iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4517

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

An Improved Approach for the Segmentation and Classification of Melanoma Tumours

Dr. S. Batmavady[1], K. Nathiya [2] (M.Tech)
[1, 2]*Dept of Electronics and Communication Engineering Pondicherry Engineering College, Pondicherry*

Abstract: ***Malignant melanoma is the deadliest form of all skin cancers. This cancer develops when an unrepaired DNA damage is caused to the skin cells which in-turn makes the skin cells to increase rapidly and form malignant tumours. These tumours originate in the pigment-producing melanocytes in the basal layer of the epidermis. Its diagnosis is crucial, if not detected in an early stage. The proposed work comprises of three steps: In the first step, the lesion region is segmented using edge detection based segmentation. In the second step, different types of color, border and texture features are extracted from dermoscopic images based on distinguished structures. In the third step, extracted features are fed to the classifier to classify melanoma using Radial Basis Function classifier. These three steps are used to differentiate the benign tumours from malignant categories. Qualitatitive and quantitative analysis are analysed. Sensitivity, Specificity and Accuracy parameters are greatly enhanced by the use of the improved algorithms.***

Keywords: Feature Extraction, Melanoma, Neural Network Classification, Pre-processing, Radial Basis Function, Segmentation.

I. INTRODUCTION

Melanoma is a form of skin cancer that occurs in the melanocytes, which are cells in the outer layer of the skin (the epidermis). Melanocytes produce the skin coloring or pigment known as melanin, which gives skin color as tan or brown color. Melanoma develops when melanocytes transforms into malignant tumour, become abnormal, grow uncontrollably and strongly get into surrounding tissues. Melanoma may affect only the skin, or it may spread through the blood or lymph system to other organs and bones. Dermoscopy is a contactless skin imaging technique which shows a magnified view of the skin surface and subdermal structures [1]. Dermoscopy images have played a significant role in increasing the survival rate of patients by assisting the early diagnosis of malignant melanoma. The main stages of the dermoscopic image analysis include preprocessing, segmentation, feature extraction and classification. Pigmented skin lesions are usually examined by dermatologists using the “ABCD” rule [2], which analyzes the Asymmetry, Border irregularity, Color variation and Different structures of a lesion. Based on the “ABCD” rule, many classification techniques have been developed on different datasets of dermoscopy images. In [3], Ganster et al. extracted a set of features from the segmented lesion images, including shape, border gradient and color. A K-Nearest Neighbor (KNN) classifier was finally trained on these features to deliver a sensitivity of 87% with a specificity of 92%. In [4], features descriptive of color, texture and shape were extracted. Support Vector Machine (SVM) was trained to classify the lesions into benign or melanoma. In [5], Backpropagation Neural Network was trained in order to classify the tumours using the set of features descriptive of color, border and texture. It achieves sensitivity of 83 % specificity of 95 % and accuracy of 91%. In this paper, an improved algorithm is proposed for the segmentation and classification of tumours as benign and malignant. The rest of the paper is organized as follows. Section II describes preprocessing. Section III describes segmentation. In Section IV, feature extraction is presented. Section V deals with classification. Section VI describes the Experimental Results. Section VII concludes the work.

II. PREPROCESSING

A. Hair Removal

In Dermoscopy images, if hair occurs on the skin surface, it will be apparently visible in the dermoscopic images, which partially covers the lesion region. Thus, hair can obstruct reliable lesion detection and feature extraction, resulting in unsatisfactory classification results. To detect and exclude the hair from the lesion, first, the hair is segmented form the lesion [6]. To accomplish this task, a set of 84 directional filters are used. These filters are constructed by subtracting a directional Gaussian filter (in one axis sigma of Gaussian is high and in other axis sigma is low) from an isotropic filter (sigma is higher in both axes). Later, these filters are applied to the dermoscopy images. After segmenting the hair mask, the image is reconstructed to fill the hair gap with actual pixels. To reconstruct the image, the system scans for the nearest edge pixels in 8 directions, such as top, bottom, left, right and 4

diagonals. These 8 edge pixels of hair region are found and the mean value of these 8 pixels is stored as pixel value of hair pixel. Using this method, the hairs are detected and removed.

B. Noise Removal

In noise removal process, Gaussian function in Gaussian filter (smoothing filter) is used for blurring the image. This helps in removing the undesirable high frequency components.

III. SEGMENTATION

Segmentation is the process of partitioning an image into several segments. The purpose of segmentation is to evaluate the lesion area accurately. The lesion area is segmented from the background of the skin. In proposed work, the lesion area is extracted using Sobel edge detection. The Sobel operator is used in image processing, particularly within edge detection algorithms. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations. On the other side, the gradient approximation produces relatively crude particularly for high frequency variations in the image.

The gradient is a vector, whose components measure how fast the pixel values are getting modified with distance in the *x* and *y* directions. Thus, the components of the gradient may be found using the following approximation.

$$\frac{\delta f(x,y)}{\delta x} = \Delta x = \frac{f(x+dx,y)-f(x,y)}{dx} \quad (1)$$

$$\frac{\delta f(x,y)}{\delta y} = \Delta y = \frac{f(x.y+dy)-f(x,y)}{dy} \quad (2)$$

where dx and dy measure distance along the x and y directions respectively. In discrete images, one can consider dx and dy in terms of numbers of pixel between two points. dx = dy = 1 (pixel spacing) is the point at which pixel coordinates are (i,j) thus,

$$\Delta x = f(i+1,j) - f(i,j) \quad (3)$$

$$\Delta x = f(i,j+1) - f(i,j) \quad (4)$$

In order to detect the presence of a gradient discontinuity, one could calculate the change in the gradient at *(i,j).*This can be done by finding the following magnitude measure,

$$M = \sqrt{(\Delta x)^2 + (\Delta y)^2} \quad (5)$$

and the direction θ is given by

$$\theta = \arctan\left(\frac{\Delta y}{\Delta x}\right) \quad (6)$$

The different operators is correspond to convolving the image with the following masks,

$$\Delta x = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \ \Delta y = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} \quad (7)$$

When a mask of odd size is used, based on the center pixel (i,j), the operators provide an estimate of pixel values. The Sobel edge operator masks are given as,

$$\Delta x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \ \Delta y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

(8)

The operator determines the gradient of the image intensity at each point, giving the direction of the largest possible increase from light to dark and the rate of change in that direction. The outcome therefore shows how "abruptly" or "smoothly" the image changes at that point and therefore how likely it is that part of the image represents an edge, as well as how that the edge is likely to be oriented. At each image point, the gradient vector points to the direction of largest possible intensity increase, and the length of the gradient vector corresponds to the rate of change in that direction. This implies that the result of the Sobel operator at any image point which is in a region of constant image intensity is a zero vector and at a point on an edge is a vector which points across the edge, from darker to brighter values. When compared to the other system in segmentation, the Sobel Edge Detector gives better segmentation results by extracting the tumour accurately especially in edges, because the lesion region in enclosed with the edges from the background skin.

IV. FEATURE EXTRACTION

Feature extraction performs an part of the identification systems. The purpose of feature extraction is to simplifying the amount of resources required to describe a large set of data perfectly. To categorize the images into benign or malignant, the extracted features are fed to the classification stage. Therefore, to characterize the type of lesions, a parametric approach has been considered. Three features are extracted, such as colour, border and texture.

A. Color Features

Color is considered as important feature descriptor [7] for the classification of melanoma. If the particular region is affected then the color changes effectively. Relative color histograms in different color spaces are constructed to identify the melanoma. The color component is represented by the mixture coefficients of RGB colors. It provides a high correlation among these three color channels. The extracted colour features are: Mean and Standard Deviation.

B. Texture Features

Texture feature [8] is used to estimate the structure, orientation, roughness, smoothness, or regularity of various regions in a set of images that enables us to distinguish between different objects. In order to quantify the texture present in a lesion, a set of statistical texture is based on the gray level co-occurance matrix (GLCM) [9]. The extracted texture features are: Energy, Contrast, Correlation, Autocorrelation, Entropy, Sum Entropy, Difference Entropy, Inverse Difference moment, Homogeneity.

C. Border Features

Border features [10] are used to analyze the border area of the lesion whether it has regular or irregular borders. The border features are based on the measurement of colour and texture features. The extracted border features are: Average, Variance, Sum Variance, Difference Variance.

V. CLASSIFICATION

The extracted features are fed to the classification stage in order to classify the tumour whether it is benign or malignant. In proposed work, the Radial Basis Function Neural Network is used. It is based on the Artificial Neural Networks [11]. The Radial Basis Function Neural Network is basically a stacked of artificial neurons as layers. Radial Basis Function Neural Network has three layers, which are the input, hidden and output layers. The input layer is where input variables are supplied to the network, the hidden layer is between the input and output variables are extracted (or learned) and the output layer is where the computed dependent and target dependent are used to obtain network error for iteratively updating the parameters of the network. The input layer has input variables x_1, x_2....x_D, hidden layer with K neurons, and output layer with M neurons. The suitable number of neurons in hidden layers is determined heuristically, while the number of neurons in output layer depends on the task. Generally, for regression problems, the number of neurons in the output layer is the number of dependent variables. The main idea of the algorithm is that the errors of the hidden layers are specified and updated by the Radial Basis Function Neural Network of the output neurons errors. Radial basis function neural networks are trained with supervised learning algorithms. The Radial basis function operates in forwards pass and Backward pass. In the forward pass, inputs attributes supplied to the network are propagate from the input layer to the output layer through the hidden layer, and the network outputs are computed. The difference between the actual outputs of the network and the desired outputs are used to obtain error terms. In the Backward pass, if the error signal is generated at the output, the error signal is propagated backwards and weights are adjusted at the middle layer in order to reduce the error. After adjusting the weights in hidden layer, the process is continued from the hidden layer instead of training the whole process. This implies that RBFNN with one hidden layer can approximate any function. As a result they are called universal approximators. Perceptrons also

have the capability of universal approximation but only RBFNN possess the ability of optimum approximation. Even if structurally they are less complicated than feed forward back propagation networks they can achieve better arbitrary functions approximations with only one hidden layer. Also, it has better results when many training vectors are available. In radial basis function neural network, mean squared error (MSE) function is calculated using equation 9.

$$MSE = \sum_{m=1}^{M} (t_m - y_m)^2 \tag{9}$$

Where, t is the desired (target) output, y is the actual (computed) output, m indexes output neurons and M is the number of output neuron. The weights of the output-hidden layer (w_{ml}) and corresponding bias weights (w_{mb}) are updated using equations 10 and 11 respectively.

$$w_{ml}(i+1) = w_{ml}(i) + \eta \Delta_m O_l + \alpha[\delta w_{ml}(i)] \tag{10}$$

Where, η is the learning rate, Δ_m is the hidden-output layer error signal for neuron m, O_l is the output of hidden neuron , α is the momentum rate, δw_{ml} is the previous weight change and i is the iteration index.

$$w_{mb}(i+1) = w_{mb}(i) + \eta \Delta_m O_b + \alpha[\delta w_{mb}(i)] \tag{11}$$

Where, O_b is the hidden-output bias neuron and it is set to 1. δw_{mb} is the previous weight change for the hidden-output bias neuron. The hidden-output error signal for neuron m, Δ_m, is calculated using equation 12.

$$\Delta_m = O_m(1-O_m)(t_m - O_m) \tag{12}$$

The output-hidden layers weights (w_{lk}) are updated using equation 13.

$$w_{lk}(i+1) = w_{lk}(i) + \eta \Delta_l \cdot_k + \alpha[\delta w_{lk}(i)] \tag{13}$$

Where, Δ_l is the error signal of hidden layer neuron, δw_{lk} is the previous weight change for the hidden-hidden bias neuron. Δ_l is obtained using the equation 14.

$$\Delta_l = O_l(1-O_l)\sum_{l=1}^{L} w_{ml}\Delta_m \tag{14}$$

Using these equations, the error signals are identified and updated in the hidden layer. Generally, hidden neurons are centered on the training data using some criteria or schemes. Common schemes include random selection of data points, orthogonal least squares, clustering, etc. Hidden neuron activations can be computed with Equation 15, using the Gaussian function.

$$\psi_k(x) = \exp\left(\frac{\|x - x_k\|}{2\sigma_k^2}\right) \tag{15}$$

Where, x_k and σ_k are the center and width (spread constant) of the Gaussian basis function, respectively. Using equation 15, the process is continued from the hidden layer instead of training the whole process and the output is obtained using Radial Basis Function Neural Network.

VI. EXPERIMENTAL RESULTS

The images are collected from Medical Images - PH2 database [12] which makes the dermoscopy images of skin cancer freely available for the research purpose. PH2 database contains 360 dermoscopy images. Among 360 images, 240 images are benign tumour images and 120 images are malignant tumour images. Using PH2 database, the performance metrics like sensitivity, specificity and accuracy are calculated.

In training phase, 290 images are trained using classification algorithm. 290 images are divided into two sets of images. First set is for benign tumour images and it contains 210 images. Second set is for malignant tumour images and it contains 80 images.

In testing phase, 36 images are used. Out of 36 images, 23 images are benign tumour images and 13 images are malignant tumour images. Features are extracted for those images and these features are used as an input layers in the classification stage.

A. Segmented and classified Results (Benign Tumour)

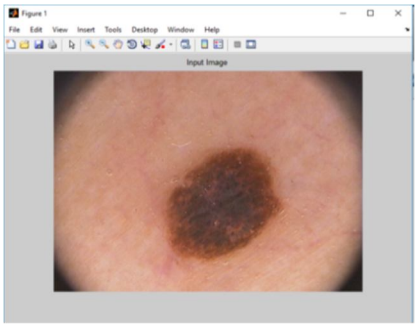


Fig. 1 Input Image

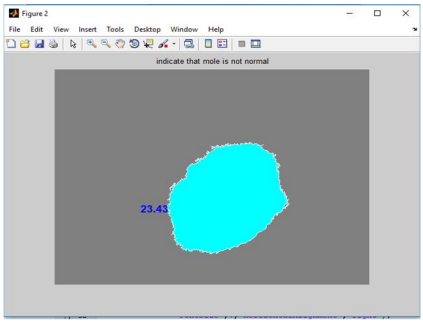


Fig. 2 Diameter value indication of mole

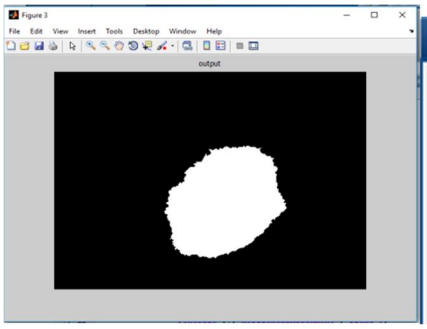


Fig. 3 Segmented output

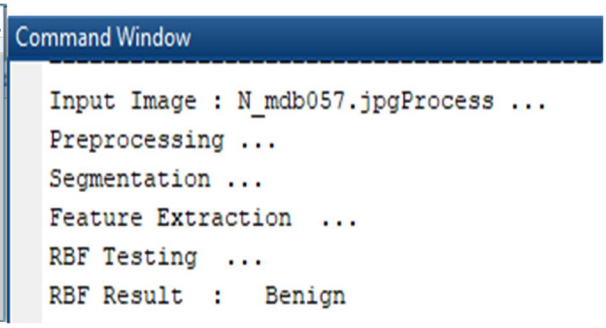


Fig. 4 Classification Output

B. Segmented and Classified Results (Malignant Tumour)

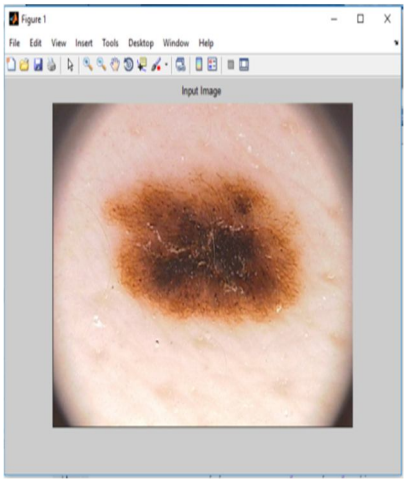


Fig. 5 Input image

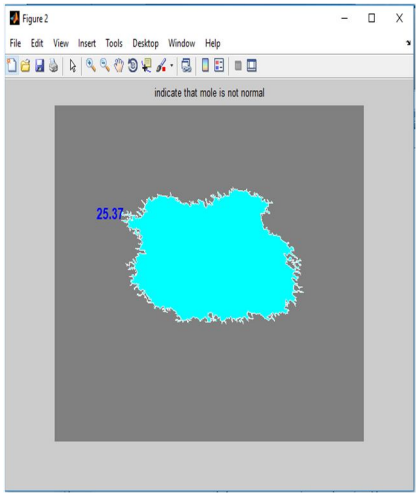


Fig. 6 Diameter indication of mole

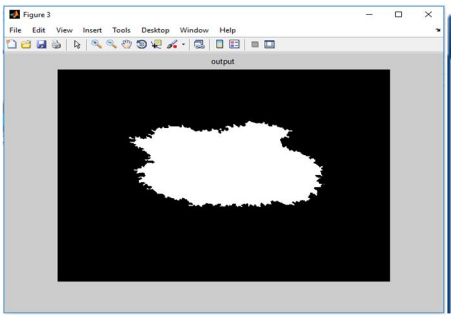


Fig. 7 Segmented Output

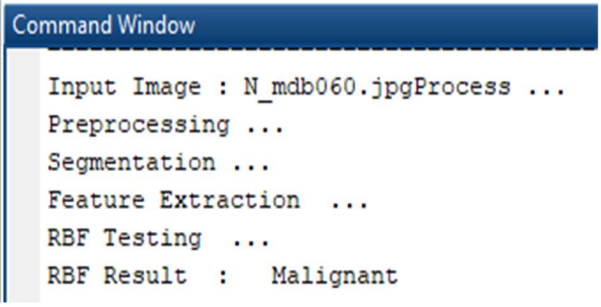


Fig. 8 Classification Output

In Fig 2 and 6, The value indicates the diameter of the mole. The diameter of the normal mole is 6mm. If the diameter of the mole is greater than 6mm, then the value indicates that the mole is abnormal.

Table 1 Confusion Matrix Of Proposed System

	PREDICTED YES	PREDICTED NO	
ACTUAL YES	TP=12	FN=1	13 (36 %)
ACTUAL NO	FP=1	TN=22	23 (64 %)
	13 (36%)	23 (64 %)	

Table 1 shows that in proposed system, 36 images (23 benign tumour and 13 malignant tumour) are used for testing images. From 23 benign images, 22 benign tumours are correctly identified as benign tumour and 1 benign tumours are incorrectly identified as malignant tumour. From 13 malignant images, 12 tumours are correctly identified as malignant tumour and 1 tumours are incorrectly identified as benign tumour.

Table 2 Classification Results

CLASSIFIERS	SEN %	SPEC %	ACC %
FNN	70.00	95.00	87.00
KNN	51.67	97.50	82.22
SVM	75.00	94.17	87.78
BPF	83.33	95.00	91.11
Proposed work	92.00	96.00	94.44

Table 2 shows that the performance metrics in the proposed system (Radial Basis Function Neural Network) is better than the other systems.

VII. CONCLUSION

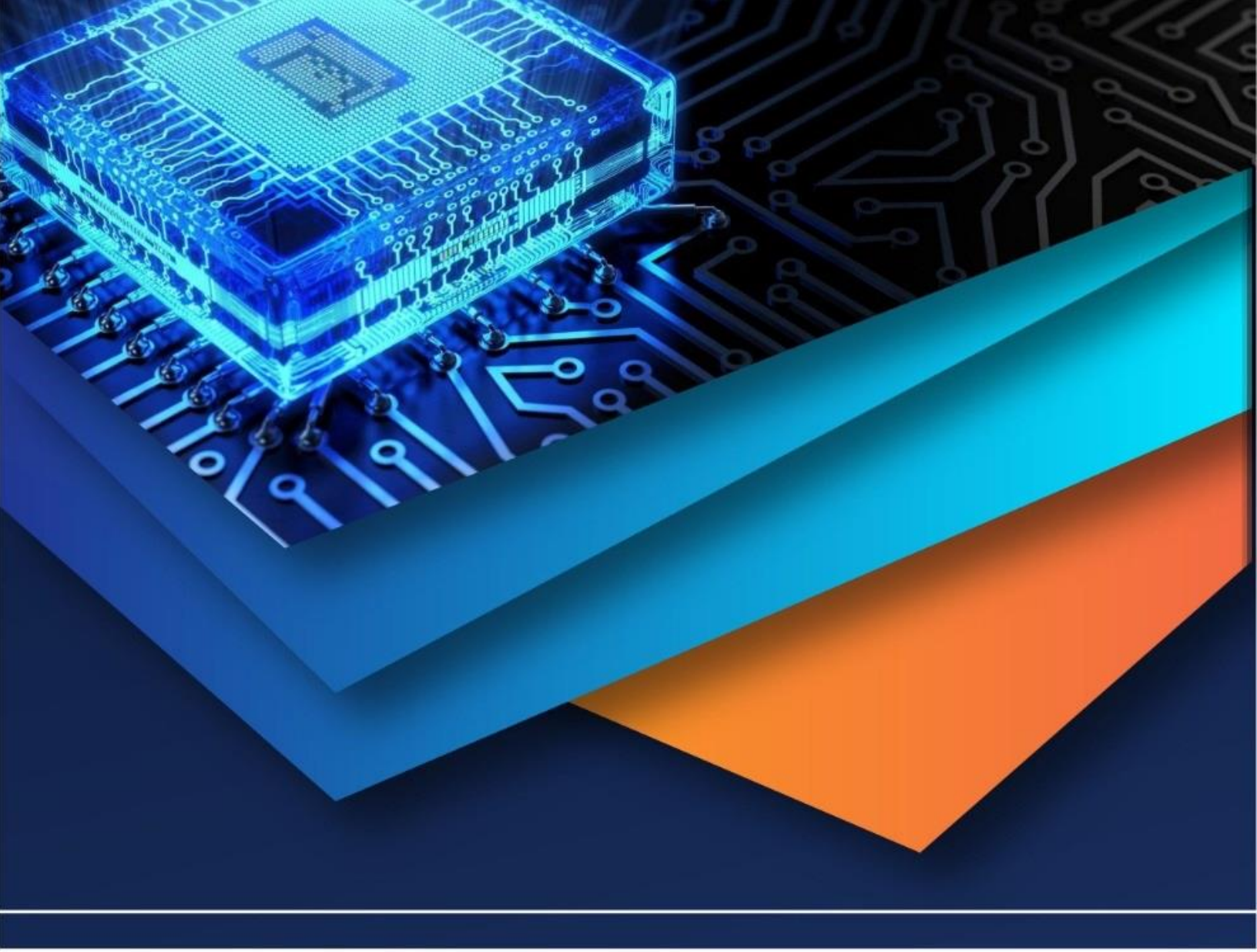
An improved skin cancer detection system is proposed in this paper. Computer based skin cancer detection is more advantageous to patients, by which patients can detect the skin cancer without going to hospital or without the help of a doctor. It saves the patients by early detection. The diagnosing methodology uses Digital Image Processing Techniques and Radial Basis Function Neural Network for the classification of Malignant Melanoma from benign melanoma. Dermoscopic images were collected form the PH2 database and they are processed by Sobel Edge detection in segmentation stage and GLCM technique is employed to extract the set of features from the segmented images. These features are used to classify the image as Malignant or Benign. This classification of proposed system improves the performance metrics like Accuracy, Sensitivity and Specificity when compared with the existing System.

VIII. FUTURE SCOPE

The proposed Radial Basis Function Neural Network classifier identifies the tumour as benign or malignant. After identifying the tumour as malignant, further work focuses on identifying the stages of malignant melanoma in order to detect whether it is in early stage or later stage.

REFERENCES

[1] G. Argenziano, H.P. Soyer, V. De Giorgi, D. Piccolo, P. Carli, et al., "Interactive atlas of dermoscopy,"EDRA Medical Publishing(http://www. dermoscopy.org), 2000.
[2] W.Stolz, A.Riemann, A.B.Cognetta, etal., "ABCDruleofdermatoscopy - a new practical method for early recognition of malignant-melanoma," European Journal of Dermatology, vol. 4, no. 7, pp. 521–527, 2014.
[3] H. Ganster, A. Pinz, R. Rohrer, et al., "Automated melanoma recognition," IEEE Transactions on Medical Imaging, vol. 20, no. 3, pp. 233-239, 2001.
[4] X. Yuan, Z. Yang, G. Zouridakis, et al., "SVM-based texture classification and application to early melanoma detection," Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE. pp. 4775–4778, 2006.
[5] Fengying Xie, Haidi Fan, Yang Li, Zhiguo Jiang, Rusong Meng, and Alan C. Bovik, "Melanoma Classification on Dermoscopy Images using a Neural Network Ensemble Model," IEEE Transactions on Medical Imaging, pp. 1-11, 2016.
[6] Q.Abbas, I.F.Garcia, M.E.Celebi, and W.Ahmad, ''A feature-preserving hair removal algorithm for dermoscopy images,'' Skin Res. Technol., vol. 19, no. 1, pp. e27–e36, 2013.
[7] M.E. Celebi, H.A. Kingravi, B. Uddin, et al., "A methodological approach to the classification of dermoscopy images," Computerized Medical Imaging and Graphics, vol. 31, no. 6, pp. 362-373, 2007.
[8] K. Korotkov, R. Garcia, "Computerized Analysis of Pigmented Skin Lesions: A Review," Artificial Intelligence in Medicine, vol. 56, no. 2, pp. 69-90, 2012.
[9] D.A.Clausi, "Ananalysis of co-occurrence texture statistics as a function of grey level quantization," Canadian Journal of remote sensing, vol. 28, no. 1, pp. 45-62, 2002.
[10] M.E. Celebi M, H.A. Kingravi, H. Iyatomi, et al., "Border detection in dermoscopy images using statistical region merging," Skin Research and Technology, vol. 14, no. 3, pp. 347-353, 2008.
[11] B. Kusumoputro, A. Ariyanto, "Neural network diagnosis of malignant skin cancers using principal component analysis as a preprocessor," Neural Networks Proceedings, 2008. IEEE World Congress on Computational Intelligence, pp. 310–315, 2008.
[12] T. Mendonca, P. M. Ferreira, J. S. Marques, et al. "PH2-A dermoscopic image database for research and benchmarking," in 35th Annual International Conference on Engineering in Medicine and Biology Society (EMBC), pp. 5437-5440, 2013.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY