Skin Cancer Detection using GLCM Matrix Analysis and Back Propagation Neural Network Classifier

Rashi Goel
Galaxy Global Group of Institutions,
Ambala

Saranjeet Singh
Galaxy Global Group of Institutions,
Ambala

ABSTRACT

The presented work here is focussed on extraction of features inclusive of contrast, correlation, homogeneity, entropy, radius, standard deviation and perimeter etc for exact detection of the cancer stage and the post treatment progress could be estimated by the direction and dimensional analysis of parameters. A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. A back propagation neural network is suggested for classification of different class of skin cancer by providing sufficient no. of training images to the classifier. The classification results are given in result section of the paper.

Keyword

Skin Cancer, Neural Network, Back Propagation

1. INTRODUCTION

The skin properties like skin dryness, fungus and allergic symptoms of skin layer may led to starting symptoms of malignant melanoma skin cancer. The correct identification of skin spots based on certain features are the key steps in detecting the skin cancer disease in advance. The affected skin texture profile correlation with malignant melanoma skin cancer is discussed in the proposed in this work. In the existing scenario, the skin images are analyzed in frequency domain. However, it is observed that the skin color in texture images does not vary over a wide range. Hence, the histogram profile of the skin texture remains almost flat. We have shifted the skin texture analysis towards the gray level profile analysis.
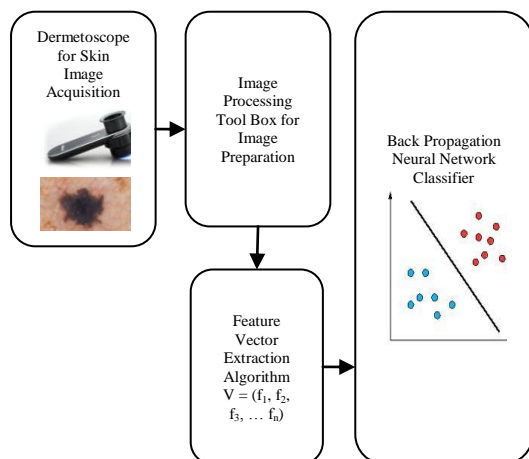

Fig 1.1: Block Diagram of Proposed Work

The gray color profile of the skin texture may give fair idea about the skin sensitivity and is a new emerging skin texture analysis tool. In the proposed work, skin gray color profile has been taken as the input parameter in order to ascertain the skin profile.

2. RELATED WORKS

An automated system for detection and classification of one of the skin four types of skin cancersis proposed here: Melanoma , Basal cell carcinoma, actinic Keratosis, Squamous cell carcinoma .There are a certain features of these types of skin cancers, which can be extracted using proper feature extraction algorithm. [1]. Different algorithms (segmentation and characterization) are used for classification of pigmented skin lesion from a macroscopic image. a new system for characterizing digital images of skin lesions has been presented.. [2]. A scheme for automated detection of skin diseases by analysing the texture recognition techniques based on gray level co-occurrence matrix (GLCM) is discussed here and wavelet decomposition matrix(WDM) and various types of classifiers are used [3]. The characteristic features of the test and the reference images and analysed the skin diseases using texture analysis are extracted. Texture analysis is one of the fundamental aspects of human vision by which we differentiate between surfaces and objects. [4]. Segmentation of skin lesion from the surrounding skin in the dermoscopic images by using Neural Network segmentation algorithm. Different segmentation techniques were applied to the dermoscopic images to segment the skin lesions and evaluated with 3 different metrics, namely sensitivity, accuracy and border error. Segmentation performance shows that Neural Network based lesion segmentation has high sensitivity, accuracy and less border error [5]. A study onthe past and present technologies for skin cancer detections along with their relevant tools is carried out in details. Then it goes on discussing briefly about features, advantages or drawbacks of each of them. discussed the mathematics preliminary required to process the image of skin cancer lesion using proposed scheme. [6]. A technique for early detection skin cancer problem is proposed. The diagnosing methodology uses Digital Image Processing Techniques and Artificial Neural Networks for the classification of Malignant Melanoma from other skin diseases. Dermoscopic images were collected and they are processed by various Image processing techniques. The cancerous region is separated from healthy skin by the method of segmentation. [7]. The detection of melanoma based on region growing segmentation and the ABCD rule used for the detection of malignancy of pigmented skin lesion isdiscussed. [8]. A method for detecting the border and identifying the incidence and propagation of cancer by analyzing the variations of the RGB spectrum of lesion skin images using novel Six Sigma threshold and region connectivity concepts is presented in this paper. [9]. An automated method for melanoma

diagnosis is applied on a set of dermoscopy images. Features extracted are based on gray level Co-occurrence matrix (GLCM) and Using Multilayer perceptron classifier (MLP) to classify between Melanocytic Nevi and Malignant melanoma [10].

3. TEXTURE IN SKIN IMAGES

Identifying the perceived qualities of texture in an image is an important first step towards building mathematical models for texture. The intensity variations in an image Table 3.1: Formulas of Texture Properties which characterize texture are generally due to some underlying physical variation in the scene (such as pebbles on a beach or waves in water).

Table 3.1: Formulas of Texture Properties

Contrast	$$\sum_{r=1,c=1}^{row,col} (i-j)^2 P_{i,j}$$
Correlation	$= \frac{1}{\sigma i . \sigma j} \sum_{r=1,c=1}^{row,col} (i-\mu i) \ (j-\mu j) \ P_{i,j}$
Homogeniety	$$\frac{1}{(1+(i-j))} \sum_{r=1,c=1}^{row,col} P_{i,j}$$
Energy	$$\sum_{r=1,c=1}^{row,col} P_{i,j}^2$$

Texture involves the spatial distribution of gray levels. Thus, two-dimensional histograms or co-occurrence matrices are reasonable texture analysis tools. Texture in an image can be perceived at different scales or levels of resolution. For example, consider the texture represented in a brick wall. At a coarse resolution, the texture is perceived as formed by the individual bricks in the wall; the interior details in the brick are lost.

4. ALGORITHM

The skin images are acquired by using the dermetascope. The images acquired in jpeg format i.e. RGB color format (24-bit). The RGB images are converted into gray color using the following transformation:

Gray = 0.2989 * R + 0.5870 * G + 0.1140 * B (4.1)

A gray level co-occurrence matrix (glcm) is extracted using the following steps:

• Quantize the image data. Each sample on the echogram is treated as a single image pixel and the value of the sample is the intensity of that pixel. These intensities are then further quantized into a specified number of discrete gray levels as specified under Quantization.
• Create the GLCM. It will be a square matrix N x N in size where N is the Number of levels specified under Quantization. The matrix is created as follows:
• Let s be the sample under consideration for the calculation.
• Let W be the set of samples surrounding sample s which fall within a window centered upon sample s of the size specified under Window Size.
• Considering only the samples in the set W, define each element i,j of the GLCM as the number of times two samples of intensities i and j occur in specified Spatial relationship (where i and j are intensities between 0

and Number of levels-1).
The sum of all the elements i, j of the GLCM will be the total number of times the specified spatial relationship occurs in W.
• Make the GLCM symmetric:
• Make a transposed copy of the GLCM
• Add this copy to the GLCM itself This produces a symmetric matrix in which the relationship i to j is indistinguishable for the relationship j to i (for any two intensities i and j).
• Divide each element by the sum of all elements The elements of the GLCM may now be considered probabilities of finding the relationship i, j (or j, i) in W.
Now compute contrast, correlation, energy, entropy and homogeneity from the glcm matrix. The mean intensity can be computed from the gray image by summing method.

$$Mean\ Intensity = \frac{1}{(rwo\ x\ col)} \sum_{i=1}^{row} \sum_{j=1}^{col} F_{(i,j)} (4.2)$$

Area and perimeter are computed by converting the gray image into binary image using the Otsu algorithm. The standard deviation 'σ' of radii is given by:

$$Standard\ Deviation\ (\sigma^2)$$
$$= \frac{1}{row\ x\ col} \sum_{r=1,c=1}^{row,col} (\mu - F(r,c))^2 (4.3)$$

Where µ is the mean intensity and F(r,c) is the pixel intensity at location (r,c).

5. AREA AND PERIMETER COMPUTATION

The area and perimeter of the object of interest in the image frames are computed by using the segmentation and boundary extraction techniques. The input image is segmented on color basis using the k-means clustering technique.

Convert Image from RGB Color Space to L*a*b* Color Space. The L*a*b* color space (also known as CIELAB or CIE L*a*b*) enables us to quantify these visual differences. The L*a*b* space consists of a luminosity layer 'L', chromaticity-layer 'a*' indicating where color falls along the red-green axis, and chromaticity-layer 'b*' indicating where the color falls along the blue-yellow axis. All of the color information is in the 'a*' and 'b*' layers. We can measure the difference between two colors using the Euclidean distance metric.

K-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. K-means clustering requires to specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. Since the color information exists in the 'a*b*' space. Use K-means to cluster the objects into three clusters using the Euclidean distance metric. Label Every Pixel in the Image Using the Results from K-MEANS For every object in our input, K-means returns an index corresponding to a cluster. Label every pixel in the image with its cluster index. Using pixel labels, separate objects in image by color, that will result in different color segments.
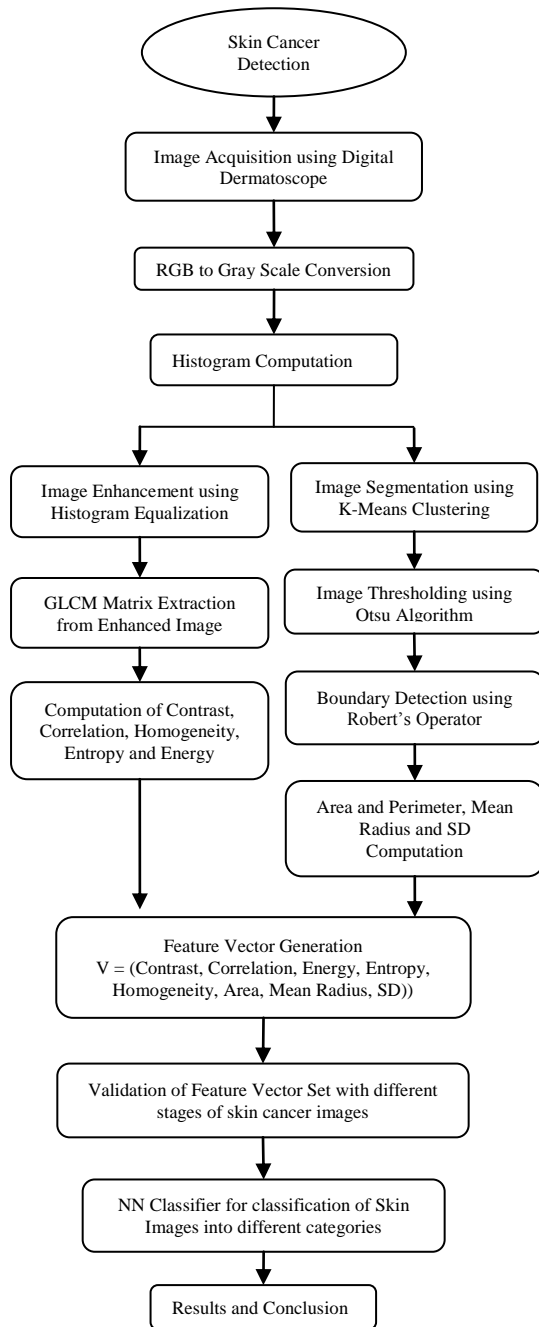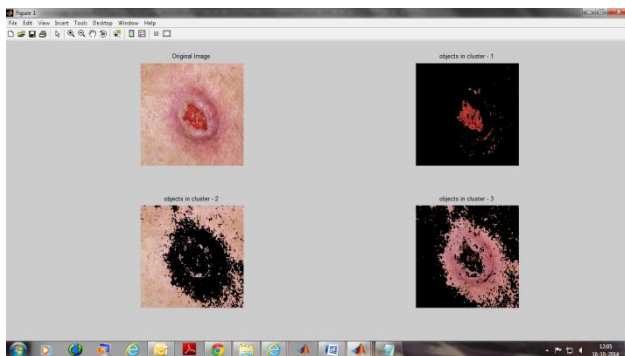
Fig. 5.1: Flow Diagram of Proposed Work

Fig 5.2 :Original Image Converted in to Segmented Image

The segmented image is converted to gray image. The gray image then is converted to binary image using the otsu algorithm as binarization technique. Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that either fall in foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum.

Otsu suggested minimizing the weighted sum of within-class variances of the foreground and background pixels to establish an optimum threshold. Since minimization of within-class variances is tantamount to the maximization of between-class scatter, the choice of the optimum threshold is made based on minimum within class variance.

The Otsu method gives satisfactory results when the numbers of pixels in each class are close to each other. The Otsu method still remains one of the most referenced thresholding methods The area and perimeter are computed by counting the total no. of pixels on the spot body itself and the no. of pixels on boundary of the spot respectively.
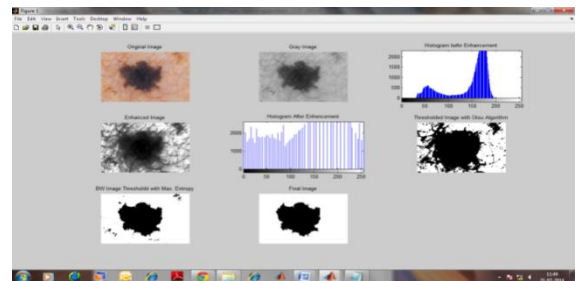

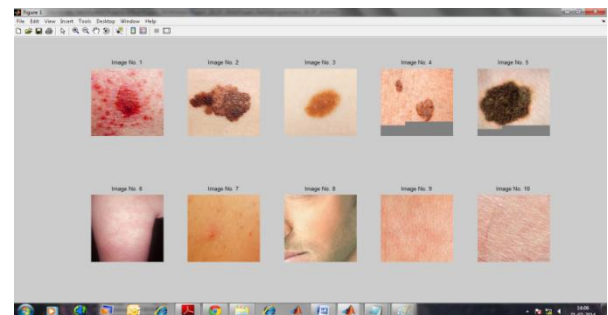
Fig. 5.3: Skin Image at different image processing steps

Fig. 5.4 Meleanoma and normal skin images

6. BACK-PROPAGATION NEURAL NETWORK CLASSIFIER (BPNN)

A back propagation neural network is trained using the following structure:

- No. of Input Neurons = 10;
- Training Class = 4
- Targets: 0.9   Class-1: Normal Skin
- Targets: 0.7   Class-2:
- Targets: 0.4   Class-3:
- Targets: 0.1   Class-4:
- **Total No. of Images   = 50**
- **No. of Classes   = 4**
- Normal Skin Images =   = 25   Class - 1
- Skin Cancer – 1   = 8   Class - 2
- Skin Cancer – 2   = 8   Class - 3
- Skin Cancer – 3   = 9   Class – 4

10 images are taken classification purposes and out which 5 images for skin cancer and 5 images for non-cancer are used.

Table 6.1: Extracted Parameters of Skin Images

Image No.	Contrast	Correlation	Energy	Homogeneity	Entropy	Area	Perim.	Mean R	SD	Threshold
1	2.36	0.72	0.03	0.62	5.75	0	0	18.33	10.77	0.83
2	0.62	0.93	0.06	0.79	5.93	0.32	0.02	101.5	25.1	0.67
3	2.36	0.73	0.03	0.62	5.9	0	0	7.02	1.99	0.91
4	4.41	0.49	0.02	0.53	5.14	0.2	0.08	31.79	14.28	0.74
5	0.85	0.9	0.05	0.77	5.64	0.09	0.01	75.74	13.34	0.91
6	2.51	0.71	0.04	0.65	5.3	0.12	0.01	154.5	23.67	0.88
7	0.71	0.92	0.05	0.76	5.85	0.04	0.01	35.55	11.5	0.96
8	5.58	0.35	0.02	0.5	5.48	0.1	0.02	27.3	2.58	0.9
9	2.33	0.73	0.03	0.61	5.96	0	0	7.34	2.84	0.85
10	2.36	0.72	0.03	0.62	5.27	0.03	0.01	49.29	19.83	0.7
11	0.83	0.9	0.05	0.75	5.23	0	0	6.34	2.52	0.97
12	0.5	0.94	0.06	0.8	5.86	0.2	0.01	77.39	8.92	0.8
13	0.78	0.91	0.05	0.78	5.75	0.06	0.01	37.35	7.33	0.94
14	1.31	0.85	0.05	0.77	5.9	0.16	0.04	54.05	21.11	0.83
15	1.9	0.78	0.04	0.67	5.69	0.13	0.02	80.34	33.6	0.86
16	1.7	0.81	0.04	0.66	5.67	0.07	0.01	60.61	13.5	0.93
17	0.51	0.94	0.07	0.84	5.93	0.01	0	16.81	7.95	0.75
18	0.63	0.93	0.06	0.79	5.96	0.18	0.02	63.47	26.25	0.74
19	4.11	0.53	0.03	0.55	5.97	0.15	0.03	74.35	41.3	0.85
20	1.15	0.87	0.04	0.71	5.61	0.03	0.02	9.62	3.84	0.97
21	3.75	0.57	0.03	0.56	5.91	0	0	8	1.69	0.88
22	1.43	0.84	0.04	0.67	5.77	0.23	0.03	67.59	21.69	0.51
23	2.24	0.74	0.03	0.63	5.95	0.1	0.04	54.03	29.63	0.72
24	1.13	0.87	0.04	0.7	5.72	0	0	7.87	3.21	0.91
25	3.36	0.62	0.03	0.58	5.62	0.05	0.02	38.18	17.81	0.67
26	2.88	0.67	0.03	0.58	5.96	0	0	8.31	2.19	0.63
27	1.19	0.86	0.04	0.71	5.98	0.21	0.04	94.3	40.4	0.75
28	2.39	0.73	0.03	0.61	5.55	0.13	0.02	66.75	23.87	0.86
29	3.2	0.63	0.03	0.59	5.32	0.09	0.02	57.76	20.71	0.9
30	1.18	0.86	0.04	0.7	5.6	0.16	0.04	161.86	66.38	0.75

The upper image set is for cancerous image and the lower image set is normal skin or non-cancer image set as shown in fig.5.4. The images are applied to the BPNN classifier algorithm. The graph below shows the fair categorization between cancer and non-cancer images.
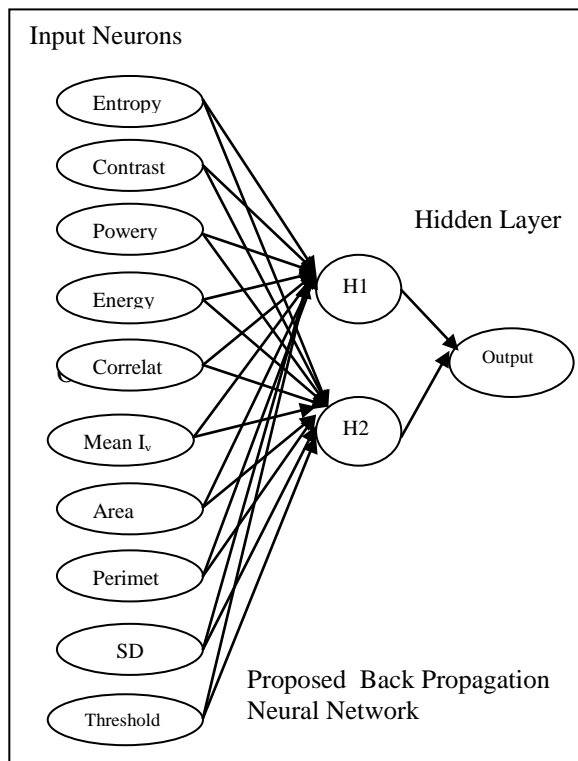

Fig. 6.2: Proposed Back Propagation Neural Network

7. RESULTS

For classificication of skin cancer images, 50 no. of images are taken for training and 10-10 images for testing purposes are chosen.

Table 7.1: Classification of Skin Cancer Images

S. No.	Image Description	No. of Images	Success Rate
1	Normal Skin Image	25	25
2	Skin Cancer-1	8	8
3	Skin Cancer-2	8	8
4	Skin Cancer-3	9	9

Classification Results

Table 7.2: Normal Skin Images with Success Rate

S. No.	Image Description	No. of Images	Success Rate	% Success
1	Normal Skin Image	4	4	100
2	Normal Skin Image	7	7	100
3	Normal Skin Image	10	10	100
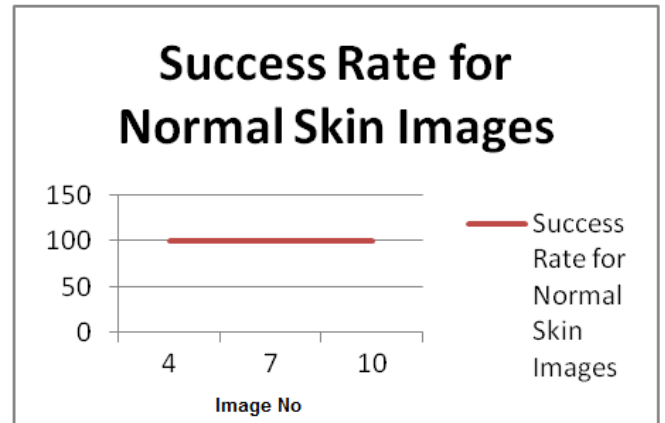
Graph 7.3 : Image Number Vs Success Rate For Normal Skin Images
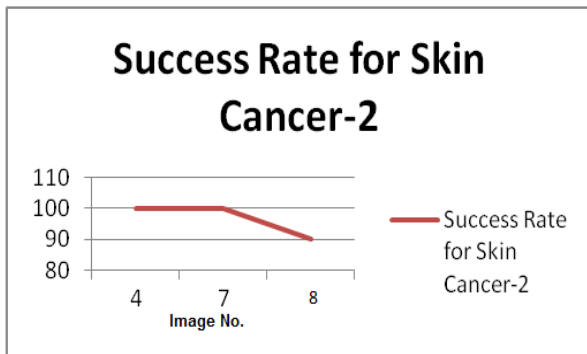
Table 7.4 : skin cancer-1 Images With Success Rate

S. No.	Image Description	No. of Images	Success Rate	% Success
1	Skin Cancer – 1	4	4	100
2	Skin Cancer -1	7	6	85
3	Skin Cancer - 1	8	8	100
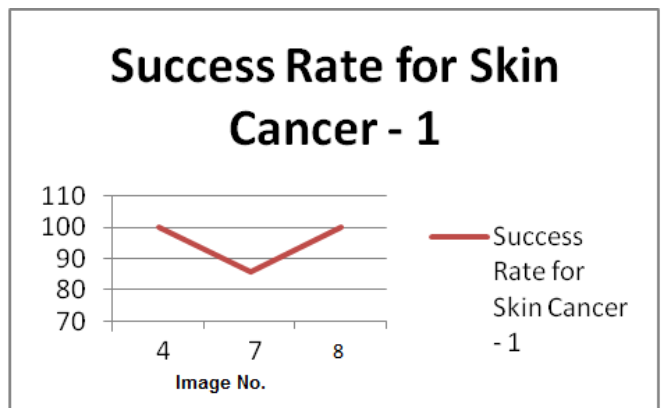
Graph 7.5: Image Number Vs Success Rate For Skin Cancer -1 Images

**Table 7.6: Skin Cancer-2 Images with Success Rate**

S. No.	Image Description	No. of Images	Success Rate	% Success
1	Skin Cancer – 2	4	4	100
2	Skin Cancer -2	7	7	100
3	Skin Cancer -2	8	7	87.5

Graph 7.7 : Image Number Vs Success Rate For Skin Cancer-2

Table 7.8 : Skin Cancer -3 Images With Success Rate

S. No.	Image Description	No. of Images	Success Rate	% Success
1	Skin Cancer – 3	4	4	100
2	Skin Cancer -3	7	6	85
3	Skin Cancer – 3	8	8	100
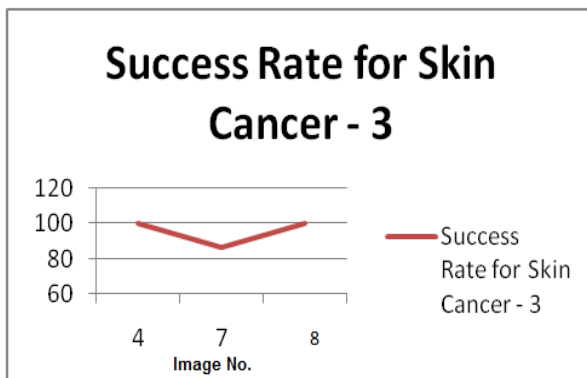
Graph 7.9 : Image Number Vs Success Rate For Skin Cancer -3 Images.

8. CONCLUSION

The algorithm discussed presents fair level of detection of skin disorder vide the parameters as discussed in results table. A validation exercise is required for correlating the parameters with different stages of skin cancer so as to ascertain the ranges the parameters to respective cancer stage. Once the ranges are confirmed, the algorithm may be developed on a real time embedded machine e.g. dermetoscope that can itself suggest the skin disorder level.

9. REFERENCES

[1] Maurya R, Surya K . S , Maurya K .A and Ajeet ," GLCM and Multi Class Support Vector Machine based Automated Skin Cancer Classification,"IEEE journal,vol 12,pp 444-447, 2014.

[2] Messadi M, Cherifi H and Bessaid A," Segmentation and ABCD rule extraction for skin tumors classification," Journal of Convergence Information Technology (JCIT), Volume9, Number2, March 2014.

[3] Parekh R ,"Using Texture Analysis for Medical Diagnosis," IEEE Computer Society, pp 28-37, april-june 2012.

[4] Raja J.V.C, Jeyaprakash M," Skin Disease Diagnosis Using Texture Analysis," International Journal of Advanced Research in Computer Science and Software Engineering, IJARCSSE, Volume 4, Issue 1, pp 275-278, January 2014.

[5] Rajam P.J.J, Thasneem H.A.A," Detection of Skin Lesions in Dermoscopic Images," International Journal of Recent Development in Engineering and Technology, Volume 2, Special Issue 3,pp 193-198, February 2014.

[6] Ramteke S.N and Shweta V.J ," Analysis of Skin Cancer Using Fuzzy and WaveletTechnique – Review & Proposed New Algorithm," International Journal of Engineering Trends and Technology (IJETT) ,Volume 4, Issue 6, june 2013.

[7] Rani N ,Nalam M and Mohan A," Detection of Skin Cancer Using Artificial Neural Network," International Journal of Innovations & Advancement in Computer Science IJIACS , Volume 2, Issue 1 January 2014.

[8] Smaoui N and Bessassi S," A developed system for melanoma diagnosis," International Journal of Computer Vision and Signal Processing (IJCVSP), vol 3, issue 1,2013.

[9] Sethumadhavan G and Sankaran S," Border Detection and Cancer Propagation on Spectral Bands of Malignant Melanoma using Six Sigma Threshold," IEEE Journal, ACIS International Conference on Computer and Information Science, vol 5,2009.

[10] Sheha A.M, Mabrouk S.M and Sharawy," Automatic Detection of Melanoma Skin Cancer using Texture Analysis," International Journal of Computer Applications (0975 – 8887) Volume 42– No.20, March 2012.