PROJECT REPORT

A CONVOLUTIONAL NEURAL NETWORK BASED SKIN DISEASE DIAGNOSIS

A PROJECT REPORT

Submitted by

AJU K TOLY (CHN16CS006)

AKHILNATH G (CHN16CS009)

AKSHAY T K (CHN16CS011)

ARJUN R (CHN16CS031)

to

The APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree

of

Bachelor of Technology

in

Computer Science and Engineering

Department of Computer Engineering

College of Engineering, Chengannur, Kerala -689121

July, 2020

DECLARATION

We undersigned hereby declare that the project report "A CONVOLUTIONAL NEURAL NETWORK BASED SKIN DISEASE DIAGNOSIS", submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by us under supervision of Mrs. Angel Thankam Thomas, Assistant Professor. This submission represents our ideas in our own words and where ideas or words of others have been included, We have adequately and accurately cited and referenced the original sources. We also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Place: CHENGANNUR

Date:19/07/2020

AJU K TOLY

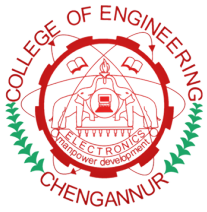
AKHILNATH G

AKSHAY T K

ARJUN R

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ENGINEERING CHENGANNUR , KERALA



CERTIFICATE

This is to certify that the report entitled **"A CONVOLUTIONAL NEURAL NETWORK BASED SKIN DISEASE DIAGNOSIS"** submitted by **AJU K TOLY, AKHILNATH G, AKSHAY T K, ARJUN R** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Department of Computer Science and Engineering, College of Engineering, Chengannur, Kerala is a bonafide record of the project work carried out by them under my/our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Ms. Angel Thankam Thomas
Assistant Professor
Project Guide

Ms. Shiny B
Assistant Professor
Project Co-ordinator

Dr. Smitha Dharan
Head of Department

ACKNOWLEDGEMENT

We are greatly indebted to God Almighty for being the guiding light with his abundant grace which strengthened me throughout the endeavour with confidence.

We express our heartfelt gratitude towards **Dr. Jacob Thomas V**, Principal, College of Engineering Chengannur for extending all the facilities required for this project. We would also like to thank **Dr. Smitha Dharan**, Head Of the Department, Computer Science and Engineering, for providing constant support and encouragement.

Now we extend our sincere thanks to our project coordinators **Smt. Shiny B** and **Smt. Angel Thankam Thomas** , Assistant Professor in Computer Engineering for guiding us in our work and providing timely advices and valuable suggestions.

Last but not the least, we extend our heartfelt gratitude to our parents and friends for their support and assistance.

ABSTRACT

Skin disease diagnosis or dermatology is one of the most difficult field. Several tests are to be carried out so as to decide upon the skin disease the patient may be facing. This takes time based on the diagnosis prediction. So there comes the need of a system which can diagnose the skin diseases without any of these constraints and provide better results in just fraction of seconds. Here an automated image based system using convolutional neural network for recognition of skin diseases is proposed. Training data set are required for various skin diseases. Skin images are filtered to remove unwanted noise and pigments. Feature extraction is done to bring the image to our required dimension.The output from feature extraction unit is fed into classifier unit.The input attributes like shape and texture etc are taken from the user input and fed into classifier.The classifier used is softmax classifier which is based on probability.The testing data and training data are compared based on probability. Each disease is assigned a unique class label. If the disease belongs to one of the training data set then the software predicts what disease it is otherwise a message reporting the image does not match with the dataset.

Contents

List of Figures

Chapter 1

INTRODUCTION

Skin is an important sense organ in protecting the body against pathogens and water loss. Insulation, temperature regulation, sensation, vitamin D synthesis and vitamin B protection are the other functions of skin. It acts as first line of defense against foreign particles and anti agents entering into the body. It is a shield. The outermost layer is epidermis. Dermatological diseases are the most common. Proper knowledge and experience in dermatology is required for accurate diagnosis. Precautious and preventive measures can control skin disease to a large extent. Early detection of skin diseases is the very need of the hour. Skin disease can impact on quality of life for patients. Some diseases are malignant and need proper care and attention. As skin is one of the sensory organ, atmost care is required .

Environmental factors together with the other causes are the main factors of skin diseases.Easily available detection schemes are to be made for early diagnosis of skin disease. Here the proposed paper provides an approach to detect various kinds of the diseases like psoriasis, melanoma, not melanoma, ringworm and measles. The user gives input of the skin disease image, which then the system processes, does feature extraction using CNN algorithm and use softmax image classifier to diagnose diseases. If no disease is found, that does not match with the dataset then disease not found message is displayed. Human skin shows wide variety of colours from the darkest brown to the lightest pinkish - white. Natural selection is the reason why humans exhibit large colour variations. Skin pigmentation in humans are caused due to the amount of ultraviolet radiation (UVR) penetrating the skin, controlling its biochemical effects. Melanin is the factor that determines the colour of human skin. The system identifies the colour of the skin to distinguish the disease. Skin colour varies for various diseases. Based on the skin colour and various other attributes like border, shape and texture the disease can be found out.

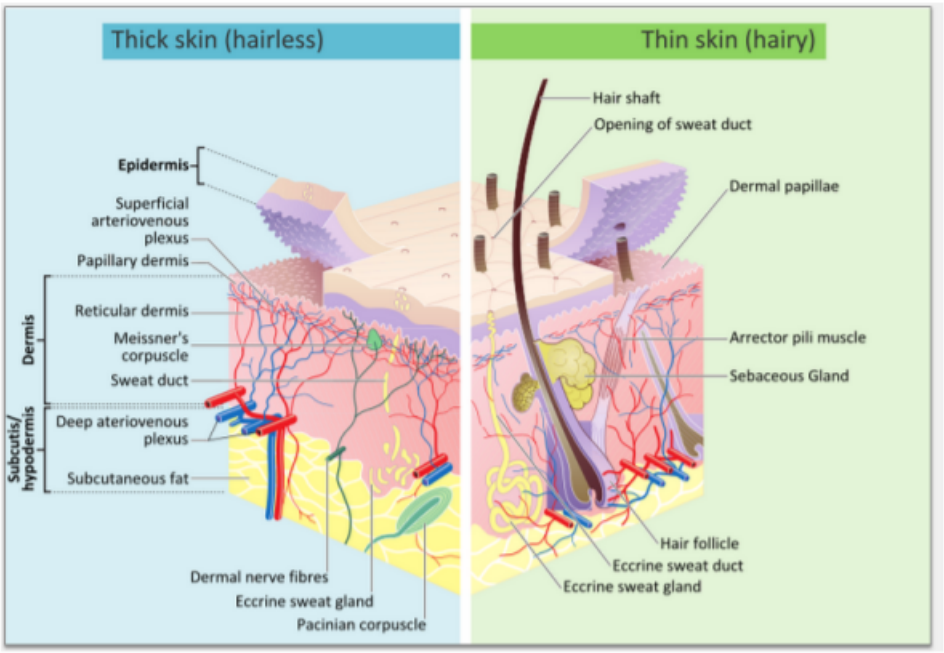


Figure 1.1: Structure of skin

1.1 Problem Definition

The early detection of skin diseases are very important as skin is one of the largest organ in the human body. Prevention of diseases and early diagnosis is of prime importance in today's scenario as the lifestyle have changed. Today's lifestyle has paved way to new diseases. Without proper care and safety, it may result in leading to chronic diseases. There is a need of system that can diagnose the skin diseases without any extensive tests being conducted.

1.2 Objective

The objective of the project is two-fold; to develop a convolutional neural network to diagnose various skin diseases like acne, psoriasis, warts etc. that overcomes the cons of existing methods and provide better results and accuracy and to develop an efficient and dependable system for dermatological disease detection that can be used as a reliable real time teaching tool for medical students in the dermatology stream.

1.3 Scope

As discussed in the problem definition, the medicine field using latest technologies has been growing at a very rapid rate. Detection of diseases through advanced technologies and soft-

wares can be implied as a modification for the existing medical technology applications. The human brain can be modelled using neural networks. The proposed system, to an extent provides a kind of skin disease diagnose using convolutional neural network. This methodology makes it possible to diagnose the user input images and classify them within the set of data provided during training. The scope of skin disease diagnosis lies in the broader applications of CNN were various human-machine interactions is made possible for better user convenience.

Chapter 2

LITERATURE SURVEY

2.1 Dermatological Disease Detection Using Image Processing and Machine Learning

In order to diagnose skin disease the system utilizes a two staged approach. In this approach the Image Processing for identification is the first stage and Machine Learning is the second stage. The system acts as an effective learning tool, aiding verification of the results as it have access to clinical data. The machine learning data repository provides the training data set. By using Computer Vision and Machine Learning algorithms the system have achieved better accuracies. The system is capable of detecting six diseases, namely Psoriasis, Seborrheic Dermatitis, Lichen Planus, Pityriasis Rosea, Chronic Dermatitis, and Pityriasis Rubra Pilaris.

The type of skin disease can be identified by considering the numerous features extracted from the image and this is identified by Computer Vision in the first stage. In the second stage the system uses various Machine Learning techniques to refine the classification of the image. The second stage of prediction is made available to the medical professionals who have access to various histopathological attributes like exocytosis, hyperkeratosis, acanthosis, parakeratosis and other attributes.To obtain better classification of the disease, the system takes these attributes as input from the user

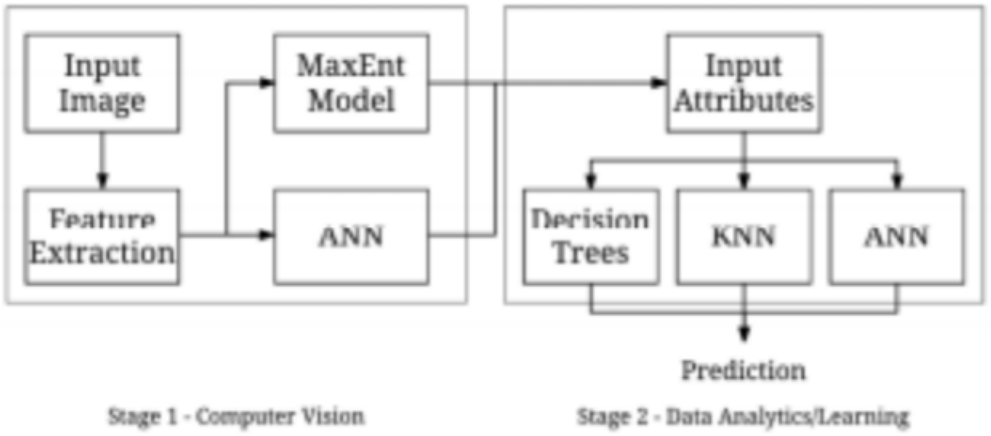


Figure 2.1: Block Diagram

Using Decision Trees, Neural Networks and kNN (kth Nearest Neighbour) model, the system tested and trained using the dataset obtained from the repository. To convert input image to grey scale image, sharpening filter, median filter, smooth filter, binary mask, RGB extraction, histogram and sobel operator the algorithms are useful. Before converting to grayscale image the RGB values are extracted from the image. To sharpen the details of the infected region, sharpening filter is applied to the grayscale image. After sharpening the image, to remove the noise from the image the Median Filter is used. The smoothing filter is used to replace each pixel with the mean values of its neighbours, including itself. By using mean filtered image and distribution of color of binary image ,binary image is generated. Inorder to extract average colour code of the infected area from the binary image, YCbCr is used.To detect the edge of the infected area Sobel operator is applied to binary image.

The inflicted region's histogram is obtained. RGB color space of the image is converted into HSV color space.A feed forward ANN with back propagation is also tried along with the Maximum Entropy model for the same features which are extracted. The ANN consisted of one input layer, two hidden layers and an output layer.

The mobile application based on the principles as described above produce better results than any application. Combining the two stages increases the accuracy, making the application an efficient, accurate and good system for dermatological disease detection. This can be used as a good teaching tool in medical field. As an add on, the application can also be used by the common user as it have been able to achieve fairly accurate detection rate by Computer Vision techniques alone.

Advantages:

1. The system produced an accuracy of upto 95 percent.
2. The system is capable of detecting 6 of the most common occurring diseases like psoriasis.
3. The system can be used by common users also.

Disadvantages:

1. The system suffers from inaccuracies when it is tasked with detection of diseases on varying skin colors.
2. Only 6 diseases can be detected.

2.2 Skin Disease Recognition Using Texture Analysis

The system is based on the texture analysis using neural network. As the complexity and number of features of the disease increases, the disease diagnosis and recognition becomes difficult. This led to a computer aided diagnosis and recognition system. The system involves steps like image processing, image feature extraction and classification of data using artificial neural network. The ANN is capable of learning the patterns of symptoms of particular diseases and provide faster diagnosis. This helps the patients to do the treatment for the skin disease immediately based on the visible symptoms.

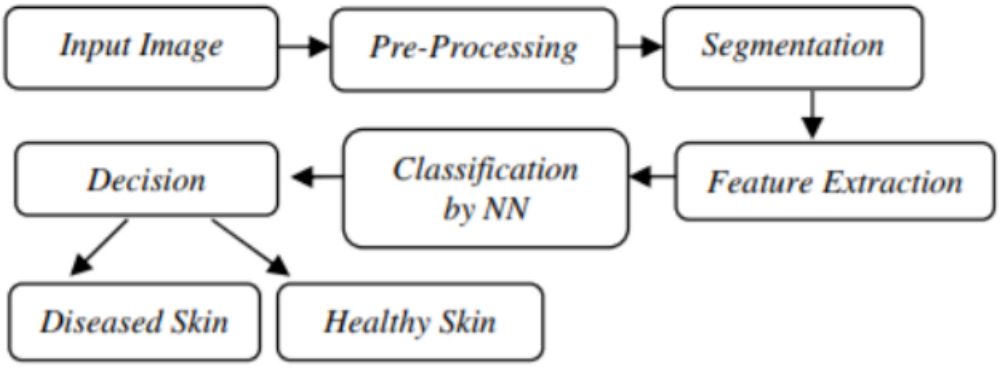


Figure 2.2: Block Diagram of the System.

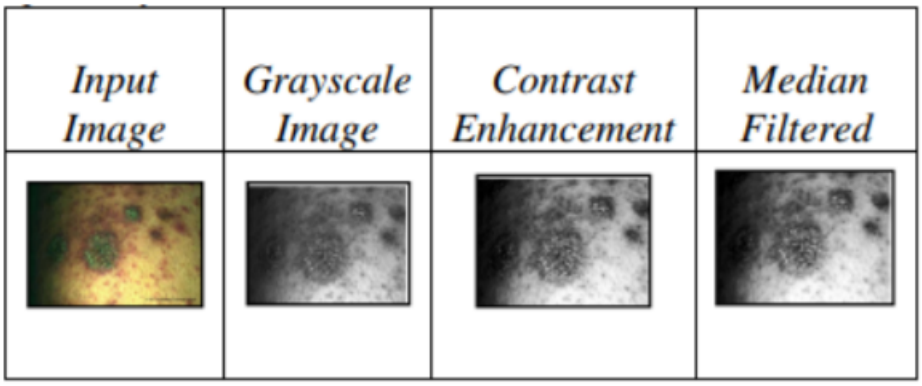


Figure 2.3: Processing of images

The block diagram involves two sections: Image Processing and Classification. The architecture involves the following steps: Image acquisition, Pre-Processing, Segmentation, Feature Extraction, Classification using neural network and decision phase is used to find whether the skin is affected with disease or not. Feature extraction is used to differentiate between each input patterns . Texture analysis using Grey Level Co-occurrence Matrix (GLCM) provide the basis of feature extraction. GLCM maps the grey level co-occurrence probabilities into different angular directions. The extracted features from GLCM includes Energy, Correlation, Homogeneity and Contrast:

Neural network classifier is used for segmenting the image. The input layer, hidden layer and output layer are the layers in feed forward multilayer network. Back propagation algorithm is used and the signal will flow in the forward direction. If both outputs do not match , the output of the network being compared with desired output, an error signal will be generated. To reduce the error, the signal is propagated backwards and weights are adjusted. At the beginning of training, the hidden layer and output layer are initialized. The process continuous until the error is zero.

Advantages:

1. Segmentation results can easily be adjusted or refined through the parameter of color sensitivity.
2. Robust against noise.
3. Flexible segmentation.

Disadvantages:

1. Shape accuracy comparatively low.
2. Pattern similarity between different disease may lead to less accuracy.

2.3 An Automated Computer Aided Diagnosis of Skin Lesions Detection and Classification for Dermoscopy Images

Skin cancer is a deadly disease nowadays. So, early detection and prevention are essential. Skin cancer is defined as the uncontrolled and unpredictable growth of cells in the skin. The main causes of skin cancer is the over exposure of ultraviolet radiation from sunshine, genetic defects and food habits. There are mainly three types of skin cancer such as Basal cell carcinoma (BCC), Melanoma, and Squamous cell carcinoma (SCC). The non-melanomas were BCC and SCC. To classify the skin lesions in accurate manner an automatic Computer-Aided Diagnosis (CAD) for dermoscopy images is needed. The flow diagram gives the classification model for skin cancer such as Melanoma, Nevus, Basal cell carcinoma and Seborrhoeic keratosis which can classify both Melanocytic skin lesions (MSLs) and Non-melanocytic skin lesions (NoMSLs) .

Pre-processing

The image of the affected part is given as the input to the system. The image contains hair that must be eliminated in the pre-processing. It is removed using Dull Razor software. The median filter was used for removing thick hair after which circular averaging filter of size r is used for smoothing.

Border detection method

After pre-processing, convert the RGB image to grayscale image and identify the pixels whose intensities are less than 200 of the image. Identified pixels of the image with disk size 3r are dilated to include border of the microscope areas that are not selected. K-means clustering is used to segment the region.

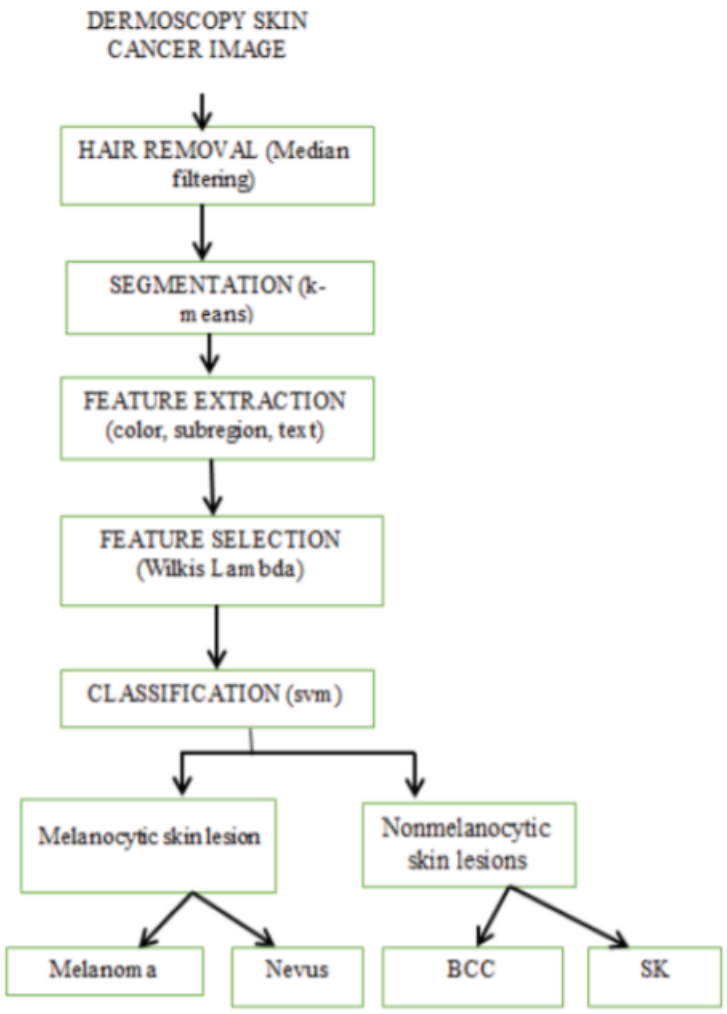


Figure 2.4: Digrammatic Representation of the System.

Classification

Segmenting the lesion of skin image into three regions-normal skin, peripheral, tumor after detecting the border of the skin lesion is the final step.The SVM classifier is used for classification. It first classifies the Melanocytic skin lesion from Nonmelanocytic skin lesions and classifies the melanoma from nevus for MSLs and for basal cell carcinoma from seborrhoeic keratosis in NoMSLs. The first step classifier (MN-BS) is used to distinguish MSLs from NoMSLs and second step classifier distinguish melanoma from nevus (M-N) and distinguish basal cell carcinoma from seborrhoeic keratosis (B-S). The M, N, B, S stands for melanoma, nevus, basal cell carcinoma and seborrhoeic keratosis for classification model by using SVM classifier. This method for classification has achieved better accuracy

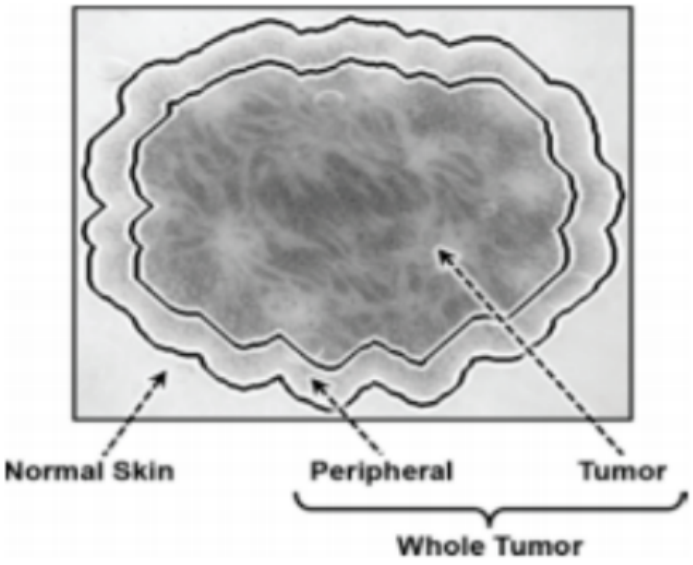


Figure 2.5: Four regions of skin lesions

Figure 2.6: Implementation results by SVM classifier for melanoma

Advantages:

1. An accuracy of 96.8% is obtained using this method.
2. It is used for the classification of skin lesions such as Melanoma, Basal cell carcinoma (BCC), Seborrhoeic keratosis (SK) and Nevus.

Disadvantages:

1. The method works well which are all having high differentiation between lesion

and skin of the image, but fails when two regions overlap. 2. Segmentation of the skin lesion in Dermis layer of the skin is not possible.

2.4 Dermatological Disease Detection using Image Processing and Artificial Neural Network

The results of image pre-processing and user inputs like liquid type, liquid color, elevation, duration, feeling, gender and age are fed into the system. The training and testing purpose of the feed forward artificial neural networks (ANN) are based on these features. Diagnosis can be made effective by using artificial neural networks (ANN) as knowledge base. Pre-processing of skin images followed by segmentation and feature extraction are the two phases in the system. Colour skin images and 8 different image processing algorithms are applied on to the system. The training is done in such a way that the user input values along with colour skin image extracted features are given to a feed forward back propagation artificial neural network to identify the dermatological disease.

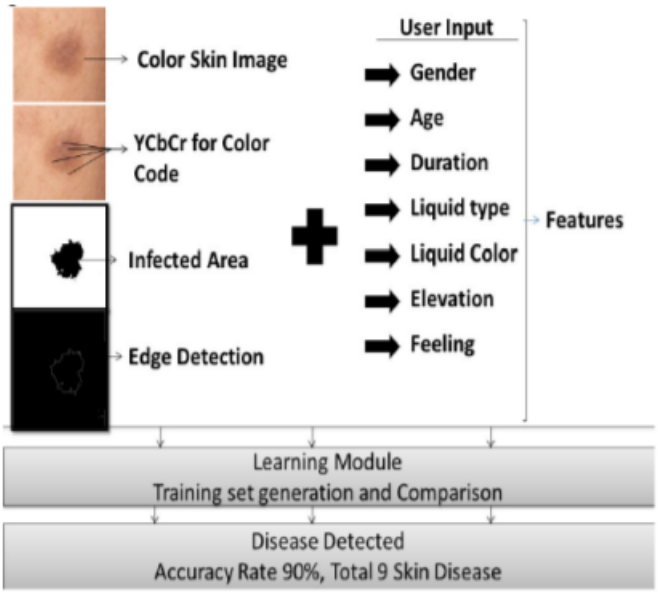


Figure 2.7: System architecture.

The eight different algorithms used are grey image, sharpening filter, median filter, smooth filter, binary mask, histogram, YCbCr and sobel operator. The algorithms are implemented sequentially. Grayscale image of the infected area is obtained first. Sharpening filter is applied to the grayscale image to get sharper details of the infected area. To remove noise, median filter

is used. 5*5 matrix is used as median filter and smoothing filter is applied after media filter. From the mean filtered image, the binary image is obtained and colour distribution of binary image is represented by histogram. To extract average colour code of the infected area from the binary image, YbCr is used. To detect edge of the infected area, sobel operator is used .

Feed-forward back-propagation neural network training is the final step. Ten fold cross validation is used for validation and testing of the system. To avoid the overlapping of the testing data and training data and making the system testing results viable and dependable, the system uses cross validation technique. The feed forward back-propagation neural network is trained with 10 different features, in which three are from image features and seven are from user inputs. In the system, 100 neurons are used in the hidden layer to get the best result from the system.

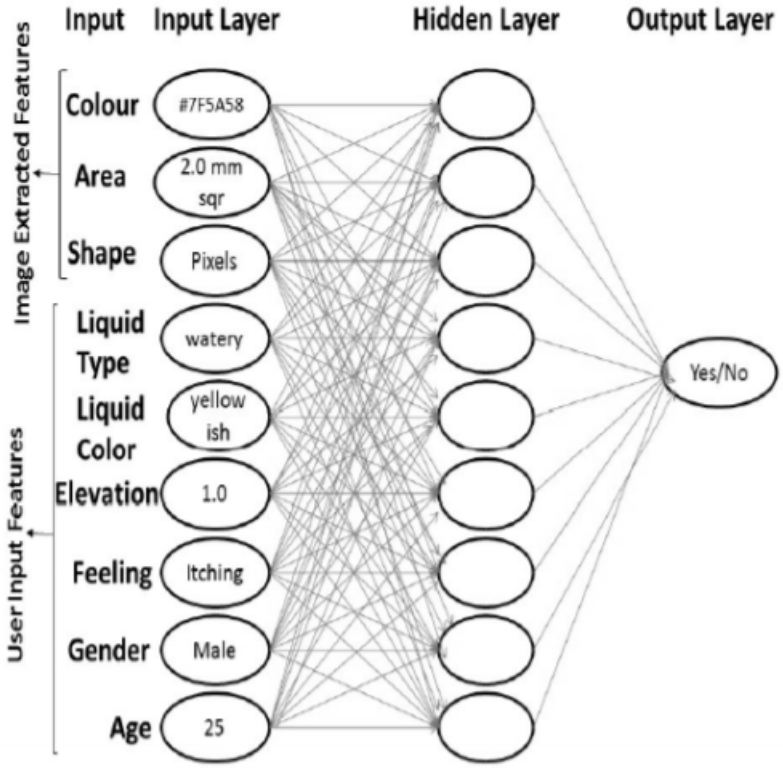


Figure 2.8: Feed forward neural network.

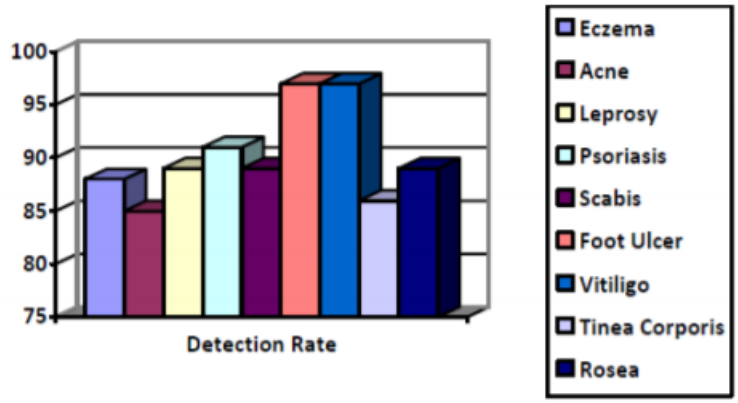


Figure 2.9: Detection rate of different skin diseases.

Low elevation in the infected area of diseases like foot ulcer, vitiligo and psoriasis are very high like 97%, 97% and 91% whereas the diagnosis of diseases that has low elevation are comparatively low which is between 85-88diseases, the system shows good accuracy rate for foot ulcer and vitiligo and the system did not perform well for acne and tinea corporis.

Advantages:

1. An accuracy of 90% is obtained using this method. 2. 9 different types of diseases like eczema, acne,leprosy,psoriasis,scabies,foot ulcer, vitiligo, tinea corporis and pityriasis rosea could be identified.

Disadvantages:

1. The system could not give better results for acne and tinea corporis.
2. Applicable only to diseases that has low elevation in the infected area.

2.5 Dermatological Disease Detection using Image Processing and Artificial Neural Network.

Skin diseases are among the most common health problems worldwide. In this article we proposed a method that uses computer vision based techniques to detect various kinds of dermatological skin diseases. We have used different types of image processing algorithms for feature extraction and feed forward artificial neural network for training and testing purpose. The system works on two phases- first pre-process the colour skin images to extract significant

features and later identifies the diseases.

Advantages:

1. Successfully detects 9 different types of dermatological skin diseases.
2. An accuracy rate of 90%.

Disadvantages:

1. Could have detected more diseases.

2.6 Detection of Vitiligo Skin Disease using LVQ Neural Network

Digital image processing is a combination of various algorithms and technique to process different types of images. It is applied in various types of image to process and get a valuable outcome from the image. The Digital image processing is the experimented on image to extract different features of the image. This paper provides the idea which is used to detect the affected area of the Vitiligo disease with help of image captured by camera and classified the affected area from non-affected area in image. Vitiligo is the deep rooted skin disease which is depigmentation of the skin in which human skin starts losing or loss of pigment from the skin. The certain portion of the skin of body became white patches. The Vitiligo is visible in dark skin persons because of some genetic problem or environmental issues. Here, the learning vector quantization neural network is used to classify Vitiligo image in affected vs. non-affected region to detect disease.

Advantages:

1. The implementation of LVQ neural network gives very good accuracy of 92.22%.
2. kappa value of 0.810 which is very good for proposed technique.

Disadvantages:

1.Can detect only Vitiligo disease.

2.7 Analysis and Classification of Human Skin Diseases

Most common skin diseases like skin cancers, leprosy etc are untreated and mostly causes death. Skin cancer has more cure rate if detected and treated early. The basic means of detecting these skin diseases is through visual inspection followed by biopsy and pathological examination. If the physician finds the appearance of lesion doubtful then normally visual inspection method is used for diagnosis but all malignant lesions are not identified through visual inspection. Now, there are no generally accepted tools that physician can use to immediately find the skin disease in the clinic. Most form of visual inspection could help to prevent misdiagnosis of BCC and other types of skin diseases. Previous work suggests that electrical impedance may distinguish skin cancer from other tissue. The electrical impedance of a tissue depends on its structural characteristics as well as its chemical composition. Studies have shown a wide degree of variation in the bio-electric properties between tissue and cells of body. The studies have shown differences in the electrical impedance of the skin as a result of irritation, allergic reaction, location, sex, age and hydration. A clinical study has also shown significant differences between affected skin and normal skin. Such clinical study is known as impedance measurement and based on a comparison of four indexes: magnitude, phase, real part and imaginary part index.

Advantages:

1.used for physiological measurements emotional disorders or lie detection.

2.used in diagnosis of early stage skin diseases like melanoma skin tumours Non-melanoma

3. Efficiency of system to diagnose normal and affected skin of different skin disease is nearly 75%.

Disadvantages:

1. Measurement of skin impedance limited to the frequency range 100 Hz to 1 MHz.

2. Not highly efficient.

2.8 Image Analysis Model For Skin Disease Detection: Framework

Skin disease is the most common disease in the world. The diagnosis of the skin disease requires a high level of expertise and accuracy for dermatologist, so computer aided skin disease

diagnosis model is proposed to provide more objective and reliable solution. Many researches were done to help detect skin diseases like skin cancer and tumor skin. But the accurate recognition of the disease is extremely challenging due to the following reasons: low contrast between lesions and skin, visual similarity between Disease and non-Disease area, etc. This paper aims to detect skin disease from the skin image and to analyze this image by applying filter to remove noise or unwanted things, convert the image to grey to help in the processing and get the useful information. This help to give evidence for any type of skin disease and illustrate emergency orientation.

Advantages:

1. Used as an effective, low cost solution for skin diseases.
2. Helps in initial diagnoses.

Disadvantages:

1. Focuses only on four types of most common skin diseases Acne, Psoriasis, Melanoma, Heat Rash.

Chapter 3

METHODOLOGY

3.1 Deep Learning

Deep learning is an artificial intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in artificial intelligence (AI) that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Also known as deep neural learning or deep neural network.

Deep learning has turned applications that previously required vision expertise into engineering challenges solvable by non-vision experts. It also opens a new range of possibilities to solve applications that have never been attempted without a human inspector

STRENGTHS OF DEEP LEARNING

• No Need for Feature Engineering Feature engineering is the process of extracting features from raw data to better describe the underlying problem. A deep learning algorithm will scan the data to search for features that correlate and combine them to enable faster learning.
• Best Results with Unstructured Data Unstructured data is hard to analyze for most machine learning algorithms. Deep learning algorithms can be trained using different data formats, and still derive insights that are relevant to the purpose of its training.
• No Need for Labeling of Data Getting good-quality training data is one of the biggest problems in machine learning because data labeling can be a tedious and expensive job. With deep learning, the need for well-labeled data is made obsolete as deep learning algorithms excel at learning without guidelines.
• Efficient at Delivering High-quality Results Once trained correctly, a deep learning brain can

perform thousands of repetitive, routine tasks within a shorter period of time than it would take a human being.

Some other important advantages are ;

• Designed for hard-to-solve applications
• Requires less data and computing power
• More reliable and easier to configure
• Identifies defects in milliseconds, supporting high-speed applications and improving throughput.

3.2 Convolutional Neural Networks

The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. Convolution is a specialized kind of linear operation. Convolutional networks are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers.

CNNs are regularized versions of multi-layer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "fully-connectedness" of these networks makes them prone to overfitting data. Typical ways of regularization include adding some form of magnitude measurement of weights to the loss function. CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble more complex patterns using smaller and simpler patterns.

CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were handengineered. This independence from prior knowledge and human effort in feature design is a major advantage.

3.3 Depthwise Seperable Convolution

A depthwise separable convolution benefits from the same characteristic as spatially separable convolutions, being that splitting the kernels into two smaller ones yields the same result with fewer multiplications, but does so differently. Effectively, two operations are performed

in depthwise separable convolutions – sequentially:

•Depthwise convolutions;

•Pointwise convolutions.

3.3.1 Depthwise convolutions

As we've seen above, normal convolutions over volumes convolve over the entire volume, i.e. over all the channels at once, producing a WidthxHeightx1 volume for every kernel. Using N kernels therefore produces a WidthxHeightxN volume called the feature map. In depthwise separable convolutions, particularly the first operation – the depthwise convolution – this does not happen in that way. Rather, each channel is considered separately, and one filter per channel is convolved over that channel only.

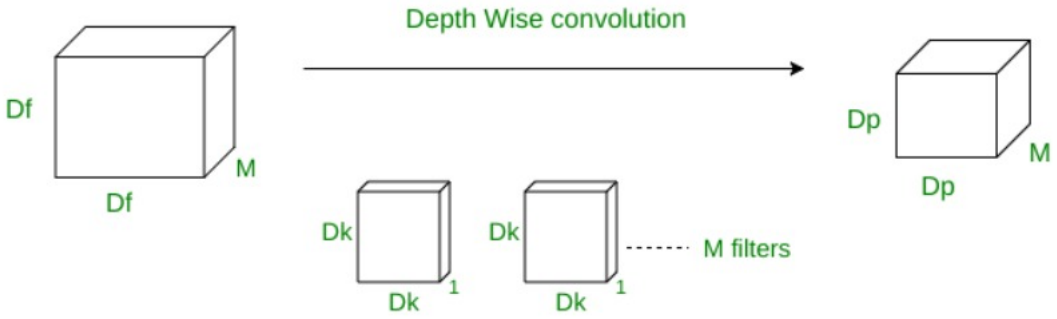


Figure 3.1: DepthWise Convolution.

3.3.2 Pointwise convolutions

From the intermediate result onwards, we can then continue with what are called pointwise convolutions. Those are filters of 11 pixels but which cover all the M intermediate channels generated by the filters, in our case M=3. And since we're trying to equal the original convolution, we need N of them. Remember that a convolution over a volume produces a Some Width x Some Height x 1 volume, as the element-wise multiplications performed over three dimensionsresult in a one-dimensional scalar value. If we would thus apply one such pointwise filter, we would end up with a Hfm x Wfm x 1 volume. As the original convolution produced a Hfm x Wfm x N volume, we need N such pointwise filters.

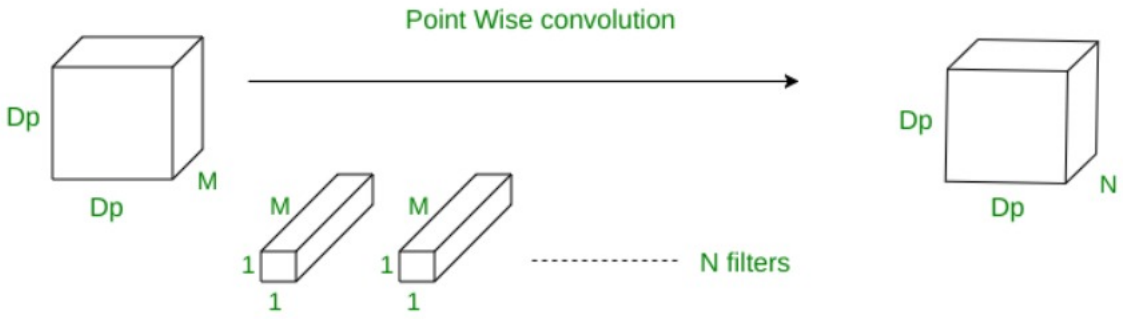


Figure 3.2: PointWise Convolution.

3.4 MobileNet

MobileNet is a CNN architecture model for Image Classification and Mobile Vision.There are other models as well but what makes MobileNet special that it very less computation power to run or apply transfer learning to.This makes it a perfect fit for Mobile devices,embedded systems and computers without GPU or low computational efficiency with compromising significantly with the accuracy of the results.It is also best suited for web browsers as browsers have limitation over computation,graphic processing and storage.

Chapter 4

PROPOSED SYSTEM

The proposed system uses convolutional neural network. The user gives input of the skin disease which the system processes, does feature extraction using CNN algorithm and use softmax classifier to diagnose the disease. If no disease is found, the system produces a negative result meaning not in the dataset. The architecture contains two major parts feature extraction unit and classification unit. Feature extraction unit will enhance the image by removing noise and unwanted parts of the skin. Initially the images are pre-processed and converted to a standard size. Then the image is given as an input to the first layer of the network. Convolutional Neural Network is applied onto it until high-level features such as colour, shape and texture are obtained from it. Basically the system consists of two modules:

1. Feature Extraction Module
2. Classifier Module.

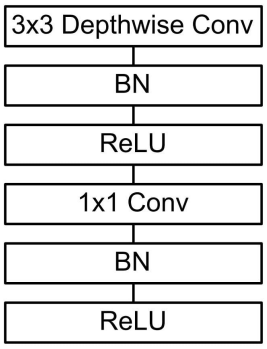


Figure 4.1: Architecture of the system.

Feature Extraction Module: The module consists of performing operations like convolution, max pooling and ReLu. Based on the requirements, this layer may extend. The objective of the convolution operation is to extract the high-level features such as shape and texture, from the

input image. Max Pooling returns the maximum value from the image covered by the Kernel. Max Pooling also performs as a Noise Suppressant. It removes the noises and also performs de-noising along with dimensionality reduction. ReLU is an activation function that has strong biological and mathematical operations.

Classifier Module: The module consists of dense layer, dropout layer and softmax layer. Dropout layer is a technique used to improve over-fit on neural networks.During prediction, the dropout layer is deactivated. Dense layer is followed by a non-linear activation.The system can be broadly categorized into following major phases: Pre-processing the images, testing and training.

Phases of the system

Pre-processing phase: This phase includes image acquisition and pre-processing. In image acquisition, the images are acquired either through camera or through locally stored device. A high quality image is required for the system implementation.

Testing and training phase: It includes data storage and classifier unit. The Data storage component is used to maintain testing and training data images. Training dataset is required for supervised learning. Testing dataset is the images acquired during image acquisition. The classifier identifies the type of skin disease. Softmax classifier used here is the last layer of the network that yields actual probability of each label.

4.1 Software Requirements

The proposed work is done in Spyder software and OpenCV. The datasets for the diseases like melanoma, not melanoma, psoriasis, measles and ringworm are collected from various sources. Initially the dataset consists of 500+ images. The dataset is partitioned into testing and training samples. The dataset is trained and accuracy is found out. The user input is given which is an image.

Tensorflow

Developed by Google, it is an open-source library. Here it is used as a backend for Keras as it is useful for numerical computations and calculations.Tensorflow architecture works in three parts: Preprocessing the data, build the model and train and estimate the model

Colab

Colaboratory or Colab is a Google research project created to help disseminate machine learning education and research. It's a Jupyter notebook environment that requires no setup to use

and runs entirely in the cloud.

It allows you to write and execute Python in your browser, with

- Zero configuration required
- Free access to GPUs
- Easy sharing

4.2 Depthwise Seperable Convolution

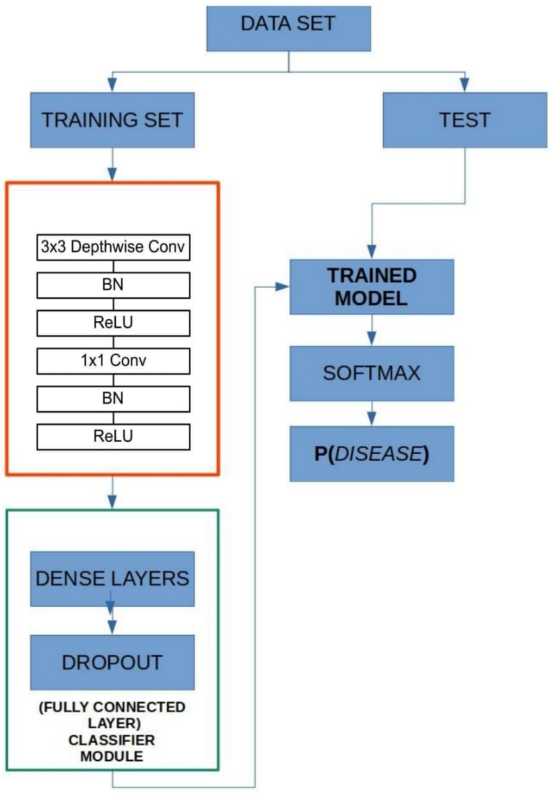


Figure 4.2: Mobile Net Data Flow Diagram.

MobileNet is a CNN architecture model for Image Classification and Mobile Vision. There are other models as well but what makes MobileNet special that it very less computation power to run. MobileNets, a family of mobile-first computer vision models for TensorFlow, designed to effectively maximize accuracy while being mindful of the restricted resources for an on-device or embedded application. MobileNets are small, low-latency, low-power models parameterized to meet the resource constraints of a variety of use cases. They can be built upon for classification, detection, embeddings and segmentation similar to how other popular large scale models,

such as Inception, are used.MobileNets are a family of mobile-first computer vision models for TensorFlow, designed to effectively maximize accuracy while being mindful of the restricted resources for an on-device or embedded application.

In this project, We use MobileNet,which has 28 layers by default.We specified Dense and dropout.The class sensitivities are also specified.The model is initialized with pre-trained weights.This project is a 4 class classification deep learning convolutional model which accepts image in 224 x 224 dimensions and 3 channels (RGB).The output prediction of this model is done by the help of global average pooling(GAP) layers to minimize overfitting by reducing the total number of parameters in the model and dense layer activated with softmax,hense if any case of partial matching to another class,we can detect it by evaluating the values given by each node in layer. Currently we are taking the most largest of the values and index it to get the disease class.

Batch Normalization

Batch normalization is a technique for training very deep neural networks that standardizes the inputs to a layer for each mini-batch. This has the effect of stabilizing the learning process and dramatically reducing the number of training epochs required to train deep networks.

ReLu Layer

Rectified Linear Unit (ReLU) transform function only activates a node if the input is above a certain quantity, while the input is below zero, the output is zero, but when the input rises above a certain threshold, it has a linear relationship with the dependent variable. The main aim is to remove all the negative values from the convolution and maxpooling. All the positive values remain the same but all the negative values get changed to zero. The fully connected layer consists of dense layer, dropout layer and softmax classifier. The image after feature extraction is flattened to convert the image into one dimensional array for classification.

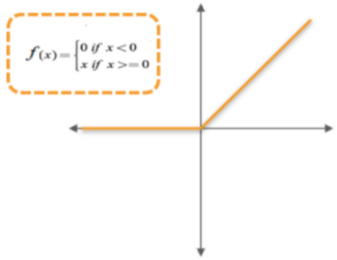


Figure 4.3: ReLu

Dense Layer

A dense layer is just a regular layer of neurons in a neural network. Each neuron receives input from all the neurons in the previous layer, thus densely connected. The layer has a weight matrix W, a bias vector b, and the activations of previous layer. The dense layer takes in parameters the number of input neurons and the activation function

Dropout Layer

Dropout is a a technique used to tackle Overfitting. The Dropout method in keras.layers module takes in a float between 0 and 1, which is the fraction of the neurons to drop.Dropout consists in randomly setting a fraction rate of input units to 0 at each update during training time, which helps prevent overfitting.

Softmax classifier

The purpose of the softmax classification layer is simply to transform all the net activations in your final output layer to a series of values that can be interpreted as probabilities. The last layer takes in parameters like number of output labels and the activation called softmax.

The model once created is then trained by using model.compile function to calculate the loss incurred during training and accuracy. The model is fit to the number of epochs and batch size. Then it is further used to predict the class label. The user selects an image and it is tested.The easygui module is used to select the image. Each image is resized to a standard size.

4.3 Data Flow Diagram

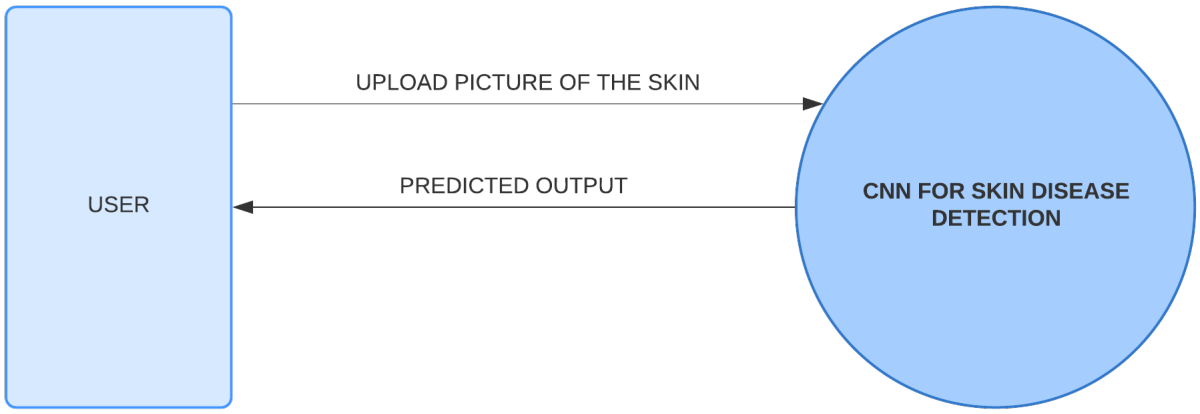


Figure 4.4: Level 0

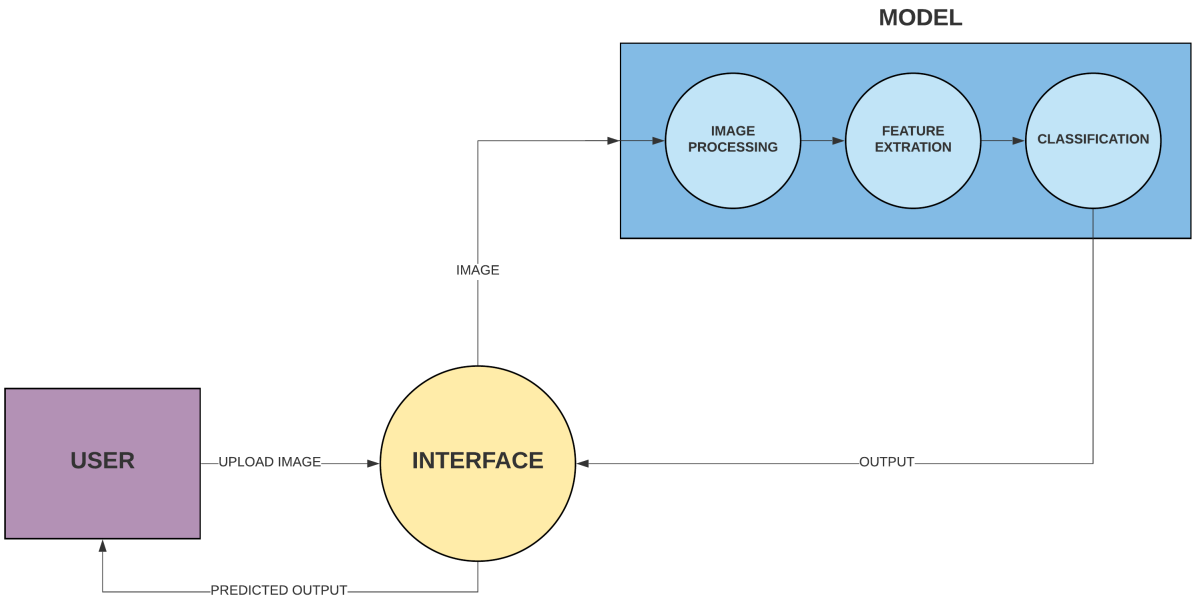
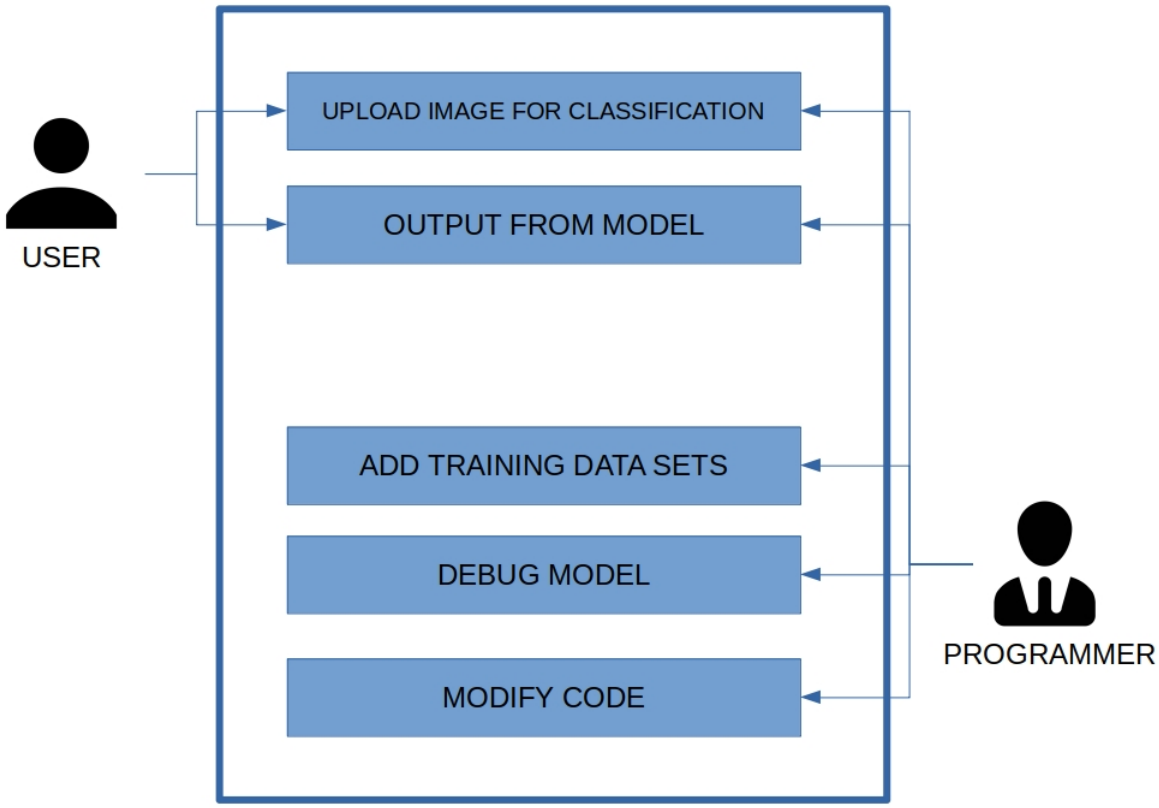


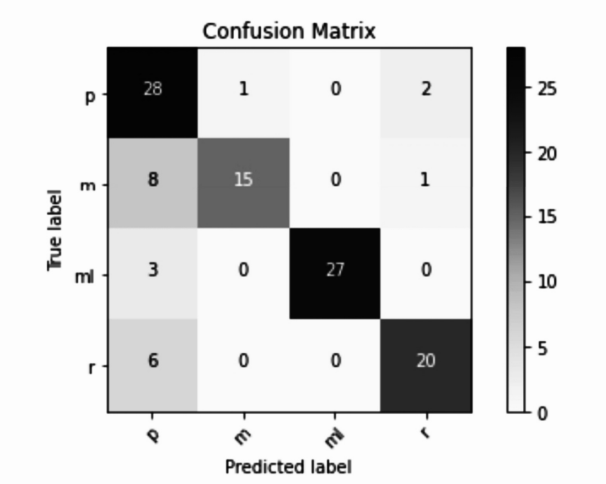
Figure 4.5: Level 1

4.4 Use Case Diagram



Chapter 5

RESULTS AND DISCUSSION



The system utilizes the MobileNet architecture. The neural network learns and extracts the features by itself.Hence unlike image processing techniques used in the related works, feature extraction is done using depthwise separable convolution. Based on the training and validation data an accuracy of 98.38 is obtained. Initially with 2000 data samples, the system recognizes 4 diseases like ringworm, measles, psoriasis, and melanoma . The use of MobileNet architecture has enabled us to create a light weight model which can be easily used on devices with even lower specs. We hope that our model is useful in the field of skin disease detection.

Chapter 6

CONCLUSION

The stated objective of the project is satisfied by the proposed design. After having carefully analysed multiple papers and evaluating their pros and cons, we were able to design a neural network based system for diagnosing skin diseases. In the system, methods which are being used to detect skin diseases exhibits better results and accuracy. The system enables the users to know what skin disease he is suffering from. This makes it possible for the users to take preventive measures to control the spreading of diseases and prevention at an early stage. Neural networks have many applications in medical field that help in early diagnosis and prevention of diseases. Convolutional Neural Networks have proved that large number of datasets can be trained within a short span of time and provide greater accuracy. Using advanced computational techniques and large dataset, the system can match the results of a dermatologist thus improving the quality standards in the area of medicine and research.

Future Work

With vast dataset, the system could recognize diseases apart from those mentioned in the work. The real time application of the above work with an android platform could help the people to analyse the disease with a fraction of seconds. The remedial measures and location of the nearby dermatologist could also be included.

REFERENCES

[1] V. B. Kumar, S. S. Kumar and V. Saboo, ”Dermatological disease detection using image processing and machine learning,” 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR) Lodz, 2016

[2] A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model

[3] Damilola A. Okuboyejo, Oludayo O. Olugbara, and Solomon A. Odunaike, An Automated Computer Aided Diagnosis of Skin Lesions Detection Classification for Dermoscopy Images” Proceedings of the World Congress on Engineering and Computer Science 2013 Vol II

[4] R. Yasir, M. Rahman, N. Ahmed, ”Dermatological Disease Detection using Image Processing and Artificial Neural Network

[5] Dermatological Disease Detection using Image Processing and Artifi-cial Neural Network R. Yasir, M. Rahman, N. Ahmed

[6]Jyotsna Anthal, Anand Upadhyay, Ashish Guptha Detection of Vitiligo Skin Disease using LVQ Neural Network

[7]P.B. Manookar, D.K.Kamat, P.M Patil Analysis and Classification of Human Skin Diseases

[8]Alaa Haddad, Shihab A Hameed Image Analysis Model For Skin Disease Detection: Framework

[9] MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam