AN AUTOMATED DIAGNOSIS OF SKIN CANCER DISEASE USING MACHINE LEARNING TECHNIQUES

U15CS801R
PROJECT WORK

Submitted by

ELAMATHI.S	***(1516102034)***
ATCHAYA.R	***(1516102014)***
DHANALAKKSHMI.M	***(1516102025)***

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING
SONA COLLEGE OF TECHNOLOGY

ANNA UNIVERSITY : CHENNAI 600 025

JULY 2020

ANNA UNIVERSITY : CHENNAI – 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"An Automated Diagnosis of Skin Cancer Disease Using Machine Learning Techniques"** is the bonafide work of ***"Elamathi.S (1516102034),*** **Atchaya.R (1516102014), Dhanalakkshmi.M (1516102025)"** who carried out the project work under my supervision.

SIGNATURE
Dr.B.SATHIYABHAMA M.Tech,Ph.D.,
HEAD OF THE DEPARTMENT
Professor
Department of Computer Science and Engineering
Sona college of Technology
Salem.

SIGNATURE
Dr.A.C.KALADEVI M.E,Ph.D.,
STAFF IN-CHARGE
Professor
Department of Computer Science and Engineering
Sona college of Technology
Salem.

Submitted for the project viva-voice examination held on 15/07/2020

_____ _____

Internal Examiner External Examiner

TABLE OF CONTENTS

ABSTRACT

The skin is the largest organ in the human body, which protects us from microbes and other pathogens. Etymologically, dermatology is the medical discipline in the analysis and prevention of skin abnormalities. The system is an automated diagnostic system unlike the conventional system involving human arbitration based on the ideology of dermatological diagnosis. The system works in two dependent steps: the first detects skin abnormalities and the second identifies skin diseases. The system works with visual inputs, i.e. high-resolution color images and patient history. The automated diagnostic system uses a modified genetic algorithm, a k - means grouping, and an SVM classifier to perform pre-processing segmentation, and feature extraction on the images respectively. To detect a skin cancer disease, the system uses a neural network to propagate artificial feedback, which is implemented using MATLAB. The system has a skin cancer detection accuracy of 98.99% and a cancer identification accuracy of 97.016% when testing diseased areas on skin images. In addition to this, various systems have been proposed to assist researchers in the automatic detection of melanoma. This investigation focuses on the algorithms used for the automated detection of melanoma in dermoscopic images through a complete analysis of the stages of the proposed methodologies. It also examines the concepts associated with skin cancer disease and describes possible future directions through open problems in this area of research.

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	EXPLANATION
ABCD	Asymmetry, Border, Color, Diameter
AMD	Advanced Micro Devices
BCC	Basal Cell Carcinoma
BLAS	Basic Linear Algebra Subprograms
CNN	Convolutional Neural Network
FCRN	Fully Convolutional Residual Network
FORTRAN	Formula Translation
GA	Genetics Algorithm
GB	Gigabyte
GLCM	Grey Level Co-occurrence Matrix
HAM	Human Against Machine
HOG	Histogram Oriented Gradients
HSV	Hue, Saturation, Value
ISIC	International Skin Imaging Collaboration
KNN	K – Nearest Neighbours
LAPACK	Linear Algebra Package
MATLAB	MATrix LABoratory
MRI	Magnetic Resonance Imaging
RBF	Radial Basis Function
RGB	Red Blue Green
ROC	Receiver Operator Characteristic
ROI	Region of Interest
SCC	Squamous Cell Carcinoma
SK	Seborrhoeic Keratosis
SRN	Simple Recurrent Network
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
VGG	Visual Geometry Group

CHAPTER 1

INTRODUCTION

1.1 Motivation

Studies shows that approximately, two in three Australians will be diagnosed with skin cancer by the time they are 70. Melanoma is the fourth most commonly diagnosed cancer in Australia. In 2019, it is estimated that 15,229 new cases of melanoma skin cancer will be diagnosed in Australia (8,899 males and 6,330 females). It is estimated that number of people died of melanoma alone in Australia is 1,725 (1,190 males and 536 females) in 2019. In 2011-2015, individuals diagnosed with the melanoma skin cancer had a 91% chance (89% for males and 94% for females) of surviving for 5 years compared to their counterparts in the general Australian population. Melanoma is usually curable when detected and treated early. Once melanoma has spread deeper into the skin or other parts of the body, it becomes more difficult to treat and can be deadly. So, earlier diagnosis improves the chances of survival. For this, technology needs to be improved on many aspects, especially in the medical field, because of its sensitivity and effect in human lives which requires accurate and objective diagnosis.

On the other hand, smartphones have become one of our lives basis, they could be used in many ways. Skin Cancer Diagnosis System application in smartphones would be very accurate and easy to use. The results obtained by implementing machine learning and artificial intelligence techniques are accurate and reliable.

1.2 Objective

The objective of this project can be summarized into the following points:

Develop machine learning application that in general, has the ability to:

- Determine the affected areas in the image
- Determine the disease in specified region
- Detect skin cancer in people having varied skin colour
- Be applicable locally in Salem
- With improved accuracy and performance
- Reduce the time consumption during processing

CHAPTER 2

SYSTEM ANALYSIS

2.1 Literature Survey

Melanoma is a serious form of skin cancer that begins in cells known as melanocytes While it is less common than basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), melanoma is far more dangerous because of its ability to spread to other organs more rapidly if it not treated at an early stage. Skin diseases are among the most common health problems worldwide, a great work has been done by many researchers to develop computer aided systems to diagnose many types of Skin Diseases, various techniques were successfully implemented, as example in 2019, a new methodology was proposed by FirozWarsi et al. to extract the two features of color and texture as a single feature, and a multilayer back propagation neural network is used to classify melanoma and non-melanoma images.

In a study done by Ayan, (2018), the performance of a CNN architecture is compared between a non-augmented dataset and augmented dataset for classification of skin lesions. They proposed that the data augmentation methods could be useful for building powerful classifiers with insuffcient data. Results showed that the network using the augmented dataset has achieved better accuracy rate than non-augmented data.

Menegola et al. (2017) proposed knowledge transfer method to enhance performance of deep learning for melanoma screening. In their study, a pre-trained model trained on the Kaggle Challenge for Diabetic Retinopathy Detection dataset. They expected that the deeper models and transfer learning from a related dataset improve the performance and leads to better results. However, their work suggest that the experimental design is sensitive to the type of skin lesions (benign or malignant).

Lopez et al. (2017) emolyed a deep-learning based approach for early detection of melanoma. Their solution is based on a modified VGGNet architecture and transfer learning technique to solve the skin lesion classification task. The proposed method achieveed a sensitivity value of 78.66% on the ISIC Archive dataset.

Z. Waheed et al. (2017) extracted the color features from HSV domain. The texture feature is extracted using the GLCM method. These features are trained and tested in the SVM classifier. Three-fold cross validation is used to perform classification. Results are obtained using the above feature set and the results are: accuracy=96, sensitivity=97 and specificity=84.

Similarly, Suganya R et al. (2016) discussed classification of skin diseases such as Melanoma, Basal cell carcinoma (BCC), Nevus and Seborrheic keratosis (SK) by using the technique support vector machine (SVM). It yields the best accuracy from a range of other techniques.

On the other hand, the spread of chronic skin diseases in different regions may lead to severe consequences. Therefore, Alam et al. (2016) proposed a computer system that automatically detects eczema and determines its severity. The system consists of three stages, the first effective segmentation by detecting the skin, the second extract a set of features, namely color, texture, borders and third determine the severity of eczema using Support Vector Machine (SVM).

In 2016, a new approach is proposed by Kumar et al. to detect skin diseases, which combines computer vision with machine learning. The role of computer vision is to extract the features from the image while the machine learning is used to detect skin diseases. The system was tested on six types of skin diseases with accurately 95%.

Recently, neural networks have found growing usage in various health-related applications (Chen et al., 2018) and medical image segmentation such as segmentation of Gliomas in multi-sequence MRI (Pereira et al. 2015), brain tumor (Havaei et al. 2015), vessels (Nasr-Esfahani et al., 2016) and left ventricle on cardiac MRI (Avendi et al. 2016). Moreover, various methods have been adopted to handle segmentation by deep learning. In (Hoft et al. 2014), semantic scene segmentation was performed by computing histograms of oriented gradients (HOG) on depth of image and feeding the result to the network. They used 1449 images for training, validation and testing processes. Jafari et al. (Jafari et al., 2016) used CNNs for lesion segmentation. They segmented melanoma lesion by drawing two types of windows around each pixel as batch, global batch and local batch. Local batch extracted local texture information while global batch revealed general texture information in the vicinity of that pixel. Then, by merging the local and global information, they trained a CNN model for melanoma segmentation of non-dermoscopic images. Employing deconvolutional neural networks is a powerful means to returning back the resolution, size and fine details which were degraded during previous convolutional layers.

Yu et al. (2017) adopt residual learning and construct a fully convolutional residual network (FCRN) for melanoma segmentation task. They upgrade the performance by employing a multi-scale contextual information integration scheme.

Yuan et al. (2017) adopt a 19-layer fully convolutional neural network for melanoma segmentation. Their primary contribution is proposing a novel loss function based on the Jaccard distance which is substituted for previous rebalanced methods.

In 2016,Pravin S et al, have develop an Image analysis system to detects skin diseases, they develop a system to be used for early detection and prevention of the skin diseases and they target 3 main diseases skin cancer, psoriasis and dermatophilosis, the disease diagnosis and classification is built on statistical parameter analysis. Statistical parameters include: Entropy, Texture index, Standard deviation, Correlation, the user of the system will able to take images of different moles or skin patches. Then the system will analyze and process the image and classifies the image to normal, melanoma, psoriasis or dermo case based extracting the image features. An alert will be provided to the user to seek medical help if the mole belongs to the atypical or melanoma category, the input images firstly passed through a median filter to remove a remove the noise, then apply the image enhancement and the statistical analysis techniques, then two-level classifier is used the first level is to specify if

the image is either normal or abnormal and the second level is to classify into specified category: Melanoma, Psoriasis or dermo, the system is classify the images with accuracy 90%.

VinayshekharBannihatti Kumar et al, then provided an approach to detect 6 different skin diseases using smartphones in 2016, they have implemented dual stage approach combines machine learning and computer vision. The computer vision consists of two stages in the first stage eight preprocessing techniques were implemented in order to extract features of the image namely converting to grey scale image, sharpening filter, median filter, smooth filter, binary mask, RGB extraction, histogram and sobel operator, the extracted features are used as input for training two different models in the second stage, these models are Maximum Entropy model and Feedforward Artificial Neural Network with two hidden layers and Softmax output layer that learned using Backpropagation learning algorithm, this stage was developed for users that couldn't access the histopathological attributes. In the Machine Learning stage the histopathological attributes entered by the user combined with the features have been extracted from the image were used as input to train three different training models, Decision Tree, Feedforward Neural Network same as in [17], and K'th Nearest Neighbor. The novel method of using a dual stage system has given very promising results in identification of skin diseases with accuracies of up to 95%, although they got high accuracy but it decreases when tested with varying skin colors.

Finally,Suneel Kumar and Ajit Singh in the same year, also develop a Computer based skin disease detection system using digital image processing techniques for the classifications of the infected skin, the unique features of the images were extracted using two algorithms HSV-histogram and SURF algorithm, then the extracted features were fed in a K-NN classifier to classify the image to normal skin or infected, 5 classes were used in which 5 shows the normal skin and 1 to 4 is showing the infected skin (i.e. 1 for bloody, 2 for burned, 3 for cancer and 4 for allergic skin), this model got good accuracy, but it only classify the images into a general classification level and do not has a further detailed disease classes, but it could be very useful in medical field to see the clear image of the infected part in the skin as well as the parts that are not visible by human eyes.

2.2 Problem Statement

They applied selective median filtering followed by un sharp masking in preprocessing. Segmentation, they used improved version of fuzzy clustering technique. Gustafson clustering followed by nearest neighbor classification in color space. The computed features are two novel shape features Dimension and contour signature. Our system has the ability of learning from misclassified tests to enhance the future accuracy of the system. Random Forest classifier is the best classifier that is able to differentiate between different types and the one which gives us the best accuracy. The system achieved accuracy in detecting and classifying types and sub-types. These segmentation have shortcomings like localization, or demarcation which may lead to unacceptable results and unavoidable errors in glaucoma diagnosis.

2.3 Proposed System

The complete methodology of our system is represented in a flowchart. The individual steps are modularized and are often autonomous and sometimes dependent on each other. The proposed method, we introduce in this application can identify different types of cancer. The entire process of project can be perform in four phases. The first process is input and pre-process. In the sample image, the background is brighter in the center than at the bottom of the image. The pre-processing step before analysis, it identifies the background uniform and then convert the image into a specific image format. This is to make the background illumination more accurate. The methods of per-processing inserted between the various-stages of the segmentation, in order to make the latter simpler and more robust. Segmentation of image is used to find objects and boundaries in images. More precisely, segmentation of image is the process of giving a label to every pixel in an image such that pixels with the same label share certain characteristics.

CHAPTER 3

SYSTEM REQUIREMENTS

3.1 Hardware Requirements

Operating System	: Windows 10
Disk	: 2.9 GB for MATLAB only, 5-8 GB for a typical installation
Processors	: Any Intel or AMD x86-64 processor

3.2 Software Requirements

Operation System	: Windows 10
Tool Used	: MATLAB
Language Used	: C
Dataset Used	: HAM10000

3.2.1 MATLAB Software Description

MATLAB a high-performance language for technical computing integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. It is a prototyping environment, meaning it focuses on the ease of development with language flexibility, interactive debugging, and other conveniences lacking in performance-oriented languages like C and FORTRAN. While Matlab may not be as fast as C, there are ways to bring it closer. We want to spend less time total from developing, debugging, running, and until obtaining results.

It is an interactive system whose basic data element is an array that does not require dimensioning. It allows you to solve many technical computing problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar non interactive language such as C or FORTRAN.

The name MATLAB stands for *matrix laboratory*. MATLAB was originally written to provide easy access to matrix software developed by the LINPACK and EISPACK projects. Today, MATLAB engines incorporate the LAPACK and BLAS libraries, embedding the state of the art in software for matrix computation. It has evolved over a period of years with input from many users. In university environments, it is the standard instructional tool for introductory and advanced courses in mathematics, engineering, and science. In industry, MATLAB is the tool of choice for high-productivity research, development, and analysis.

Its features a family of add-on application-specific solutions called *toolboxes*. Very important to most users of MATLAB, toolboxes allow you to *learn* and *apply* specialized technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. You can add on toolboxes for signal processing, control systems, neural networks, fuzzy logic, wavelets, simulation, and many other areas.

3.2.2 Graphics

MATLAB has extensive facilities for displaying vectors and matrices as graphs, as well as annotating and printing these graphs. It includes high-level functions for two-dimensional and three-dimensional data visualization, image processing, animation, and presentation graphics. It also includes low-level functions that allow you to fully customize the appearance of graphics as well as to build complete graphical user interfaces on your MATLAB applications.

3.2.3 External Interfaces

The external interfaces library allows you to write C and Fortran programs that interact with MATLAB. It includes facilities for calling routines from MATLAB (dynamic linking), for calling MATLAB as a computational engine, and for reading and writing MAT-files.

3.2.4 Array Pre-allocation

Mat lab's matrix variables have the ability to dynamically augment rows and columns. Matlab automatically resizes the matrix. Internally, the matrix data memory must be reallocated with larger size. If a matrix is resized repeatedly like within a loop this overhead can be significant. To avoid frequent reallocations, pre allocate the matrix with the zeros command.

3.2.5 Block Creation

The built-in modeling functionality provided by Simulink, you can create custom blocks and add them to the Simulink Library Browser. You can create a custom block from a MATLAB function MATLAB Function blocks enable you to use the MATLAB language to define custom functionality. These blocks are a good starting point.

- You have an existing MATLAB function that models the custom functionality.
- You find it easier to model custom functionality using a MATLAB function than using a Simulink block diagram.
- The custom functionality does not include continuous or discrete dynamic states such as masking a subsystem of other blocks, or by incorporating C, C++, or Fortran code.

3.2.6 Dataset

The HAM10000 dataset is where we will get the images needed to train our model. This is a collection of around 10,000 labelled images of skin lesions. Inmachine learning there are two different phases namely training phase and testing phase.

- **Training Phase :** In this phase, 200 images of melanoma are used for training.
- **Testing Phase :** In this phase, test images are given to the classifier and the classifier uses knowledge gained during the training phase to classify the test image.

CHAPTER 4

SYSTEM DESIGN

4.1 Block Diagram And Database Design

4.1.1 Block Diagram

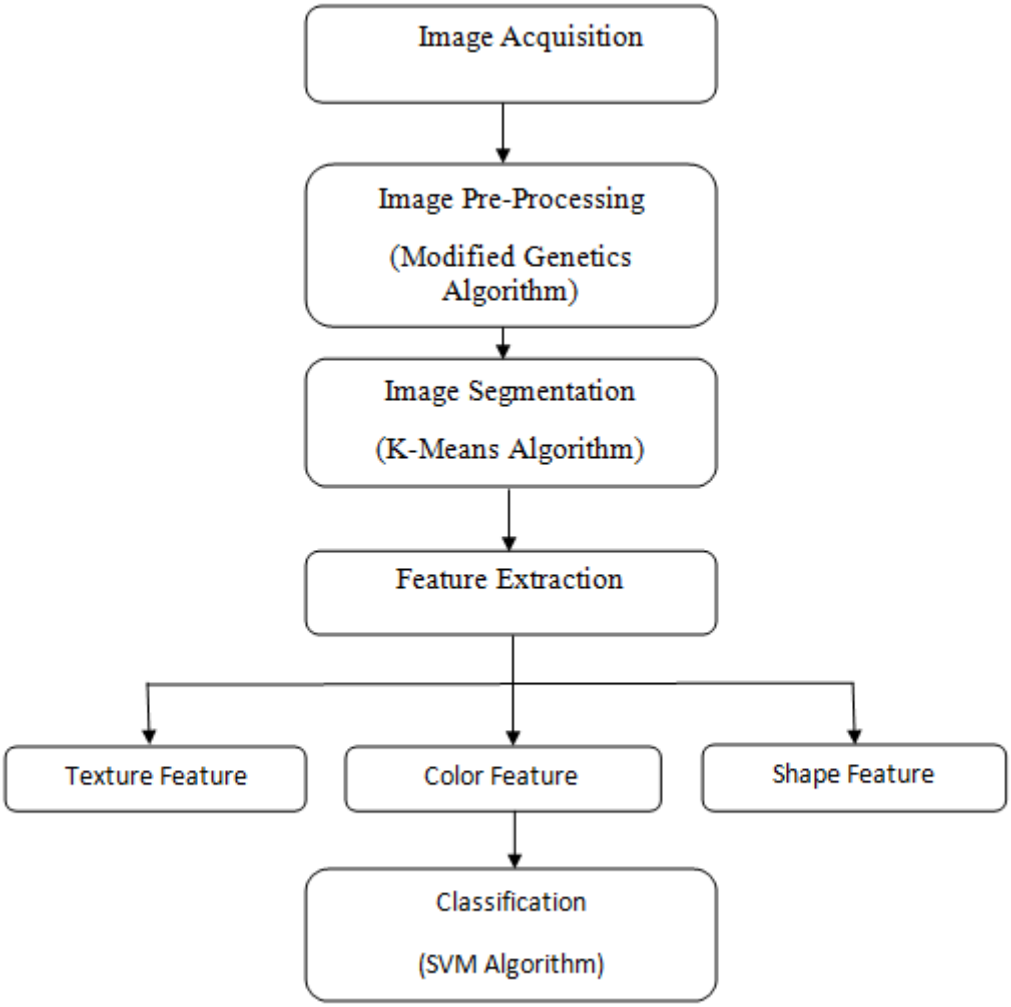


Fig 4.1.1. Block Diagram for Skin Cancer Detection

4.1.2 Database Design

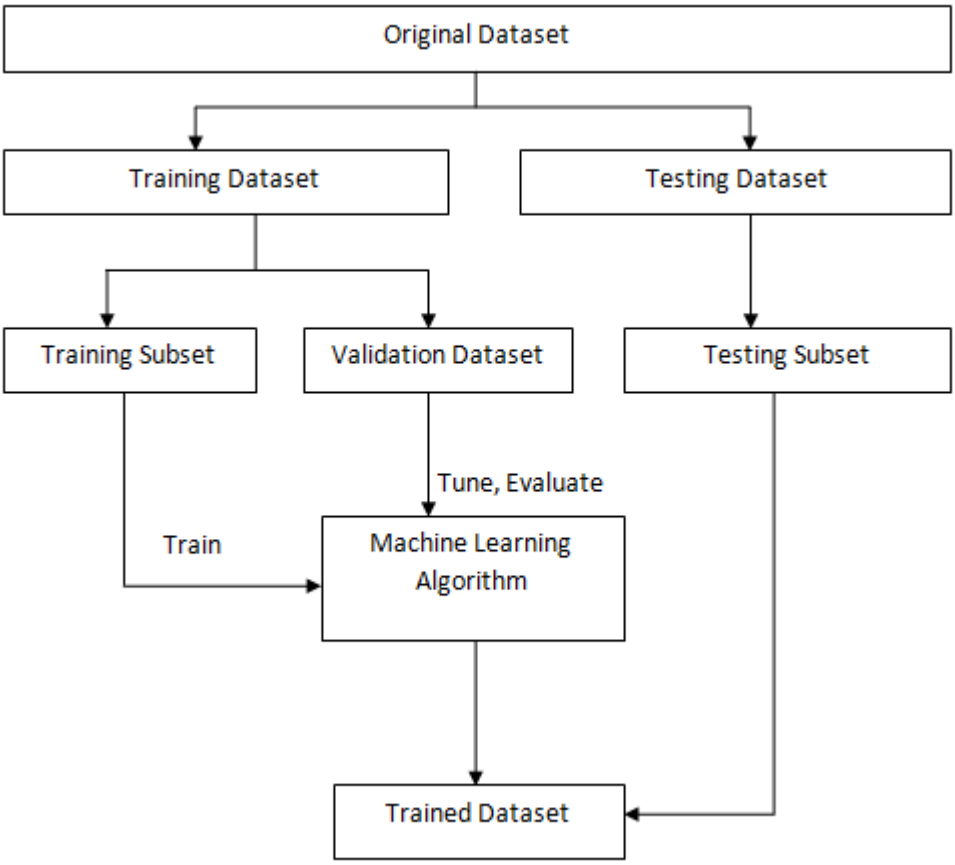


Fig 4.1.2. Block Diagram for Dataset Processing

In machine learning there are two different phases namely training phase and testing phase.

Training And Validation Set

All images were obtained from the International Skin Imaging Collaboration (ISIC) archive; most images came from the HAM 10000 Dataset [14]. This archive contains dermoscopic images of heterogeneous populations that are publicly accessible, anonymous and taken by different camera systems. Because some images from the HAM 10000 Dataset show the same lesion from different magnifications and angles, duplicates were removed by lesion ID (provided by HAM 10000 creators) so that only one image per lesion was used. This data set was supplemented with an additional 4291 images from the ISIC archive. A sufficiently large number of images per skin disease should be presented to dermatologists to attain a statistically reliable value for sensitivity and specificity. For this reason, we limited this study to the differential diagnoses of the most frequent skin diagnoses in the archive, split into five compound classes: (1) actinic keratosis (solar keratosis), intraepithelial carcinoma (Bowen

disease), squamous cell carcinoma (akiec); (2) basal cell carcinoma (bcc); (3) benign keratosis, including seborrhoeic keratosis, solar lentigo and lichen planuse-like keratosis (bkl); (4) melanocytic nevi (nv) and (5) melanoma (mel). Class division was based on the diagnostic categories set up by the HAM 10000 Dataset creators. Using these restrictions, this study used 11,444 images, 6390 of which had been biopsy verified.

Test Set

In this study, only biopsy-verified images from the HAM 10000 Dataset were used for evaluating the algorithm. To prevent selection bias for the 300 test images (60 for each of the five disease classes) from the available biopsy-verified image set, we programmed a random generator in MATLAB using C. The test set is available for downloading at (https://skinclass.de/TestSet.zip).

CHAPTER 5

SYSTEM IMPLEMENTATION

5.1 Module Description

5.1.1 Pre-Processing

Genetic Algorithm

Genetic Algorithm is stochastic search algorithms mainly inspired by the genetic process of biological organisms. It searches for the finest solution in search space sharply in an iterative order to attain a new generation from the old one. Genetic algorithms are used to find the subset of features where the chromosome bits correspond the feature included or not. Genetic algorithms are dominant search algorithms that can be useful to a wide range of problems [3]. For organizing a genetic algorithm parameter setting is employed, which remains unaffected during execution. The interesting problem here is the self-adaptive parameter adjustment of genetic algorithm. This paper proposes an approach for controlling the parameters of genetic algorithm that can be determined within the chromosome of each individual. Depending on the problem situation, the parameter values are totally dependent on the evolution mechanism. Our initial analysis demonstrate that GA has the ability to learn and assess the value of self-set parameter depending on the degree of contribution to the decree of the problem [8]. These results indicate the possible methodolgy for the development of GAs with self-adaptive parameter settings which do not require the user to modify parameter at the beginning. The basic mechanism of GA depends on various key parameters such as crossover operators, mutation operators, crossover probability, mutation probability, mechanism of selection and the size of population [13]. All these parameter have a great impact on GA's performance [12].One can thus describe the parameters quickly and later modify the parameters in a exact manner depending on the problem. A high rate of crossover and low rate of mutation might be very good in the study of new solutions for the first generations formed by the algorithm [12] [13]. Similarly while the algorithm is close to the optimal solution, it becomes unfavourable. One of the probable solution to decide the best set of parameters e in the exist in use of learning, involving an enrichment of the GA's performance. It is suggested that this improvement relies on selfadaptive parameters in GA. In this study a hypothesis of selfadaptive parameters in GA is proposed for further work. A selfadaptive approach to GA parameter based on the problem is considered in this work.The individual is based on two learning levels. (1) a genetic algorithm is applied to the knowledge of new sets of parameters which results in an improvement of the individual's adaptation to the problem to be solved. (2) reinforcement learning is used to learn the finest parameter settings. This learning occurs in communication with the problem situation. Lastly, the individual evaluate the quality of learned parameter settings [14].

An algorithm of the self-adaptive parameter approach is proposed which is a customized version of the basic genetic algorithm structure.

Step 1: generate random population

Step 2: estimate the individual population fitness value

Step 3: a new population is created- repeat steps until stopping condition is reached

Step 3a: to discover the best parameter settings for each individual use reinforcement learning depending on the problem.

Step 3b: depending on the fitness value, choose two parents from a population

Step 3c: perform search in genetic crossover space search

Step 3d: to form best fitness new offspring cross the two individuals

Step 3e: to identify the best parameter for individual based on fitness and the problem, search in genetic mutation

Step 3f: in each position of the chromosome- mutate the new offspring Step 3g: place new offspring in the population

Step 3h: develop a new set of parameters

Step 3i: based on the reinforcement location select two parameter settings

Step 4: continue by generating newly developed population

Step 5: terminate if end condition satisfies-return the best solution

Step 6: return to step [2]

The proposed method is provoked in part by the theory of independent individual. Various authors have recommended that the control of GAs' parameters could be encoded in the chromosome of every individual of the population [15]. This suggest the inclusion of a mechanism, which inserts the parameters in the individual's chromosomes, the changes of parameter values are thus completely reliant on the evolution mechanism. There are no suggested best settings for parameters to any problem. The operators used are crossover operators, mutation operators, crossover probability and mutation probability.

5.1.2 Image Segmentation

Image segmentation is the process of partitioning the digital images into multiple segments. Segmentation is the process of straight forward approach. Segmentation is used to identify the object (or) other related information. The goal of segmentation is to simplify and (or) change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects &boundaries (lines & curves) in images.

Image segmentation is an important process in many computer vision and image processing applications, because people are interested in certain parts of the image. It divides an image into a number of discrete regions such that the pixels have high similarity in each region and high contrast between regions. Properties like gray-level, color, intensity, texture, depth or motion help to recognize similar regions and similarity of such properties, is used to construct groups of regions having a specific meaning. Segmentation is a valuable tool in many fields including industry, health care, image processing, remote sensing, traffic image, content based image, pattern recognition, video and computer vision etc. A particular type of image segmentation method can be found in application involving the detection, recognition, and measurement of objects in an image. Till now many researchers have focused on gray-level image segmentation, whereas we know that color images carry most of the information.
Segmentation techniques can be classified into the following categories: Edge-based, Threshold based, Region-based, Neural Network based, Cluster-based, and Hybrid. Image segmentation based on is one of the oldest and powerful technique, since the threshold value divides the pixels in such a way that pixels having intensity value less than threshold belongs to one class while pixels whose intensity value is greater than threshold belongs to another class.

The Neural Network based image segmentation techniques reported in the literature can mainly be classified into two categories: supervised and unsupervised methods. Supervised methods require expert human input for segmentation. Usually this means that human experts are carefully selecting the training data that is then used to segment the images. Unsupervised methods are semi or fully automatic.

An unsupervised segmentation method automatically partitions the images without operator intervention. However, these architectures might be implemented using application specific a priori knowledge at design time, i.e. anatomical, physical or biological knowledge. Clustering is an unsupervised learning technique, where one needs to know the number of clusters in advance to classify pixels .A similarity condition is defined between pixels, and then similar pixels are grouped together to form clusters. Among these K-means clustering is used for segmenting cancerous skin.

5.1.3 Image Filtering

The main goal of image enhancement is to process an image so it appears more acceptable or pleasing way. The removal of noise, the sharpening of image edges and blurring effects are all some popular enhancement techniques. These operations can be achieved through the process of filtering.

Filters act on an image to change the values of the pixels depends on the method used. Each of the pixels in an image, the pixel under consideration at a given movement is called as target pixel and it is successively addressed. Filtering operations over an image are performed as a series of local neighborhood operations using a sliding-window-based principle.

5.1.4 Feature Extraction

The feature of the image is an important thing that helps for isolation of the common properties of the image and as well detecting and naming the regions. It is the primary characteristics of the image. Few are detected by visual appearance certain features are provided by the artificial methods. The luminance of a pixels and grey scale textual regions are the natural features. Image amplitude histograms and spatial frequency spectra are those examples for the artificial features. Isolation of common property in image and subsequent identification of few features image.

5.1.4 Support Vector Machine (SVM)

The proposed classification procedure followed, resulting in the final object classification. The classification results were compared to the Nearest Neighbour object-based classifier results, and were found satisfactory. The SVM methodology seems very promising for Object Based Image Analysis and future work will focus on integrating SVM classifiers with rule-based classifiers.

SVM is a supervised learning algorithm used for regression and classification problems. It is a prediction tool which uses theory of machine learning to maximize accuracy of prediction while automatically avoiding data over fitting. For a linearly separable case, SVM works by solving the following optimization problem. In this method, Logistic regression and support vector machines (SVM) were trained and explored for performing classification.

Map the data to a predetermined very high dimensional space via a kernel function Find the hyper plane that maximizes the margin between the two classes. If data are not separable - find the hyper plane that maximizes the margin and minimizes the (weighted average of the) misclassifications Particularly, kernel trick method based on radial basis function (RBF) kernel was adopted for the SVM where used for training the logistic regression classifier of machine-learning algorithms that are used for mathematical and engineering problems including for example handwriting digit recognition, object recognition, speaker identification, face detections in images and target detection.

SVM performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two categories. Among the possible hyper planes, we select the one where the distance of the hyper plane from the closest data points (the “margin”) is as large as possible. An intuitive justification for this criterion is: suppose the training data are good, in the sense that every possible test vector is within some radius r of a training vector.

Then, if the chosen hyper plane is at least r from any training vector it will correctly separate all the test data. By making the as far as possible from any data, r is allowed to be correspondingly large. The desired that maximizes the margin) is also the bisector of the line between the closest points on the convex hulls of the two data sets.

ABCD Rule

As per ABCD rule the features which we need to extract include Asymmetry Index Border Colour Index Diameter.

A-Asymmetry Index

Asymmetry Index is computed with the following equation:

AI=(A1+A2)/2Ar

Where, A1= Area of non-overlapped region along minor axis of the lesion

A2= Area of non-overlapped region along major axis of the lesion

Ar= Area of lesion Implementation

Area of lesion (Ar) can be calculated using bw area over the binary image of the segmented region. For calculating non overlapped area over axis. The segmented region is divided along the lines passing through centroid of the region Two separate areas are generated which are then adjusted so that the areas will be overlapped by flipping one area. Using XOR over the area will generate the non-overlapped region whose area is calculated using bw area function To generate area along x axis the bisection will be generated using first Gx pixels and the next Gx pixels along x axis and bisecting line on y axis. To generate area along y axis the bisection will be generated using first Gy pixels and the next Gy pixels along x axis and bisecting line on y axis. After calculating area of the regions Asymmetry index is calculated using the specified formula.

B-Border Irregularity

In order to calculate border irregularity, there are different measures such as: compactness index, fractal index, edge abruptness.

- Compact Index: Compact Index can be determined by using the following equation: CI = (P2L) = (4AL) Where, PL = Perimeter of the Lesion. AL = Area of the Lesion.
- Fractal Dimension: Fractal set is provided by the”box counting” method. It returns two variables whose differential log ratio provides the fractal dimension as the mean value along 4-7 index.
- Edge Variation: Edge variation is calculated using the following equation EI= ((Max Min) %6+2)/100; Where, Max and min are length of major and minor axis. Axis lengths are calculated using region props function.

C-Colour Index

Colour index is calculated by converting the input image to have image value by checking the presence of the following colours. Length of all the available pixels with given values is divided by total number of pixels. The presence of colour is dependent on the value of resultant not equal to zero. For each colour present the Colour Index is +1.

D-Diameter

The diameter value is said to be 5 if the diameter of lesion is greater than 6mm. For other values the diameter is one less than its actual rounded value. To calculate Diameter the region props function is used to get the minor axis length of the lesion region. Resultant value is converted into mm value and the value is assigned to diameter.

5.1.5 Evaluation Metrics

Accuracy

The accuracy obtained are briefly discussed in the form of classification performance. Four possible outcomes for binary class classification i.e. true positive, true negative, false positive, false negative.

The performance of binary class classification can be calculated in term of accuracy. Accuracy defines the number of positive cases to the total number of test case.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision / Recall

Sometimes, it's easier to evaluate a model's performance using numbers rather than relying on a library to visualize a confusion matrix.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

If the precision is equal to 100% and the recall is equal to 0% then the harmonic mean is equal to 0%. We call this value the **F1 Score**.

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

ROC / AUC

Similarly to the precision/recall curve, the Receiver Operator Characteristic (ROC) graph provides an elegant way of presenting multiple confusion matrices produced at different

thresholds. A ROC plots the relationship between the true positive rate and the false positive rate.

- True positive rate = Recall = Sensitivity = true positive / (true positive + false negative)
- False positive rate = 1 – specificity = false positive / (false positive + true negative)

5.1.6 Advantages

- ❖ To support early detection, diagnosis and optimal treatment.
- ❖ Image segmentation plays an essential role in many medical applications.
- ❖ Low SNR conditions and various artifacts makes its automation challenging.
- ❖ To achieve robust and accurate segmentation**.**

5.2 Execution Procedures

- ❖ Firstly, the image is taken as an input from the user
- ❖ Then it is processed and filtered to remove noisy and unwanted information with the help of modified genetic algorithm and gray scale filter
- ❖ Five phases are considered in a genetic algorithm.
    - Initial population
    - Fitness function
    - Selection
    - Crossover
    - Mutation

- ❖ Next, the image is segmented with the help K-means clustering algorithm. The way kmeans algorithm works is as follows:
    - Specify number of clusters *K*
    - Initialize centroids by first shuffling the dataset and then randomly selecting *K* data points for the centroids without replacement
    - Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.

- ❖ Then, compute the sum of the squared distance between data points and all centroids. Assign each data point to the closest cluster (centroid). Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.
- ❖ Finally, the method used in feature extraction is based on the border irregularities behaviour. The method explored a concept of border representation, where two

kinds of features are extracted from the border, which are valleys and crevasses geometrical forms. These two new concepts can considered as first and second level of lesion border irregularities. These features are learned by SVM classifier with linear kernel. The results obtained show a higher performances of classification compared to the literature's results using the same database. This idea of characterization of border could be implemented in a typical flow of ABCD rule method, it represents more variability than the method based on eight segments from the border used currently in ABCD rule. The challenge fixed in the third objective, which is the exploration of the border irregularities in formation of, is achieved.

CHAPTER 6

CONCLUSION AND FUTURE WORK

Conclusion

Skin cancer is the major causes of death globally and the early detection of this is very important. The computer aided cancer detection system helps the physician for cancer diagnosis. From the analysis it is examined that, edge detection operator works good for edge detection technique. The automatic detection of cancerous cells in skin possesses various challenges. The automated system should correctly identify the intended objective at the earliest. The accuracy of the automated system greatly depends on type of segmentation algorithm employed, the feature to be extracted and the classification methodology incorporated.

Future Work

The proposed modifications of the skin cancer diagnosis system are generally to increase the performance of the system, resolve the system limitations, or to increase its capability.

So, there are several suggested modifications to both the system core model and the system mobile interface:

- Increase the training data used for training the model, not only in term of quantity but also obtaining more data from different resources namely collecting data from hospitals and healthcare centers, to increase the learning model generalization.
- Apply better preprocessing techniques to resolve the images distortions.
- Apply training data of more classes, that the model will be capable to recognize and diagnose more diseases.
- Develop a cross-platform application to work on different mobile platforms, which will increase the number of system users.

REFERENCES

[1] Lopez, A.R., Giro-i-Nieto, X., Burdick, J., Marques, O., 2017. “Skin lesion classification from dermoscopic images using deep learning techniques”. 2017 13th IASTED International Conference on Biomedical Engineering (BioMed). pp. 49–54.

[2] Menegola, A., Fornaciali, M., Pires, R., Bittencourt, F.V., Avila, S., Valle, E., 2017. “Knowledge transfer for melanoma screening with deep learning”. Proceedings International Symposium on Biomedical Imaging.

[3] NishimaSachdeva(Mtech scholar), Professor Rohan Gupta ―Hybrid Approach To Investigate The Probability Of Skin Cancer By ABCD And PCA Method‖ ISSN: 2278 – 909X International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 6, Issue 4, April 2017

[4] Ayan, E.H.M.Ü., 2018. “Data augmentation importance for classification of skin lesions via deep learning”. Electric Electronics, Computer Science, Biomedical Engineerings’ Meeting (EBBT).

[5]Waheed Z, Waheed A, Zafar M, Riaz F. “An efficient machine learning approach for the detection of melanoma using dermoscopic images”. International Conference on Communication, Computing and Digital Systems (C-CODE), 8-9 March. 2017.

[6] Suganya R. (2016) “An Automated Computer Aided Diagnosis of Skin Lesions Detection and Classification for Dermoscopy Images.” International Conference on Recent Trends in Information Technology.

[7] Alam, N., Munia, T., Tavakolian, K., Vasefi, V., MacKinnon, N., &Fazel-Rezai, R. (2016) “Automatic Detection and Severity Measurement of Eczema Using Image Processing.” IEEE.

[8] Ani1 K. Jain, Fellow, IEEE, Robert P.W. Duin, and Jianchang Mao ―Statistical Pattern Recognition: A Review‖ IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 22, No. 1, January 2000

[9] Kumar, V., Kumar, S., &Saboo, V. (2016) “Dermatological Disease Detection Using Image Processing and Machine Learning.” IEEE.

[10] Chen, T., Zhang, X., Jiang, H., Asaeikheybari, G., Goel, N., Hooper, M. W., et al. (2018). “Are you smoking? Automatic alert system helping people keep away fromcigarettes”. Smart Health, 9, 158–169.

[11] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2015). “Deep convolutional neural networks for the segmentation of gliomas in multi-sequence mri”. In Internationalworkshop on brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries (pp. 131–143). Springer.

[12] MineichiKudo1 And Jack Sklansky ―A Comparative Evaluation Of Medium and Large - Scale Feature Selectors For Pattern Classifiers‖ Kybernetika — Volume 34 (1998), Number 4, PAGES 429 -434

[13] AlirezaFasih, Jean Chamberlain Chedjou ― Cellular Neural Networks-Based Genetic Algorithm for Optimizing the Behavior of an Unstructured Robot ― Transportation Informatics Group, Institute of Smart Systems Technologies, University of Klagenfurt Klagenfurt

[14] Havaei, M., Dutil, F., Pal, C., Larochelle, H., &Jodoin, P.-M. (2015). "A convolutional neural network approach to brain tumor segmentation". In International workshop onbrainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries (pp. 195–208). Springer.

[15] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S. R., Ward, K., Jafari, M. H., et al. (2016). "Vessel extraction in x-ray angiograms using deep learning. InEngineering in medicine and biology society (EMBC)", 2016 IEEE 38th annual international conference of the (pp. 643–646). IEEE.

[16] Avendi, M., Kheradvar, A., &Jafarkhani, H. (2016). "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle incardiac mri". Medical Image Analysis, 30, 108–119.

[17] Jafari, M. H., Nasr-Esfahani, E., Karimi, N., Soroushmehr, S., Samavi, S., &Najarian, K. (2016). "Extraction of skin lesions from non-dermoscopic images using deep learning". arXiv preprint 1609.02374.

[18] Yuan, Y., Chao, M., & Lo, Y.-C. (2017). "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance". IEEE Transactions on MedicalImaging.

[19] Yu, L., Chen, H., Dou, Q., Qin, J., &Heng, P.-A. (2017). "Automated melanoma recognition in dermoscopy images via very deep residual networks". IEEE Transactions onMedical Imaging, 36, 994–1004.

[20] FirozWarsia, RuqaiyaKhanama, SurajKamyab, Carmen Paz Suárez-Araujoc. "An efficient 3Dcolor-texture feature and neural network technique for melanoma detection". Elsvier, Informatics in Medicine Unlocked, 2019.

[21] Pravin S. Ambad1, A. S. Shirsat, "A Image analysis System to Detect Skin Diseases," www.iosrjournals.org, Vadgaon, Pune, India, 2016.

[22] VinayshekharBannihatti Kumar, Sujay S Kumar and VarunSaboo, "Dermatological Disease Detection Using Image Processing And Machine Learning," IEEE, 2016.

[23] Suneel Kumar and Ajit Singh, "Image Processing for Recognition of Skin Diseases," International Journal of Computer Applications (0975 – 8887), vol. 149, 2016.

APPENDIX I – SOURCE CODE

```
function varargout = modulethree(varargin)
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @modulethree_OpeningFcn, ...
                   'gui_OutputFcn',  @modulethree_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin&&ischar(varargin{1})
gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
gui_mainfcn(gui_State, varargin{:});
end
handles.output = hObject;
ss = ones(200,200);
axes(handles.axes1);
imshow(ss);
axes(handles.axes2);
imshow(ss);
function varargout = modulethree_OutputFcn(hObject, eventdata, handles)
function pushbutton1_Callback(hObject, eventdata, handles)
[FileName,PathName] = uigetfile('*.jpg;*.png;*.bmp','Pick an MRI
Image');
if isequal(FileName,0)||isequal(PathName,0)
warndlg('User Press Cancel');
else
    P = imread([PathName,FileName]);
    P = imresize(P,[200,200]);
  axes(handles.axes1)
imshow(P);title('SKIN Image');
handles.ImgData = P;
guidata(hObject,handles);
end

function pushbutton2_Callback(hObject, eventdata, handles)
if isfield(handles,'ImgData')
        I = handles.ImgData;
gray = rgb2gray(I);
level = graythresh(I);
img = im2bw(I,.6);
img = bwareaopen(img,80);
img2 = im2bw(I);
axes(handles.axes2)
imshow(img);title('Segmented Image');

handles.ImgData2 = img2;
guidata(hObject,handles);

end
```

```
function pushbutton4_Callback(hObject, eventdata, handles)
if isfield(handles,'ImgData')
        I = handles.ImgData;

gray = rgb2gray(I);
level = graythresh(I);
img = im2bw(I,.6);
img = bwareaopen(img,80);
img2 = im2bw(I);
axes(handles.axes2)
imshow(img);title('Segmented Image');
handles.ImgData2 = img2;
guidata(hObject,handles);
signal1 = img2(:,:);

[cA1,cH1,cV1,cD1] = dwt2(signal1,'db4');
[cA2,cH2,cV2,cD2] = dwt2(cA1,'db4');
[cA3,cH3,cV3,cD3] = dwt2(cA2,'db4');

DWT_feat = [cA3,cH3,cV3,cD3];
G = pca(DWT_feat);
whosDWT_feat
whos G
g = graycomatrix(G);
stats = graycoprops(g,'Contrast Correlation Energy Homogeneity');
Contrast = stats.Contrast;
Correlation = stats.Correlation;
Energy = stats.Energy;
Homogeneity = stats.Homogeneity;
Mean = mean2(G);
Standard_Deviation = std2(G);
Entropy = entropy(G);
RMS = mean2(rms(G));
Variance = mean2(var(double(G)));
a = sum(double(G(:)));
Smoothness = 1-(1/(1+a));
Volume = biovolume(double(G(:)));
Breadth = skewness(double(G(:)));

m = size(G,1);
n = size(G,2);
in_diff = 0;
for i = 1:m
    for j = 1:n
        temp = G(i,j)./(1+(i-j).^2);
in_diff = in_diff+temp;
    end
end

Dimension = double(in_diff);

feat = [Contrast,Correlation,Energy,Homogeneity, Mean,
Standard_Deviation, Entropy, RMS, Variance, Smoothness, Volume,
Breadth, Dimension];

load Trainset.mat
xdata = meas;
```

```
 group = label;
svmStruct1 = svmtrain(xdata,group,'kernel_function', 'linear');
 species = svmclassify(svmStruct1,feat,'showplot',false);

set(handles.edit5,'string',Mean);
set(handles.edit6,'string',Standard_Deviation);
set(handles.edit7,'string',Entropy);
set(handles.edit8,'string',RMS);
set(handles.edit9,'string',Variance);
set(handles.edit10,'string',Smoothness);
set(handles.edit11,'string',Volume);
set(handles.edit12,'string',Breadth);
set(handles.edit13,'string',Dimension);
set(handles.edit14,'string',Contrast);
set(handles.edit15,'string',Correlation);
set(handles.edit16,'string',Energy);
set(handles.edit17,'string',Homogeneity);

q =  getimage(handles.axes2)
q=double(q(:));
ima=max(q(:));
imi=min(q(:));
ims=std(q(:));
snr = 20*log10((ima-imi)./ims)

textLabel = sprintf('COPD = %f',snr*3);
set(handles.edit21, 'String', textLabel)

end

function pushbutton5_Callback(hObject, eventdata, handles)
load Trainset.mat
Accuracy_Percent= zeros(200,1);
itr = 100;
hWaitBar = waitbar(0,'Evaluating Maximum Accuracy with 100
iterations');
for i = 1:itr
data = meas;
groups = ismember(label,'MALIGNANT');
[train,test] = crossvalind('HoldOut',groups);
cp = classperf(groups);
svmStruct =
svmtrain(data(train,:),groups(train),'showplot',false,'kernel_function'
,'linear');
classes = svmclassify(svmStruct,data(test,:),'showplot',false);
classperf(cp,classes,test);
Accuracy_Percent(i) = cp.CorrectRate.*100;
sprintf('Accuracy of Linear Kernel is: %g%%',Accuracy_Percent(i))
waitbar(i/itr);

end


delete(hWaitBar);
Max_Accuracy = max(Accuracy_Percent);
sprintf('Accuracy of Linear kernel is: %g%%',Max_Accuracy)
set(handles.edit2,'string',Max_Accuracy);
```

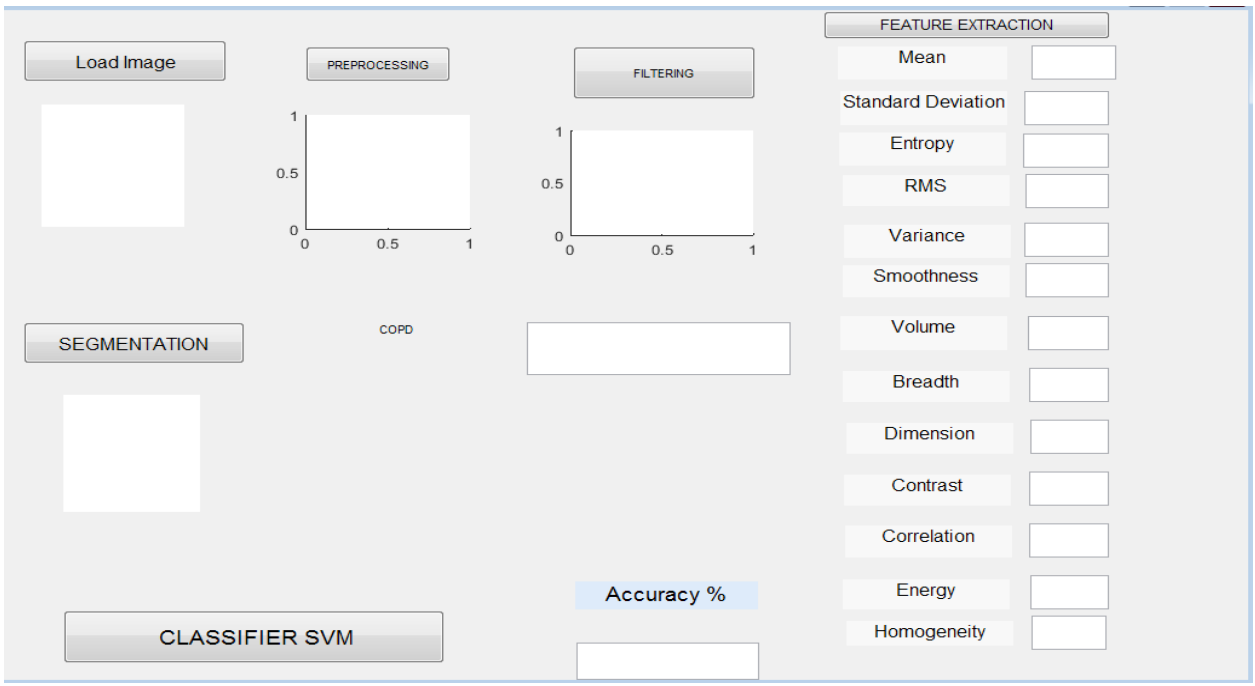
```
function edit2_Callback(hObject, eventdata, handles)
function edit2_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit3_Callback(hObject, eventdata, handles)
function edit3_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function pushbutton4_CreateFcn(hObject, eventdata, handles)
function edit5_Callback(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit6_Callback(hObject, eventdata, handles)
function edit6_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit7_Callback(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit8_Callback(hObject, eventdata, handles)
function edit8_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit9_Callback(hObject, eventdata, handles)
function edit9_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit10_Callback(hObject, eventdata, handles)
function edit10_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit11_Callback(hObject, eventdata, handles)
function edit11_CreateFcn(hObject, eventdata, handles)
```

```
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit12_Callback(hObject, eventdata, handles)
function edit12_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit13_Callback(hObject, eventdata, handles)
function edit13_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit14_Callback(hObject, eventdata, handles)
function edit14_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit15_Callback(hObject, eventdata, handles)
function edit15_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit16_Callback(hObject, eventdata, handles)
function edit16_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function edit17_Callback(hObject, eventdata, handles)
function edit17_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
endfunctionedit19_Callback(hObject, eventdata, handles)
function edit19_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function pushbutton7_Callback(hObject, eventdata, handles)
aa=getimage(handles.axes1)

x=rgb2ntsc(aa);
x(:,:,1)=histeq(x(:,:,1));
c2=ntsc2rgb(x);
```

```
rIMAGE=histeq(c2(:,:,1));
gIMAGE=histeq(c2(:,:,2));
bIMAGE=histeq(c2(:,:,3));
x=cat(3,rIMAGE,gIMAGE,bIMAGE);

axes(handles.axes3);
G=imnoise(x,'gaussian',0.0005,0.0019);
h = fspecial('average', 3);
F=imfilter(G,h);

imshow(F)
function pushbutton10_Callback(hObject, eventdata, handles)
aa=getimage(handles.axes3)

x=rgb2ntsc(aa);
x(:,:,1)=histeq(x(:,:,1));
c2=ntsc2rgb(x);

rIMAGE=histeq(c2(:,:,1));
gIMAGE=histeq(c2(:,:,2));
bIMAGE=histeq(c2(:,:,3));
x=cat(3,rIMAGE,gIMAGE,bIMAGE);

axes(handles.axes4);
G=imnoise(x,'gaussian',0.0015,0.0029);
h = fspecial('average', 3);
F=imfilter(G,h);

imshow(F)

function edit21_Callback(hObject, eventdata, handles)
function edit21_CreateFcn(hObject, eventdata, handles)
if ispc&&isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end
```

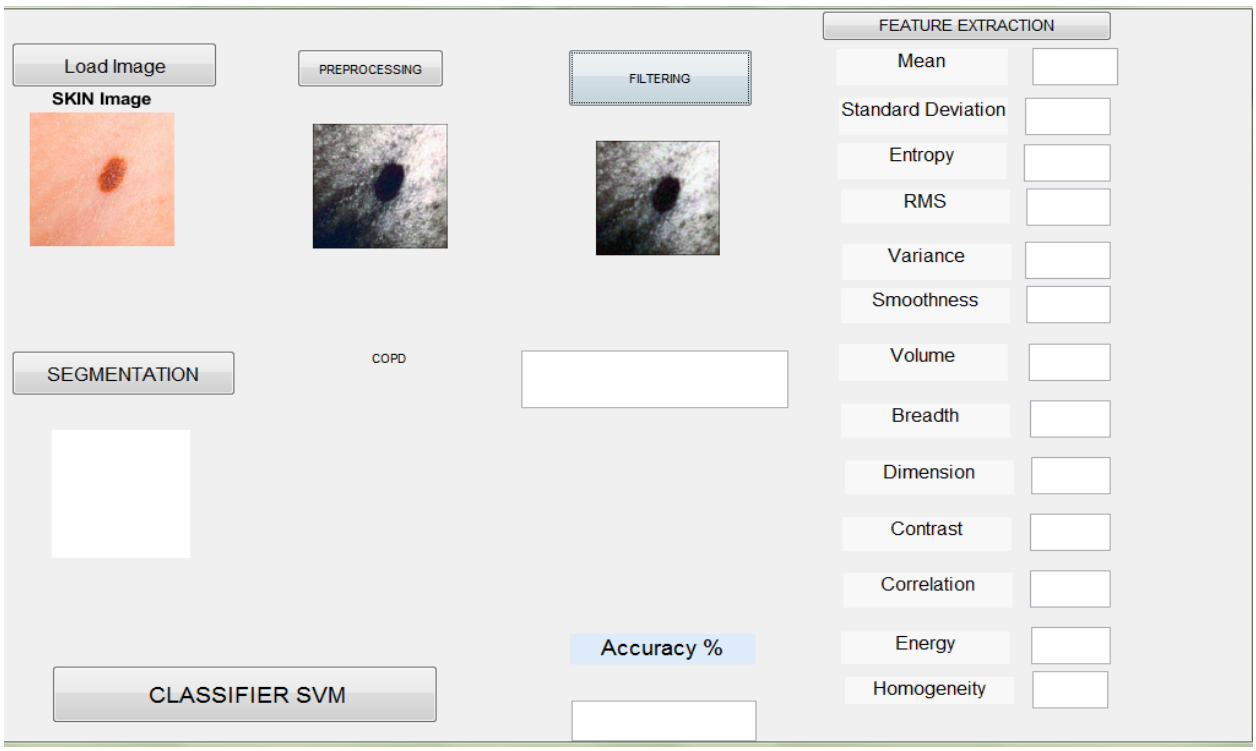
APPENDIX II – SCREENSHOTS

SIMULATION OUTPUT



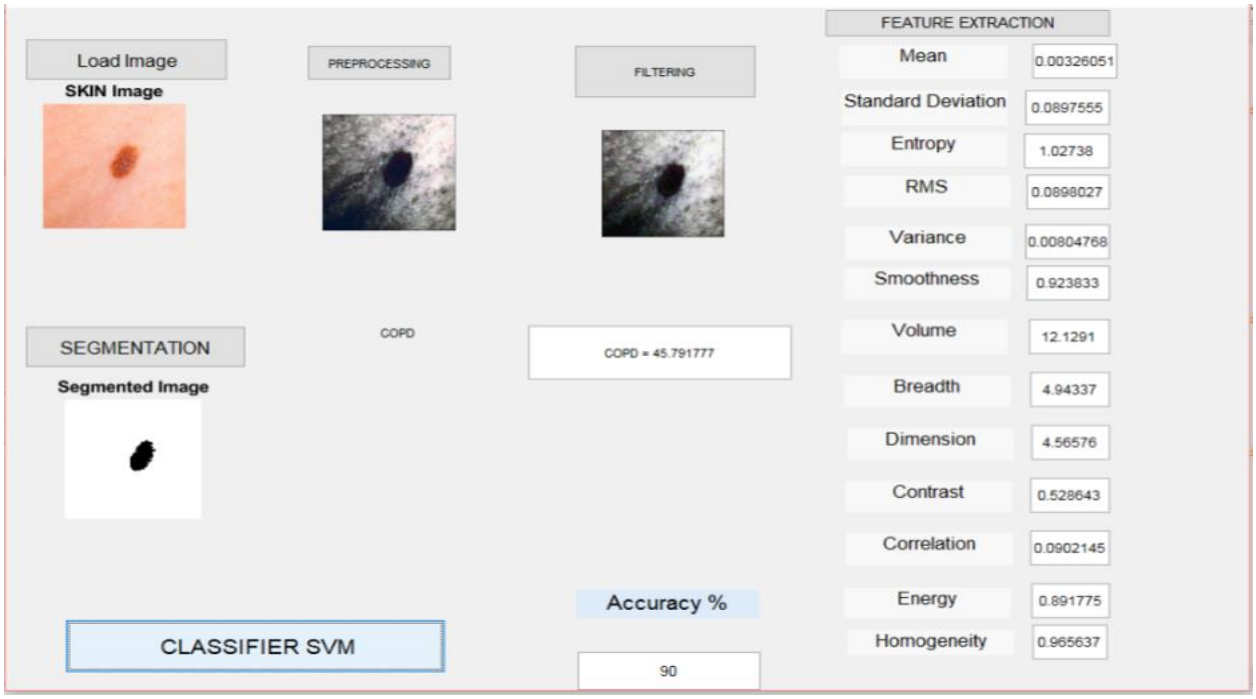
- The first step in MATLAB image processing is to understand that a digital image is composed of two or three dimensional matrix of pixels.
- Individual pixels contain a number or numbers representing what grayscale or color value is assigned to it. Color pictures generally contain three times as much data as grayscale pictures, depending on what color power to process.
- The method for conversion from color to grayscale will be demonstrated and all processing will be done on grayscale images.
- However, in order to understand how image processing works, we will begin by analyzing simple two dimensional 8-bit matrices.

PRE-PROCESSING OUTPUT



- In the pre-processing, Noise occurs while image acquisition. The noise may be due to illumination or shadows that make region of interest (ROI) appear as blurred image region.
- During this process, image enhancement such as contrast enhancement will be done.
- Affected region is taken for the analysis. This is to reduce errors in people having darker skin.
- If the image that we have is in color, but color is not important for the current application, then we can change the image to grayscale.
- This makes processing much simpler since then there are only a third of the pixel values present in the new image.
- Color may not be important in an image when you are trying to locate a specific object that has good contrast with its surroundings.

CLASSIFICATION RESULT



- For image segmentation K-means clustering algorithm is used.
- The goal of segmentation is to simplify and (or) change the representation of an image into something that is more meaningful and easier to analyze.
- Image segmentation is typically used to locate objects &boundaries (lines & curves) in images.
- Here, the SVM classifier is used to find the ROI.
- Finally, the accuracy and Mean value are calculated and produced as a result.

FEATURE EXTRACTION-MODIFIED RESULT

Fig.1 Input Image

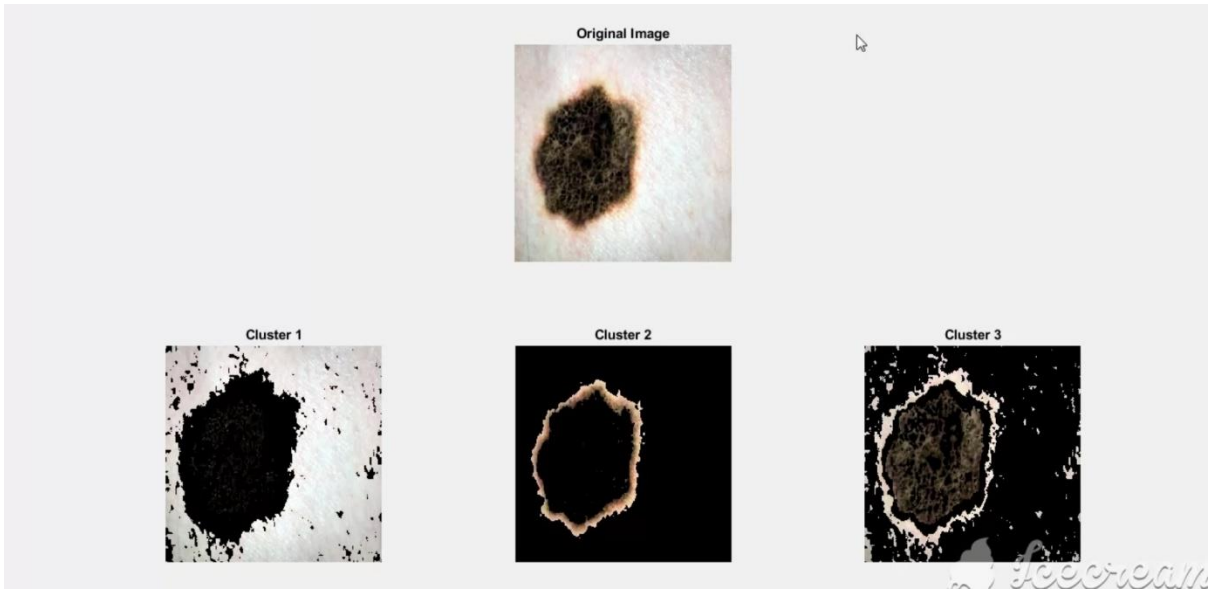


Fig.2 Segmentation process of Cancerous part

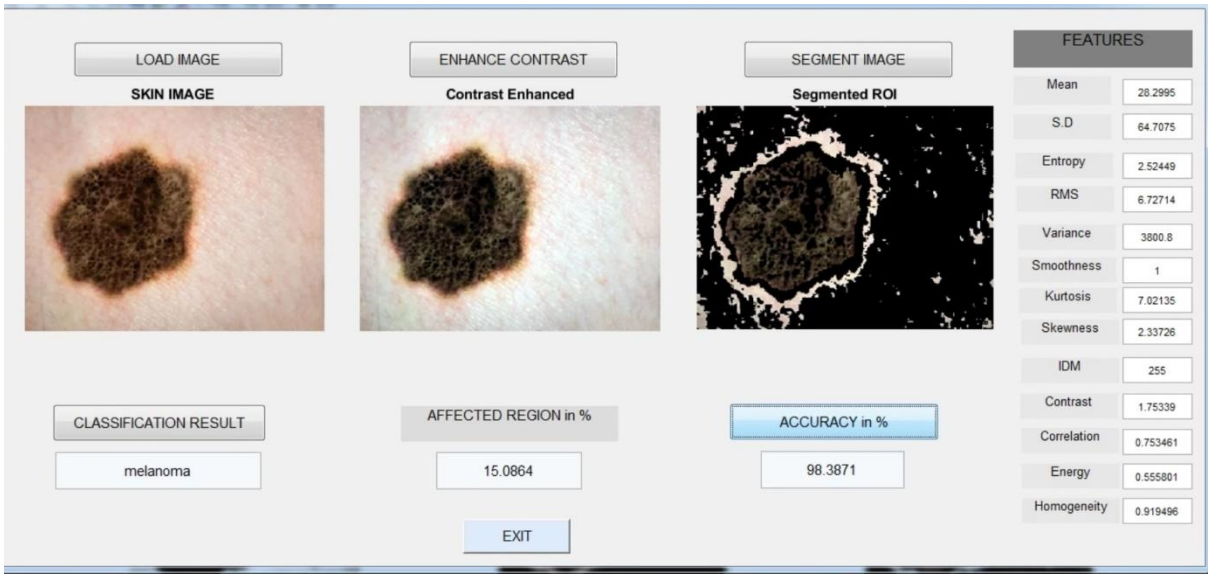


Fig.3 Classification Result

Fig.3 shows the classification result of the input image. Here for improving accuracy to 98.38%, the changes were made in the feature values. Therefore, the accuracy of the automated system greatly depends on type of segmentation algorithm employed, the feature to be extracted and the classification methodology incorporated.

APPENDIX III – JOURNAL FRONT PAGE

International Journal of Advanced Science and Technology
Vol. 29, No. 6, (2020), pp. 4115 - 4126

An Automated Diagnosis of Skin Cancer Disease Using Machine Learning Techniques

A.C.Kaladevi[1], P.Ramya[2]
Elamathi S[*], Atchaya R[*], Dhanalakkshmi M[*]
[1]Professor, [2]Assistant Professor,
[*]Bachelor of engineering
Department of CSE, Sona College of Technology
kaladeviac@sonatech.ac.in,
ramyaperumal@sonatech.ac.in

Abstract

The skin is the largest organ in the human body, which protects us from microbes and other pathogens. Etymologically, dermatology is the medical discipline in the analysis and prevention of skin abnormalities. The system is an automated diagnostic system unlike the conventional system involving human arbitration based on the ideology of dermatological diagnosis. The system works in two dependent steps: the first detects skin abnormalities and the second identifies skin diseases. The system works with visual inputs, i.e. high-resolution color images and patient history. The automated diagnostic system uses a modified genetic algorithm, a k -means grouping, and an SVMclassifier to perform preprocessing segmentation, and featureextraction on the images respectively. To detect askin cancer disease, the system uses a neural network to propagate artificial feedback, which is implemented using MATLAB. The system has a skin cancer detection accuracy of 98.99% and a canceridentification accuracy of 97.016% when testing diseased areas on skin images. In addition to this, various systems have been proposed to assist researchers in the automatic detection of melanoma. This investigation focuseson thealgorithms usedfor the automated detection of melanoma in dermoscopicimages through a complete analysis of the stages of the proposed methodologies. It also examines the concepts associated with skin cancer disease and describes possible future directions through open problems in this area ofresearch.

Keywords *Dermatology, Melanoma, Images, Accuracy.*

1. Introduction

Skin cancer will mainly affect adults and its incidence rate is higher in adult survivors. However, survivors are in danger of late effects of childhood cancer treatment, including an increased risk for subsequent malignancies, which occur more frequently during the radiation emitting fieldwork. Child skin cancer, mainly basal cell carcinoma