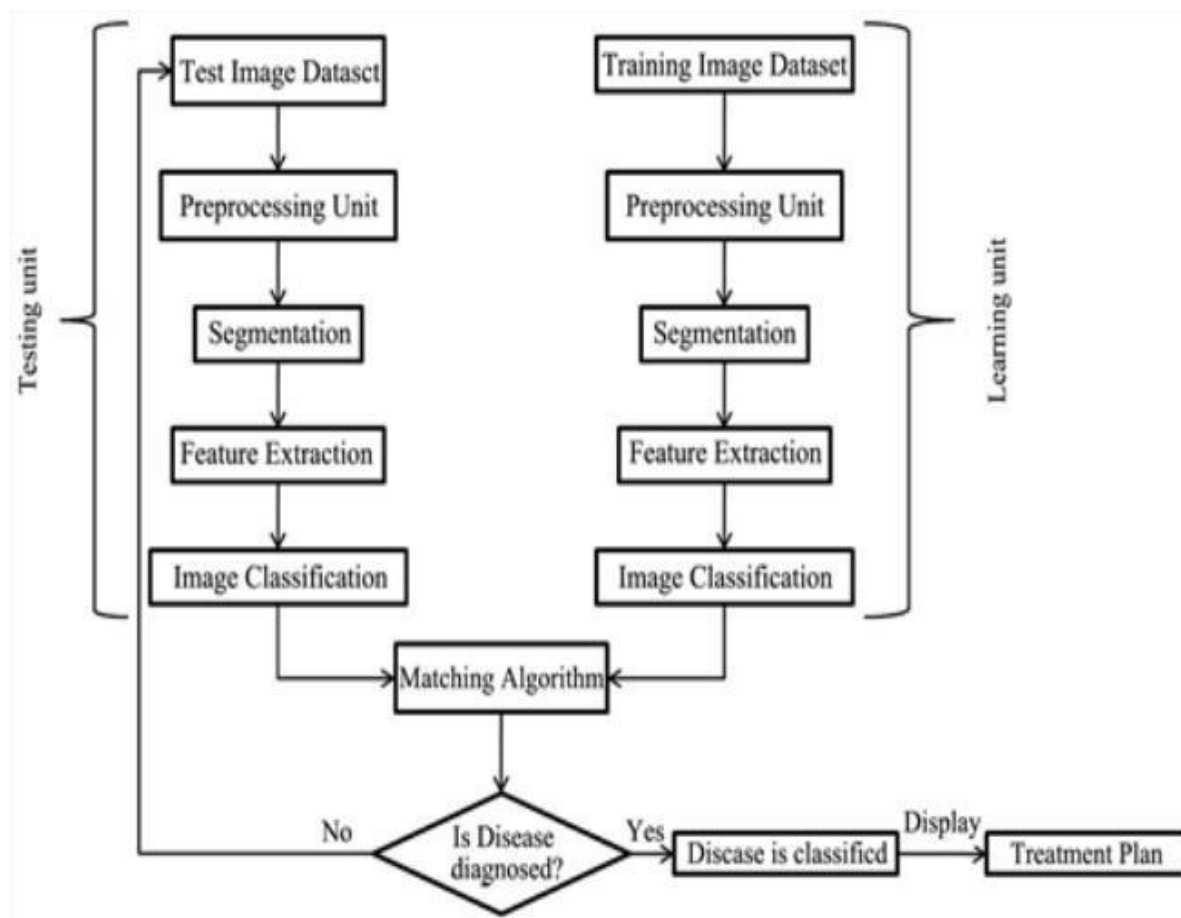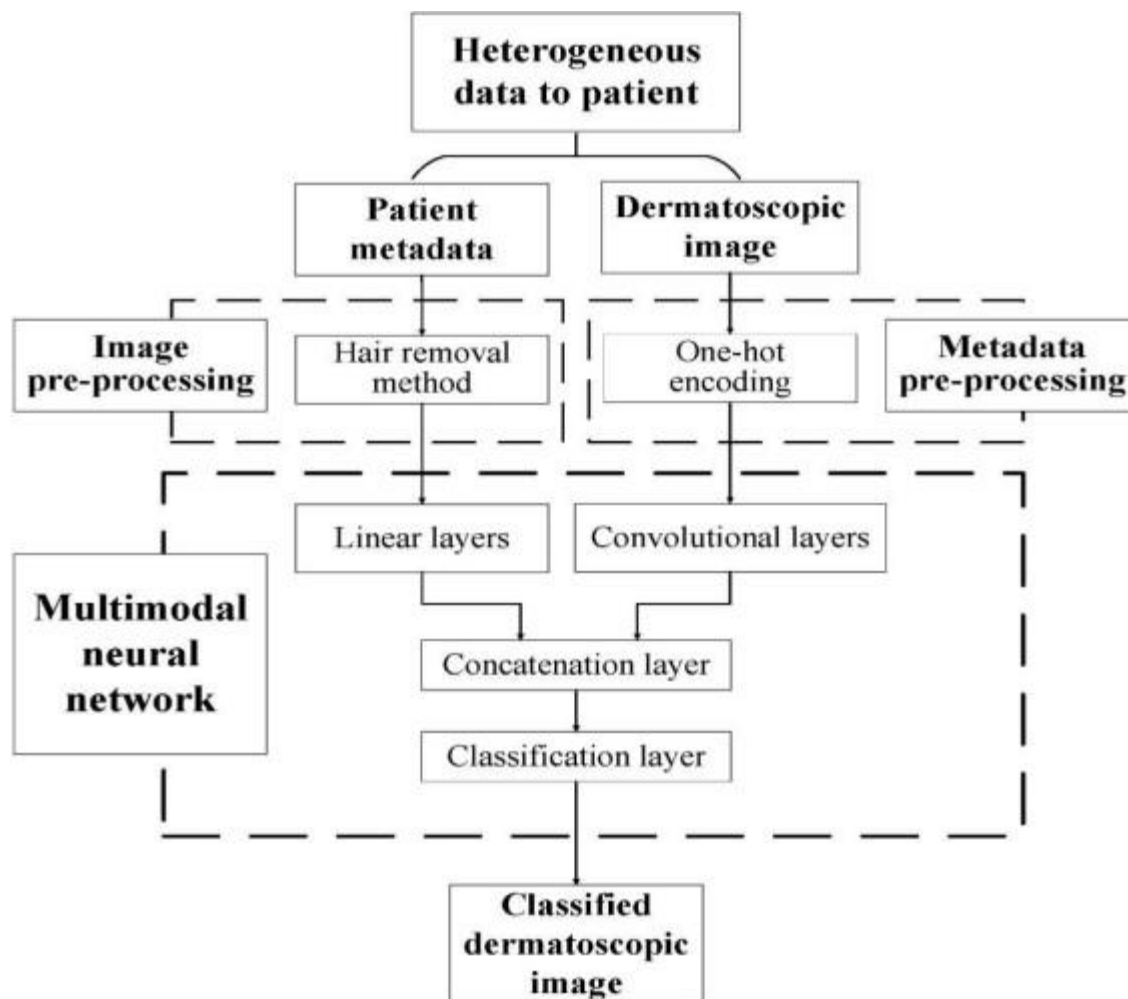

Figure 6.5 Activity Diagram

CHAPTER 7

RESULTS AND ANALYSIS

7.1 RESULT

The training is done on the consisting of 7classes of images. This is a categorical classification problem with basal cell carcinoma, melanocytic nevus actinic keratosis, intraepithelial carcinoma, benign keratinoid lesions,dermatofibroma, and vascular lesions as classes . Before training the dataset is shuffled, then 80% images are used for training and remaining 20% images are used for validation.The dataset consists of image of different dimensions.

7.2 ANALYSIS

These images are resized to common size for the purpose of giving it to the network. All these images are normalized by dividing the image pixels by 255 before giving to the network. The learning rate is initialised to 0.001. The programming is done using python language on Jupyter notebok platform.
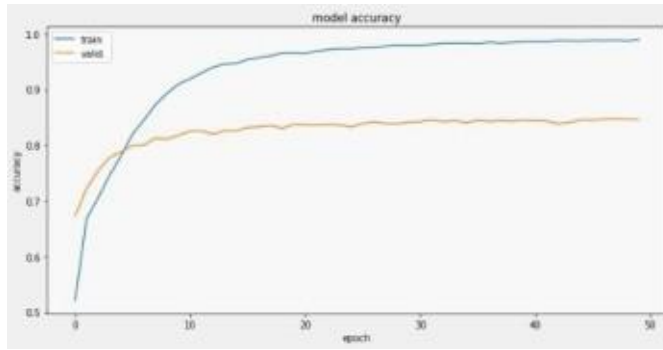
Figure7.2 MODEL ACCURACY
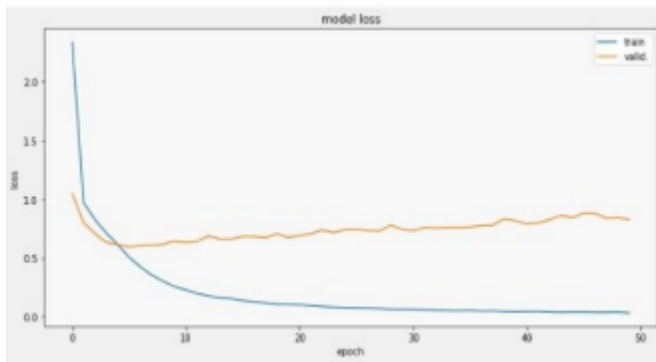
Figure 7.2 : MODEL LOSS

From the metrics DenseNet model trained with Adam optimizer obtained the accuracy greater than 95% for melanoma classification. Using this architecture prediction of new dermoscopic images is done. Depending on the type of the image the network predicts it into one of the 7 classes.

Figure 7.3 : Disease Types

CHAPTER 8

TESTING AND TRAINING

8.1 TRAINING PHASE

For each hidden layer, for each training sample, for each iteration, ignore (zero out) a random fraction, p, of nodes (and corresponding activations).

8.2 TESTING PHASE

use all activations, but reduce them by a factor p (to account for the missing

activations during training).

(a) Standard Neural Net   (b) After applying dropout.
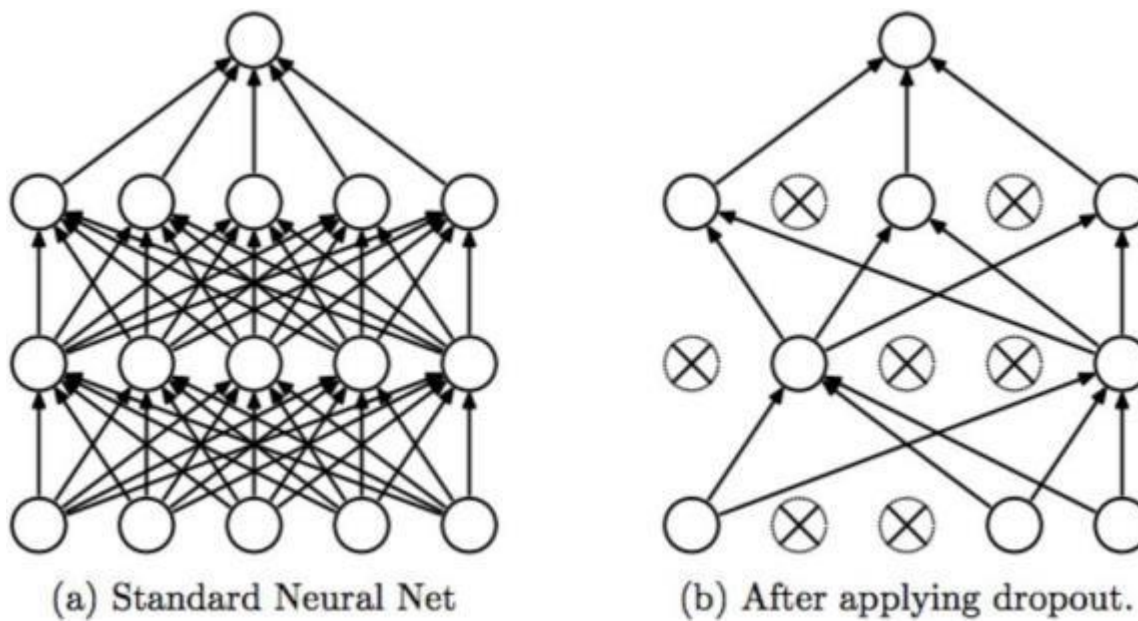
Figure 8.1 Testing Phase

8.3 SOME OBSERVATIONS:

1. Dropout forces a neural network to learn more robust features that are useful in conjunction with many different random subsets of the other neurons.

2. Dropout roughly doubles the number of iterations required to converge However, training time for each epoch is less. With H hidden units, each of which can be dropped, we have

3. 2^H possible models. In testing phase, the entire network is considered and each activation is reduced by a factor *p.*

8.4 FLATTEN LAYERS

Flatten is used to flatten the input**.** For example, if flatten is applied to layer having input shape as (batch_size, 2,2), then the output shape of the layer will be (batch_size, 4)

8.5 DENSE Layer:

The dense layer is a neural network layer that is connected deeply, which means each neuron in the dense layer receives input from all neurons of its previous layer. The dense layer is found to be the most commonly used layer in the models. In the background, the dense layer performs a matrix-vector multiplication.

8.6 FLOW DIAGRAM:

```
┌─────────────────────────┐
│         DATASET         │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│         IMAGE           │
│     PREPROCESSING       │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│         DATA            │
│     AUGMENTATION        │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│        NEURAL           │
│       NETWORK           │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│ TESTING TRAINED MODEL   │
│  WITH VALUATION DATA    │
│                         │
└─────────────────────────┘
```
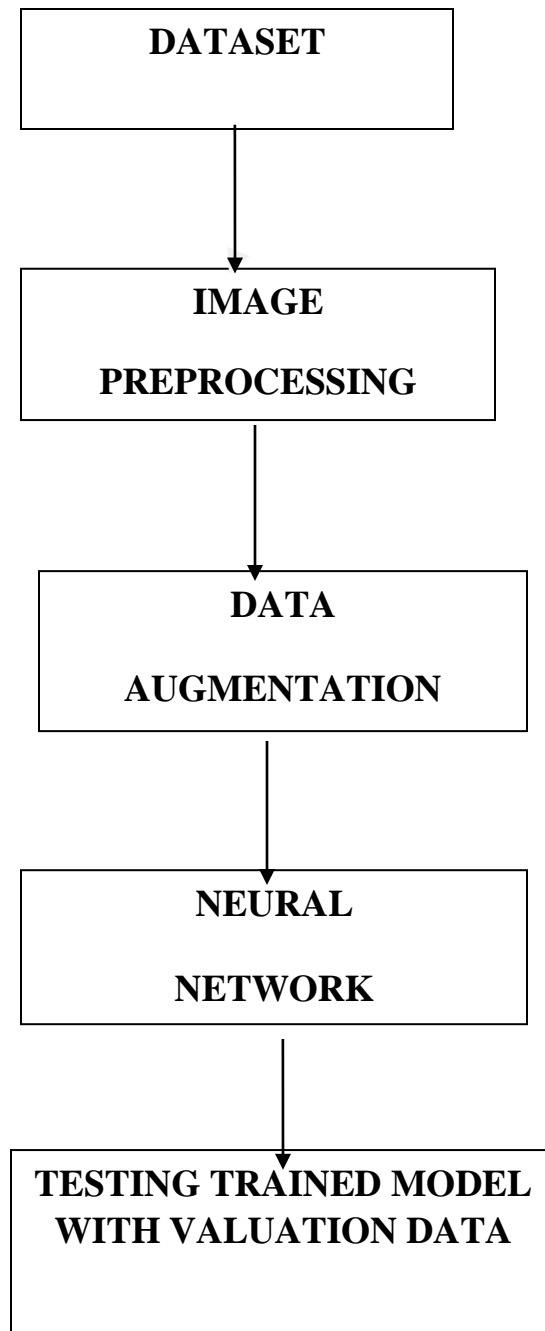
FIGURE 8.6   FLOW DIAGRAM

CHAPTER 9

CONCLUSION AND FUTURE ENHANCEMENTS

Conclusion conveys the completion and defines the limitations that are not processed. Future enhancements provide an innovation that could be made in this project.

9.1 CONCLUSION

Skin cancer continues to impact communities worldwide as a deadly disease. Early detection is important to increase the patients' survival chance as the disease is fatal. In recent years, research on deep learning models in detecting skin cancer has grown substantially, given that the models offer the concept of error-less decision-making for medical applications. Most recently, research effort has slowly progressed towards deep convolutional neural network architectures. From what has been reviewed above, it is clear that considering CNN, data generation, and augmentation aim to mitigate insufficiency of labeled data prone to overfitting and generally improving the performance of skin lesion classification in CAD systems.

9.2 FUTURE ENHANCEMENTS

Tempo, a technology for creating risk-based screening guidelines. Using an AI-based risk model that looks at who was screened and when they got diagnosed, Tempo will recommend a patient return for a mammogram at a specific time point in the future, like six months or three years. The same Tempo policy can be easily adapted to a wide range of possible screening preferences, which would let clinicians pick their desired early-detection-to- screening-cost trade-off.

APPENDIX A

SAMPLE CODINGS

```
"import os\n", "import cv2\n", "import random\n",

"import pandas as pd\n", "import numpy as np\n", "import tensorflow as tf\n",

"import matplotlib.pyplot as plt\n", "from tensorflow.keras import Input\n",

"from sklearn.model_selection import train_test_split\n",

"from tensorflow.keras.models import Model, load_model, save_model\n", "from
tensorflow.keras.layers import Input, Activation, BatchNormalization,

Dropout, Lambda, Conv2D, Conv2DTranspose, MaxPooling2D, concatenate\n",
"from tensorflow.keras.callbacks import EarlyStopping, ModelCheckpoint\n",
"from tensorflow.keras.applications.densenet import DenseNet121\n",

"from tensorflow.keras.preprocessing.image import ImageDataGenerator\n", "from
tensorflow.keras.preprocessing.image import img_to_array\n",

"from tensorflow.keras.preprocessing.image import load_img\n", "from
tensorflow.keras.layers import Dropout,Flatten,Dense,Input\n", "from
tensorflow.keras.models import Model,Sequential\n",

"from tensorflow.keras.optimizers import Adam\n", "from tensorflow.keras import
backend as K"

"cells": [
```

{

"cell_type": "code", "execution_count": 62, "metadata": {

"_cell_guid": "79c7e3d0-c299-4dcb-8224-4455121ee9b0",

"_uuid": "d629ff2d2480ee46fbb7e2d37f6b5fab8052498a", "id": "VEXED9vfSlhI",

"trusted": true

},

"outputs": [], "source": [

]

},

{

"cell_type": "code", "execution_count": 28, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"id": "4PvlEUwrl2_u",

"outputId": "70b3ef14-740a-4fdd-9f62-6841629a7bd9"

},

"outputs": [

{

"name": "stdout", "output_type": "stream", "text": [

"   image MEL NV BCC AKIEC BKL   DF VASC\n",

"0 ISIC_0024306 0.0 1.0 0.0   0.0 0.0 0.0   0.0\n",
"1 ISIC_0024307 0.0 1.0 0.0   0.0 0.0 0.0   0.0\n",
"2 ISIC_0024308 0.0 1.0 0.0   0.0 0.0 0.0   0.0\n",
"3 ISIC_0024309 0.0 1.0 0.0   0.0 0.0 0.0   0.0\n",
"4 ISIC_0024310 1.0 0.0 0.0   0.0 0.0 0.0   0.0\n",

"10015\n",

"Index(['image', 'MEL', 'NV', 'BCC', 'AKIEC', 'BKL', 'DF', 'VASC'],

dtype='object')\n"

]

}

],

"source": [

"filepath=\"yourskincancerdataset folder location\"\n", "df =
pd.read_csv(r'yourfolder/GroundTruth.csv')\n", "print (df.head())\n",

"print (len(df))\n", "print (df.columns)\n",

"labels = ['MEL', 'NV', 'BCC', 'AKIEC', 'BKL', 'DF', 'VASC']"

]

},

{

"cell_type": "code", "execution_count": null, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"id": "p00C54bpl26n",

"outputId": "30b39e84-c54d-45fb-fbe2-e52da558cfdf"

},

"outputs": [], "source": [

"df['image'] =   filepath+\"/images/\"+df['image'].apply(lambda   x: x+
'.jpg')\n",   "print (df.head())"

]

},

{

"cell_type": "code", "execution_count": null, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"id": "HZrM6SzwoUen",

"outputId": "6870061c-304f-45eb-9b9c-cfde4bb97fa0"

},

"outputs": [], "source": [

"labels   =['MEL', 'NV', 'BCC', 'AKIEC', 'BKL', 'DF', 'VASC']\n",

"label_list =[]\n",

"for i in range (len(df)):\n",

"   row = list(df.iloc[i])\n", " del row[0]\n",

"   index = np.argmax(row)\n", "   label = labels[index]\n",

"   label_list.append(label)\n", "df['label'] = label_list\n", "df=df.drop(labels,

axis=1)\n", "print (df.head())"

]

},

{

"cell_type": "code",

"execution_count": 31, "metadata": {

"id": "QyJPPxmXoUbO"

},

"outputs": [], "source": [

"def create_gen():\n",

"   train_generator = tf.keras.preprocessing.image.ImageDataGenerator(\n", "

preprocessing_function=tf.keras.applications.mobilenet_v2.preprocess_input,\n", "
        validation_split=0.1\n",

"   )\n",

"\n",

"   test_generator = tf.keras.preprocessing.image.ImageDataGenerator(\n", "

preprocessing_function=tf.keras.applications.mobilenet_v2.preprocess_input\n", "

)\n",

"\n",

"   train_images = train_generator.flow_from_dataframe(\n", "   dataframe=train_df,rescale = 1./255,\n",

"   x_col='image',\n",

"   y_col='label',\n",

"   target_size=(150,150),\n", "   color_mode='rgb',\n",

"   class_mode='categorical',\n", " batch_size=32,\n",

"   shuffle=True,\n",

"   seed=0,\n",

"   subset='training',\n"

"   rotation_rang=30,\n", "   zoom_range=0.15,\n",

"   width_shift_range=0.2,\n", "   height_shift_range=0.2,\n", "
shear_range=0.15,\n",

"   horizontal_flip=True,\n", "   fill_mode=\"nearest\"\n", ")\n",

"\n",

"   val_images = train_generator.flow_from_dataframe(\n", "

dataframe=train_df,rescale = 1./255,\n",

"   x_col='image',\n",

"   y_col='label',\n",

"   target_size=(150,150),\n", "   color_mode='rgb',\n",

"   class_mode='categorical',\n", " batch_size=32,\n",

"   shuffle=True,\n",

"   seed=0,\n",

"   subset='validation',\n", "   rotation_range=30, \n", " zoom_range=0.15,\n",

"   width_shift_range=0.2,\n", "   height_shift_range=0.2,\n", "
shear_range=0.15,\n",

"   horizontal_flip=True,\n", "   fill_mode=\"nearest\"\n", ")\n",

"\n",

"   test_images = test_generator.flow_from_dataframe(\n", "
       dataframe=test_df,rescale = 1./255,\n",

"   x_col='image',\n",

"   y_col='label',\n",

"   target_size=(150,150),\n", "   color_mode='rgb',\n",

"   class_mode='categorical',\n", " batch_size=32,\n",

"   shuffle=False\n", ")\n",

"   \n",

"return   train_generator,test_generator,train_images,val_images,test_images"

]

},

{

"cell_type": "code", "execution_count": 34, "metadata": {

"id": "1o5WmLtNoUXZ"

},

"execution_count": 35, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"id": "kQgBwQ9boUV0",

"outputId": "0211a42a-ca8e-4488-8049-eda33268440f"

},

"outputs": [

{

"name": "stdout", "output_type": "stream", "text": [

"Found 6309 validated image filenames belonging to 7 classes.\n", "Found 701 validated image filenames belonging to 7 classes.\n", "Found 3005 validated image filenames belonging to 7 classes.\n"

]

}

],

"source": [

"train_generator,test_generator,train_images,val_images,test_images=create_g en()"

]

},

{

"cell_type": "code", "execution_count": 44, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"id": "ekkmB8CvoURh",

"outputId": "69eb64b0-4acb-4c38-f94a-32c783da1c15"

},

"outputs": [

{

"name": "stdout", "output_type": "stream", "text": [

"Model: \"sequential_1\"\n",

}

],

"source": ["model2=Sequential()\n",

"mdl = DenseNet121(input_shape=(150,150,3), \n", "   include_top=False,\n",

"   weights=\"imagenet\")\n", "for layer in base_model.layers:\n",

"   layer.trainable=False\n", "model2.add(mdl)\n",
"model2.add(Dropout(0.5))\n",

"model2.add(Dense(1024,kernel_initializer='he_uniform'))\n",

"model2.add(BatchNormalization())\n", "model2.add(Activation('relu'))\n",

"model2.add(Dropout(0.5))\n", "model2.add(Dense(7,activation='softmax'))\n",

"model2.summary()"

]

},

{

"cell_type": "code", "execution_count": 45, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"id": "R-IKYZ1Evhwt",

"outputId": "eaf52d45-295d-4695-9293-f78a37ce48ae"

},

"outputs": [

{

"name": "stderr", "output_type": "stream", "text": [

"WARNING:absl:`lr` is deprecated, please use `learning_rate` instead, or use the legacy optimizer, e.g.,tf.keras.optimizers.legacy.Adam.\n"

```
]

}

],

"source": [

"OPT = Adam(learning_rate=0.001)\n", "\n",

"model2.compile(loss='categorical_crossentropy',\n", "    metrics=['accuracy'],\n",

"      optimizer=OPT)"

]

},

{

"cell_type": "code", "execution_count": 46, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/"

},

"outputId": "57982231-aa44-474c-e17b-858b3c6efd38"
```

"id": "I34Rxa5boUN9",

"outputId": "57982231-aa44-474c-e17b-858b3c6efd38"

},

"outputs": [

{

"name": "stdout", "output_type": "stream",

]

}

],

"source": [

"history = model2.fit(\n", "   train_images,\n",

"   validation_data=val_images,\n",

"   batch_size = 32,\n", "   epochs=15,\n", ")\n", "model2.save('cl.h5')"

]

},

{

"cell_type": "code", "execution_count": 76, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/", "height": 295

},

"id": "A0pNfMwdAQ4a",

"outputId": "fd695df0-7b88-42d8-c11d-61fbc104b7bc"

},

"outputs": [

{

"data": { "image/png":

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": { "needs_background": "light"

},

"output_type": "display_data"

```
}

],

"source": [ "plt.plot(history.history['accuracy'])\n",
"plt.plot(history.history['val_accuracy'])\n", "plt.title('Model Accuracy')\n",
"plt.ylabel('Accuracy')\n", "plt.xlabel('Epoch')\n",

"plt.legend(['Train', 'Validation'], loc='upper left', bbox_to_anchor=(1,1))\n",
"plt.show()"

]

},

{

"cell_type": "code", "execution_count": 77, "metadata": {

"colab": {

"base_uri": "https://localhost:8080/", "height": 295

},

"id": "cN6N6QisA22W",

"outputId": "4251b8a3-6e03-42b5-9856-bc8063153a7d"

},

"outputs": [
```

```
        ]

        },

        "metadata": { "needs_background": "light"

        },

        "output_type": "display_data"

        }

        ],

        "source": [ "plt.plot(history.history['loss'])\n",
        "plt.plot(history.history['val_loss'])\n", "plt.title('Model Loss')\n",
        "plt.ylabel('Loss')\n",

        "plt.xlabel('Epoch')\n",

        "plt.legend(['Train',            'Validation'],        loc='upper   left',
        bbox_to_anchor=(1,1))\n",        "plt.show()"

        ]

        },

        {

        "cell_type": "code", "execution_count": 71, "metadata": {

        "colab": {
```

"base_uri": "https://localhost:8080/", "height": 299

},

"id": "qurH9Nbv9PmN",

"outputId": "b2ca2c9e-ff64-4aba-9d2f-6d9318ebb9f6"

},

"outputs": [

{

{

"name": "stdout", "output_type": "stream", "text": [

"1/1   [==============================] -   0s   117ms/step\n",

"The image is of skin disease of class : Dermatofibroma\n"

]

},

{

"data": { "image/png":

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": { "needs_background": "light"

},

},

"output_type": "display_data"

}

], "source": [

"labels=['Melanoma', 'elanocytic nevi', 'Basal cell carcinoma', 'Actinic keratoses', 'Benign keratosis-like lesions', 'Dermatofibroma', 'vascular lesions']\n",

"image_path=\"yourfolderlocation/images/ISIC_0024308.jpg\"\n", "img = load_img(image_path, target_size=(150,150,3))\n",

"img = img_to_array(img)\n", "img=img/255\n", "plt.imshow(img)\n", "plt.axis(\"off\")\n",

"img = np.expand_dims(img, axis=0)\n",
"result=np.argmax(model.predict(img))\n", "label=labels[result]\n",

"print(\"The image is of skin disease of class : \"+ label)\n", "plt.title(\"INPUT

IMAGE\")\n",

],

"metadata": { "accelerator": "GPU", "colab": {

"provenance": []

},

"gpuClass": "standard", "kernelspec": { "display_name": "Python 3", "language": "python", "name": "python3"

},

"language_info": { "codemirror_mode": { "name": "ipython", "version": 3

},

"file_extension": ".py", "mimetype": "text/x-python", "name": "python", "nbconvert_exporter": "python", "pygments_lexer": "ipython3", "version": "3.6.4"

}

},"nbformat": 4, "nbformat_minor": 0

}

],

"metadata": { "accelerator": "GPU", "colab": {

"provenance": []

},

"gpuClass": "standard", "kernelspec": { "display_name": "Python 3", "language":
"python", "name": "python3"

},

"language_info": { "codemirror_mode": { "name": "ipython", "version": 3

},

"file_extension": ".py", "mimetype": "text/x-python", "name": "python",
"nbconvert_exporter": "python", "pygments_lexer": "ipython3", "version": "3.6.4"

        }

},"nbformat": 4, "nbformat_minor": 0

}

Vascular Lesions:

This picture shows that the patient has Vascular lesions which is just a simple birthmark.

```
1/1 [==============================] - 5s 5s/step
The image is of skin disease of class : vascular lesions
```


INPUT IMAGE

Elanocytic nevi

This image shows that the patient has elanocytic nevi which is a non cancerous pigment.

```
1/1 [==============================] - 0s 267ms/step
The image is of skin disease of class : elanocytic nevi
```
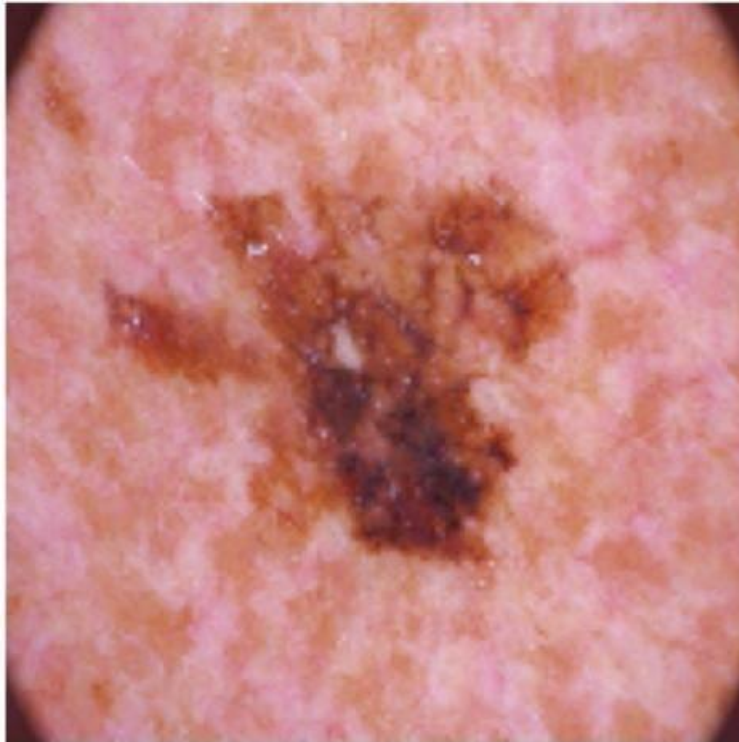

INPUT IMAGE
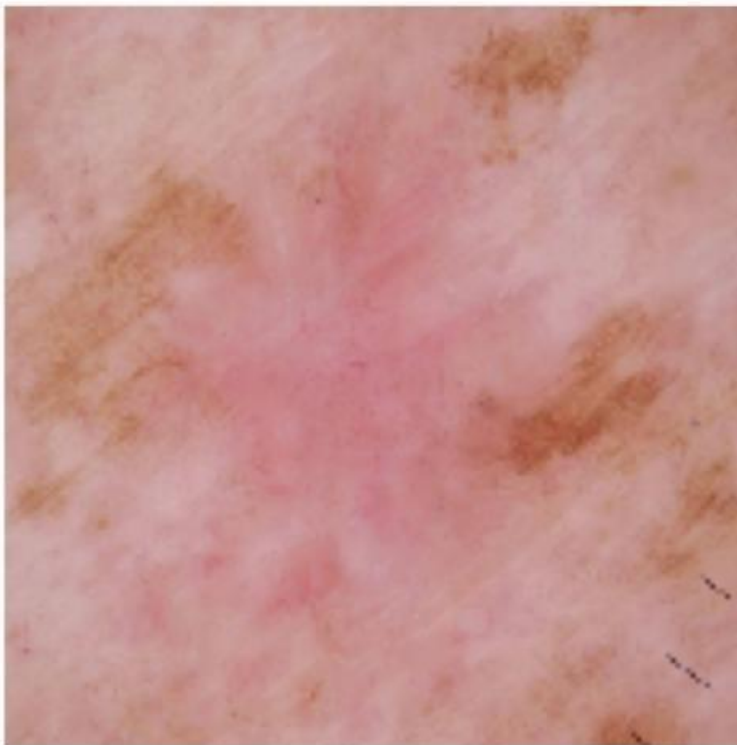
Basal cell carcinoma

This image shows that the patient has Basal cell carcinoma which is a dangeroustype of skin cancer

```
1/1 [==============================] - 0s 233ms/step
The image is of skin disease of class : Basal cell carcinoma
```

INPUT IMAGE

Dermato Fibroma

This image shows that the patient has dermato fibroma which is nothing but a scar.

```
1/1 [==============================] - 0s 215ms/step
The image is of skin disease of class : Dermatofibroma
```

INPUT IMAGE
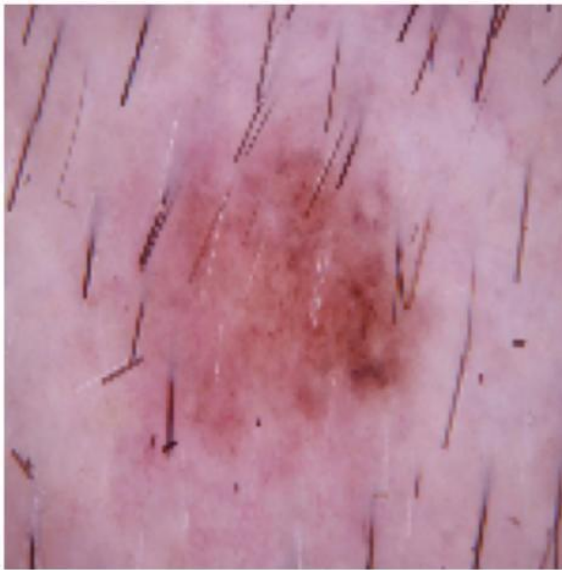
REFERENCES

[1] P. Avci, G. K. Gupta, and M. R. Hamblin , Imaging in Dermatology. Academic Press, 2016.

[2] American Cancer Society, "Cancer facts & figures 2018,"Availableathttps://www.cancer.org/ research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html, 2018, accessed in 19/05/2018.

[3] M. P. Correa, R. J. S. A. P. de Souza, A. C. A. Ferreira, A. P. Mattedi, ˆM.L.Rezende, "Estimativa do custo do tratamento do cancer de peletiponˆao-melanoma ˜ no estado de sao paulo - Brasil," ˜ Anais BrasileirosdeDermatologia, 2011.

[4] D. R. Bickers, E. Faulkner, C. Goodman, C. Gould, H. W. Lim, D. Margolis,M. A. Weinstock, "The burden of skin diseases: 2004: Ajoint project oftheamerican academy of dermatology association and the society for investigativedermatology," Journal of the American Academy of Dermatology, vol. 55, no.3, pp. 490–500, 2006.

[5] I. W. Bolliger, R. P. Dellavalle, N. E. Johns, R. J. Hay, R. Marks, L. Naldi,M.A.Weinstock, H. C. Williams, S. K. Wulf et al., "The global burden of skindiseasein2010: an analysis of the prevalence and impact of skin conditions,"JournalofInvestigati Dermatology, vol. 134, no. 6, pp.1527–1534, 2014.

[6] American Cancer Society, "Cancer facts & figures 2017," Availableat https://www.cancer.org/ research/cancer-facts-statistics/all-cancer-factsfigures/ cancer-facts-figures-2017.html, 2017, accessed in 29/04/2018.

[7] P. Bilek, O. Braun-Falco, A. B. Cognetta, M. Landthaler, T. Merkle, F. Nachbar, G. Plewig, and W. Stolz, "The abcd rule of dermatoscopy: highprospective value in the diagnosis of doubtful melanocytic skinlesions,"Journal of the American Academy of Dermatology, vol. 30, no. 4, pp. 551–559, 1994.

[8] American Cancer Society, "About basal and squamous cell skincancer,"Available at cancer.org/content/ dam/CRC/PDF/Public/8818.00.pdf, 2016, accessed in 29/04/2018.

[9] B. Stewart, C. P. Wild et al., "World cancer report 2014," Health, 2014.

[10] Instituto Nacional de Cncer Jos Alencar Gomes da Silva, "Estimativa

2018 incidłncia de cncer no brasil,"

Availableathttp://www1.inca.gov.br/inca/Arquivos/ estimativa2018.pdf, 2018,

accessedin29/04/2018.

[11] B. Barankin, and V. Prajapati "Answer: Can you identifythiscondition?" Canadian Family Physician, vol. 54, no. 5, pp. 699–699, 2008. [Online]. Available: http://www.cfp.ca/content/54/5/699

[12] H. J. Aerts, J. Bussink, P. Grossmann, B. Haibe-Kains, R. T. Leijenaar, R. Monshouwer, C. Parmar, D. Rietveld et al., E. R. Velazquez, "Decodingtumour phenotype by noninvasive imaging using a quantitative radiomicsapproach," Nature communications, vol. 5, p. 4006, 2014.

[13] H. M. Blau, A. Esteva, J. Ko, B. Kuprel, R. A. Novoa, S. M. Swetter, and S. Thrun, "Dermatologist- level classification of skin cancer withdeepneural networks," Nature, 2017.

[14] T. K. Alkasab, G. Choy, S. Do, H. Lee, J. Lee, S. Tajmir, B. A. Yeshiwas, and M. Zissen, "Fully automated deep learning systemfor boneageassessment," Journal of digital imaging, vol. 30, no. 4, pp. 427–441, 2017. [15] C. Barker, W. Scorer, M. L. Scott,

M.Soderberg, andM.E. Vandenberghe, "Relevance ¨ of deep learning to facilitate the diagnosisofher2 status in breast cancer," Scientific reports, vol. 7, p. 45938, 2017.