IMAGE PROCESSING

SKIN DISEASE RECOGNITION SYSTEM

Name: TANMAYI NANDAN

Reg No.: 17BCI0039

Abstract

This project is an endeavour to learn the concepts of image processing and to implement the basic techniques and lessons taught in the course into real world problems. The project basically tries to extract images from a huge dataset with constraints using MATLAB coding to find the desired result in minimal time. The project will be having a dataset from which the images will be extracted the patient is suffering from. The required image will be provided to the skin disease recognition system, there it will be trained i.e., the image will be converted into segments using segmentation and further it will be converted into vector which will then be compared to the vector image that is already present in the dataset. At last the image will be matched with the required image and the user will be able to find that the patient is suffering from which disease in very short amount of time. This technique can be used in real world situations which will help in providing better outcomes and save many lives.

Keywords

Image processing, skin diseases, segmentation, filtering, feature extraction

Background

There are hundreds of skin disorders that affect human beings. The most common skin infection or disease can have some symptoms that are similar, so it is important to understand the differences between them.

- Permanent skin diseases

The skin diseases or conditions that are chronic and mostly have no or very rare cure and their scars and marks or event he condition stays forever are called Permanent Skin Diseases.
Example: Vitiligo, Eczema, Psoriasis, Lupus, Rosacea etc.

- Temporary skin diseases

The skin diseases that are caused only for some short amount of time like for a week or a month and after their cure there will be no sign of them are called Temporary Skin Diseases.
Example: Acne, Hives, Warts, Cold sore, Candidiasis etc.

- Internal skin diseases

When Staphylococcus aureus bacteria make their way into the hair follicles and cause an infection, an internal skin disease is formed. This is a red, irritated lump underneath the skin. It can also be a bacterial infection in the deeper layers of the skin, which develops quickly and can spread rapidly throughout the body.
Example: Carbuncle, Cellulitis etc.

- Skin diseases for different age groups

It is quite clear that skin diseases can be of hundreds of types but not all diseases are caused to everyone in every age group. Several skin diseases are age specific and are only caused to individuals of that specific age group only and some diseases are spread over few age groups and very few are spread over all age groups.

Literature Survey

Yadav, N. presents a survey of various skin disease diagnosis systems using image processing techniques in recent times in [1]. A study of a number of skin disease diagnosis systems are done in this paper, with different methodologies and their performances. The next study [2], presented by Okuboyejo, D. A. unlike the above paper, is a specific one, that has made a system which will collate specific image results and analyse them by converting them into prototypes based on their texture and morphology. Amarathunga, A. A. L. C. uses a different technique in [3], namely data mining to detect the disease and also additionally suggests advice for patients. The system made by the research scholars has a high accuracy of detecting diseases with 85% for Eczema, 95% for Impetigo and 85% for Melanoma. Vezhnevets, V. has a different approach that he uses in [4] than the above presented papers have been followed as they present a comparative study of various results for skin detection. They have presented the paper with significant differences between each approach and their advantages and disadvantages. The authors of [5], Bourouis, A., have done a great job and is similar to the 3rd paper except more sophisticated. They present a smart phone based integrated system which can work in any place. This application is specialized towards cancer detection and uses ANN algorithm.
Maglogiannis, I. continues the trend of the [5] in [6] as well as it treads along a specific path of pigmented skin lesions, malignant melanoma vs dysplastic nevus. Features are extracted from images and a module is made based on Support Vector Machines (SVM). Feature extraction, which is made in the [6] is continued in paper [7] as well by Venkatesan, S. The extracted features are optimised using ANT Colony Optimisation, after which Genetic Algorithm is used for recognition of faces.

Paper [8], by Ahmed, I. N., is similar to the [1] paper of the survey as it includes work in the areas of segmentation and feature extraction. The skin lesions, also talked about in [6], are classified. Weiner's filtering is used to enhance parts of the images which are to be preserved. [9] by Barbedo, J. G. A., also follows [4] paper and has presented a survey of various works. But it differs in that this identifies works on plant diseases. The comparative studies include detecting, quantifying and classifying plant diseases. [10] by Mengistu, A. D., can be considered as an extended version of the [5] paper which deals with cancer. The authors have conducted an experiment which has showed that human skin cancer can be classified using SOM quite accurately rather than Naïve Bayes, ANN and KNN classifier. Mittra, A. K., & Parekh, D. R. have developed an effective solution for the recognition of diseases related to skin which they have represented in [11]. This paper has used a relatively different approach for disease detection than the rest of the papers in the survey as it has the context of health informatics.

Abuzaghleh, O. has presented a paper that strays a little from the main topic as it provides an analysis module of images which contain feature extraction, lesion segmentation and image acquisition in [12]. The experiment conducted used PH2 Dermoscopy image database for development and testing the prescribed modules. The next paper [13] by Qin, J., Burks specialises in detection of Citrus canker disease using the PCA algorithm, which is the algorithm I intend to use in the project. According to this paper, the accuracy for the canker detection is found to be 92.7%, which can definitely be considered a success. The paper [14] by Raja, C. V. J., is a generalised concept of all the papers presented above as it basically analyses various skin diseases using texture analysis and many other standard methods such as grey scale analysis, histogram equalisation, some other statistical methods such as mean, entropy, etc. Grey level Co-occurrence Matrix (GLCM) is performed by Kaur, D. on the image in [15], which is a very different one from the other papers mentioned above as it deals with image analysis in frequency domain. Various skin problems having correlation with skin texture profile has been discussed by the authors.

Proposed Work

The skin disease recognition system basically recognizes the disease from entered dataset using disease recognition system. The author created a database of 100 images for skin diseases from which the user on entering the image of skin disease he/she is suffering from will be able easily retrieve the information that is the details of that particular disease (desired output). This will help the user to efficiently receive the disease information without wasting much time.

Problem Statement

An application that can detect various skin diseases based on images which will undergo processing and use PCA algorithm to compare input with a dataset to identify the disease.

Research Framework (Methodology)

Fig 1. Database

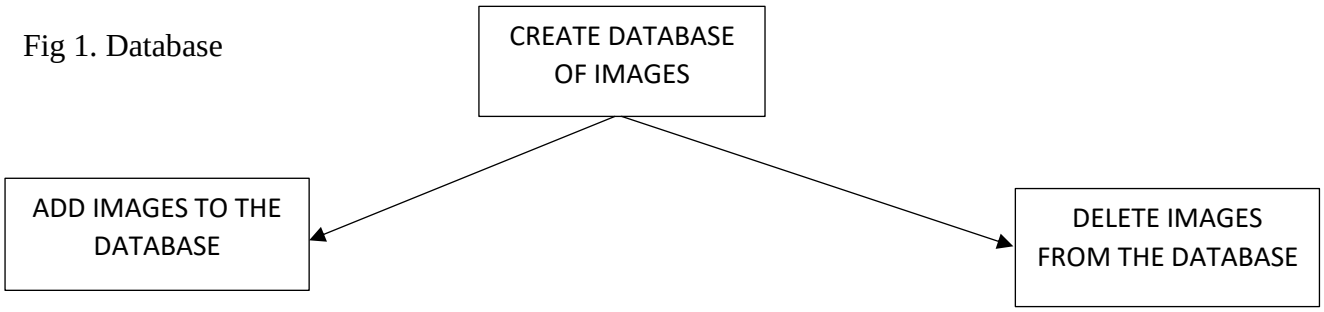
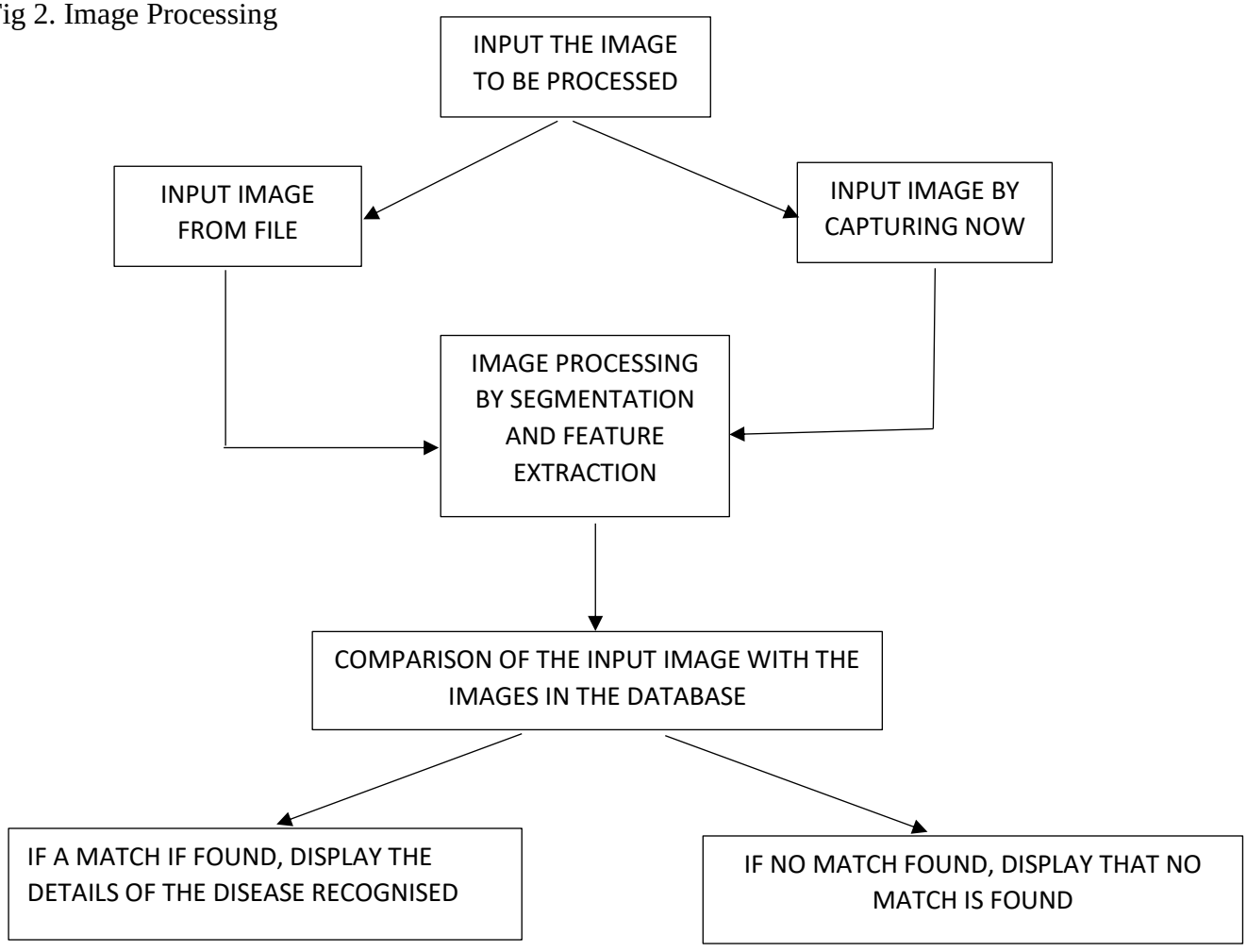


Fig 2. Image Processing



Algorithm: Principle Component Analysis (PCA)

The Principle Component Analysis (PCA) is the transform of a set of n number of input vectors, or variables, which will have the same length K formed in the n- dimensional vector into a vector **y.**

This point of view enables to form above formula but it is necessary to keep in the mind that each row of the vector **x** consists of *K* values belonging to one input.

Matrix **A** in is determined by the covariance matrix **Cx**. Rows in the **A** matrix are formed from the eigenvectors **e** of **Cx** ordered according to matching eigenvalues in descending order.

As the vector **x** of input variables is *n*-dimensional it is obvious that the size of **Cx** is *n* x *n*. The elements **Cx** (*i, i*) lying in its main diagonal are the variances of **x** and the other values **Cx** (*i, j*) determine the covariance between input variables **xi**, **xj**.

The kernel of PCA defined by above equations has some other great properties resulting from the matrix theory which can be used in the signal and image processing to achieve various goals as stated below.

These steps are the main procedure of the PCA algorithm and are generalized for most applications:

1. Consider the whole dataset comprising of m-dimensional samples ignoring the class labels.

2. Compute the m-dimensional mean vector (i.e., the means for every dimension of the whole dataset)

3. The scatter matrix or the covariance matrix of the data is computed.

4. Compute eigenvectors (e_1,e_2,...,e_m) and corresponding eigenvalues (λ_1,λ_2,…,λ_d)

5. Sort the eigenvectors by decreasing eigenvalues and choose kk eigenvectors with the largest eigenvalues to form a m×k dimensional matrix W(where every column represents an eigenvector)

6. Use this m×k eigenvector matrix to transform the samples onto the new subspace.

The mathematical equation that is finally obtained in which the m*l - dimensional vector representing each sample is represented as x, and the transformed k*l – dimensional in the new subspace is represented as y is:

y = (W ^ T) * x

The skin disease recognition system basically recognizes the disease from entered dataset using disease recognition system. A database of 20 images for skin diseases was created from which the user on entering the image of skin disease he/she is suffering from, will be able easily retrieve the information that is the details of that particular disease (desired output). This will help the user to efficiently receive the disease information without wasting much time.
The system basically consists of a large image database of skin diseases wherein a database can be created and deleted. The system can be trained so as to make the system understand the procedure of retrieving the information according to the input query. The disease recognition system finally finds the match which is similar to the image being added by the user and displays the user all the details related to that particular skin disease.

System Implementation

- **Create Database:**

```
function []=CreateDatabase
cdTrainDatabase;
while (1==1)
choice=menu('Create Database',...
' Add  an  Image',...
' Add  a  Folder',...
                  ' Exit');
if (choice ==1)
addimage;
end
if (choice == 2)
addfolder;
end
if (choice == 3)
cd ..;
clc;
close all;
return;
end
end
End
```

- **Delete Database:**

```
function [ ] = DeleteDatabase(  )
disp('Please dont delete in between');
```

```
cdTrainDatabase
while (1==1)
choice=menu('Delete DataBase',...
   'Delete an Image',...
                'Delete a Folder',...
                'Exit');
if (choice ==1)
ChooseFile=imgetfile;
delete(ChooseFile);
end
if (choice == 2)
delfolder=uigetdir('E:\ufd\TrainDatabase','Delete Folder');
rmdir(delfolder,'s');
end
if (choice == 3)
cd ..
clc;
close all;
return;
end
End
```

- **Eigen Face Core:**

```
function [m, A, Eigenfaces] = EigenfaceCore(T)

disp('Creating Eigen Images');
m = mean(T,2);
Train_Number = size(T,2);

A = [];
for i = 1 : Train_Number
temp = double(T(:,i)) - m;
    A = [A temp];
end

L = A'*A;
[V D] = eig(L);

L_eig_vec = [];
for i = 1 : size(V,2)
if( D(i,i)>1 )
L_eig_vec = [L_eig_vecV(:,i)];
end
end

Eigenfaces = A * L_eig_vec;
save train
End
```

- **Disease Recognition Code:**

```
function [ OutputName ] = FaceRec(m, A, Eigen_imgs)
cdTestImage;
while (1==1)
choice=menu('Skin Disease Recognition',...
      'Input Image From File',...
                'Capture Now',...
'Recognition',...
                'Exit');
if (choice ==1)
try cd TestImage;close all; end;
ChooseFile = imgetfile;
capcha = imread(ChooseFile);
         %capcha = imcrop(capcha,[180,20,280,380]);
imshow(capcha);
saveimage(capcha);
end
if (choice == 2)
try cd TestImage;close all; end;
capturenow;

end
if (choice == 3)
OutputName=Recognition(m, A, Eigen_imgs);
       n=((OutputName+1)/2);

im=imread('InputImage.jpg');
cd ..;
img=strcat('TrainDatabase\s',int2str(n),'\1.png');
SelectedImage=imread(img);
subplot(121);
imshow(im)
title('Test Image');
subplot(122),imshow(SelectedImage);
title('Equivalent Image');
disp('Disease No');
disp(int2str(n));

end

if (choice == 4)
clc;
close all;
return;
end
end
```

- **Main:**

```
close all
```

```
clear all
clc
while (1==1)
choice=menu('Disease Recognition System',...
                'Create Database of Skin Disease Images',...
                'Delete DataBase',...
                'Train System',...
                'Disease Recognition',...
                'Exit');
if (choice ==1)
CreateDatabase;
end

if (choice == 2)
DeleteDatabase;
end

if (choice ==3)
        [m, A, Eigen_imgs]=Trainit();
end
if (choice == 4)
if exist('train.mat');
load train;
end
FaceRec(m, A, Eigen_imgs);
end

if (choice == 5)
clear all;
clc;
close all;
return;
end
End
```

- **Train Database:**

```
function T = TrainDatabase(TrainDatabasePath)
 no_folder=size(dir([TrainDatabasePath,'\*']),1)-
size(dir([TrainDatabasePath,'\*m']),1)-2;
T = [];
disp('Loading Images');
for i = 1 : no_folder
stk = int2str(i);
disp(stk);
stk = strcat('\s',stk,'\*png');
folder_content = dir ([TrainDatabasePath,stk]);
nface = size (folder_content,1);
disp(nface);
for j = 1 :  nface
str = int2str(j);
str = strcat('\',str,'.png');
```

```
str = strcat('\s',int2str(i),str);
str = strcat(TrainDatabasePath,str);
img = imread(str);
img = rgb2gray(img);

    [irowicol] = size(img);

temp = reshape(img',irow*icol,1);
    T = [T temp];
end

end

End
```

- **Train It:**

```
function [m, A, Eigen_imgs]=Trainit()

clear all
clc
close all

TrainDatabasePath =
uigetdir('C:\Users\TanmayiNandan\Desktop\facerec_PCA\TrainDatabase',
'Select training database path' );
T = TrainDatabase(TrainDatabasePath);
[m, A, Eigen_imgs] = EigenfaceCore(T);
End
```

Results and Discussion

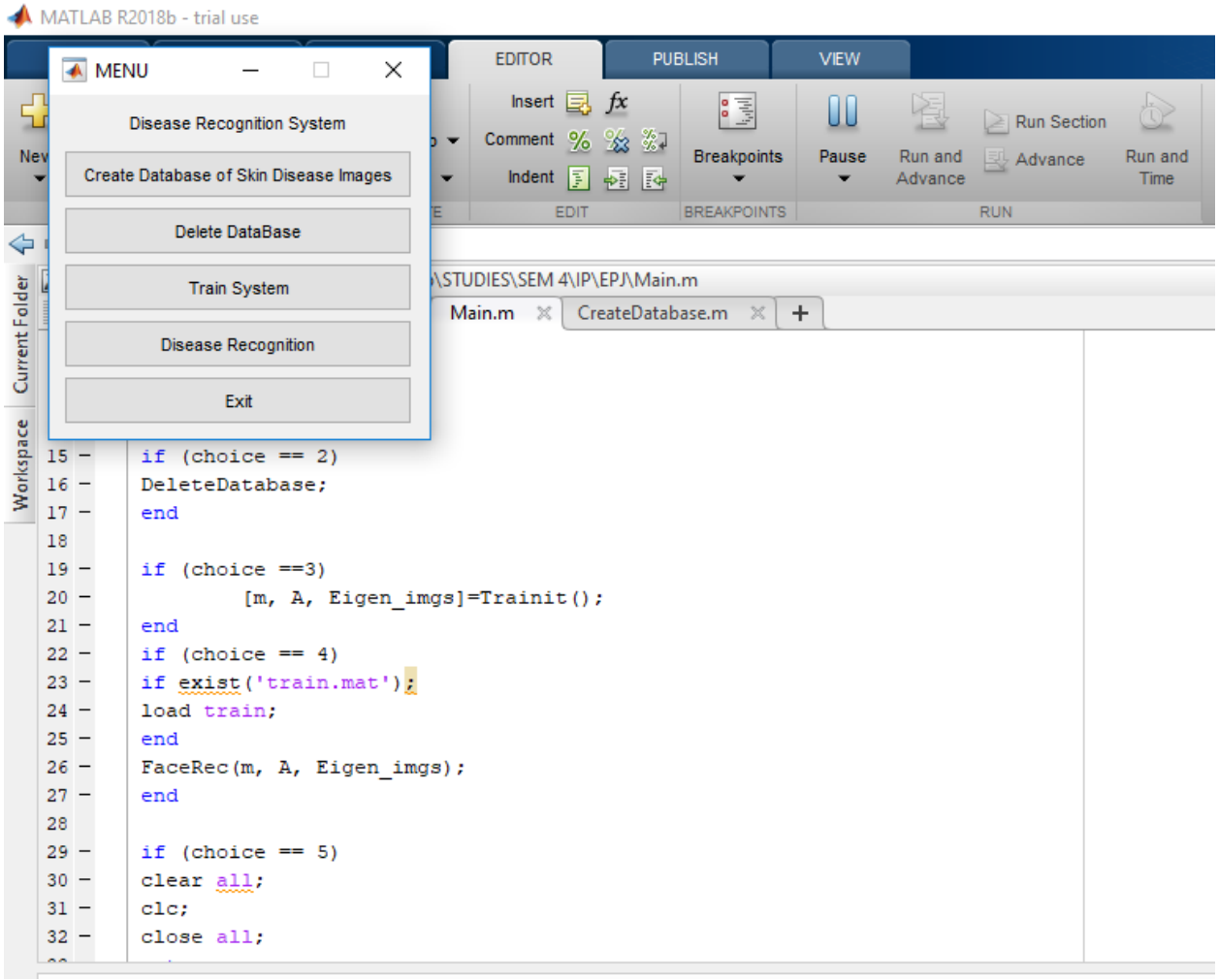


Fig 1.1. Main menu

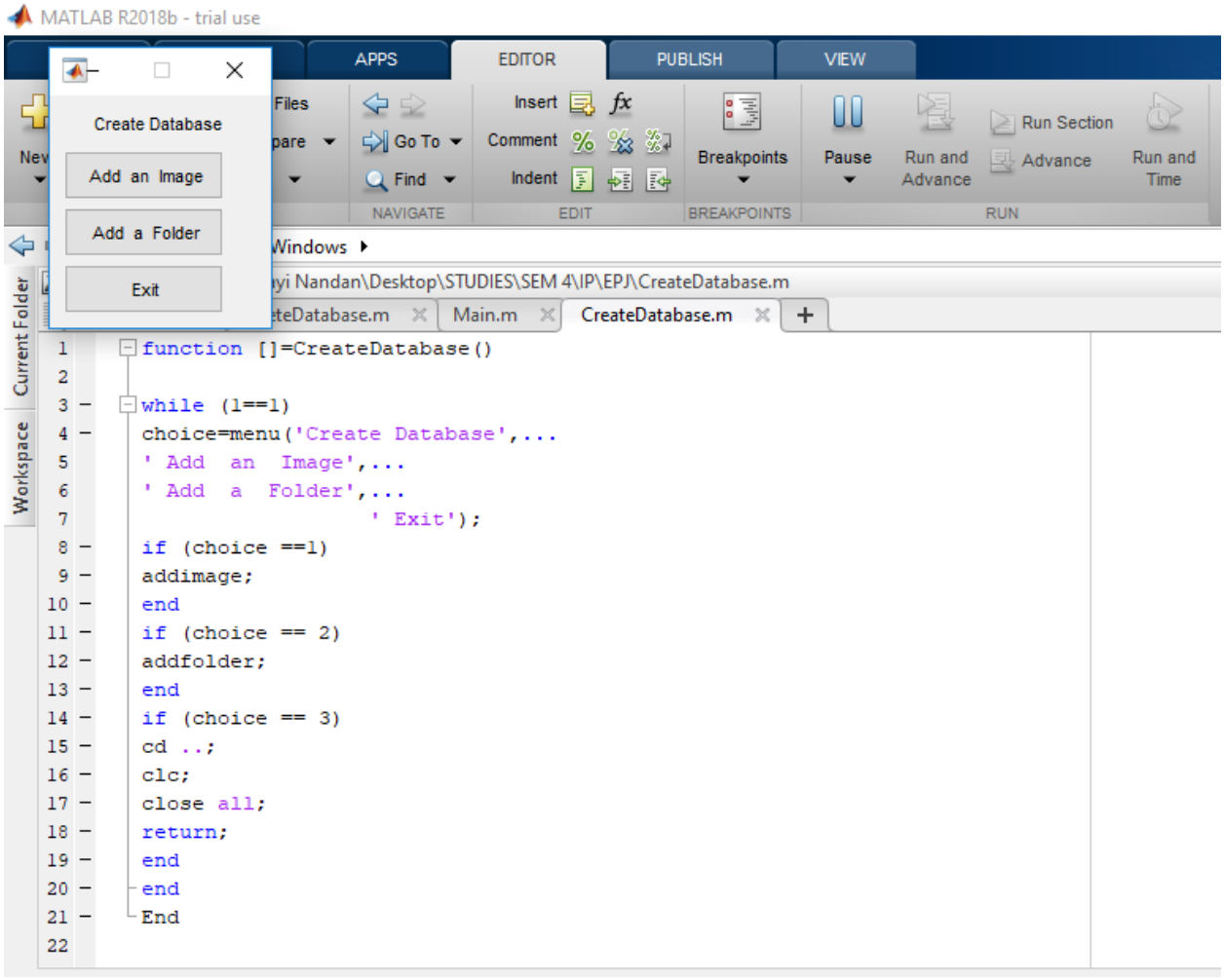


Fig 1.2. Creation of Dataset

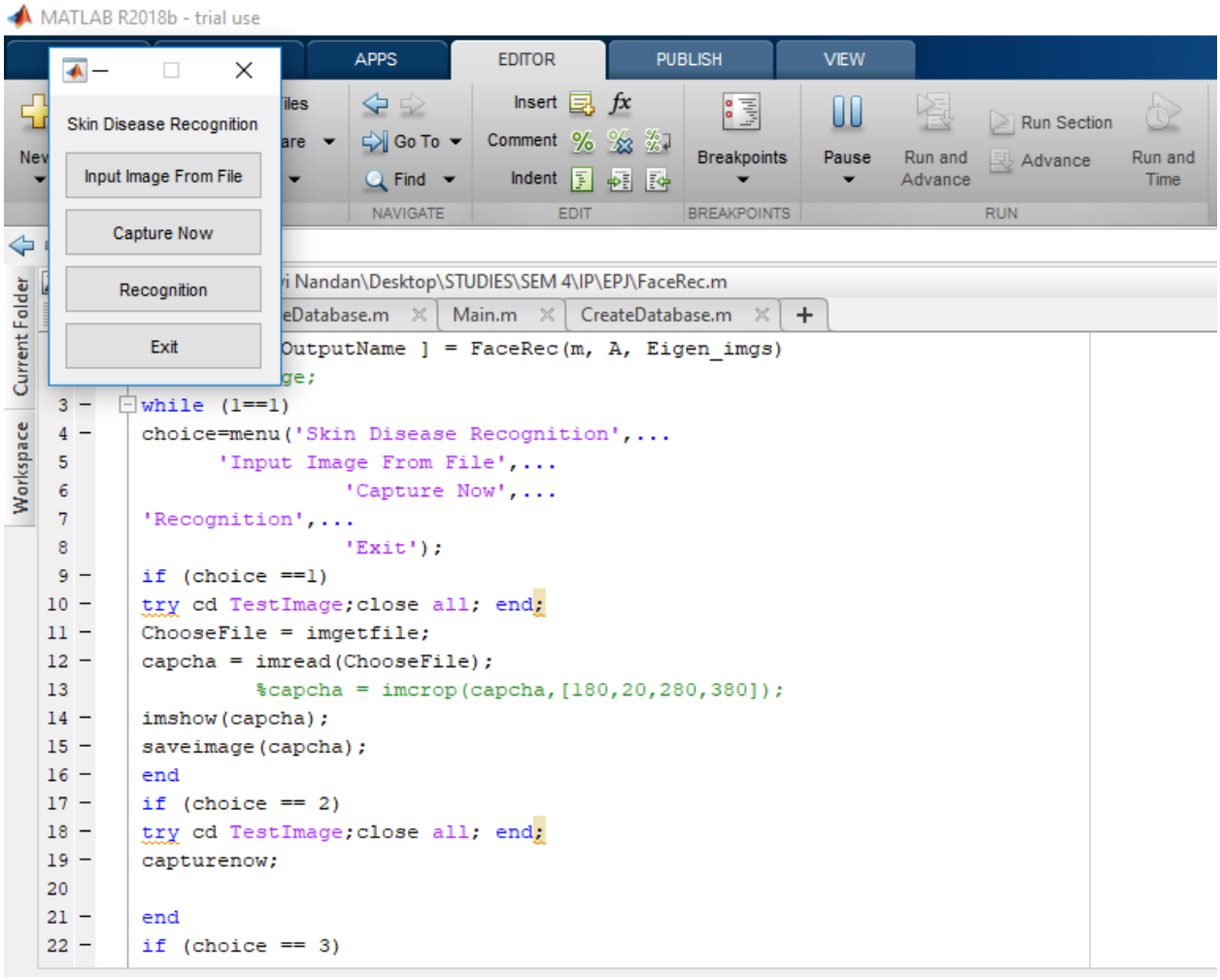


Fig 1.3. Input image

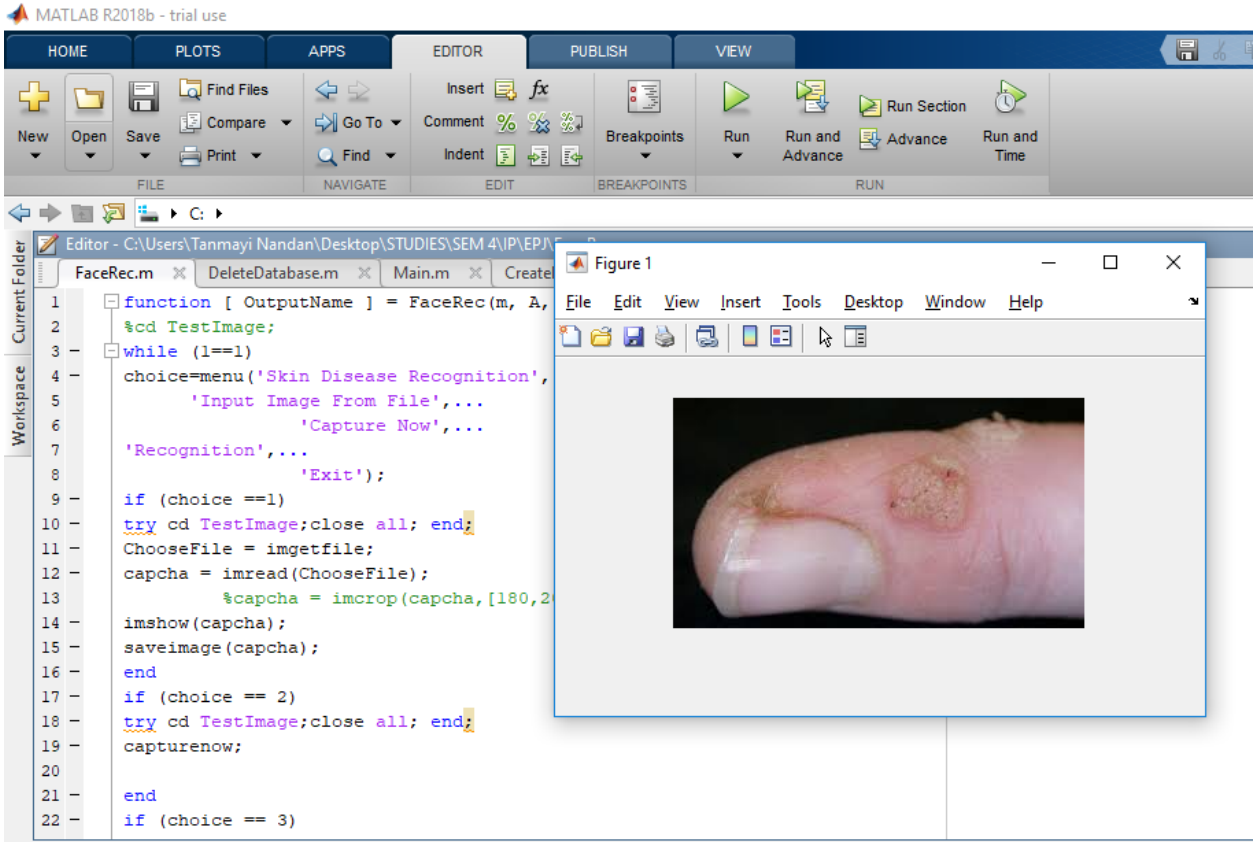


Fig 1.4. Image with skin disease

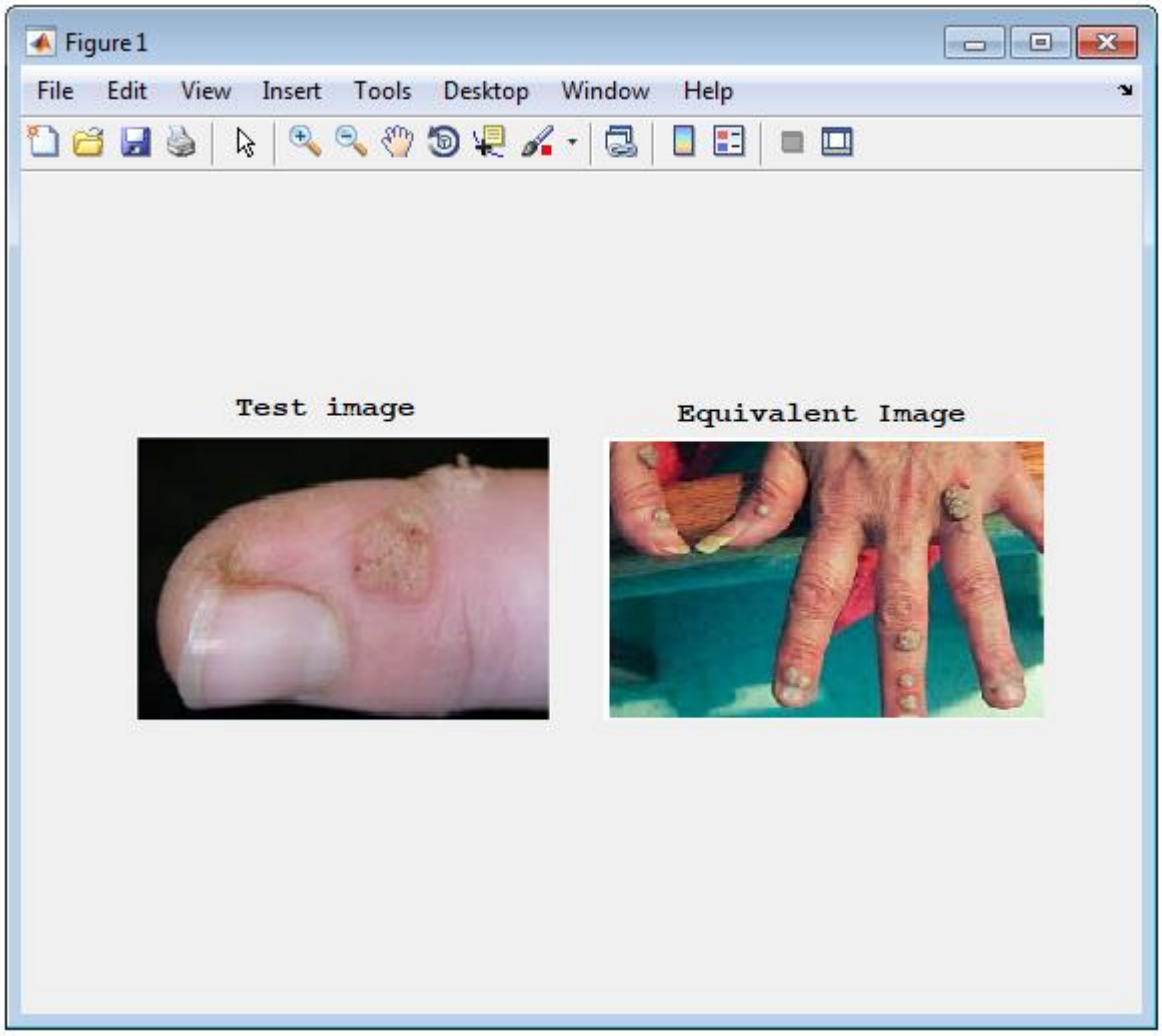


Fig 1.5. The output obtained

A test image when given as input, is successfully segmented and the required feature is extracted, using Principle Component Analysis (PCA). This is them compared with the dataset to find the images of the given test disease. The two images are then displayed in the same window for the user to understand and analyse.

Conclusion and Future Work

A Computer based skin disease detection system is proposed. The diagnosing methodology uses Digital Image Processing Techniques for the classification of infected skin. The unique features of the enhance images were extracted using PCA. Based on the features, the images were classified as infected skin and normal skin. This methodology has got good accuracy also. By varying the Image processing techniques and Classifiers, the precision can be improved for this system. Though a lot of hardships could be encountered, medical science is in need of techniques such as these. For any infected part of a person's skin or even infections which can't

be seen with naked eyes, a clear image of the affected area which are collected by the authors will be helpful for the medical sciences.

The technique Pattern recognition mainly originates from the need for automated machine recognition of signals, images and objects or any decision-based approach, on the basis of the set of features. The goal of Pattern recognition is to predict the correct level corresponding to given feature set based on a better knowledge obtained through training. The pattern recognition can well understood by considering an example: the human being can easily identify the gender based on the face while machine can't, so the aim is to train the machine by considering various features as: facial expression, facial bone structure, hair length and others. After training, the machine can easily identify the required class the new test object belongs. From this above example, it is quite obvious that the heart of pattern recognition system is feature extraction and classification.

As mentioned in the introduction, the skin disease detection problem is still a much-investigated problem; many authors have proposed techniques to solve it by fixing one or more parameters of the problem, but a solution of P considering all of them has never been given. There is a camera-specific colour-based method able to recognize skin disease in different light conditions and proposed a database of camera behaviours to complete it. The use of a normalized colour space, in this case the rg normalized colour space, is interesting because it allows isolating skin locus with simple quadratic functions. Also, for normalized colour space, the rg normalized colour space again (in the following simply rg), is the most effective to extract with success a skin locus. This is because it is as little as possible dependent on the illuminant. In addition, it is. affirmed that an optimum filter for skin disease detection will have the same performance even working in different colour spaces.

Some authors have given suggestions regarding the solution of the P problem by using the combination of the various techniques so as to improve the results of the already proposed solutions and help cover its defects. Some other solutions that can be obtained for face detection includes combining a colour-based filter method with a shape identification which improves the effectiveness of the solution. In a gain and also in a static prefilter on RGB space is used too: with this last kind of filters it is easier and more natural to remove zones that surely are non-skin areas (pixels too inclined to black, to green or to blue etc.). At last the author mentions Lee et al.who proposed an elliptical boundary for skin locus using a Gaussian model and six chromatic spaces who proposed a Bayesian network approach instead.

References

[1] Yadav, N., Yadav, N., & Narang, V. K. (2016). Skin Diseases Detection Models using Image Processing: A Survey. International Journal of Computer Applications (0975–8887) Vol, 137.

[2] Okuboyejo, D. A., Olugbara, O. O., & Odunaike, S. A. (2013, October). Automating skin disease diagnosis using image classification. In Proceedings of the World Congress on Engineering and Computer Science (Vol. 2, pp. 850-854).

[3] Amarathunga, A. A. L. C., Ellawala, E. P. W. C., Abeysekara, G. N., & Amalraj, C. R. J. (2015). Expert system for diagnosis of skin diseases. International Journal of Scientific & Technology Research, 4(01), 174-178.

[4] Vezhnevets, V., Sazonov, V., & Andreeva, A. (2003, September). A survey on pixel-based skin colour detection techniques. In Proc. Graphicon (Vol. 3, pp. 85-92).,

[5] Bourouis, A., Zerdazi, A., Feham, M., & Bouchachia, A. (2013). M-health: skin disease analysis system using Smartphone's camera. Procedia Computer Science, 19, 1116-1120.

[6] Maglogiannis, I., Zafiropoulos, E., & Kyranoudis, C. (2006, May). Intelligent segmentation and classification of pigmented skin lesions in dermatological images. In Hellenic Conference on Artificial Intelligence (pp. 214-223). Springer, Berlin, Heidelberg.

[7] Venkatesan, S., & Madane, S. S. R. (2010). Experimental research on identification of face in a multifaceted condition with enhanced genetic and ANT colony optimization algorithm. International Journal of Innovation, Management and Technology, 1(5), 2010-0248.

[8] Ahmed, I. N., & Chaya, P. (2014). Segmentation and classification of skin cancer images. International Journal of Advanced Research in Computer Science and Software Engineering, 4(5).

[9] Barbedo, J. G. A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. SpringerPlus, 2(1), 660.

[10] Mengistu, A. D., & Alemayehu, D. M. (2015). Computer Vision for Skin Cancer Diagnosis and Recognition using RBF and SOM. International Journal of Image Processing (IJIP), 9(6), 311-319.

[11] Mittra, A. K., & Parekh, D. R. (2011). Automated Detection Of Skin Diseases Using Texture. International Journal of Engineering Science and Technology (IJEST) ISSN.

[12] Abuzaghleh, O., Faezipour, M., & Barkana, B. D. (2015, May). A comparison of feature sets for an automated skin lesion analysis system for melanoma early detection and prevention. In Systems, Applications and Technology Conference (LISAT), 2015 IEEE Long Island (pp. 1-6). IEEE.

[13] Qin, J., Burks, T. F., Kim, M. S., Chao, K., & Ritenour, M. A. (2008). Citrus canker detection using hyperspectral reflectance imaging and PCA-based image classification method. Sensing and Instrumentation for Food Quality and Safety, 2(3), 168-177.

[14] Raja, C. V. J., & Jeyaprakash, M. (2014). Skin Disease Diagnosis Using Texture Analysis. International Journal of Advanced Research in Computer Science and Software Engineering, IJARCSSE, 4(1).

[15] Kaur, D., & Sandhu, P. (2013). Human skin texture analysis using image processing techniques. Int. J. Sci. Res.(IJSR), 17-20.