Melanoma Detection in Dermoscopic Images using Naïve Bayes Classification Algorithm

Deval E. Chavan[1] Prof. Karbhari V. Kale[2]
[1,2]Department of Computer Science & Information Technology
[1,2]Dr. Babasaheb Ambedkar Marathwada University,Aurangabad (MH),India

Abstract— *Malignant melanoma is the most dangerous form of human skin cancers and its occurrence has been rapidly increasing. Early detection of cancer could be very crucial as its detection in the early stage can be beneficial to cure it. Computer Aided Diagnosis (CAD) can be beneficial in early detection of cancers for dermatologists. In this paper, we add up a particular approach for the detection of cancer the usage of Naïve Bayes classification. In this ABCD parameters are extracted for the cancer detection. The approach became evaluated on ISIC dataset of 141 dermoscopic images. CAD could give dermatologists a nearer take a look at suspicious skin lesions. This can be beneficial to dermatologists to locate suspicious lesions in an early level. So, intend is to broaden laptop aided-diagnostic gadget which is non-invasive approach for the detection of melanoma at early stage.*

Keywords— Melanoma, Dermoscopy, Computer Aided Diagnosis (CAD), ABCD rule, Naïve Bayes

I. INTRODUCTION

Skin cancer is a most prime public health problem, with over 5 million newly identified cases in the United States every year [1]. Melanoma is one of the maximum deadly kinds of skin cancer, previously accountable for over nine,000 deaths a year in the United States by myself [2], and over 10,000 predicted deaths in 2016 [3]. As melanoma happens at the pores and skin floor, it is amenable to detection by means of simple visual exam. Indeed, maximum melanomas are first visible by using sufferers, now not physicians [4]. However, unaided visual inspection by using professional dermatologists is associated with a diagnostic accuracy of about 60%, meaning many capability curable melanomas are not detected until more advanced degrees [5]. To improve diagnostic performance and decrease melanoma deaths, dermoscopic has been brought, which is an imaging approach that eliminates the floor reflection of skin, allowing deeper layers to be visually more desirable. Previously most organizations have proposed pc aided diagnosis structures to become aware of melanomas in dermoscopic photos. These structures use numerous features, inclusive of color, form, and texture, to characterize the photos. These functions are given in scientific algorithms inclusive of the ABCD rule [6]. Computer Aided-Diagnosis (CAD) structures were evolved to assist dermatologists of their recurring exercise. Melanoma is a cancerous lesion inside the pigment-bearing basal layers of the epidermis and is the most deadly shape of pores and skin cancer, yet it is also the maximum treatable, with a therapy rate for early-stage melanoma of almost one hundred%. Therefore, there's a need to expand computer-aided diagnostic structures to facilitate the early detection of cancer.

The utility of photograph processing for diagnostic motive is a non-invasive technique. There is big interest inside the prospects of automated images analysis technique for photograph processing, both to provide quantitative records approximately a mole, which can be relevance for the medical, and as a early caution device. In order to obtain an effective way to discover pores and skin cancer at an early level without appearing any unnecessary pores and skin biopsies, virtual images of melanoma skin lesions were investigated. The dermoscopy is a non-invasive method for popularity of pores and skin surface structures, served each for magnification (as much as 400 instances) and evaluation of images of coetaneous tissue. The dermatoscope in discern a suggests the only form, consisting of a convex lens with angulations of 20° and surrounded via various LED of polarized mild, utilized by dermatologists with a view to view the image manually with value of about four hundred times. Figure b indicates a dermoscopic photograph which consists of a virtual digital camera with a dermatoscope connected to it, used to take pix at high magnitude. A digital dermoscopy machine which includes a dermoscopic photograph coupled to a computer, which lets in one to experiment images of skin lesions and also manipulate them by using the use of image processing software program. The foremost causes of Melanoma include sunburn, elevated numbers of uncommon moles; depressed immune structures; a own family history (in 10%, some having mutations in genes); truthful pores and skin and a preceding cancer history, girl has better survival rate as compared to male.

The paper is prepared as follows. In Section II, literature survey is supplied. In Section III A, we describe the proposed methodology, in Section III B; we supply the implementation of proposed technique. Section IV represents result and Section V gives conclusion.

II. LITERATURE SURVEY

Celebi et al. summarized all of the studies in this field within the beyond 30 years and supplied future steering for scientific image evaluation. Most studies in this vein revolves round analysis of pores and skin lesion photos taken the usage of dermatoscope (dermoscopic photographs) and falls below 3 extraordinary classes: mathematical modeling based on certain capabilities of the lesion, fuzzy-logic primarily based structures, and neural network primarily based structures [7].

Ercal et al. described a basic neural network classifier that extracts the asymmetry, border irregularity and coloration capabilities of a photo and fed them to a feed ahead neural community. However, because of the constrained number of capabilities extracted, the device may want to most effective acquire between 70-eighty% type accuracy [8].

Gniadecka, et al. Proposed a method that used Raman spectroscopy and neural networks for detecting pores and skin cancers. They focused a laser beam on the skin lesion to excite the molecules in the lesion. The

scattering impact of the molecules within the pores and skin lesion reasons frequency shifts inside the meditated Raman spectra. They educated a neural network with the meditated beam's frequency traits, and had been able to get right sensitivities. However, Raman spectrometers are not broadly to be had and are very pricey, and as a result are not often used by dermatologists. Although an awful lot paintings has been achieved inside the area of neural community based class of dermoscopic pics, there may be but to be a classifier this is accurate, sensible, and wellknown enough to have a real-world effect [9].

Jaleel and Saleem defined a neural network based totally classifier that did no longer use any of the ABCD features but relied on capabilities extracted from the two-D wavelet transformation of the photos. The pattern size used in their type changed into small (less than 21 snap shots) and there was no mention of the overall performance or sensitivity finished via the gadget. Smaller training and trying out pattern sizes typically result in over-becoming, wherein the learning machine has a tendency to alter to precise random quirks of the training facts that cannot be generalized to large samples [10].

Stanley et al. proposed a fuzzy common sense based totally coloration histogram analysis technique for skin lesion dedication. However, vast color changes in melanoma pores and skin lesions arise most effective in advanced tiers. Depending absolutely on the coloration histogram alone will now not assist in early detection. Fuzzy class strategies also have the tendency to over-match due to the absence of gaining knowledge of. Since fuzzy common sense uses extra superior strategies to discover lesions, it's miles definitely superior to a easy components. However, not like a machine mastering based device, the accuracy of the gadget does not enhance after the initial device parameters are chosen [11].

Stoeker et al. proposed an automatic classifier that quantified sure capabilities of the lesions and applied it to a formulation. If the end result of the formulation is above a sure threshold, the lesion becomes classified as malignant. Otherwise, the lesion turned into labeled as benign. Although this formulaic technique turned into capable of reap a sensitivity of above 80%, this machine had no way of learning from enjoy with new lesions and for that reason is inferior to even a popular visible inspection [12].

Fatima.R. Et al. Introduces a multi-parameter extraction and type system to resource an early detection skin most cancers cancer [13].

Fasihi, et al. Describes utilization of morphologic operators for segmenting a image and wavelet analysis to extract the function which ends up in to better melanoma analysis device [14].

Alcon, J.F. Et.Al.Has used pigmented skin lesion snap shots, captured the use of digital digital camera for automated cancer analysis with an accuracy of eighty five%, sensitivity of ninety three% and specificity of 69% [15].

Patwardhan Jain, Y. K. Specializes in the development a skin cancer screening device that can be used by non-experts to classify ordinary from abnormal cases, using feature calculation and class strategies. Here functions are extracted the use of wavelet remodel in which because the classification is completed using synthetic neural networks [16].

S. V. Dhawan et al. Makes use of wavelet transformation based totally pores and skin lesion pix type device which makes use of a semantic illustration of spatial frequency facts incorporates within the skin lesion images [17].

According to the literature, image processing is important for diagnostic parameter for melanoma. Some literature defines significance of machine learning algorithms in cancer detection. Most of the methods are sturdy, reliable, computer-aided diagnostic device for studying the texture in lesions of the pores and skin. Different strategies use one of kind methods for extraction of characteristic set as well as very last class in terms of gift or absent of the cancer in the image.

III. METHODOLOGY

The method used for Melanoma Skin Cancer Detection the usage of Image Processing is as proven in Fig. 1. The entered image of the skin lesion from ISIC 2016 dataset. MATLAB 2016a Software is used for test. Image processing techniques and Naïve Bayes classification algorithm is used for implementation. The input image is then pre-processed to decorate the photograph satisfactory. The binary masks are used for image segmentation. The RGB coloration space is used. The segmented lesion-image is given to the feature extraction block which consists of lesion vicinity analysis for its geometrical capability and ABCD capabilities. The extracted features are provided to the next step i.e. classification stage which classifies the skin lesion as cancerous or non-cancerous by the usage of Naïve Bayes.

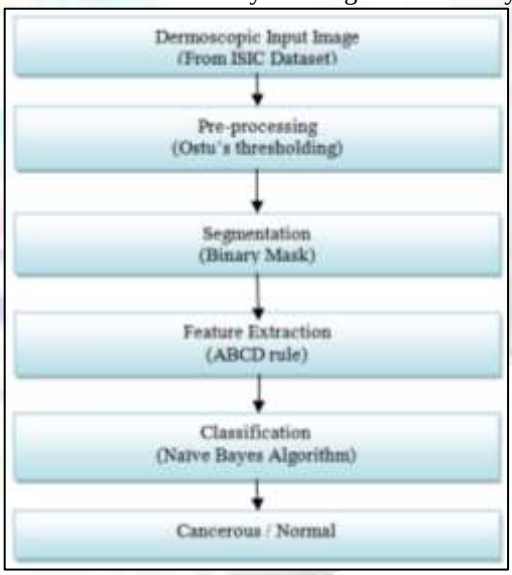


Fig. 1: Methodology

The ABCD rule of dermoscopy was the first method of the melanoma used in dermoscopy. This model was described by Stolz and colleagues in 1994. For the calculation of ABCD score the criteria of asymmetry (A), abrupt cut-off of the pigment pattern at the border (B), different colors (C), and different structural components (D) are assessed to yield a semi-quantitative score. The ABCD rule has been proved to be useful in:

1) Early detection of malignant melanoma
2) Discrimination between benign and malignant lesions
3) Selection of lesions for excision in patients with numerous typically appearing nevi

4) Monitoring of certain malignant nevi that for cosmetic or patient preference were not removed

IV. METHODOLOGY IMPLEMENTATION

Depending upon the implementation of the ABCD rule different degrees of accuracy can be expected, for example there exists a variety of algorithms in order to find the axis of symmetry of images with different relative precision and robustness. The same applies to Border/Edge detection schemes and to determination of abruptness of pigment change at the border of a lesion. Color and Diameter parameters can be calculated with less uncertainty. Our first approach is to use the simplest algorithms available that nonetheless produce fairly accurate results when compared to human visual assessment. The steps that were programmed are those regarding the pre-processing of the image (steps 1-5), with the main calculations following (steps 6-10):

1) Input of the image in RGB color format.
2) Transformation of the image in Gray level format.
3) Image histogram equalization to increase contrast.
4) Converting image into black & white binary format using a fixed manually introduced threshold.
5) Edge & internal pattern detection using Prewitt method with fixed manually introduced threshold.
6) Lesion bisection into two (orthogonal) axes–axes determination by moment's method, principal components analysis, gradient & orientation histogram, etc. or some very simple sampling technique.
7) Asymmetry scores a calculation by image overlapping.
8) Division of lesion in 8 segments and determination of pigment change in each segment for calculating the B score (0-8 points).
9) Determination of presence of 6 basic colors inside the lesion (white, red, light brown, dark brown, blue-gray, black) giving 1 point to each for the calculation of the C score.
10) Calculation of the lesion diameter based upon the edge of the lesion found on step 5, and determination of the D score accordingly.
11) Then we have to train and test dataset using Naïve Bayes algorithm and classify the image as cancerous or normal.

Especially steps 6 & 8 can be implemented using from very sophisticated up to very simple algorithms and this is expected to greatly influence the future development effort. More specifically, at this point the focus is on step 9 determination of presence of 6 basic colors inside the lesion (white, red, light brown, dark brown, blue-gray, black) giving 1 point to each for the calculation of the C score. Step 11 is essential to classify the image.

V. RESULT

In this work, we have implemented ABCD rule and feature made simplified look at on pores and skin most cancers. We chose ABCD rule because it is quite simple and fast algorithm, useful for the effective and automatic detection of melanoma. We have evolved a code for ABCD rule in MATLAB 2016. We have applied pc aided image processing techniques. The consequences obtained have validated the capacity to classify cancerous and non-cancerous lesions. This computer aided detection has verified to be quick, spontaneous and cost powerful. The purpose of the challenge is to help studies and development of algorithms for computerized analysis of melanoma, a deadly shape of skin cancer, from dermoscopic pics. 141 images of melanoma are taken from ISIC 2016 database. The International Skin Imaging Collaboration (ISIC) is a global attempt to enhance cancer prognosis, which has these days started efforts to combination a publicly handy dataset of dermoscopic photos. The photos are screened for each privacy and nice warranty. The overarching purpose of this assignment was to provide a picture from the ISIC Archive to support improvement of automated cancer prognosis algorithms from dermoscopic images.

A. Step-1: Input Image from ISIC Database

Fig. 2: Original image: (a) Original image of true positive (melanoma) (b) Original image of true (mole) negative

B. Step-2: Creating Binary Mask

First resizing of input image to (256 X 256) length. Then put into effect in-built command imfilter with the multidimensional filter out of 25 X 35 size to alter depth of image. Then convert photo into grayscale. Then using Otsu's thresholding technique convert grayscale photo into binary picture the Otsu's method, which chooses the edge to decrease interclass variance of the black and white pixels.

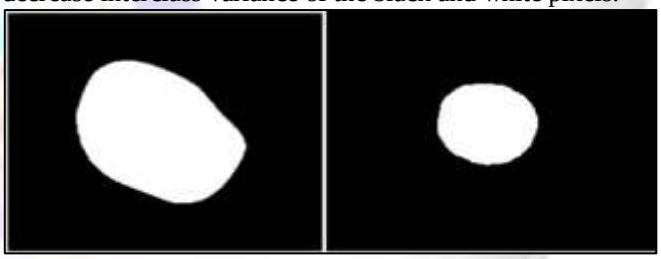
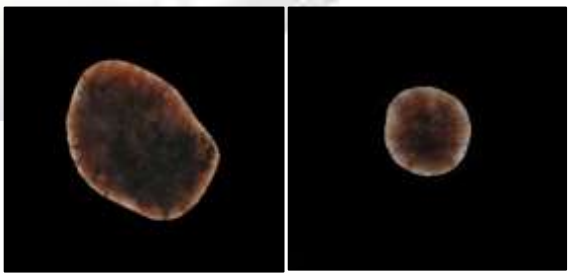


Fig. 3: (a) Binary mask of true positive sample (melanoma) using RGB (b) Binary mask of true negative sample (mole) using RGB

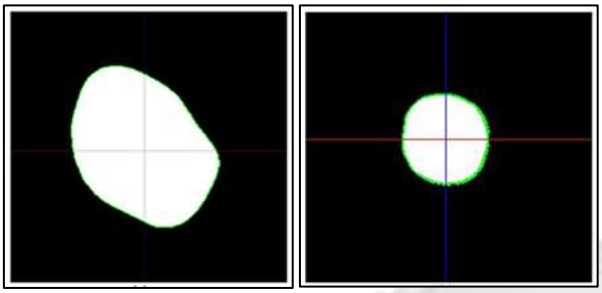
C. Step-3: Segmentation of lesion from original image using binary mask



(a) (b)

Fig. 4: (a) Lesion segmentation from true positive (melanoma) (b) Lesion segmentation from true negative (mole)

D. Step-4: Calculation of geometric center of segmented lesion



(a) (b)

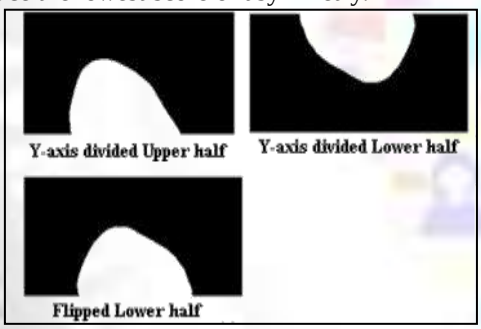
Fig. 5: (a) geometric center of segmented lesion from true positive (melanoma) (b) Geometric center of segmented lesion from true negative (mole)

E. Step-5: Images Feature Extraction

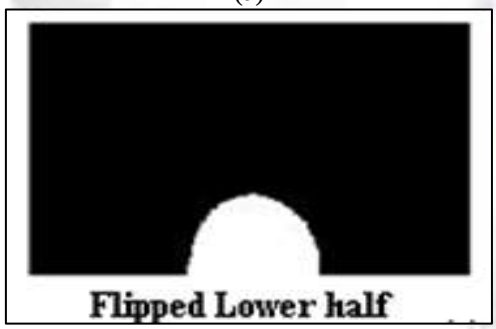
This step targets to extract traits of the resulting photos segmentation through algorithms that allow extracting the records precisely. After the segmentation, 4 parameters have been extracted which belongs to the ABCD rule, they may be: asymmetry, border, coloration and lesion diameter. These parameters had been calculated for dermoscopic photographs of skin lesions.

1) Asymmetry Parameter:

Asymmetry: asymmetry of the calculation was based on the method proposed by Stolz in 1994, where the image of the lesion is divided into two segments perpendicular to each other (horizontal and vertical) and, positioned on the lesion to produce the lowest score of asymmetry.



(a)



(b)

Fig.6 (a) Y axis divided of binary image for true positive (melanoma) (b)Y axis divided of binary image for true negative sample (mole)

The calculation of asymmetry gives an analysis of the injury at a geometry level. So it become proposed a place take a look at for analysis of the lesions, wherein the quantity of black pixels (which corresponds to the vicinity of the lesion) belonging to every axis (on each the superb and the terrible), are counted. Then the distinction in the quantity of pixels on both sides is calculated. A variation of ± zero.25% is taken into consideration inside the distinction between the correspondent pixels, considering in the segmentation manner, there is a opportunity to be counted false pixels within the border. If each axes are asymmetrical, its rating is two. If the asymmetry is in best one of the axes its rating is 1, and if there may be no asymmetry its rating is 0. Most melanomas have a rating of 2, in place of benign nevi and melanocytes nevi.

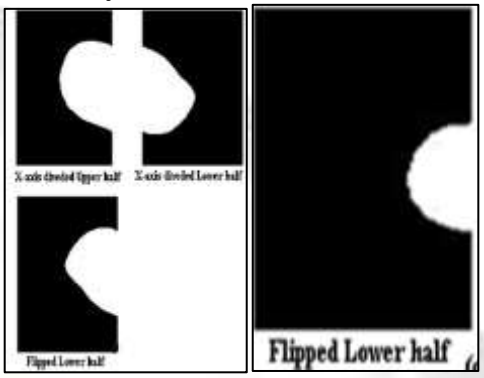


Fig. 7: (a) X axis divided of binary image for true positive (melanoma) (b) X axis divided of binary image for true negative (mole)

2) Border:

The lesion is divided into 8 equal parts, as shown in Figure 8(a)(b). The score is found by the analysis of eight parts, i.e., the amount of black pixels belonging to each slice is counted and this amount was compared to amount obtained in other slices. A variation of ±20% was considered. For each slice presenting a difference greater than the established here 1 point is summed to the total score.

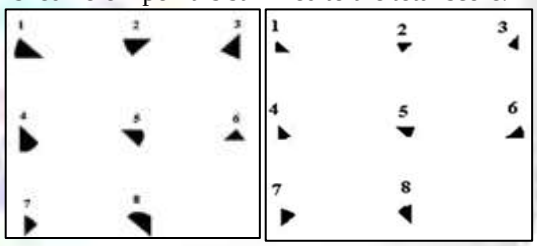


Fig. 8: (a) Lesion is divided into eight equal parts of binary image for true positive (melanoma) (b) Lesion is divided into eight equal parts of binary image for true negative sample (mole)

3) Color:

In the ABCD rule the observed colors indicating malignancy are black, white, blue-gray, light brown, dark brown and red. A range of possible colors for each of these six tones was used as a matter of evaluation in this test, as presented in tableI. The colors observed were only counted if it occupies an area of 0.1% of the total lesion area, since the variation is estimated as tolerance factor for each channel R, G and B. This variation was ± 10 color levels for each channel. The presences of each shade of gray add more value in a score where the minimum is 1 and the maximum is 0.

Sr. No.	Color	RGB	Rgb
01	White	255,255,255	1.0,1.0,1.0
02	Black	0,0,0	0.0,0.0,0.0
03	Red	255,0,0	1.0,0.0,0.0
04	Light-Brown	205,133,63	0.80,0.52,0.25

05	Dark-Brown	101,67,33	0.40,0.26,0.13
06	Blue-Gray	0,134,139	0.0,0.52,0.54

Table 1: RGB description of ABCD colors

4) Diameter:

Values above 6 mm are alarms of malignancy, and the score is based on this length. To obtain this parameter, the amount of pixels in the width and height of the lesion were determined, therefore obtaining the average amount which is converted in millimeters. For each 1mm step 1 point is added to the score of the diameter.

(a) (b)

Fig. 9: (a) The result using naïve bayes for true positive (melanoma) sample using RGB (b) The result using naïve bayes for true negative sample (mole) sample using RGB

F. Evaluation Matrix

To quantitatively compare the borders drawn by dermatologists with the computer derived borders, different metrics have been utilized. The three statistical measurements of sensitivity, specificity and accuracy are applied.

$$\text{Sensitivity} = TP/(TP + FN),$$
$$\text{Specificity} = TN/(TN + FP),$$
$$\text{Accuracy} = (TP + TN)/(TP + FP + FN + TN)$$

Where,

TP = true positive, TN = true negative,

FP = false positive, and FN = false negative, respectively.

Sr. No.	Parameter	Naïve Bayes
01	Sensitivity(SE)	88.63%
02	Specificity(SP)	93.50%
03	Accuracy	91.73%

Table 2: Classification Result using Naïve Bayes

We got 91.73% accuracy using Naïve Bayes classification algorithm.

VI. Conclusion

It has been tested on software MATLAB 2016a is a powerful platform for the early detection and analysis of malignant cancer through the usage of photo processing strategies and device algorithms. Only some strains of compactly written code are sufficient in an effort to application the desired algorithms. We have carried out ABCD rule & pick out ABCD rule because, it is quite simple and efficient set of rules, purposeful for computerized detection of cancer. We have developed a code for ABCD rule in MATLAB 2016a. Within of the DIP, the work goal became covered and finished namely: the pre-processing, the segmentation and the characteristic extraction (asymmetry, border, color and diameter) from skin lesions dermoscopic photographs. The outcomes received have demonstrated the capacity to categorize cancerous and non-cancerous lesions. This pc aided detection has verified to be quick, spontaneous and cost powerful. In this paper, we carried out Naïve Bayes machine learning algorithm is used for type of mole as cancerous or normal

The result obtained using a ISIC 2016 dataset. The performance of the system gives Sensitivity (SE) = 88.63% and Specificity (SP) = 93.50% and Accuracy 91.73%. Color features derived from other color space RGB.

Acknowledgement

The dataset used in this paper is acquired from ISIC which is publically available dataset. Authors would like to thank Dr. Ashish R. Deshmukh, Professor and Head of Dermatology Department, MGM College and Hospital, Aurangabad(MS) for his selfless support. We would also like to thank Department of Computer Science and Information Technology and SAP Lab, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad(MS) for providing the infrastructure and necessary support for carrying out the research.

References

[1] Rogers HW, Weinstock MA, Feldman SR, Coldiron BM.: "Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population", 2012. JAMA Dermatol 2015; 151(10):1081-1086.

[2] Cancer Facts & Figures 2014, website American Cancer Society, 2014.

[3] Siegel, R.L., Miller, K.D., and Jemal, A.: "Cancer statistics, 2016," CA: A Cancer Journal for Clinicians, volume 66, number 1, pp. 7-30, 2016.

[4] Brady, M.S., Oliveria, S.A., Christos, P.J., Berwick, M., Coit, D.G., Katz, J., Halpern, A.C.: "Patterns of detection in patients with cutaneous melanoma." Cancer. 2000 Jul 15;89(2):342-7.

[5] Kittler, H., Pehamberger, H., Wolff, K., Binder, M.: "Diagnostic accuracy of dermoscopy". In: The Lancet Oncology. vol. 3, number 3, pp. 159.165. 2002.

[6] W.Stolz,A.Riemann,and A.B.Cognetta, "ABCD rule of dermatoscopy: A new practical method for early recognition of malignant melanoma," Eur. J. Dermatol., volume. 4, pp. 521–527, 1994.

[7] Celebi, M. E., et. al "Advances in skin cancer image analysis", Computerized Medical Imaging and Graphics, Volume. 35, Number 2, pp. 83-84, 2011.

[8] Ercal, Fikret, et al. "Neural network diagnosis of malignant melanoma from color images", IEEE Transactions on Biolmedical Engineering, Vol.41, Number 9, pp. 837-845, 1994.

[9] Gniadecka, M, et al. "Melanoma diagnosis by raman spectroscopy and neural networks: structure alterations in proteins and lipids in intact cancer issue", J Invest Dermatology, Vol. 122, Number 2, pp. 443-449, 2004.

[10] Jaleel, A, Sibi Salim and Ashwin R B. "Artificial neural network based detection of skin cancer", Internal journal of advanced reseach in electrical, electronic and instrumentation engineering, Vol. 1, Issue. 3, pp. 200-205, 2012.

[11] Stanley, R. Joe, Randy Hays Moss and Chetna Aggarwal. "A fuzzy based hostogram analysis technique for skin lesion descrimination in dermatology

clinical images", Computerized Medical Imaging and Graphics : the Official Journal of the Computerized Medical Imaging Society, Vol. 27, Number 5, pp. 387-396, 2003.

[12] W V Stoecker, W. W. Lee, and R.H Moss. "Automatic detection of asymmetry in skin tumors", Computerized Medical Imaging and Graphics, Volume. 16, Issue.3, pp. 191-197, 1992.

[13] Fatima R., Khan, Mohd. Z. A., Govardhan, A and Kashyap, D. "Computer Aided Multi-Parameter Extraction System to Aid Early Detection of Skin Cancer Melanoma," International Journal of Computer Science and Network Security, Volome.12, Number10, pp.74-86, Oct.2012.

[14] Fassihi, N., Shanbehzadeh, J., Sarafzadeh, A., and Ghasemi, E., "Melanoma Diagnosis by the Use of Wavelet Analysis based on Morphological Operators," Proc. of International Multi Conference of Engineers and Computer Scientists 2011, Vol. I, Hong Kong, pp. 1-4, March 16-18, 2011.

[15] Alcon, J. F., Ciuhu, C., Kate, W. ten, Heinrich, A., Uzunbajakava, N., Krekels, G., Siem, D. and Haan, G. de., "Automatic Imaging System With Decision Support for Inspection of Pigmented Skin Lesions and Melanoma Diagnosis," IEEE Journal of Selected Topics in Signal Processing, Vol. 3, Number 1, Feb. 2009, pp. 14-25.

[16] Patwardhan Jain, Y. K. and Jain, M., "Comparison between Different Classification Methods with Application to Skin Cancer," International Journal of Computer Applications, Vol. 53, Number11, pp. 18-24, Sept. 2012.

[17] S. V., Dhawan, A. P. and Relue, P. A., "Classification of melanoma using tree structured wavelet Transforms," Computer Methods and Programs in Biomedicine, Vol. 72, pp. 223-239, 2003