IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Computer aided diagnosis of skin cancer

Dr. Manjula Vasant Kiresur
Dept. of ECE
RNS Institute of Technology
Bengaluru
India

Faseeha Tarannum
Dept. of ECE
RNS Institute of Technology
Bengaluru
India

Abstract- Melanoma is one of the most increasing cancers since past decades. For accurate detection and classification, discriminative features are required to distinguish between benign and malignant cases. Computer vision can play important role in medical image diagnosis and it has been proved by many existing systems. Considering and analysing features of cancer image, which includes Asymmetry, Border irregularity, Compact index, Fractal dimension, Color variation and diameter is popular technique of analysing the patients with skin cancer. To extract and analyse such features, image segmentation plays important role for automatic skin cancer detection system. In this paper we propose the image segmentation scheme based on support vector machine and Convolutional neural network classifier the validated results showed a very encouraging performance with a sensitivity 80.48%, a specificity of 76.19% and accuracy of 98.36% of SVM classifier using linear kernel.

Keywords: *Disease, medical image, image processing, digital image, diagnosis, skin.*

I. INTRODUCTION

In the present world of electronics and communication, there is a huge demand for the applications with medical and hospitals. Under the fast elaboration of the internet and biomedical techniques, image data in the form of features, images, extraction and selection through, the transform technique provides an excellent features for the kernels functions and over these classification has become a challenging task. Today the most important problem in health and medicine context is cancer. Earlier diagnosis and timely treatment are very effective to improve and survival so image processing as a decisive tool can help the physician to diagnose cancer early. Mechanism for image processing is a simple and non invasive method to detect cancer cells so that it accelerate earlier diagnosis and increase rate of survival. Earlier diagnosis and timely treatment lead to improve and survive cancer patients. As treatments of cancer are based on interference methods such as surgery, radiotherapy and chemotherapy, the studies showed that new technologies such as image processing have been successful for diagnosis and classification of cancers [1]. Melanoma is less common than some other types of skin cancer, but it is more likely to grow and spread. Skin tumor such as other tissue tumors may be malignant or benign. Their nature and status are very different in skin cancer so that they may soft or hard, loosing or moving, shallow or deep as respects their shape and size may not be consistent [2]. Cancer especially melanoma is a skin disease which thousands of patient die in world annually. 40 -50 % of diagnosed cancers are related to skin and malignant melanoma is the most aggressive type of skin cancer that is very fatal [3]. It is important to note that melanoma is a treatable disease if it is diagnosed early. Studies showed that earlier diagnosis leads to improve 90-95% of patients [4]. Earlier diagnosis of melanoma can dramatically prevent the death cause by malignant cancers. There are two major problems to diagnose the disease: 1) because of ignoring skin lesions or lack of access to dermatologist, skin lesions change from benign to malignant 2) because of similar symptoms, skin lesions are not diagnosed correctly. Recently many systems and algorithms were designed to diagnose malignant or benign lesions by dermoscopy.

II. LITERATURE SURVEY

Various research works have been done using image processing and computer vision to detect melanoma and benign in most cases as shown in table 1

Table 1: Literature survey table

Ref	Kind of class	Features	Classifier	Accuracy
[5]	Melanoma And Benign	Structural and textural	SVM	86.07%
[6]	Melanoma And Benign	Boundary and geometric	SVM	84%
[7]	Melanoma And Benign	GLCM	Neural network	60%
[8]	Melanoma And Benign	Shape, texture and color	Neural network	73%
[9]	Melanoma and benign	Sparse coding feature	SVM	93.10%

III. METHODOLOGY

Our proposed approach is comprised of two phases

Training phase and testing phase as shown in fig 1 and 2. In our proposed method leverages a collection of classifiers which are trained at various resolutions to categorize into

melanoma and non-melanoma skin cancer. In testing phase, trained classifiers are applied on new images.

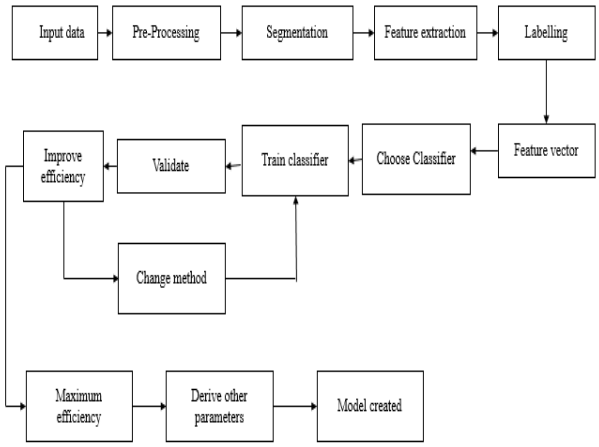

Fig1: Block diagram of training and validation

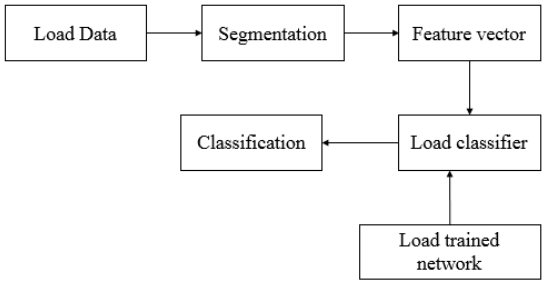

Fig 2: Block diagram of testing

A. Image Acquisition

Acquisition of skin cancerous image is an important procedure. Because the acquired image will undergo the rest of the process in the system, which may not be achievable or obtained result will not be a reasonable one. Hence the acquired image should have healthy in resolution and clarity. The images for this research were obtained from online skin diseases gallery https:// www. ISIC.com/. A total of 400 image samples are used. An example of an acquired images is as shown in fig 3.

Fig 3: Sample Acquired Image

B. Image Pre-Processing

Images collected for skin cancer detection are not suitable for direct application of classification algorithm. Skin images may contains different unwanted artifacts such as fine hair, air bubbles, and noise. All images in the dataset need to be demonised. Different literature uses different methods for this purpose. Median filter is methods which is used in most of the literature hair removal and de-noise of skin images [6]. The output of the pre-processing steps is as shown in the figure 4

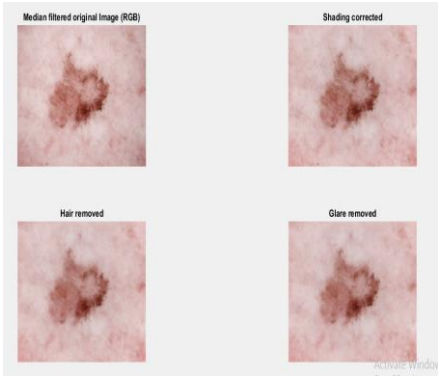

Fig 4: Pre-processing output images

C. Image Segmentation

Image segmentation is a process of partitioning an image into multiple parts to identify an object and filter out relevant information. In skin cancer detection technology, after pre-processing of a skin image, it is necessary to segment out the region of interest. Efficient segmentation of skin image can improve the performance of classification algorithm used in different literature are discussed.

1. Watershed Segmentation

The method of watershed is based on a concept of visualizing an image in three dimension where two domain constructs spatial coordinates and third dimension constitutes intensity. Hence three points are considered one belonging to the local minima where if a pixel is placed at a locations of these points, would fall to a single local minimum. Known as catchment or watershed catchment basin or watershed. The boundaries constituting watershed are equal to the numbers of minima in image. Watershed segmentation is shown in fig 5.

Steps: 1. .Read in the color image and converts it to the grayscale image.
2. Find the gradient magnitude of the grayscale image.
3. Apply morphological operators by computing the following:
a) Compute opening: defined by erosion followed by dilation.
b) Compute opening by reconstruction: this is obtained by calculating the complement of the opening image.
c) Compute closing: defined by dilation followed by erosion.
d) Compute opening-closing: it is obtained by opening and closing operation.
e) Compute opening-closing reconstruction: this is obtained by calculating the complement of the opening closing image.
f) Compute binary map segmentation using otsu thresholding method.

2. Otsu segmentation method

Otsu's thresholding method [7,8] 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 as shown in fig 6.

Steps 1. Convert to grey scale.
2. Filter using a Gaussian filter with a high enough sigma to blur
3. Convert to binary picture using Otsu's threshold method
4. Transform logical binary image into numeric values.
5. Fill holes.
6. Keep the largest area.
7. Compute the boundaries.

3. *Modified Otsu segmentation*

The modified Otsu algorithm calculates the derivation and demonstrates the monotonicity of class between variance first, and then uses quartered search method to search optimal threshold value according to the monotonicity of class between variance and traversing the region of gray level after quartered search for calculating the maximum class-between variance [9]. This method can find reliable optimal thresholding as shown in fig 6, compare to watershed and Otsu based segmentation method, conduct image segmentation and improve the processing speed.

Steps: 1. Convert image into the light intensity channel.
2. Apply morphological opening / closing by reconstruction
3. Apply Ostu thresholding.
4. Remove area that touch the border.
5. Keep the larger area
6. Compute boundaries
The time consumed for 3 different types of segmentation is as shown in table 2

Table 2: Time consumed foe 3 methods of segmentation.

Segmentation methods	Watershed Method	Otsu method	Modified Otsu method
Time	6.131 sec	9.862sec	5.761 sec

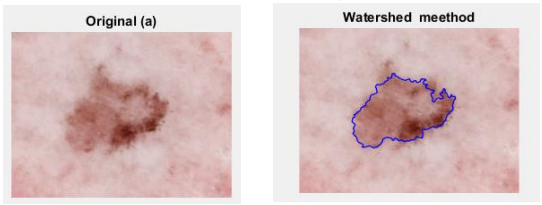

Fig 5: watershed segmentation

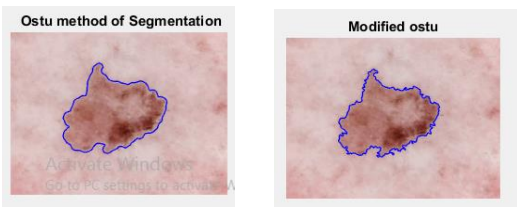

Fig 6: segmentation of Otsu and Modified Otsu methods

D. *Feature Extraction*

No machine learning algorithm can work without pre-defined features set. Define feature is the most important prerequisite for using any image classification algorithm. In this section, we describe different features used for skin cancer detection. The type of features can be broadly divided into following categories.

Border Feature

1. Feature Border Border feature are one of the most important as they determine the degree of irregularity of the lesion from the normal skin region. The feature border value is calculated using the formula

Border_ diff = mean (abs (r_bin_border-r_bin_hull))

2. Feature irregularity

This feature is used to determine the swelling of the lesion. The feature irregularity is calculated using formula

$$\text{Irregularity}= \frac{(\text{minor axis length} * \text{major axis length})}{2*\pi*(\text{minor axis length}^2 * \text{major axis length}^2)}$$

3. Feature roundness

This feature gives the roundness score of the lesion. The feature roundness is calculated using formula

Roundness = $(4*\pi * \text{area}) / (\text{perimeter}^2)$

RGB Feature

In this method of feature extraction, an RGB image is considered and separated into 3 color channels and obtain features like the color, chromaticity, range of individually of R,G ,B color channel in the lesion. This helps classifier to differentiate between the melanoma and non-melanoma cancer affected region based on the pixel values in each color channel. A brief introduction to each of the features of RGB is shown in table 3

Table 3: List of RGB feature extracted

Features	Explanation	Values
Relative chromaticity	Defined as the quality of color as determined by dominant wavelength	2.2428
Minimum value of RGB	This feature returns the minimum value of each color channel within the lesion	1.5313
Maximum value of RGB	This feature returns the minimum value of each color channel within the lesion	8.9217
Range of RGB	Gives the range of each color channel within lesion	1.7954

Mean value of RGB	This function returns the mean value of individual color channel	6.8353
Standard deviation of RGB	Standard deviation is the amount of variation in a given dataset. This function returns the standard deviation of the each color channel in the lesion	7.8135
Skewness of RGB	This function returns the skewness of the each color channel in the lesion	6.8304

Geometric Feature

It is the combination of machine learning and computer vision for recognition tasks. The main objective is to find features of geometric form to classify an object. The set of features is given to classifier, so that it can learn and be able to classify the objects into its respective groups. The geometric features extracted for this paper is density, aspect ratio and area of the bounding.

1. Feature density

It is the measure of compactness of an object. The density measure is given by the formula

Density_ score = Area of lesion /Area of convex hull

2. Feature aspect ratio

Aspect ratio is the comparative association among the image thickness and highness. The measure of aspect ratio is given by formula

Aspect ratio _score = s. Minor Axis length /s.Major Axis length.

3. Feature region of the bounding box

This gives the proportion of region of bounding box to region of the lesion. The region of bounding box is given by formula

Area_ ratio = region of the lesion / region of bounding box.

Once all features are extracted for the dataset given, it is then given to the classifier so that the classifier can train and test the images and later classify them into the correct class. All the geometric features is as shown in table 4

Table 4: Geometric features

Features	Values
Feature density	0.0514
Feature aspect ratio	0.333
Region of bounding box	0.1667

E. Skin cancer image Classification

The computer based skin cancer detection technology consists of image acquisition, pre-processing of the image to remove artifacts, hair and bubbles, image segmentation and finally segmented image classification into melanoma and non-melanoma. Most of the literature for skin cancer detection uses machine learning classification algorithm at the last step of skin cancer detection. In this paper we are using two types of classifiers

a. Support Vector Machine

Support vector machine is a supervised nonlinear classifier which constructs an optimal n-dimensional hyper plane to separate all the data points in two categories. Support vector machine tries to assign each data point in the training dataset to a class by maximizing the separate hyper plane in between them. SVM is also called as maximum margin classifier. While training, SVM tries to maximize the distance of the both the hyper plane from the separating hyper plane. SVM uses the kernel to transform the feature space to higher dimension when the data are not linearly separable in original feature space.

Classification

Algorithm

Algorithm 1 gives the flow of Testing, Training and classification procedure used in the proposed work.
Algorithm 1: Identification of skin image
Input: Different types of skin image
Output: Identification of skin image as benign and malignant

Method: SVM classifier

Train Phase

Start
Step 1: Input the skin image
Step 2: Perform pre-processing of image using median filter.
Step 3: Perform segmentation for the primary detection.
Step 4: Extract Border, RGB and Geometric features of segmented image.
Step 5: Store all the features as a feature vector in train library.
End.

Test Phase

Start
Step 1: Input the skin image.
Step 2: Perform pre-processing of image using median filter.
Step 3: Perform segmentation for primary detection.
Step 4: Extract Border, RGB and Geometric features of segmented image.
Step 5: Compare the test image features with the database features.
Steps 6: Identify the input test image using SVM classifier as benign or malignant.
End

Parameters	Formulae	Results
Sensitivity	TP / (FN+TP)	1
Specificity	TN / (FP+TN)	1
Precision	TP / (FP+TP)	1
Accuracy	TP+TN/(FP+FN+TP+ TN)	98%

b) Neural Network Classifier

The heaviness of the inputs are adjacent to train the perception with supervised learning. In neural network learning system, in development the designs that are to be predictable are well-known and the input ideals of the training set are categorized with number one outputs. Previously, the weight initialization have been done by means of random values. In turn the perceptron is delivered with every training set.

The anticipated productivity is obtained by relating the outcomes from the perceptron to every input set. The modification of weights is not achieved if the output is correct. If incase the output is wrong, the result have to be resolute by differentiating the patterns as the result would like to be, and weights have to be adjusted towards the wanted result based on the active inputs. More than two classes are represented using a perceptron structure and a learning rule is defined for the network using the multi-layer perceptron or multi-layer feed forward network model. Each level that has been allocated into input levels, the unknown layer and the output layer gives the input to the succeeding one. The structure required to identify the non-linearly separable classes is given by the extra layers.

Classification

Algorithm

Algorithm 1 gives the flow of Testing, Training and classification procedure used in the proposed work.

Algorithm 1: Identification of skin image
Input: Different types of skin image
Output: Identification of skin image as benign and malignant
Method: Neural Network classifier
Train Phase
Start
Step 1: Input the skin image
Step 2: Perform pre-processing of image using median filter.
Step 3: Perform segmentation for the primary detection.
Step 4: Extract Border, RGB and Geometric features of segmented image.
Step 5: Store all the features as a feature vector in train library.
End.

Test Phase

Start
Step 1: Input the skin image.
Step 2: Perform pre-processing of image using median filter.
Step 3: Perform segmentation for primary detection.
Step 4: Extract Border, RGB and Geometric features of segmented image.
Step 5: Compare the test image features with the database features.
Steps 6: Identify the input test image using SVM classifier as benign or malignant.
End

III. Results and Discussions

The system has educated from training data, for the organization the test data is provided, then the outcome is created by test data for test data to measure correctness of expectation and then with a conclusion that whether the person is distress as of the benign or malignant melanoma.

SVM CLASSIFIER OUTPUT

Classifier SVM is selected and trained with feature vector. The classifier efficiency is calculated based on confusion matrix as shown in fig 7

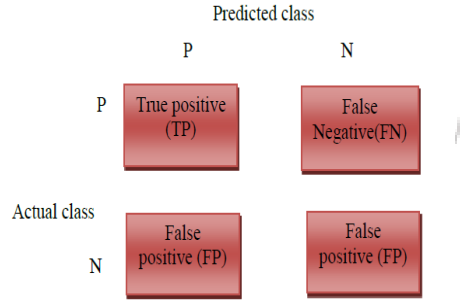

Fig 7: Confusion Matrix

Table 5: Validation results of SVM

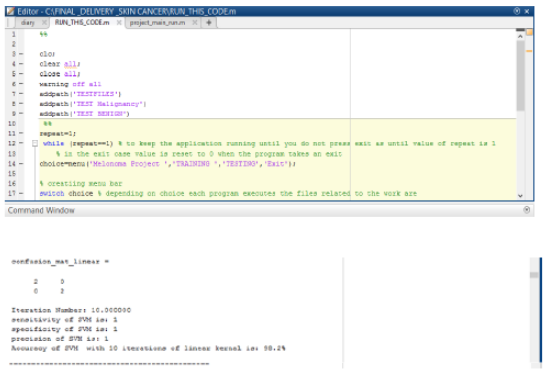

Fig 8: Evaluation matrix obtained using Matlab

Neural Network Classifier

The classification is carried out using neural network classifier as unconventional classifier the database contained images of benign and malignant melanoma.

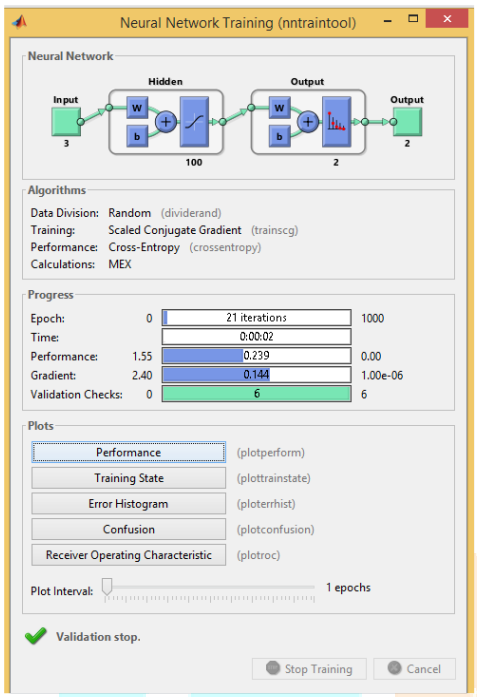

Fig 9: Graphical user interface of neural network

The classifier efficiency is calculated based on confusion matrix as shown in fig 10

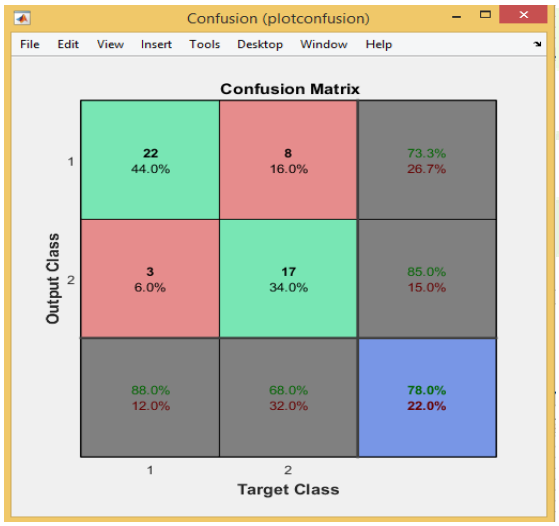

Fig 10: Confusion matrix of neural network

Table 6: Validation results of neural network

Parameters	Formulae	Results
Sensitivity	TP / (FN+TP)	88%
Specificity	TN / (FP+TN)	68%
Precision	TP / (FP+TP)	73%
Accuracy	TP+TN/(FP+FN+TP+ TN)	78%

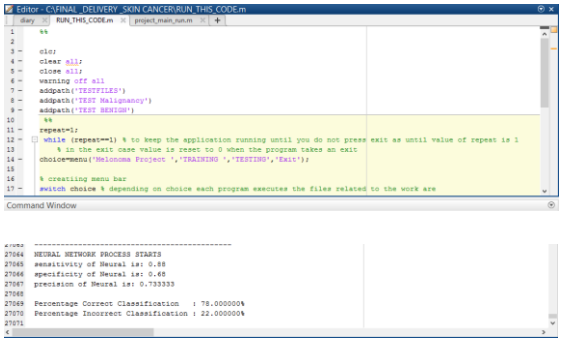

Fig 11: Evaluation matrices obtained using Matlab

Test Results

The simulation of Support vector machine to test and training images provide classification of the cancer, which uses the built in features extracted from an images is acts as an input. In the resultant image, the region of cancer present is shown as a highlighted region with the help of segmentation as shown in fig 8 and it provides the type of the cancer. Simulation had done and the accuracy, sensitivity and specificity is as shown in table 5.

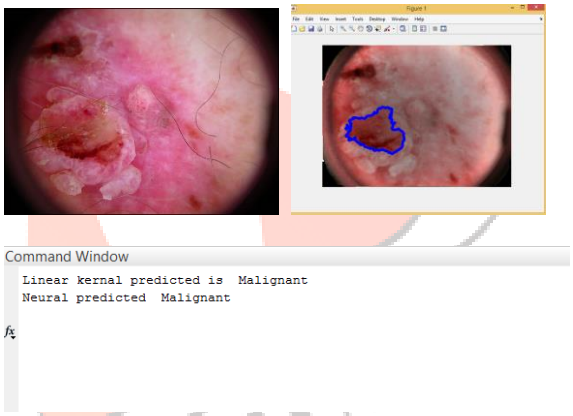

Fig 8: Test image of melanoma type skin cancer

Table 7: Performance of the method used

Sensitivity	Specificity	Accuracy
80.48%	76.19%	98.36%

IV. Conclusion

In past decade, medical imaging as non-invasive tool with high accuracy and performance has been attracted. Therefore mechanisms for processing image can be used to diagnose and treat disease without invasive activities. The aim of this study was to diagnose benign and melanoma. The most important step to segment image with high accuracy. Hence modified Otsu segmentation method is used it is giving accurate segmented region compare to watershed and Otsu method of segmentation. Border-based features, RGB based features and Geometric features is coded successfully using matlab code. In our investigation there is 98.34% accuracy in svm where as only 78% accuracy is found to be in neural network. Therefore SVM technique is said to be superior than neural network. Hence

high performance is achieved. Finally melanoma or benign is found or not is decided based on classification, results and performance.

V. References

[1]. D. Glotsos, et al. "A multiclassifier system for the characterization of normal, infectious, and cancerous prostate tissues employing transrectal ultrasound images" , computer methods and programs in biomedicine, vol 97 .no. 1. pp. 53-61, 2010

[2]. T. Dwyer, et al. " The use of spectrophotometry to estimate melanin density in caucasins" , cancer epidermiology and prevention biomarkers, vol 7, no.3 , pp-203-206, 1998.

[3] MP. Staples, et al . " non-melanoma skin cancer in Australia the 2002 national survey and trends since 1985," medical journal of Australia, vol 184, no.1, pp-6-10, 2006.

[4] M.A . Mahmoud, A. A. Jumaily and M.S . Takruri, "Wavelet and curvelet analysis for automatic identification of melanoma based on neural network classification". International Journal of computer information systems and industrial management 2013.

[5] Faouzi Adjed, Syed jamal safdar Garde Zi et.al "Fusion of structural and textural features for melanoma recognition",IET Journals ISSN 1751-9632, 2017..

[6] Yading Yuan, Ming Chao, et.al. "Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance", IEEE transaction on medical imaging, DOC 10,1109, 2017

[7] D. Csabai , K. Szalai , M. Gyngy, " Automated classification of common skin lesions using bio inspired features", IEEE international ultrasonic symposium proceedings (987-1-4673-9897-8), 2016.

[8] J. Abdul Jalul, Aswin. R.B " Computer aided detection of skin cancer", International conference on circuits, power and computing technology (978-1-4673-4922-2) ICCPCT-2013..

[9] M.A. Farooq, M.A.M .Azhar et.al "Automatic lesion detection system for skin cancer classification using SVM and Neural Network DOI 10.1109, October 2016.

[10]. R. Suganya, " An automated computer aided diagnosis of skin lesions detection and classification for dermoscopy images", april 2016.

[11]. Gonzalez, R.C, Woods, R.E: Digital image processing, 2nd edn, prentice hall (2002).

[12]. Sonka, Hlavac, Boyle: Digital image processing and computer vision, India edn. CENGAGE learning (2007).

[13]. Guang-zhou, yun-zhen, et.al, " Modified Otsu algorithm using quartered search method" , vol 978, no-1 pp-413-4, 2016

[14]. R.J. Hijmans and J.Van etten, " Raster Geographic analysis and modelling with raster data, " R Package version 2.0-12, Jan 12, 2012 [online]. Available:http://CRAN.R-project.org/package=raster.

[15] G.Argenziano, H. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, and M. Binder," Dermoscopy of pigmented skin lesions: Results of consensus meeting via the Internet Journal of the American Academy of Dermatology",vol 48,pp 679-69, 2003.