Automated diagnosis of skin cancer

Nagaraj[#1], Dr. Prakash Pattan[#2]
Department of Computer Science and Engineering
Faculty of Engg. and Technology, Sharnbasva University, Kalburgi
PDA College of Engineering, Kalburgi
nagarajmh86@gmail.com, prakashpattan@gmail.com

***Abstract*— Melanoma, Squamous cell carcinoma and Basal cell carcinoma skin cancers are the most common form of cancers in humans. Among these Melanoma is the most dangerous and deadliest one. In which survival rate of patients is very low, if the disease is not diagnosed at the early stage [1]. Therefore, the early diagnosis of skin cancer is very much important. In general, doctors diagnose the cancer by visual inspections or using dermoscopic images. In manual based methods some of the details of affected area of skin may not be observed and can become reason for wrong diagnosis. But automated dermoscopic digital image processing and computer vision technique with the aid of artificial intelligence can overcome this problem. This paper proposes such a method and has yielded good results in identification of skin cancer. In this paper, dermoscopic images are filtered using 84-directional filters to remove hairs from the surface of suspected skin area. The lesion part of the skin is segmented by K-Means clustering technique. Gray Level Co-Occurrence Matrix (GLCM) features are extracted from the segmented part of image. Then the presence of cancer will be detected using SVM classifier.**

***Keywords*— Skin cancer, Melanoma, K-Means, GLCM, SVM**

I. INTRODUCTION

Human Cancer seems as a dangerous disease which is caused mainly by genetic instability and accumulation of multiple molecular alternations. There are many types of cancer among which skin cancer are most common. Skin cancer is a malignant tumor, able to invade surrounding tissues and metastasize (or spread) to other parts of the body, but it depends on the type of skin cancer, and how it's treated. The biological description about the layers of the human skin as shown in fig.1.1 [7]. Skin cancer occurs when errors or mutations occur in the DNA of skin cells. The changes cause the cells to develop crazy and structure a mass of malignant growth cells. Skin malignant growth starts in your skin's top layer: the epidermis. The epidermis is a slight layer that gives a defensive front of skin cells that your body consistently sheds.

The epidermis contains three main types of cells:

Squamous cells: These cells lie just below the outer surface and function as the skin's inner lining.

Basal cells: These cells sit beneath the squamous cells and produce new skin cells. As these cells move upward, they become flattened squamous cells.

Melanocytes: Melanocytes produce melanin. This is the pigment that gives skin its normal color. They are located in the lower part of your epidermis. Melanocytes produce more melanin when you are in the sun to help protect the deeper layers of your skin.

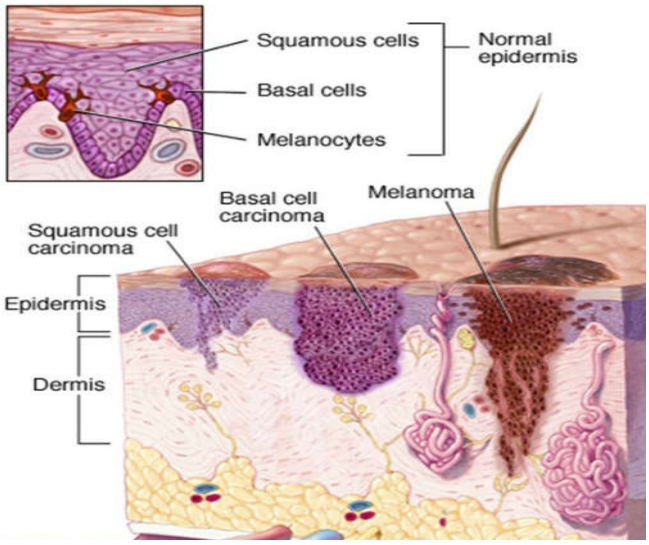


Fig.1.1 Layers of the skin [Wikipedia]

There are mainly two types of skin cancers. The below fig. 1.2 shows the classification of skin cancer cells. Melanoma is more severe than Non-melanoma, but it can be curable, if detected early. Non-melanoma does not spread to other parts of the body, but Melanoma spreads to other parts of the body.

Skin malignant growth is the most widely recognized type of disease in the United States. In 2018, an estimated

1,735,350 new cases of cancer will be diagnosed in the United States and 609,640 people will die from the disease [8]. It is evaluated that one American passes on consistently from skin disease. Every year there are more new instances of skin malignancy than the consolidated frequency of tumors of the bosom, prostate, lung and colon. The sun is the essential wellspring of over the top bright (UV) radiation, which is the reason for most skin diseases. World health organization (WHO) estimates that death due to cancer will increase to 13.1 million in 2030.

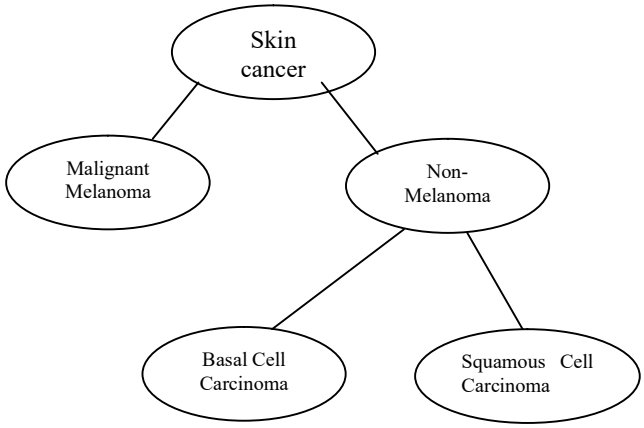


Fig. 1.2 classification of skin cancer cells

II. LITERATURE SURVEY

Enakshi Jana, Dr.Ravi Subban*, S. Saraswathi (2017) [1], "Research on Skin Cancer Cell Detection using Image Processing". in this paper author has proposed three types of skin cancer, among which Melanoma is the most dangerous one in which survival rate is very low if it is not detected at an early stage. So an early and quick identification of skin malignant growth can spare the patient's life. In this median filter used for pre-processing. K-means clustering is used for segmentation. GLCM, Wavelet Transform & Color Features are used for Feature extraction. In this paper, an extensive literature survey of current technology is made for skin cancer detection, among which SVM and Adaboost produces the best results. Finally it uses Artificial Neural Network (ANN) & Support Vector Machine (SVM) classifier for classification.

Sophia Lobo and Pallavi M S (2018) [2], "Predicting Protein in Cancer diagnosis using Effective Classification and Feature Selection Technique", in this paper diverse information mining strategies are used that are expected to predict the protein causing cancer. Fuzzy C-Means clustering is applied for lesion segmentation. Brute Force Algorithm is used to extract crucial features from the genes dataset. Finally the classification based on the K Nearest Neighbours Algorithm (KNN).

Zahra Waheed, Madeeha Zafar, Farhan Riaz (2017) [3], "An Efficient Machine Learning Approach for the Detection of Melanoma using Dermoscopic Images" in this paper author presents first pre-processes the skin image by using mean filter. Then Otsu's Thresholding is applied for lesion segmentation. This paper mainly focuses on two feature sets, color and texture. It is found that color feature gives better than texture in terms of accuracy. Both color and texture features gives accuracy of 96%. Highlight extraction is finished by utilizing Gray Level Co-Occurrence Matrix (GLCM). After feature extraction support vector machine (SVM) classifier is used for classification.

Tamanna Tabassum Khan Munia, Md Nafiul Alam, Jeremiah Neubert (2017) [4]," Automatic Diagnosis of Melanoma Using Linear and Nonlinear Features from Digital Image", in this paper author have proposed Morphological operation and guided filter for pre-processing the skin image. A combination of Otsu's Thresholding & K-Means clustering for segmenting the skin lession. Feature extraction is done by using colour Features, texture Features & ABCD rule. Finally classification is done by SVM, k-NN, decision trees & random forest. This model predicts the diagnosing skin cancer with an accuracy of 89.7%.

M.A. Rahman, M.T. Haque, C. Shahnaz, S.A. Fattah (2017) [5], "Skin Lesions Classification Based on Color Plane-Histogram-Image Quality Analysis Features Extracted from Digital Images", in this paper author has used Median, Gaussian filter & Dull-Razor algorithm. Color Thresholding method in Lab Domain is applied for lesion segmentation. Feature extraction is done by using Histogram based analysis in RGB, HSV, YCbCr & NTSc color planes & Image Quality Analysis(IQA). Finally it uses k-NN & SVM classifier for classification

Youssef Filali, Assia Ennouni, Abdellah Aarab (2017) [6], "Multiscale approach for skin lesion analysis and classification", This paper suggests a new approach for automatic segmentation and classification of skin lesion for dermoscopic images. The segmentation is based on a pre-processing, using the color structure - texture image decomposition to decompose a textured image into texture and geometrical components. Geometrical component is used in the lesion segmentation and the texture component is used to extract the lesion texture features. K-Means clustering is applied for lesion segmentation. Feature extraction is done by using GLCM. Finally it uses SVM classifier for classification.

III. METHODOLOGY

In this paper, a systematic step-wise method is proposed to detect skin cancer. It is shown in the form of block diagram in Fig. 3.1.

A. Skin Lesion Detection

Th The proposed system has mainly 5-stages for early detection of skin cancer.

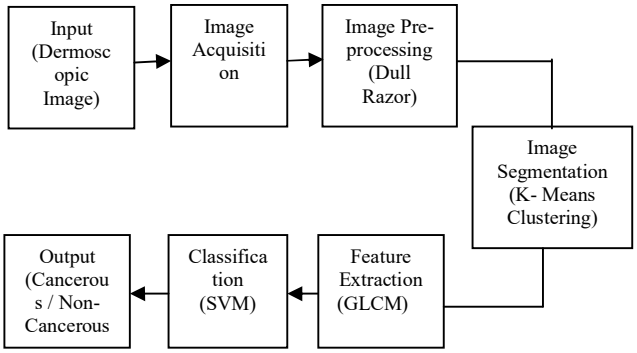


Fig. 3.1 Block diagram for an automated early detection of skin cancer

1) Image Acquisition

Image acquisition is said to be the first step in skin image processing. In practice, various methods are adopted for skin image acquisition, namely, photography, dermoscopic, multispectral imaging, laser-based enhanced diagnosis, optical coherence tomography, ultrasound imaging & magnetic resonance imaging that each of them has their own advantages and limitations. In this paper, dermoscopic image acquisition method is used for image acquisition.

2) Pre-Processing

The main aim of pre-processing is to improve the quality of the image by removing the unwanted artifacts like hairs, air bubbles and scratches from the back ground of the image for further processing. Pre-processing involves mainly three types: Image Enhancement techniques, Image Restoration techniques and Hair removal, but the image has to be processed in two steps. First hair has to be removed by 84-directional filters and after that Wiener filtering is applied to remove additional noises present in the image shown in figure 3.2(e). Wiener filtering is best suited for poisson noise. The presence of hair affects on classification results. Figure 3.2 (a) shows a sample image captured by dermoscopic. Here 84 directional filters are applied to the input image. Then a hair mask is created which is helpful to find the position of hair shown in figure 3.2(b). Finally reconstruct the original image as shown in figure 3.2(c).

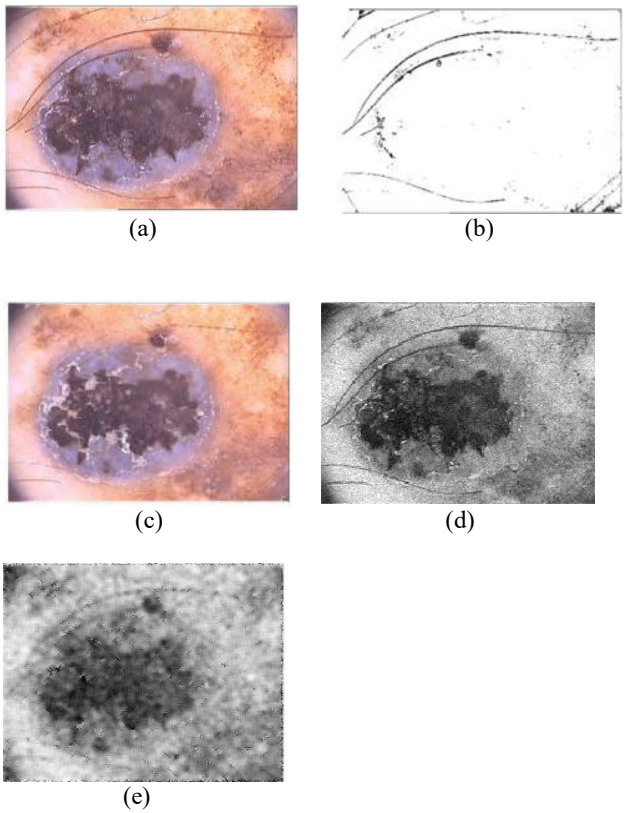
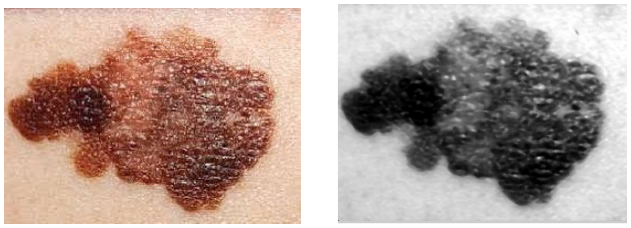


Figure 3.2(a) Dermoscopic image with hair (b) Hair mask (c) Reconstructed images without hair (d) Poisson noise (e) Wiener filter

In contrast enhancement, try to increase image clarity and obtain better performance. It 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 as shown in fig. 3.3.



(a)Original image) *(b) After contrast enhancement*

Fig. 3.3 Contrast Enhancement

3) Segmentation

Segmentation is that the method of partitioning a digital image into multiple segments.. In segmentation, the exact border of the lesion is marked and extracted the informative features from the lesion. The segmented image is then grouped in regions with similar characteristics for simplified analysis. In this paper K-Means clustering segmentation were used.

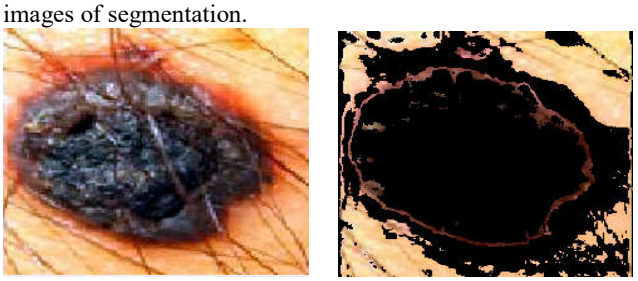
K-Means clustering:

In machine learning, k-means is an unsupervised clustering algorithm used to cluster given data in different similar groups. K-means is used to extract the region of interest from the background. The k-means algorithm works in two steps:

Step: 1 To compute mean of each cluster, and

Step: 2 To compute the distance of each point from each cluster by computing its distance from the corresponding cluster mean.

In this method, the grey-scale image pixels are clustered into two groups. One group is labelled with zero and the other group is labelled with one. These two groups form a mask matrix, which is a black and white version of the original image. White is labelled with one and represents the skin lesion, while black is labelled with zero and represents the normal skin. After that, the mask matrix is multiplied pixel by pixel using Grey-scale/color image which results only the skin lesion part. The below Fig. 3.4 shows the some sample images of segmentation.



(a) (b)

Fig. 3.4 Segmentation by K-Means: (a) Original image (b) After segmentation

4) Feature extraction

After segmentation step, features from segmented regions are extracted. Feature extraction is to extract the parameters of image to characterize the dermatological features and performing the diagnosis based on these parameters. Gray Level Co-occurrence Matrix (GLCM) as one of the popular method to extract the image features. GLCM is used for texture extraction, by using this it will extract only the skin part and other regions will not consider. In feature extraction process the disease part values of the human skin cancer are extracted and the features like mean, standard deviation, entropy, RMS, smoothness, kurtosis, skewness etc are calculated for classification and recognition purpose.

5) Classification

Classification is the final step in computerized analysis to estimate whether the skin lesion is cancerous or non-cancerous. Finally classification is done by SVM classifier. By using Support Vector Machine (SVM) classifier, we classify the given image as cancerous or non-cancer, based on GLCM features.

Support Vector Machine (SVM) classifier:

SVM is a supervised non-linear classifier. Based on the extracted features, SVM model predicts the given image as cancerous or non-cancerous. The purpose of SVM is to create hyper plane that separates two classes with maximum gap between them. SVM is also called as maximum margin classifier. In our planned system, output of GLCM is given as input to SVM classifier that takes coaching knowledge, testing knowledge and grouping information that classifies whether or not given input image as cancerous or non-cancerous. SVM classifier will give better performance among various classification models.

IV. CONCLUSION

Melanoma is the most dangerous and deadliest one, if not detected at an early stage. So an early detection of skin cancer can save the patient's life. To decrease the cost and increase the accuracy of the detection process an automated melanoma detection system is needed. In this paper, we present an automatic system for detection of pigmented skin lesions and melanoma diagnosis various steps are involved. First step is pre-processing in which filtering the image for hair removal and noises. Segmenting the skin lesion by K-Means clustering and feature extraction by Gray Level Co-Occurrence Matrix (GLCM). Finally it classifies the given data set into Cancerous and Non-Cancerous by SVM classifier based on GLCM features.

References

[1] Enakshi Jana, Dr.Ravi Subban and S. Saraswathi, "Research on Skin Cancer Cell Detection using Image Processing", IEEE International Conference on Computational Intelligence and Computing Research, 2017, Puducherry.

[2] Sophia Lobo and Pallavi M S, "Predicting Protein in Cancer diagnosis using Effective Classification and Feature SelectionTechnique", International Conference on Communication and Signal Processing,IEEE, April 3-5, 2018, India, PP. 0156-0159.

[3] Zahra Waheed, Madeeha Zafar, Amna Waheed and Farhan Riaz, "An Efficient Machine Learning Approach for the Detection of Melanoma using

Dermoscopic Images", International Conference on Communication, Computing and Digital Systems, IEEE, 2017, Pakistan, PP.316-319.
[4] Tamanna Tabassum Khan Munia, Md Nafiul Alam, Jeremiah Neubert, "Automatic Diagnosis of Melanoma Using Linear and Nonlinear Features from Digital Image", IEEE, 2017, PP. 4281-4284.
[5] M.A. Rahman, M.T. Haque, C. Shahnaz, S.A. Fattah, W.P. Zhut, "Skin Lesions Classification Based on Colour Plane-Histogram-Image Quality Analysis Features Extracted from Digital Images", IEEE, 2017, Bangladesh, PP. 1356-1359.
[6] Youssef Filali, Assia Ennouni, My Abdelouahed Sabri and Abdellah Aarab, "Multiscale approach for skin lesion analysis and classification", International Conference on Advanced Technologies for Signal and Image Processing, IEEE, May 22-24, 2017, Morroco, PP. 1-6.
[7] The biological description of human skin layer, Available: "https://www.discoverplasticsurgery.com/skin-cancer-richmond-va", September-2018.
[8] Cancer Statistics, National cancer institute, Available: https://www.cancer.gov/about-cancer/understanding/statistics", April-2018.