

Image Processing for Recognition of Skin Diseases

Suneel Kumar
Department of Computer
Science and Engineering,
B.T.K.I.T. Dwarahat, India

Ajit Singh, PhD
Department of Computer
Science and Engineering,
B.T.K.I.T. Dwarahat, India

ABSTRACT

It is an extremely bulky process to predict a disease based on the visual diagnosis of cell type with precision or accuracy, especially when multiple features are associated. If we get the information about the dead skin which is not visible by naked eyes well in time then we can easily prevent the further spreading of disease on the other part of body. One of the major problems coming in the medical field is that doctors are not able to detect that infected part which is not visible by naked eyes and therefore they only operate the visible infected part of the skin and this may cause a major problem like cancer or any dangerous disease in the future. Skin cancer classification system is developed and the relationship of the skin cancer image across different type of neural network is established. The collected medical images are feed into the system, and using different image processing schemes image properties are enhanced. Useful information can be extracted from these medical images and pass to the classification system for training and testing using MATLAB image processing toolbox for detection of dead skin.

Keywords

SURF, HSV-histogram, KNN, image enhancement, feature extraction

1. INTRODUCTION

Skin diseases are one of the most common diseases in humans and its incidence is increasing dramatically. Skin cancer is very common diseases. Therefore early diagnosing is a crucial issue for patient. However, only experienced doctor is able to classify the skin cancer from other skin diseases. Thus the computer based skin cancer detection is necessary to provide recommendation for non-specialized user. It is well-known that early finding and treatment of skin cancer can reduce the mortality and morbidity of patients. Digital Dermoscopy is widely considered as one of the most cost effective means to identify and classify skin-cancer. An automatic medical images analysis system has usually three stages: (1) Proper Enhancement, (2) feature extraction and selection (3) Classification. The proper enhancement is the most important, since it affects the precision of the subsequent steps. Supervised enhancement is somewhat easy to implement by varying its parameters for variety of lesion shapes, sizes, and colors along with diverse skin types and textures. During the past few years, various contributions have been made in literature regarding the application of pattern recognition techniques for dead skin diagnosis in cell level.

T.H. Lau and A. Jumaily [1] developed an automatically skin cancer classification system and the relationship of skin cancer image across different type of neural network were studied with different type of processing. The collected images were feed into the system, across different image processing procedure to enhance the image properties. Then

the normal skin was removed from the skin affected area and the cancer cell left in the image. Useful information could be extracted from this image and pass to the classification system for training and testing. Zhou et al. [2] developed a Gradient vector flow (GVF) based algorithm to segment a variety of 2-D imagery. However, due to the compromise of internal and external energy forces result. A mean shift based GVF segmentation algorithm that can successfully locate the correct border. MSGVF is developed so that when the contour reaches equilibrium, the various forces resulting from the different energy terms are balanced. In additional, smoothness constraints of image pixels were kept so that over or under segmentation can reduced. Experimental result on publicly accessible dataset of dermoscopy and optic disc images demonstrate that the proposed method effectively detects the border of the objects of interest. Martins et al. [3] proposed the technique for detection of mass on digitized mammograms. They used K-means clustering algorithm for image segmentation and gray level co-occurrence matrix to describe and analyze the texture of segmented structures in the image. The classification of these structures was achieved through Support Vector Machines, which separate them into two groups; using shape and texture descriptors: masses and non-masses A. Bhardwaj and J.S. Bahtia [4] implemented image segmentation method in MATLAB followed by STOLZ and TDS algorithm. An image segmentation method have used to classifying the current state of melanoma lesions. It further includes various steps for analysis; preprocessing which have used for image enhancement, removal of noise, hair. Then the boundary of the skin lesion is detected with the help of initial segmentation. The lesions have investigated to find the parameters using feature extraction property. The fetched parameter values have used to calculate STOLZ algorithm and by feeding the output of the STOLZ algorithm in TDS, Detection of the current stage of melanoma can be identified.

2. SYSTEM FRAMEWORK

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. The pattern recognition system for this classification task is given by the flowchart; as shown in Fig.1. Here, the proposed Pattern recognition system is

developed with the help of 55 numbers of cultured cell images.

3. IMAGE ENHANCEMENT

Image Enhancement is the process of adjusting digital image so, that the result are more suitable for display or further image analysis. For example, you can remove noise, sharpen or brighten of digital images, making easier to identify key feature.

Methods:

- Filtering with morphological operation.
- Histogram equalization.
- Noise removal.
- Linear contrast adjustment
- Median filtering. Etc.

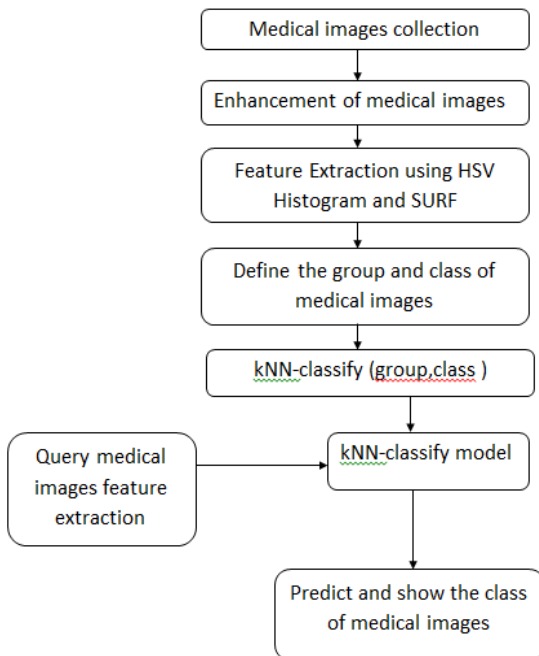


Fig.1. Proposed System framework



Fig.2. Skin images dataset

3.1 Histogram equalization

Histogram equalization is used to enhance contrast. It is not necessary that contrast will always be increase in this. There may be some cases were histogram equalization can be worse. In that case the contrast is decreased.

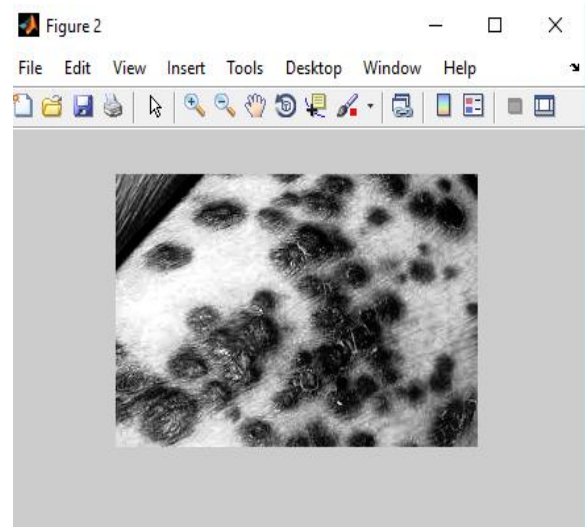


Fig.2. Histogram equalization

4. FEATURE EXTRACTION

In the field of Image processing and pattern recognition, the feature extraction is a technique used for collecting various features from image and dimensionality reduction. Here, I use two different feature extraction methods (i.e. HSV-histogram feature extraction and SURF) to extract the various features from enhanced images.

4.1 HSV-histogram

In HSV color space for histogram generation each pixel contributes either its Hue or its Intensity. The number of constituents in the feature vector produced based on Hue is given by

$$N_h = Round(2pMULT_FCTR) + 1$$

Here MULT_FCTR defines the quantization level for the Hues. We typically choose a value of 8.

The number of constituents signifying gray values is

$$N_g = Round\left(\frac{I_{max}}{DIV_FCTR}\right) + 1$$

Here I_{max} is the maximum value of the Intensity commonly 255, and DIV_FCTR defines the number of quantized gray levels. We normally choose DIV_FCTR = 16.

4.2 SURF feature

The SURF algorithm has three main portions (a) interest point detection (b) local neighborhood description (c) matching.

SURF uses blob detector based on Hessian matrix to find point of interest. The determinant of the Hessian matrix is used as a degree of local change nearby the point and points are chosen where this determinant is highest.

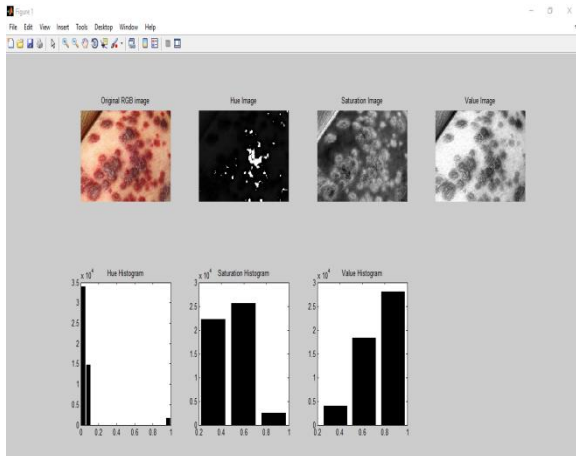


Fig.3. HSV-histogram

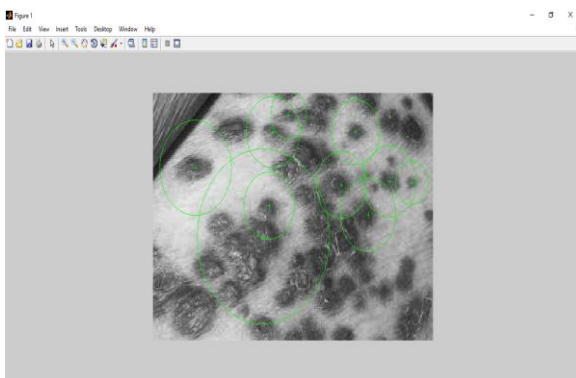


Fig.4. SURF feature of gray scale image

5. MACHINE LEARNING

Machine learning is a multidisciplinary field of study that mainly concerned with the design of algorithms which allow computers to learn. In machine learning the classification is a task of identifying, to which class the new observation belongs. An algorithm which implements the classification is mainly called as the classifier.

5.1 K-Nearest Neighborhood (KNN) Classifier

K-nearest neighbor algorithm is a technique for classifying data based on the closest training examples in the feature space. Before using the KNN the protocol should be followed, i.e. given as: First the dataset is divided into a testing and training set. For each row in the testing set, the 'K' nearest training set objects is found, and the classification of test data is determined by majority vote with ties are broken at random. If there are ties for the K_{th} nearest vector then all the instances are included in the vote [5].

The way the KNN classifier works is, first by calculating the distances between the testing data vector and all of the training vectors using a particular distance calculation methodology which is given as follows:

Considering the case of two input variable; the Euclidean distance between two input vectors p and q is computed as the magnitude of difference in vectors i.e. $p - q$, Where both the data are having 'm' dimensions i.e. $p = (p_1, p_2, \dots, p_m)$ and $q = (q_1, q_2, \dots, q_m)$.

The Euclidean distance between 'p' and 'q' is found to be

$$D(p, q) = |p - q|$$

$$= \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_m - q_m)^2}$$

The KNN classifier takes the test instance 'x' and finds the K-nearest neighbors in the training data and assigns 'x' into the class occurring most among the K neighbors.

6. RESULTS AND DISCUSSION

Initially, the image enhancement method is used to enhance the cells from image backgrounds. The results obtained during enhancement for the sample cancer image is given in figure 5.

After enhancement, the required textures features (i.e. HSV histogram based features, SURF features) are extracted from each enhance images. From these feature extraction techniques, a total of 52 features cells in which 32 features of HSV-histogram and 20 features of SURF were obtained. Similarly for 55 numbers of images the same HSV histogram based features and SURF features extraction procedure was followed to make a feature dataset.

The dataset obtained mainly consist of the number of instances which are in the rows and the feature vectors are in the column. The size of the dataset is found to be 55×52 . Further KNN-classifier was applied to make a model.

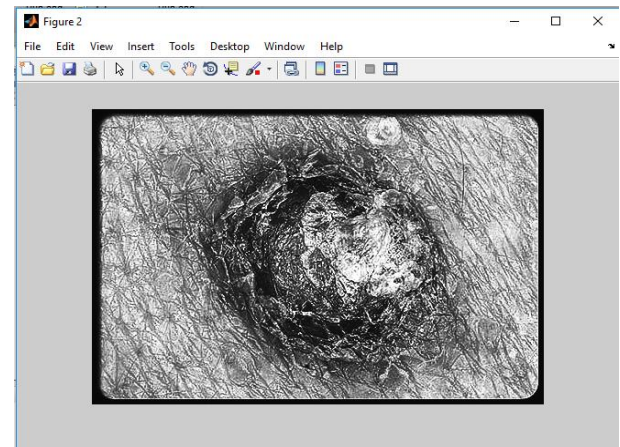


Fig.5. Enhancement of cell image

After using of KNN-classifier the following result was obtained

```
ClassificationKNN
PredictorNames: {1x52 cell}
ResponseName: 'Y'
ClassNames: [1 2 3 4 5]
ScoreTransform: 'none'
NumObservations: 55
Distance: 'euclidean'
NumNeighbors: 1
```

This model uses to predict the class of the images by using command predict (model name, data for check). Here we use the multiple classes 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).

ans =
3
3
5
3
4
3
4
2
1
5
5
5

7. CONCLUSION

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 HSV-histogram and SURF. 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. Despite having some difficulty, these techniques are very helpful in medical science. The data we have

collected will be helpful in medical field to see the clear image of the infected part in the skin as well as that part which are not visible by human eyes.

8. REFERENCES

- [1] T.H. Lau and A.A. Jumaily, "Automatically Early Detection of Skin Cancer." 978-0-7695-3879-2/09 IEEE
- [2] Huiyu Zhou, Xuelong Lu, Gerald Schaefer, M. Emre Celebi, Paul Miller, "Mean shift based gradient vector flow for image segmentation." Computer Vision and Image Understanding, Volume 117, Issue 9, September 2013, Page 1004-1016 ELSERVIER
- [3] Martins LDO, Junior GB, Silva AC, Paiva ACD, Gattass M, "Detection of Masses in Digital Mammograms using K-means and Support Vector Machine." Electronic Letters on Computer Vision and Image Analysis 8(2):39-50.
- [4] A. Bhardwaj and J.S. Bhatia, "An Image Segmentation Method for Early Detection and Analysis of Melanoma." 2279-0853, p-ISSN: 2279-0861. Volume 13.IOSR 2013.
- [5] S. Wold, K. Esbensen, P. Geladi, "Principal component analysis." Chemometrics and intelligent laboratory systems 2(1):37-52.
- [6] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing Using MATLAB", Third Edition Tata McGraw Hill Pvt. Ltd., 2011