International Journal of Electronics and Communication Engineering (IJECE)
ISSN(P): 2278-9901; ISSN(E): 2278-991X
Vol. 3, Issue 3, May 2014, 1-10

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

STATISTICAL ANALYSIS OF SKIN CANCER IMAGE –A CASE STUDY

SANTOSH ACHAKANALLI[1] & G. SADASHIVAPPA[2]

[1]M.Tech Student, Department of Digital Communication Engineering, R.V. College of Engineering, Bangalore, Karnataka, India

[2]Professor, Department of Telecommunication Engineering, R.V. College of Engineering, Bangalore, Karnataka, India

ABSTRACT

Skin cancers account for more than 40% of all malignancies all over the world, and the incidence continues to rise. This increase is attributed to environmental exposure, principally sunlight. Melanoma is considered the most dangerous type of skin cancer. Early detection and diagnosis depends mainly on important issues, accuracy of feature extracted and efficiency of classifier method. The aim of this paper is to assess what is the role of Image processing in dermoscopy analysis and to determine features of melanoma Images for statistical analysis. The different stages of detection involves- collection of dermoscopic images, filtering the images for removing hairs and noise, segmenting the images using Maximum Entropy Threshold, feature extraction using Gray Level Co-occurrence Matrix (GLCM), and c1assification using Artificial Neural Network (ANN).

KEYWORDS: Artificial Neural Network, Feature Extraction, Melanoma, Skin Lesion

INTRODUCTION

Melanocytes are cells of neural crest origin that migrate during fetal development to multiple sites in the body, principally the skin. Positioned along the basement membrane at the dermoepidermal junction, these cells are exposed to carcinogenic stimuli that result in malignant transformation to become melanoma. Melanoma is the deadliest form of cancer affecting the skin. It arises from cancerous growth in pigmented skin lesion. It is a gradually spreading condition which begins in the melanocytes in skin [1]. Melanocytes are the pigments giving color to the skin. It generally starts as a small lesion later spreads to the other skin areas. If the skin cancer is not diagnosed at initial stages, it can cause death of the patient. So early detection is unavoidable.

Dermoscopy is a non-invasive examination technique based on the use of incident light and oil immersion to make possible the visual examination of sub surface structures of the skin. The rate of detection of melanoma using dermoscopy is higher than detection only with unaided observation [8]. The diagnostic accuracy of dermoscopy also depends on the expertism of the doctors, So that the computer based image processing technique will help for automatic diagnosis which is essential tool for less experienced doctors. It considered to be a "Reference model'' system where doctors take into consideration the information provided by computer before making decision. In this paper a Computer Aided Diagnosis of skin cancer is proposed.

This proposed methodology uses both image processing and Artificial neural network for classification. The similarities among skin lesions make the diagnosis of malignant melanoma, a difficult task. Hence feature extraction using Gray Level Co-occurrence matrix and other features are extracted from image for statistical analysis.

LITERATURE

Various kinds of techniques have been proposed for diagnosis of melanoma. The dermoscope method was first described on 1987 [9]; it facilitates non-invasive diagnosis process based on using of incident light, oil immersion and a magnifier. But its accuracy is still mainly depends on physician experience.

In 2003, M. Wiltgen, A. Gerger and J. Smolle used a method of tissue counter analysis (TCA), which is based on the partition of the image into square elements of equal size where the features, are calculated out of each square element. The features, based on grey level histogram and co-occurrence matrix, allow the differentiation of homogeneous and high contrast tissue areas. The highest classification accuracy obtained by this approach was 92.7%. Jain, Y. K. [10] focuses on the development a skin cancer screening system that can be used by non-experts to classify normal from abnormal cases, using feature detection and classification techniques. The features are extracted using wavelet transform were as the classification is done using neural networks. Fatima, R. [11] introduces a multi-parameter extraction and classification system to aid an early detection skin cancer melanoma Fassihi, N. [12] utilizes morphologic operators in segmenting a d wavelet analysis to extract the feature which culminated in to better melanoma diagnosis system. Alcon, J.F. [13] has used pigmented skin lesion's images, acquired using consumer digital camera for automatic melanoma diagnosis with an accuracy of 86%, sensitivity of 94% and specificity of 68%.Patwardhan, S. V. [14] uses wavelet transformation based skin lesion images classification system which utilizes a semantic representation of spatial frequency information contains in the skin lesion images.

In this proposed method we systematically utilizing the above mentioned research to detect and analyze skin cancer severity. We are proposing a neural network based classifier based cancer severity quantification of skin cancer image. The proposed system will not only extract features from GLCM but along with three other features like 'mean','Skewness' and 'Kurtosis', which leads to overcome different problem cited by above literature.

METHODOLOGY

The proposed methodology of melanoma detection is shown in Figure 1. The method uses the steps of Image acquisition, Pre-processing, feature extraction, feature selection, classification and then evaluation.

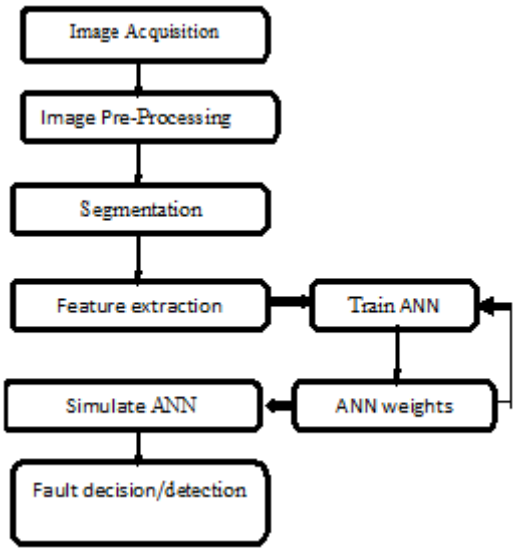

Figure 1: Automated Diagnosis Block Diagram

Image Acquisition

Image acquisition in image processing can be defined as the action of retrieving an image from some source, usually a camera, so it can be passed through whatever processes need to occur afterward. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible.

Image Pre-Processing

The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing. First, image resizing to adjust images to Fixed scale (512*512) so that it supports the classification with clear and accurate features. After that, convert images from RGB to grey level where the features are based on grey level co-occurrence matrix.

Image Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels) The goal of segmentation is to simplify and/or change the representation

Figure 2: Original Cancerous Image

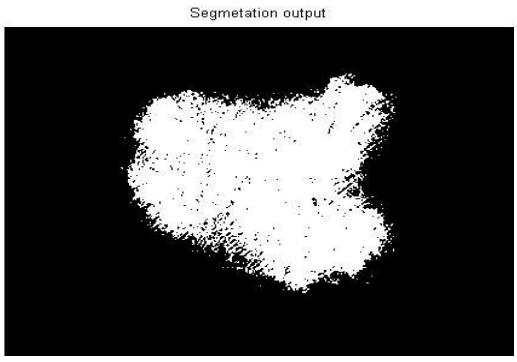

Figure 3: Segmented Image

of an image into something that is more meaningful and easier to analyze.[15][16]. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics, using thresholding method for segmentation.

Feature Extraction

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 features extracted based on GLCM are: contrast, energy, homogeneity, correlation Along with other statistical parameters such as mean, skewness, and kurtosis.

Classification Using Artificial Neural Network

The segmented image is classified based on the extracted features, viz. the texture and color. The design of classifiers were investigated in many studies, most of them being Artificial Neural Networks. Classifier is used for classifying Malignant Melanoma from other skin diseases. Based on the computational simplicity Artificial Neural Network (ANN) based classifier is used [17]. Neural Network is able to solve highly complex problems due to the nonlinear processing capabilities of its neurons. In this system, a feed forward multilayer network is used. Back propagation (BPN) Algorithm is used for training. 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. In BPN algorithm, signal flow will be in forward direction. 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. In BPN, weights are initialized randomly at the beginning of training. There will be a desired output, for which the training is done with Supervisory learning. During forward pass of the signal, according to the initial weights and activation function used, the network gives an output and is compared with desired output. If both are not same, an error occurs.

During reverse pass, the error is back-propagated and weights of hidden and output layer are adjusted. The whole process then continues until error is zero. The network is trained with known values. After training, network can perform decision making. In this proposed methodology, Seven Features were given as input to a multilayer feed forward network as shown in figure 4. There is one hidden layer with four hidden neurons. Output layer with one output neuron, which gives output as 0 or 1, 0 represents non-cancerous and 1 represents cancerous.

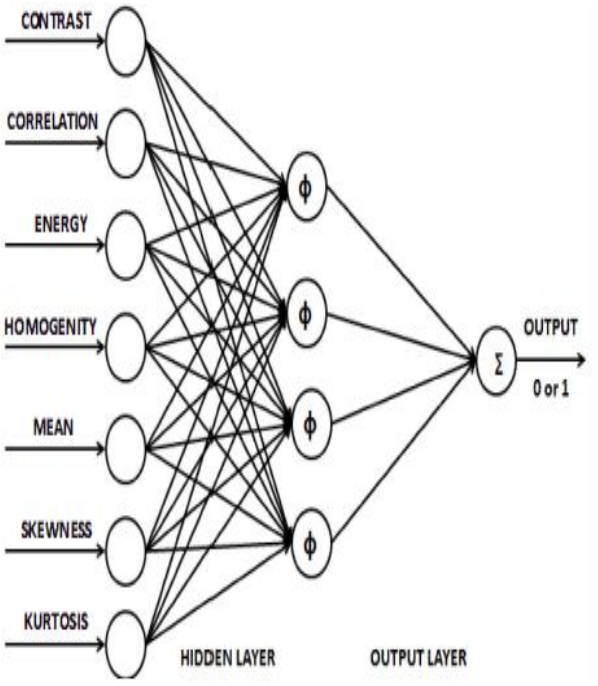

Figure 4: Structure of Artificial Neural Network

Matlab is used for implementation including image preprocessing and for artificial neural network classification. The neural network is trained using the known statistical values of melanoma images. The statistical values of around 60 melanoma cancerous images and non-cancerous images are calculated and given to the classifier, which will give output as either 0 or 1.

STATISTICAL PARAMETERS

The statistical parameters used are mean, skewness, and kurtosis, along with GLCM parameters contrast, energy, homogeneity, correlation.

The mean will give the average value of the pixels of image,

The skewness will provide the information of asymmetry distribution of pixels and kurtosis will give peakedness value of the pixels. The Contrast returns a measure of the intensity contrast between a pixel and its neighbor over the whole image and homogeneity returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. These parameters are defined by the following functions:

$$\text{mean}(\mu)=\frac{\sum_{x=1}^{M}\sum_{y=1}^{N} I(x,y)}{M \times N} \tag{1}$$

$$Energy(e) = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} I^2(x,y) \tag{2}$$

$$Skewness = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y)-\mu)^3}{M\times N\times\sigma^2} \tag{3}$$

$$Kurtosis(k) = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y)-\mu)^4}{M\times N\times\sigma^4} \tag{4}$$

The gray-level co-occurrence matrix (GLCM) or gray-level spatial dependence matrix based calculations fall under the category of second-order statistics. Haralick et. al. [18] suggested a set of 14 textual features which can be extracted from the co-occurrence matrix, and which contain information about image textural characteristics such as homogeneity, contrast and entropy.

A gray level co-occurrence matrix (GLCM) contains information about the positions of pixels having similar gray level values. It is a two-dimensional array, P, in which both the rows and the columns represent a set of possible image values. A GLCM Pd [i, j] is defined by first specifying a displacement vector d = (dx, dy) and counting all pairs of pixels separated by 'd' having gray levels i and j. Pd [i, j] = n_{ij}, where nij is the number of occurrences of the pixel values (i, j) lying at distance d in the image. The co-occurrence matrix P_d has dimension n x n, where n is the number of gray levels in the image. From this co-occurrence matrix we can derive the following statistics as texture features.

$$\text{Contrast} = \sum_{i\,j=1}^{n} P_d(i-j)^2 \tag{5}$$

$$\text{Homogeneity} = \sum_{i\,j=1}^{n} \frac{P_d}{1+|i-j|} \tag{6}$$

GLCM correlation returns a measure of how correlated a pixel is to its neighbor over the whole image. Correlation is 1 or -1 for a perfectly positively or negatively correlated image.

RESULTS

In this proposed system, many Dermoscopic test images were collected from Internet and hospitals. They were undergone

Filtering using Fast Median Filtering. After that, Filtered images were segmented by Maximum Entropy Threshold Segmentation. Feature Extraction technique used is GLCM. Feature extraction was done in MATLAB software. The obtained Features were given as inputs to Neural Network., which gives an output of 0 or 1. Zero represents non-cancerous condition and one represents cancerous or malignant condition. The neural network is designed using Matlab software.

For classification, 60 cases were considered. The ANN Classifier classified the given data into cancerous and Noncancerous. Among that, 37 were classified as cancerous and 33 non-cancerous. There were 6 misclassifications. The Accuracy of this proposed system is 88 %.

The following steps explain the proposed methodology results:

Step 1: To acquire image from the patient

Figure 5: Acquired Melanocyte Image

Step 2: Preprocessing: In this step test image is resized to 512x512and removal of hair (if any)

Figure 6: Resized Melanocyte Image (512x512)

Step 3: Segmentation is done based on thresholding in order to get the pixels of cancerous part

Figure 7: Segmented Image to Get Cancerous Part

Step 4: Converting image to gray level in order to obtain statistical parameters from the grey level co-occurrence matrix of the image.

Figure 8: Grey Scale Converted Image

Step 5: Obtaining statistical parameters of Grey scale image and feeding to neural network same result shown in table 1.

RESULTS VALIDATION

The obtained results were validated with Diagnosis results Prepared by doctors, using the conventional diagnosing Procedures. Among the 60 images given for classification, 39 were Cancerous and 31 non-cancerous according to the Doctor's results. It was taken as the reference for validation.

The ANN based classifier gives the output of 37 cancerous and 33 non-cancerous conditions. There were 12 misclassifications. The obtained results show that the methodology has an accuracy of 88%. The error in classification will get reduced, as the number of sample images taken for classification is increased. By improving the image processing techniques and algorithms for training network, the accuracy of classification can be increased.

Table 1: Statistical Parameters

Mean	Skewness	Kurtosis	Correlation	Contrast	Energy	Homogeneity	Output Of ANN	Conclusion
153.977471	-.974981	3.356303	-0.091121	0.1023	0.3485	0.9541	0	Non-Cancerous
204.668808	-.054156	2.324035	0.072144	0.0697	0.8972	0.9855	1	Cancerous
93.829270	-.087883	2.313674	0.511248	0.2278	0.1590	0.8920	0	Non-Cancerous
140.654838	0.399535	3.463117	0.179639	0.9470	0.4936	0.8742	1	Cancerous
205.097557	-.062998	2.311113	0.077073	0.0691	0.8967	0.9856	1	Cancerous
192.157	-0.61	-l.571	0.9839	0.1659	0.4588	0.9722	1	Cancerous
247.692	-4.753	20.836	0.8567	0.2406	0.9091	0.9844	0	Non-Cancerous

CONCLUSIONS

A statistical analysis using image processing for skin cancer detection system is proposed. It proves to be a better diagnosis method than the conventional Biopsy method. This image processing technique is more advantageous to patients, by which patients can identify the skin cancer without going to hospital or without the help of a doctor. It saves a lot of time for patients. The diagnosing Methodology uses Statistical features of image and Artificial Neural Networks for the classification of Malignant Melanoma from noncancerous. Dermoscopic images were collected and they are processed by various Image processing techniques. Cancerous region is separated from healthy skin by the method of segmentation. The unique features of the segmented images were extracted using GLCM. Based on the features, the images were classified as Malignant or non-melanoma. This methodology has got good accuracy of 88%. By varying the Image processing techniques and training algorithms of ANN, the accuracy can be improved for this system.

REFERENCES

1. Ho Tak Lau and Adel AI-Jumaily, "*Automatically Early Detection of Skin Cancer: Study Based on Neural Network Classification* ", International Conference of Soft computing and Pattern Recognition, IEEE, pp 375-380, 2009
2. Fikret Ercal, Anurag Chawla, William V. Stoecker, Hsi- Chieh Lee, and Randy H.Moss, "*Neural Network Diagnosis of Malignant Melanoma From Color Images*", IEEE Transactions on Biomedical Engineering. vol. 41, No. 9,1994
3. Mang Cao, Kyung A Kang, and Duane F. Bruley "*Skin Cancer Detection Based on NIR Image Analysis* ", IEEE, Biomedical Engineering Conference, 1997., Proceedings of the 1997 Sixteenth Southern pp. 431-436, 1997
4. Pankaj Agrawal, S.K.Shriwastava and S.S.Limaye, "*MATLAB Implementation of Image Segmentation Algorithms*", IEEE Pacific Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference, pp.68-73.

5. Asadollah Shahbahrami, Jun Tang "*A Colour Image Segmentation algorithm Based on Region Growing*", IEEE Trans. on Consumer Electronics Euromicro, pp 362-368, 2010.
6. J Abdul Jaleel, Sibi Salim, Aswin.R.B, "*Computer Aided Detection 0f Skin Cancer*", International conference on circuits, power and computing technology,2013.
7. Mariam, A Sheha, Mai, S. Mabrouk, Amr Sharawy, "*Automatie Detection of Melanoma Skin Cancer using Texture Analysis* ", International Journal of Computer Applications, Volume 42, 2012.
8. Kittler et al.H. Kittler, K. Pehamberger, K. Wolf and M. Binder, "*Diagnostic accuracy of dermoscopy, The Lancet Oncology3*"2002, pp.159–165.
9. McGovern TW, Litaker MS. Clinical predictors of malignant pigmented lesions, "*A comparison of the Glasgow seven-point checklist and the American Cancer Society's ABCDs of pigmented lesions*", J Dermatol Surg Oncol.1992,18:22–6.
10. Jain, Y. K. and Jain, M., "*Comparison between Different Classification Methods with Application to Skin Cancer,*" International Journal of Computer Applications, Vol. 53, No.11, pp. 18-24, Sept. 2012.
11. Fatima, R., Khan, Mohd. Z. A., Govardhan, A. and Kashyap, D. "*Computer Aided Multi-Parameter Extraction System to Aid Early Detection of Skin Cancer Melanoma*," International Journal of Computer Science and Network Security, Vol.12, No.10, pp.74-86, Oct. 2012.
12. Fassihi, N., Shanbehzadeh, J., Sarafzadeh, A., and Ghasemi, E.,"*Melanoma Diagnosis by the Use of Wavelet Analysis based on Morphological Operators*," Proc. of International Multi Conference of Engineers and Computer Scientists 2011, Vol. I, Hong Kong, pp. 1-4, March 16-18, 2011.
13. Alcon, J. F., Ciuhu, C., Kate, W. ten, Heinrich, A., Uzunbajakava, N.,Krekels, G., Siem, D. and Haan, G. de., "*Automatic Imaging System With Decision Support for Inspection of Pigmented Skin Lesions and Melanoma Diagnosis*," IEEE Journal of Selected Topics in Signal Processing, Vol. 3, No. 1, Feb. 2009, pp. 14-25.
14. Patwardhan, S. V., Dhawan, A. P. and Relue, P. A., "*Classification of melanoma using tree structured wavelet Transforms*," Computer Methods and Programs in Biomedicine,Vol.72, pp. 223-239, 2003.
15. Linda G. Shapiro and George C. Stockman (2001): "*Computer Vision*", pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3
16. Barghout, Lauren, and Lawrence W. Lee. "*Perceptual information processing system*." Paravue Inc. U.S. Patent Application 10/618,543, filed July 11, 2003.
17. Guoqiang Peter Zhang, "*Neural Networks for Classification*", IEEE Transactions On Systems, Man, And Cybernetics-Part C: Applications And Reviews, Vol. 30, 2000
18. Haralick R. M., Shanmugum K. and Dinstein I. (1973), '*Texture features for image classicification*', IEEE Transactions on System, Man and. Cybernetics, SMC-3(6), Brodatz Textures: A photographic Album for Artists & Designers, New York: Dover, 1966.