A study of digital image processing in skin cancer diagnostics

Aparna Atul Phatak

#Department of Computer Applications
Pratibha College of commerce and Computer Studies,Chinchwad,Pune.
gateaparna@gmail.com

Abstract*— In today's medical field image processing has very important role in each step of clinical processes, from diagnostic, treatment planning, surgical up to follow up studies. With the recent advances in the field of computer science and technology, enhancement in the interpretation of the medical images has contributed to the early diagnosis of various diseases. The application of image processing for diagnostics purpose is a non-invasive technique. The current study is aimed to identify emerging trends in digital image processing for diagnostics of skin cancer. Computerized Diagnostic Methods being developed to avoid biopsies as well as to decrease the time span required for diagnostics.

***Keywords*—** Skin Cancer, ABCD Parameter, ANN, GLCM, Segmentation.

I. INTRODUCTION

Due to uncontrolled growth of cells in human body leads to the cancer. Normally cells grow and divide to form new cells which are necessary for growth of a human body. When the cells in body become old or get damages, they usually die and new cells take their place. But when new cells from the body doesn't need them, the damaged cells don't die as they should. This extra cell often forms a mass of tissue called a growth or tumor. Tumors on the skin can be benign (not cancer) or malignant Melanoma (cancer). Tumors on the skin can be benign (not cancer) or malignant Melanoma (cancer). Benign growths are not as harmful as malignant growths The sign of skin cancer often starts as the change of color in the skin. They are usually mixed color (pink, red, and brown). There are three types of skin cancer that occurred. They are- Basel cell cancer, squamous cell carcinoma and malignant melanoma tumor. The first two does not spread quickly, but the third one spreads quickly. Melanoma is much less common than basal cell and squamous cell skin cancer, but it is far more dangerous than the other two types. However, it is much more dangerous if it is not found early. It causes the majority (75%) of deaths related to the skin cancer. Worldwide, doctors investigate about 160,000 new cases of melanoma per year.

Melanoma diagnosis is difficult and needs sampling and laboratory tests. Melanoma can spread out to all parts of the body through lymphatic system or blood. The main problem to be considered dealing with melanoma is that, the first affliction of the disease can pave the way for future ones. Laboratory sampling often causes the inflammation or even spread of lesion. So, there has always been lack of less dangerous and time-consuming methods. In order to achieve an effective way to identify skin cancer at an early stage without performing any unnecessary skin biopsies, digital images of melanoma skin lesions have been investigated. To achieve this goal, feature extraction is considered as an essential-weapon to analyze an image appropriately. Computer based diagnosis can improve the speed of skin cancer diagnosis which works according to the disease symptoms. The similarities among skin lesions make the diagnosis of
malignant cells a difficult task. But, there are some unique symptoms of skin cancer, such as: Asymmetry, Border irregularity, Color variation and Diameter. Those are popularly known as ABCD parameters. ABCD parameters. Asymmetry, Border irregularity, Colour, Diameter. Asymmetry is one half of the tumor does not match the other half. Border Irregularity is the
unevenness of images. Colour intensity change in the lesioned region is irregular. Malignant melanoma is having a diameter greater than 6mm.

This paper gives the introduction about Skin cancer and features of skin cancers. It also gives an idea about the Computer based Skin cancer detection system. The paper also, describes the Automatic Skin cancer Detection system and various methods involved in the system.

II. METHOD

Following stages are involved in Skin Cancer Diagnostics

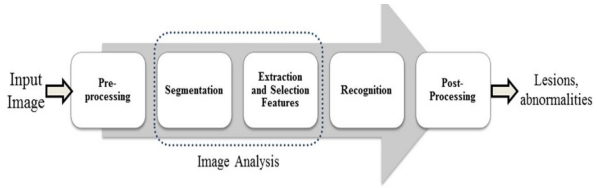

Fig 1: Diagnostics Block Diagram

A. Image Acquisition (input image)

Image acquisition is defined as the process of capturing or retrieving an image from a camera, so it can be passed through various processes need to occur later. Image acquisition in image processing is the first step in the workflow sequence since, without an image, no further process is possible.

B. Pre-processing

Usually the image consists of noises in the form of hairs, bubbles etc. These noises cause
Inaccuracy in classification. Pre-processing is done to removes the noise, fine hair and bubbles in the image. For smoothing image from noise, median filtering is used. Median filtering is a common step in image processing.

Median filtering is used for minimizing the influence of small structures like thin hairs and isolated islands of pixels like small air bubbles.

C. Segmentation

Segmentation removes the healthy skin from the image and finds the region of interest. Usually the cancer cells remains in the image after segmentation. Thresholding often provides an easy and convenient way to perform this segmentation on the basis of the different intensities or colours in the foreground and background regions of an image. The input to a Thresholding operation is typically a gray scale or colour image. After segmentation, the output is a binary image. Segmentation is accomplished by scanning the whole image pixel by pixel and labelling each pixel as object or background according to its binarized gray level.

Fig 2: Cancerous Image

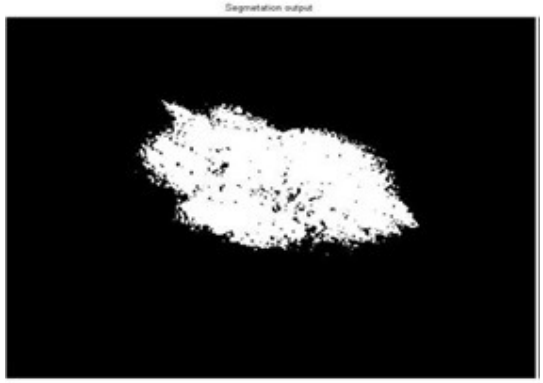

Fig. 3: Segmented Image

D. Extraction and Selection

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. All feature extracted in this study are based on texture analysis using GLCM. The GLCM is a powerful tool for image feature extraction by mapping the grey level co-occurrence probabilities based on spatial relations of pixels in different angular directions. The feature extracted based on GLCM is: Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, energy, Entropy, Homogeneity, Maximum probability, Sum of squares Variance, Sum

average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation 1,Information measure of correlation [3],Inverse difference homogenous (INV), Inverse difference normalized (INN) and Inverse difference moment normalized Selection of features is step where only the most significant features are selected to increase the detection accuracy using fisher score ranking that calculated according to equation (1). The Fisher score ranking technique calculates the difference, described in terms of mean and standard deviation, between the positive and negative examples relative to a certain feature.

$$R_I = \frac{|\mu_{ip} - \mu_{in}|}{|\sigma_{ip} - \sigma_{in}|} \quad \text{.....} \quad (1)$$

Where, Ri is the rank of feature i, the bigger the Ri, the bigger the difference between the values of positive and negative examples relative to feature i.By fisher score ranking equation and a co-occurrence matrix Cd one can draw out some important features for texture classification. According to the input dermoscopy images and fisher score ranking, 12 features are selected where the highest score features seems to be the most effective, these features are Contrast, Correlation, Cluster Prominence, Dissimilarity, Homogeneity, Difference variance, Difference entropy, Information measure of correlation 1,Information measure of correlation 2,Inverse

difference homogenous (INV), Inverse difference normalized (INN) and Inverse difference moment normalized.

E. Recognition

Artificial Neural Networks (ANN) is used for recognition stage. The neural network classifier structure consists of Input layer, Hidden layer and Output layer. The hidden and output layer adjusts weights value based on the error output in classification. The output of the network is compared with desired output. If both do not match, then an error signal is generated. This error is propagated backwards and weights are adjusted so as to reduce the error . The modification of the weights is according to the gradient of the error curve, which points in the direction to the local minimum. In BPN, weights are initialized randomly at the beginning of training. There will be a desired output, for which the training is done. Supervisory learning is used here. The aim of this network is to train the net to achieve a balance between the ability to respond correctly to the input patterns that are used for training [2]. During forward pass of the signal, according to the initial weights and activation function used, the network gives an output. That output is compared with desired output. If both are not same, an error occurs.

Error = Desired Output - Actual Output

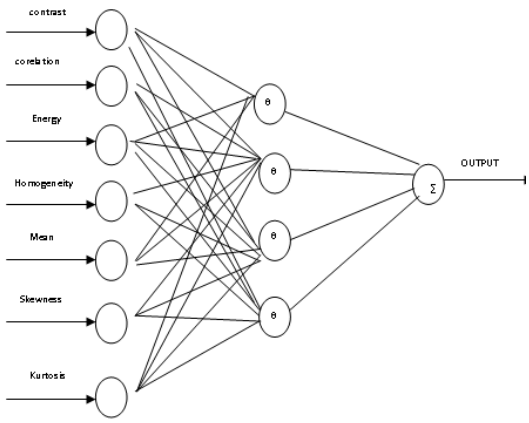

Fig 4: Artificial Neural Network Structure

F. Post *Processing*

Post-processing is done to enhance the shape and edges of image. In addition, contrast enhancement can sharpen the image border and improve the accuracy for segmentation.

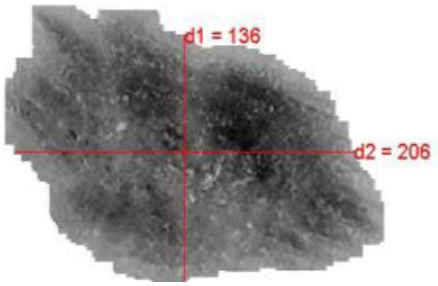

Fig. 5: Lesion with Diagonals to Calculate Symmetry

G. Medical Feature Analysis

Melanoma skin cancer grows asymmetrically and presents a high variegation and some differential structures as a consequence of the irregular distribution of melanin across the lesion. This particular feature leads the classification of melanomas. Medical diagnostic algorithms based on the presence or absences of these features, such as the ABCD rule of dermoscopy, were developed to aid dermatologists [7].

- **Asymmetry (A):** Cancerous lesions are checked for symmetry. If the lesion is Symmetric then it is benign (non-cancerous) and 0 value is assigned to A. For Cancerous (Melanoma) cases the lesion area is asymmetric and value 1 is assigned to A.
- **Border Irregularity (B):** Most of the cancerous lesions edges are ragger, notched or blurred. Its value ranges 0 to 8.To verify Border irregularity the parameter called Compact Index is calculated (CI), it is measurement of the most popular form of barrier unanimous 2D objects. This can be determined by using the following equation:

$$CI = \frac{P_L^2}{4\pi A_L}$$

- **Color(C):** Cancerous skin lesion's pigmentation is not uniform. The presence of up to six known colors must be detected - white, red, light brown, dark brown, blue-gray, and black giving 1 point to each for calculation of C score. Its value ranges 0 to 6. The above mentioned colors are described in RGB color space as follows in Table

TABLE 1: RGB DESCRIPTION OF COLOR PARAMETERS

Colour	RGB	rgb
White	255,255,255	1.0,1.0,1.0
Black	0,0,0	0.0,0.0,0.0
Red	255,0,0	1.0,0.0,0.0
Light-Brown	205,133,63	0.80,0.52,0.25
Dark- Brown	101,67,33	0.40,0.26,0.13
Blue-Gray	0,134,139	0.0,0.52,0.54

In the above table, The RGB column describes colors using integers in the 0-255 range (8bits), while the third rgb column indicates colors in normalized form of colors in RGB column in the range of 0.0-1.0 using floating point representation (rgb = RGB/255).In this proposed algorithm, all the pixels constituting the lesion are scanned and the Euclidian distance in "rgb" coordinates D is calculated between their color and the six reference colors mentioned above.

For example, for the first color, white r1g1b1 = [1.0, 1.0, 1.0] and thus:

D = SQRT ((r-r1)^2 + (g-g1)^2 + (b-b1)^2)
= SQRT ((r- 1.0)^2 + (g-1.0)^2 + (b-1.0)^2).

- **Diameter (D)**: Melanoma cancerous lesions are greater than 6mm wide. Differential structures with at least five patterns are relevant for specific types of lesions. Any other growth of a mole should be of concern. Its value ranges 0 to 5. To find the diameter, first find centroid then draw line from all the edge pixels to the pixel edges through the midpoint and averaged.

III. CONCLUSION

The skin cancerous detection and diagnosis using image processing, neural networks and medical features such as ABCD parameters is implemented. It proves to be a better diagnosis method than the conventional Biopsy method used by dermatologist. This image processing technique is having more advantageous to patients. Artificial Neural Networks is used for the classification of Malignant Melanoma from non-cancerous. The cancerous region is separated from healthy skin by the method of segmentation Based on the features; the images were classified as Cancerous and Non-cancerous. By varying the Image processing techniques and Classifiers, the accuracy can be improved for this system.

IV. REFERENCES

[1.] Ingrid Scholl , Til Aach , Thomas M. Deserno, Torsten Kuhlen " Challenges of medical image processing" Comput Sci Res Dev (2011) 26: 5–13

[2.] Bareqa Salah, Mohammad Alshraideh, Rasha Beidas, and Ferial Hayajneh "Skin Cancer Recognition by Using a Neuro-Fuzzy System," Cancer Inform. 2011; 10: 1–11.

[3.] Mariam A.Sheha, Amr Sharawy , Mai S.Mabrouk "Automatic Detection of Melanoma Skin Cancer using Texture Analysis," International Journal of Computer Applications (0975 – 8887) Volume 42– No.20, March 2012 ,23-26

[4.] Dr. J. Abdul Jaleel, Sibi Salim, Aswin.R.B "Artificial Neural Network Based Detection of Skin Cancer" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 1, Issue 3, September 2012, 200-205

[5.] Damilola A. Okuboyejo, Oludayo O. Olugbara, and Solomon A. Odunaike ",Automating Skin Disease Diagnostic using Image Classification. ," Proceedings of the World Congress on Engineering and Computer Science 2013 Vol II WCECS 2013, 23-25

[6.] Md.Amran Hossen Bhuiyan, Ibrahim Azad, Md.Kamal Uddin "Image Processing for Skin Cancer Features Extraction", International Journal of Scientific & Engineering Research Volume 4, Issue 2, February-2013, 1-6

[7.] Santosh Achakanalli & G. Sadashivappa " skin cancer detection and diagnosis using image processing and Implementation using neural networks and ABCD parameters", International Journal of Electronics,Communication & Instrumentation Engineering Research and Development (IJECIERD) Vol. 4, Issue 3, Jun 2014, 85-96

[8.] https://www.melanomacenter.com