JES
Journal of Engineering Sciences

Vol 13, Issue 01, Jan /2022
ISSN NO:0377-9254

The Melanoma Skin Cancer Identification and Classification Using Machine Learning Based Support Vector Machine

C.Nikhileswar reddy, M.Tech Student, Tadipatri Engineering College, Dept.Of.ECE
V.Narasimhulu,Asst.Professor, Tadipatri Engineering College, Dept.Of.ECE
R.Sreenivasulu, Assoc.Professor, Tadipatri Engineering, College Dept.Of.ECE
P C Praveen Kumar, Assoc.Professor, Tadipatri Engineering College, Dept.Of.ECE

Abstract: Melanoma skin cancer detection at an early stage is crucial for an efficient treatment. Recently, it is well known that, the most dangerous form of skin cancer among the other types of skin cancer is melanoma because it's much more likely to spread to other parts of the body if not diagnosed and treated early. The non-invasive medical computer vision or medical image processing plays increasingly significant role in clinical diagnosis of different diseases. Such techniques provide an automatic image analysis tool for an accurate and fast evaluation of the lesion. The steps involved in this study are collecting dermoscopy image database, preprocessing, segmentation using thresholding, statistical feature extraction using Gray Level Co-occurrence Matrix (GLCM), Asymmetry, Border, Color, Diameter, (ABCD) etc., feature selection using Principal component analysis (PCA), calculating total Dermoscopy Score and then classification using Support Vector Machine (SVM). The results show that the achieved classification accuracy is 92.1%.

Keywords: Melanoma skin cancer, Image processing, Features, Principal component analysis, Support vector machine.

I. INTRODUCTION

A. Background and Motivation

In the recent 3 decades Melanoma incidence rates have been increasingly high, though most people diagnosed with skin cancer have higher chances to cure, Melanoma survival rates are lower than non-Melanoma skin cancer. Melanoma skin cancer (MSC) can occur on any skin surface, and its incidence has continued to rise over the past two decades in many regions of the world. In men, it's often found on the skin on the head, on the neck, or between the shoulders and the hips while, in women, it's often found on the skin on the lower legs or between the shoulders and the hips [1]. It's rare in people with dark skin and when it does develop in people with dark skin, it's usually found under the fingernails, under the toenails, on the palms of the hands or on the soles of the feet [1].

Many research papers using image processing have been proposed for identification of melanoma skin cancer because of its non-invasiveness and it has increasingly become an efficient diagnostic tool for medical images accurate interpretation, and therefore early and appropriate treatment can be administered to the patient.

. It offers a powerful diagnostic tool to the specialists after the images acquisition immediately, such as
quantifying changes in patient case over time, providing a set of images for teaching and demonstrating, quick comparison of images and, also, might be economically beneficial to the hospitals [2].

B. Contribution

In this study, we proposed an efficient method for detection, features extraction, and classification of melanoma skin cancer using image processing technique applied on suspected melanoma lesion of dermoscopy images. the classification system uses the SVM to classify the lesions. The proposed method shows high accuracy in determining the type of skin lesion whether it is benign or malignant which will be very beneficial for diagnosis of melanoma skin cancer efficiently.

C. Paper Organization

The rest of this paper is organized as follows: Section II describes related work on skin cancer image extraction and classification. Section III explains the procedure of our proposed method. Section IV describes the results of the proposed technique for melanoma images classification. In Section V we conclude the paper with future work.

II. Related Works

Melanoma skin cancer (MSC) detection using non-invasive methods such as image processing techniques became one of the attractive and demanding research in the recants few years. Wiltgen, et al. uses a method of tissue counter analysis (TCA), which is based on partitioning the whole image into square elements of equal size and then features are calculated from these square elements of the image. The features, based on GLCM (Grey level co-occurrence matrix) and gray level histogram, allow the differentiation of homogeneous and high contrast or luminous tissue areas. The highest accuracy of classification obtained by this approach was 92.7% [2].

Fatima, et al. introduced a Multi-Parameter Extraction and Classification System (MPECS) to aid an early detection skin cancer melanoma. The system is based on the extraction of 21 feature from the detected image using six phase approach. After the extraction of these features, a statistical analysis is performed [3].

Patwardhan, et al. provided a method which is based on the use of wavelet transformation based tree structure model for evaluation and the classification of skin lesion images into melanoma and dysplastic nevus. The proposed tree structure model will utilize the semantic representation of the extracted spatial-frequency information contained in the skin lesion images including textural information [4].

Doukas, et al. developed a smart phone based system to store the captured images of skin areas, extract a region of interest and then perform a self-assessment of the images. The system uses a mobile application to acquire and identify the moles in skin images and classify them as melanoma, nevus and benign lesions based on their brutality. The system implemented using 11 classifiers and the experimental result shows that the Support Vector Machine (SVM) has the highest accuracy of 77.06%, then the Multilayer Perceptron of 75.15% [5].

For segmentation of skin lesion in the input image, existing systems either use manual, semi-automatic or fully automatic border detection methods. The features to perform skin lesion segmentation used in various papers are shape, color, texture, and luminance. Many border detection methods are reported in the literature [5, 6].

histogram thresholdin include methods Some, global thresholding on optimized color channels followed by morphological operations, Hybrid thresholding [7,8, 9]. The ABCD ruleof dermoscopy, suggests that asymmetry gives the most prominent among the four features of asymmetry, border irregularity, color, and diameter. Several studies have been carried out on quantifying asymmetry in skin lesions. In Some techniques, the symmetry feature is calculated based on geometrical measurements on the whole lesion, e.g. Symmetric distance and circularity [10]. Other studies proposed the circularity index, as a measure of irregularity of borders in dermoscopy images [11, 12, 13].

III. procedure

In this paper, we proposed an image processing based system to detect, extract and classify the lesion from the dermoscopy images, the system will help significantly in the diagnosis of melanoma skin cancer. More specifically, we proposed a new method to extract the lesion regions from digital dermoscopy images which will be discussed in the next section, where block diagram is shown in Figure. 1.

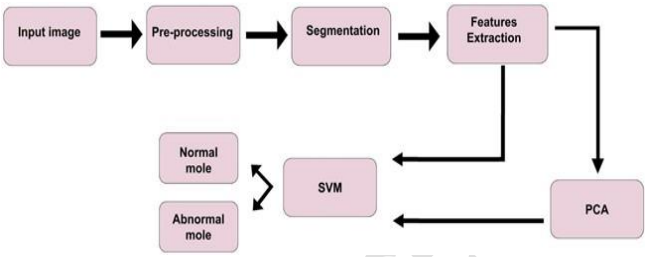

Figure. 1 The proposed system block diagram.

A. Image Database

The database was generated by collecting images from different websites [14,15,16] with known category (Normal, /Melanoma). These websites are specified for melanoma skin cancer.

B. Preprocessing

This step includes Converting the RGB acquired skin image to gray image, Contrast enhancement, Histogram modification and, Noise Filtering. Contrast enhancement and histogram modification are proposed since some of the acquired images are not homogenous due to incorrect illumination during the image acquisition. While the histogram modification techniques such histogram equalization is used to enhance the contrast of the image and, therefore, making the segmentation more accurate.
Noise filtering using median filter is implemented to reduce the impact of hair cover on the skin in the final image used for classification.
Figure. 2 shows the output image from the preprocessing stage.

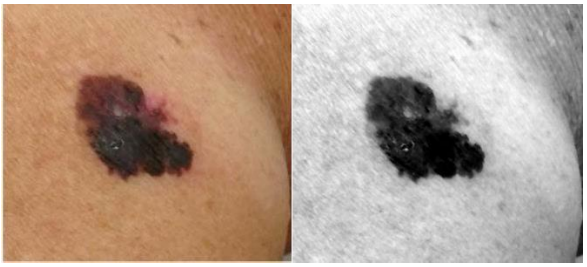

Figure. 2 Image before and after preprocessing process.

C. Segmentation

The second stage after preprocessing is detecting and segmenting the region of interest (ROI) which represents the lesion region. The segmentation stage includes steps: Image thresholding, image filling, image opening, converting extracted region to gray level, and then performing histogram equalization to the extracted gray level image.

We use the Otsu thresholding method since the ROI is homogenous and, therefore, the thresholding becomes dynamic depending on the histogram of the enhanced image. After that, image filling is applied to remove background pixels from inside the detected object and, therefore, make the ROI clear. Image opening is used to remove the extra background pixels which represent a part of non-ROI and, also, to smooth the contour of the object's boundary and breaks narrow isthmuses and eliminates thin protrusions. Finally, the extracted region is cropped then converted to a gray level image and the histogram image is calculated. The results of segmentation process are shown in figure 3.

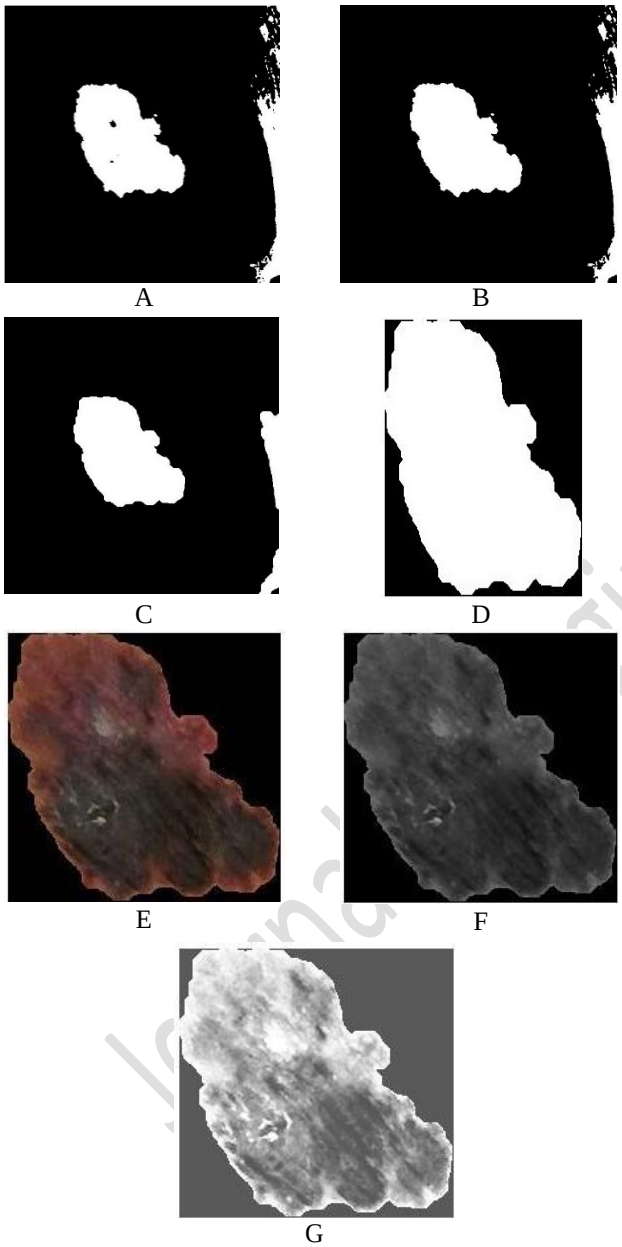

Figure. 3 The result of image segmentation process. A: image after Otsu thresholding, B: image after filling, C: image after opening, D: The ROI mask, E: the segmented ROI, F: the final image after conversion to gray level and G: Histogram equalization to the Image in F.

D. Features extraction

After extracting the lesion (ROI) in the segmentation stage, the predefined features will be extracted from the ROI for classification. The selected features are shape, color and various texture features. Since these images have some statistical texture features, we use one of the common algorithm to extract such features which is Gray Level Co-Occurrence Matrix (GLCM) [17]. In addition, the dermoscopy features (ABCD) are important in distinguishing skin lesion types. We combined these features to get a good classification results for distinguishing the benign from the malignant skin lesions [18, 19]. The feature extraction process includes 4 phases as follows:

1. Phase 1:
   In this phase, we deal with the original image in RGB format, which contains three channels of colors, Red, Green, and Blue. Color feature extracted from image shown in Figure.3- E by calculating the density of specific colors in the lesion image.

2. Phase 2:
   In this phase, we deal with the binary, where the features of Asymmetry, border irregularity, and circulation are obtained from the binary image as it is shown in Figure.3-D. TDS features are calculated with parameters as Asymmetry, Border irregularity, color and diameter

3. Phase 3:
   in this phase, we deal with a lesion image in grayscale image. Energy, correlation, homogeneity and contrast features are obtained by applying gray-level co-occurrence matrix (GLCM) on the gray level image of the lesion as it is shown in Figure.3-F.

4. Phase 4:
   in this phase, we deal with the histogram equalized image, where the features of entropy, skewness, kurtosis and mean are obtained as it is shown in Figure.3-G.

E. Principal Component Analysis (PCA)

The feature extracted from the 4 phases above (contrast, skewness, kurtosis, energy, mean, standard deviation, circulation, energy, correlation, homogeneity and TDS value) are fed into PCA. Since some of the features maybe ineffective on accuracy beside to the time required for accurate classification. PCA is used to reduce the number of features, due to the different units in the feature set. The PCA [20] uses the correlation matrix instead of the covariance matrix. After implementing this operation and calculation of eigenvalues and variances, set of main components is obtained which are arranged based on their ability to distinguish between benign and malignant lesions.

To determine the number of features which are lead to the best classification result, we store the features and their efficiency were checked during classification. Finally, we have selected the best 5 features with maximum efficiency as follows: TDS, mean, standard deviation, energy, and contrast respectively.

F. Classification using SVM

SVM is one of the most common machines learning algorithm that can be used in data classification [18]. It's based on the concept of decision planes that define decision boundaries, a decision plane separates the objects to distinguish classes. The simplest SVM plane is linear whenever the data can be linearly separated. But the data here is not linearly separable. therefore, kernel SVM with radial basis function [21] was used to classify the data into benign or malignant. In the training stage, the classifier model was built using cross validation procedure to find the optimize parameters of the hyperplane to avoid biasing with overfitting [22].

The full set of features (11 features) and the selected 5 features using PCA are fed into the SVM model which is used to classify the image into binary classes benign and malignant.

IV. RESULTS

The proposed method, using the SVM based on the selected features from PCA, was successful in classifying the extracted lesion ROI. The results of the SVM classifier shows accuracy of 92.1% with the full set of features (11 features) as shown in Figure. 4, and the same accuracy is achieved with the selected 5 features using the PCA as shown in Figure .5. These results explain the performance of the classifier model in discriminating the skin lesion into benign and malignant.

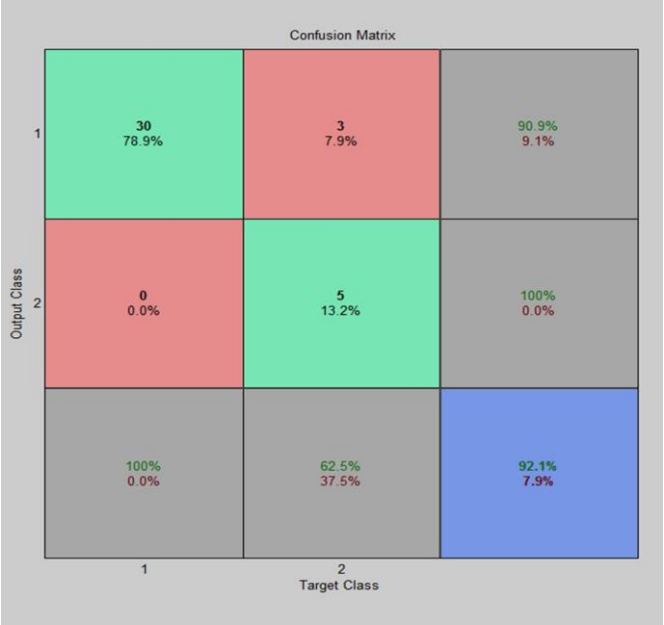

Figure. 4 Confusion Matrix without PCA.

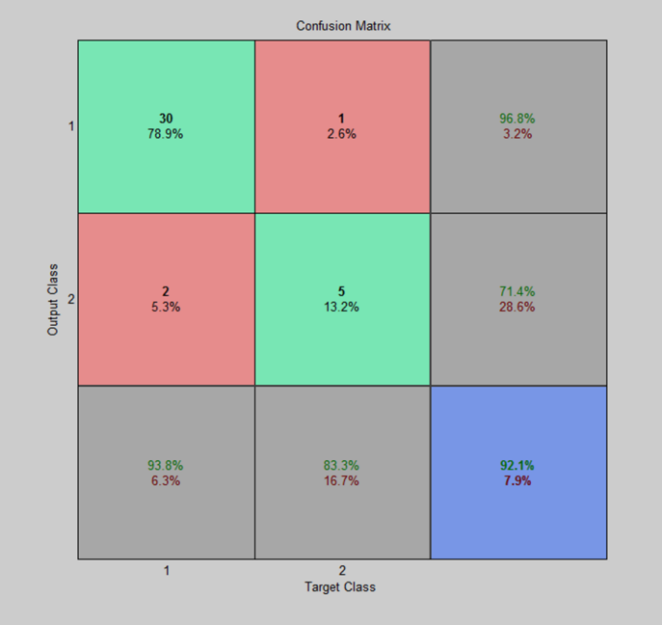

Figure. 5 Test Confusion Matrix of Support Vector Machine with selected 5 features using PCA.

V. Conclusion and Future Work

In this study, we presented a powerful tool for detection, extraction, and classification of skin lesion using PCA and SVM.

We conclude that the same accuracy is achieved when the set of the features selected by PCA or the entire set of features are used, but with lower computational complexity.

The future work on the skin cancer detection system can be more accurate and efficient where the system can be implemented in the stand-alone mobile application, and, therefore, make the system more reliable and practical.

REFERENCES

[1] Landis, Sarah H., et al. "Cancer statistics, 1999." CA: A cancer Journal for Clinicians 49.1 (1999): 8-31.

[2] Wiltgen, Marco, A. Gerger, and Josef Smolle. "Tissue counter analysis of benign common nevi and malignant melanoma." International journal of medical informatics 69.1 (2003): 17-28.

[3] Fatima, Ruksar, Mohammed Zafar Ali Khan, and K. P. Dhruve. "Computer Aided Multi-Parameter Extraction System to Aid Early Detection of Skin Cancer Melanoma." International Journal of Computer Science and Network Security 12.10 (2012): 74-86.

[4] Patwardhan, Sachin V., Atam P. Dhawan, and Patricia A. Relue. "Classification of melanoma using tree structured wavelet transforms." Computer methods and programs in biomedicine 72.3 (2003): 223-239.

[5] Doukas, Charalampos, et al. "Automated skin lesion assessment using mobile technologies and cloud platforms." Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE. IEEE, 2012..

[6] Santosh Achakanall & G. Sadashivappa. Skin Cancer Detection and Diagnosis using Image Processing and Implementation using Neural Network and ABCD parameters. International Journal of Electronics,Communication& Instrumentation EngineeringResearch and Development (IJECIERD) ISSN(P): 2249-684X; ISSN(E): 2249-7951,Vol. 4, Issue 3, Jun 2014, 85-96.

[7] Hall, Per N., Ela C. Claridge, and Jonathan D. Morris Smith. "Computer aided diagnosis of melanoma." Melanoma Research 5 (1995): 19.
[8] Celebi, M. Emre, et al. "Lesion border detection in dermoscopy images." Computerized medical imaging and graphics 33.2 (2009): 148-153.
[9] Iyatomi, Hitoshi, et al. "Quantitative assessment of tumour extraction from dermoscopy images and evaluation of computer-based extraction methods for an automatic melanoma diagnostic system." Melanoma research 16.2 (2006): 183-190.
[10] Garnavi, Rahil, et al. "Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels." International Journal of Medicine and Medical Sciences 1.2 (2010): 126-134.
[11] Stolz, W., et al. "Abcd rule of dermatoscopy-a new practical method for early recognition of malignant-melanoma." European Journal of Dermatology 4.7 (1994): 521-527.
[12] Ng, Vincent TY, Benny YM Fung, and Tim K. Lee. "Determining the asymmetry of skin lesion with fuzzy borders." Computers in biology and medicine 35.2 (2005): 103-120.
[13] She, Zhishun, Y. Liu, and A. Damatoa. "Combination of features from skin pattern and ABCD analysis for lesion classification." Skin Research and Technology 13.1 (2007): 25-33.
[14] Kwait Incp, "Intranational Skin Cancer Collaboration: Melonoma Project"; https://isic-archive.com/#images 2016.
[15] SkinVision BV,"Skin Cancer Picture", https://skinvision.com/mole-images/normal, 2016.
[16] SkinVision BV,"Skin Cancer Picture", https://skinvision.com/mole-images/melonoma, 2016.
[17] Maglogiannis, Ilias, and Charalampos N. Doukas. "Overview of advanced computer vision systems for skin lesions characterization." IEEE transactions on information technology in biomedicine 13.5 (2009): 721-733.
[18] Bono, Aldo, et al. "The ABCD system of melanoma detection." Cancer 85.1 (1999): 72-77.
[19] 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 SurgOncol.1992,18:22–6.
[20] Vasan, K. Keerthi, and B. Surendiran. "Dimensionality reduction using Principal Component Analysis for network intrusion detection." Perspectives in Science8 (2016): 510-512.
[21] Al-Quran, Hiam Hamad. 3D modeling and classification of microcalcification patterns in breast computed tomography. Diss. University of Massachusetts Lowell, 2014.
[22] Alquran, Hiam, et al. "Enhancement of 3D modeling and classification of microcalcifications in breast computed tomography (BCT)." SPIE Medical Imaging. International Society for Optics and Photonics, 2014.

JES
Journal of Engineering Sciences
View publication stats

Vol 13, Issue 01, Jan /2022
ISSN NO:0377-9254