International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 10, October 2017
www.ijarse.com
IJARSE
ISSN 2319 - 8354

An Enhanced Classification Algorithm for Breast Cancer Density using Feature Extraction and Segmentation Techniques

S. Manimekalai[1], R Suguna[2]

[1,2]Assistant Professor in Computer Science

Theivanai Ammal College for Women (Autonomous), Villupuram

ABSTRACT

This research paper focuses on early detection of breast cancer, which plays an important role in dropping the associated morbidity and humanity rates. This feature of image analysis can also be used in medical application for early analysis of any sickness using image processing. Now-a-days, cancer is a common diseases affected by all the age groups of people. Its early recognition and cure is very significant. It can be done by using feature extraction and segmentation methods. This technique is cost successful as it does not use any costly instruments and hence can be used by every person. It's also a time saving move toward. Currently great attention is there in the aspects of routine image analysis method for image processing, to supply quantitative information about an abrasion, which is applicable to the medical, as well as a tool for its early warning. The proposed method provides survey techniques to mechanically diagnose skin cancer by using various images of different risks. The proposed method used ABCD feature extraction method and OTSU segmentation method to identify the cancer cells in advance. In this method breast cancer is identified using affected skin color. The data sets are collected from BreakHis repository. So it is a better approach to detect the cancer at a near the beginning stage.

Keywords*: *Breast cancer, region features, density classification

I. INTRODUCTION

Breast cancer is the most general malignancy in women and cause many deaths. Early detection is the best solution in arranging to avoid mastectomy, reduce the probability to return, and decrease the rate of mortality. The mammography test allows to detect and characterize lesions in the breast, so it is important that women are aware of this disease and do auto exams frequently, and after of a specific age it is recommended doing regular mammography exams. There are many techniques for detecting breast lesions, like ultrasonography and magnetic resonance imaging, but mammography is the more common choice. Mammography is a very useful test; however, the similarity between normal and pathologic patterns may cause difficult diagnosis.

Breast Cancer affects one in eight women in their lifetime, a survey says. It is a dangerous tumor that begins in cells of the breast and gets into the surrounding tissues as well. Both men and women are affected by this disease, but the symptoms are more in women. United States cancer statistics showed that only in the United States the number of deaths caused by breast cancer crossed forty thousand in one year. The stats showed that this disease is very common in women and there is plenty of work being done and to be done in this field to get control over such a deadly disease. Medical research targeting breast cancer is not new and its roots go back into 16th century. Due to the lack of communication and advancement in medical field this disease kept on taking the edge on humans and still considered one of the most deadly diseases of all the times. Recent advancement in medical field and more precisely the involvement of information technology in the medical field introduces a new diagnostic mechanism called Medical Image Processing [1]. Medical Image Processing is not only limited to, cancer disease, instead it has helped greatly in the diagnosis of different kinds of diseases and it is evident through statistics. With the help of image processing techniques it has become easier to detect tumor from an infected breast and diagnose breast cancer. Early detection can help in proper diagnosis and treatment resulting in minimizing the risk of the most unwanted outcome of this disease.

Mammography: Mammography is a diagnostic tool for the examination of the human breast. These examinations are recorded as specialized images which are then observed by radiologists for any possible abnormality.

The paper is organized as follows: Section 2 discussed about Review of Literature in the field of Breast Cancer Identification. Proposed methodology is detailed in Section 3. Result is discussed in Section 4. Finally Conclusion is discussed in Section 5.

II. REVIEW OF LITERATURE

Breast cancer is the top cancer in women worldwide and is increasing particularly in developing countries where the majority of cases are diagnosed in late stages. There are many literature reviewed in this following,

Rajvi Parikh, Dr Hitesh shah 2013**,** In this paper, computer diagnostic tools enable objective judgments by making use of quantitative measures. The basic three steps are there to achieve the results i.e. 1) image processing 2) Feature extraction 3) Classification.

Steps

1. It deals with noise reduction artifacts removing,
2. It deals with extracting variety of information from the processes image for accurate detection and
3. It deals with results that say various types of skin lesions. In this paper we are showing the process of it and also discussed some clinical diagnosis methods which are being incorporated with the tool for detecting the type of lesion.

D. Brown, I. Craw, J. Lewthwaite 2001, The transformation of RGB to HSV is invariant to high intensity at white lights, ambient light and surface orientations relative to the light source and hence, can form a very good choice for skin detection methods.

Albiol, A., Torres, L., Delp, E, 2001, The conversion from RGB to HSI or HSV will take time and expensive. Moreover, if there is a lot of fluctuation in the value of the color information (hue and saturation), pixels with small and large intensities are not considered. In the case of YCbCr color space, transformation and efficient separation of color and intensity information is easy as compared to HSI or HSV. HSV based skin color detection and segmentation is elaborately described.

H C Chen et al, 2006, Image segmentation techniques can be differentiated into the following basic concepts: pixel oriented, contour-oriented, region-oriented, model-oriented, and color-oriented and hybrid. Color segmentation of image is a crucial operation in image analysis and in many computer vision, image interpolation, and pattern recognition system. The performance of color segmentation may significantly affect the quality of an image understanding system.

S. Gundimada, L. Tao, and V. Asari, 2004, Some colour spaces have their luminance component separated from the chromatic component, and they are known to possess higher discriminality between skin pixels and non-skin pixels over various illumination conditions. Skin colour models that operate only on chrominance subspaces such as the Cb- Cr.

III. PROPOSED METHODOLOGY

The segmentation is the most important stage for analyzing image properly since it affects the accuracy of the subsequent steps. However, proper segmentation is difficult because of the great varieties of the lesion shapes, sizes, and colors along with different skin types and textures. In addition, some lesions have irregular boundaries and in some cases there is a smooth transition between the lesion and the skin. To address this problem, several algorithms have been proposed. They can be broadly classified as thresholding, edge-based or region-based methods. In this thesis, three methods of segmentation have been discussed. The methods are:

- Otsu's method.
- Gradient Vector Flow (GVF)
- Color Based Image Segmentation Using K-mean Clustering.

In automated diagnosis of skin lesions, feature extraction is based on the so-called ABCD-rule of dermatoscopy [8],[9]. ABCD represents the asymmetry, border structure, color variation, and dermatoscopical structure so called diameter of the lesion and define the basis for a diagnosis by a dermatologist [7].

The proposed method used ABCD feature extraction method and OTSU segmentation method to identify the cancer cells in advance. In this method breast cancer is identified using affected skin color.

The process of proposed breast cancer detection method using affected cancel skin cells is shown in Fig.1.

Main features of the proposed scheme are as follows. First, the region skin color model and morphology processing in the segment skin area in the image is used.

Two channels of Cb and Cr are selected to calculate the range of skin color. Secondly, the proposed scheme uses an algorithm to decompose the skin area into two parts of hemoglobin and melanin. Thirdly, location histogram is used to measure the intensity of hemoglobin and melanin.

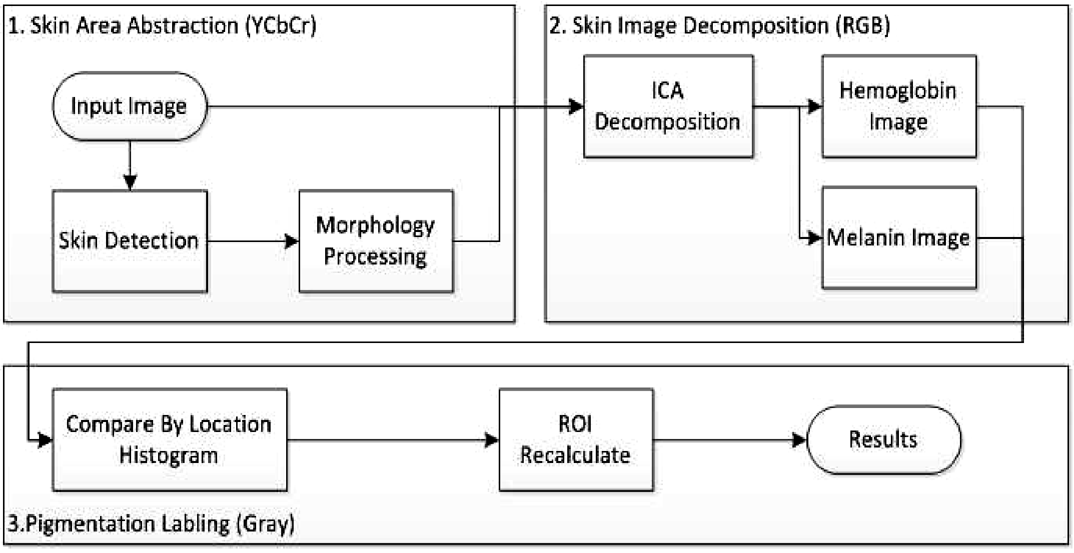


Fig. 1 Proposed method

IV. AFFECTED SKIN AREA SEGMENT

The illumination influences mainly on the skin color but the chromaticity little influences on it. For these reasons that skin colors are closely distributed on chromaticity in YCbCr space although they are different. Therefore, our scheme uses the Gaussian distribution to describe the skin color in the CbCr channels. To collect two hundred of skin pictures that contain hand, arm, and face in a number of illumination environments and analyze color distributions of them for the Gaussian skin color model[10].

Skin lesion imaging methods

The goal of any imaging methodology used in dermatology is to diagnose melanoma in early stages, because it depends an effectiveness of treatment. Investigations show, that early diagnosis is more than 90% curable and late is less than 50%. The diagnosis and successful treatment are often supplemented with permanent monitoring of suspicious skin lesions.

Doctor's diagnosis is reliable, but this procedure takes lots of time, efforts. These routines can be automated. It could save lots of doctor's time and could help to diagnose more accurately. Besides using computerized means there are good opportunity to store information with diagnostic information in order to use it for further investigations or creation of new methods of diagnosis.

ALGORITHM:

Step1: Start the process.

Step2: Get the input skin lesion image.

Step3: Create four connected graphs using horizontal and vertical pixels.

Step4: Calculate image slope to find edge pixels. Image gradient provides the information about pixel intensity variation in horizontal and vertical direction.

Step5: Produce pairs using four connectivity graphs.

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 10, October 2017
www.ijarse.com
IJARSE
ISSN 2319 - 8354

Step6: Group the pairs as regions based on pixel intensity similarity.

Step7: Combine regions based on merging predicate using the equation (1), (2) & (3).

$$P(R,R') = \left|\overline{R'a} - \overline{Ra}\right| \le \sqrt{b^2(R) + b^2(R')} \text{ ------(1)}$$

$$b(R,R') = \sqrt{b^2(R) + b^2(R')} \text{------- (2)}$$

$$b(R) = g\sqrt{1/2Q|R|\ln|K_{|R|}|/\delta} \text{ ------- (3)}$$

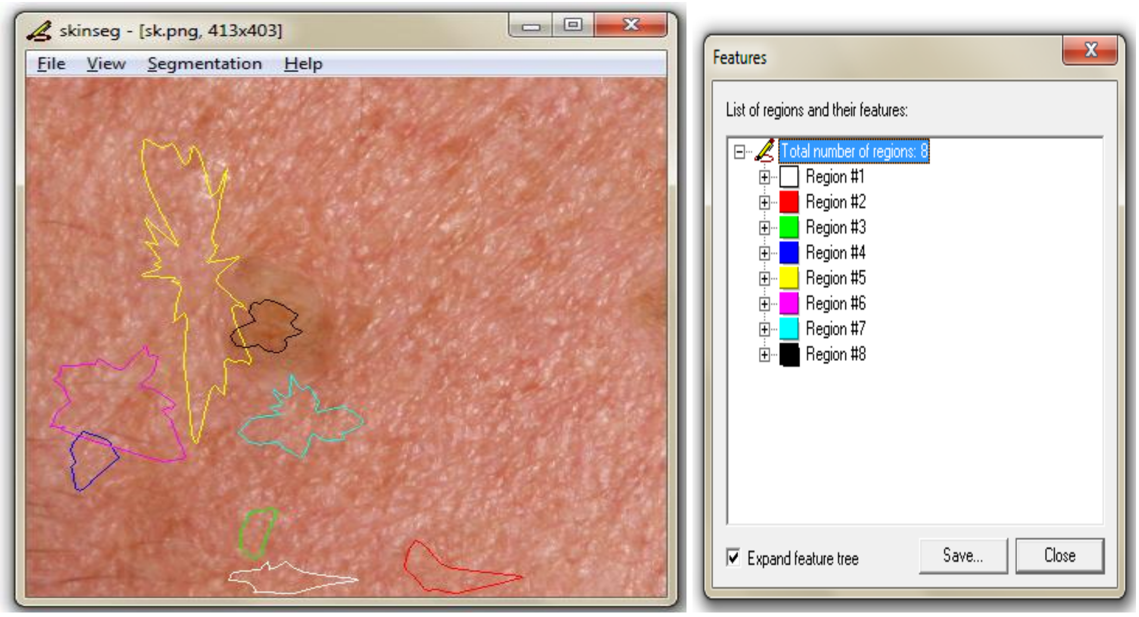
CbCr channel YCbCr is a common and important color space and it is a way of encoding RGB information. To obtain YCbCr color space from RGB color space, as follows

$$\begin{aligned} Y &= 0.229 \times R + 0.587 \times G + 0.114 \times B \\ Cr &= (R - Y) \times 0.713 + 128 \\ Cb &= (B - Y) \times 0.564 + 128 \end{aligned} \qquad (1)$$

Generally and absolutely, it can find that color spaces of RGB and HIS are not good for color clustering algorithm, as shown in Fig. 2, but Cb and Cr spaces are good for clustering. This scheme designs the Gaussian skin color model in YCbCr space. The different skin color has been mainly influenced by illumination, but a little influence from chromaticity.

V. RESULT AND DISCUSSION

In the following images of Fig.2 are presented all steps of this algorithm are in a Cancer, which are explained in the methodology and also an example of segmentation and extraction of pectoral muscle in a region.



IJARSE
ISSN 2319 - 8354

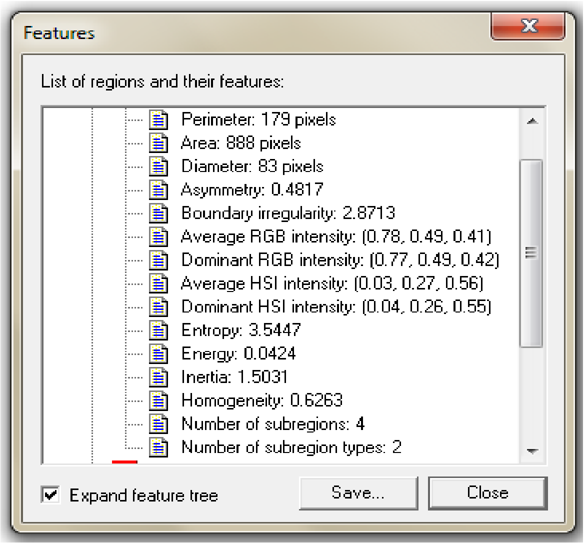


Fig 2. Segmented Cancer image

The last image shows that it eliminates necessary some artifacts that can interfere in the breast density classification. As used in the previous steps, the use of labels allows eliminate those eight regions and fifteen result presents the final result.

Density Estimation Results

All the characteristics used to classify breast density. All features extracted where normalized between -1 and 1, using the following equation:

$$\boxed{y = \frac{(ymax - ymin) * (x - xmin)}{(xmax - xmin) + ymin}}$$

Where *ymin* and *ymax* stand for the desired range for the new values and *xmin* and *xmax* stand for the smallest and largest feature value that is normalized, respectively.

A piecewise comparison is generated between the original image applied to the segmentation mask and the spatially varying threshold model for that particular image. If the intensity of the pixel at (x,y) in the masked original image is greater than the intensity value of the pixel at the same location in the spatially varying threshold model, then that pixel is classified as being radio dense tissue and will appear as white (‘1’). Otherwise, the pixel is classified as radiolucent tissue and will appear as black (‘0’) [11].

These images that are segmented between radio dense and radiolucent tissue are shown in Fig 3, the calculated percentages of radio dense tissue of all images for each of the spatially varying thresholds used along with a squared error between these results and the validated percentages of radio dense tissue from the Toronto method and illustrates this comparison.

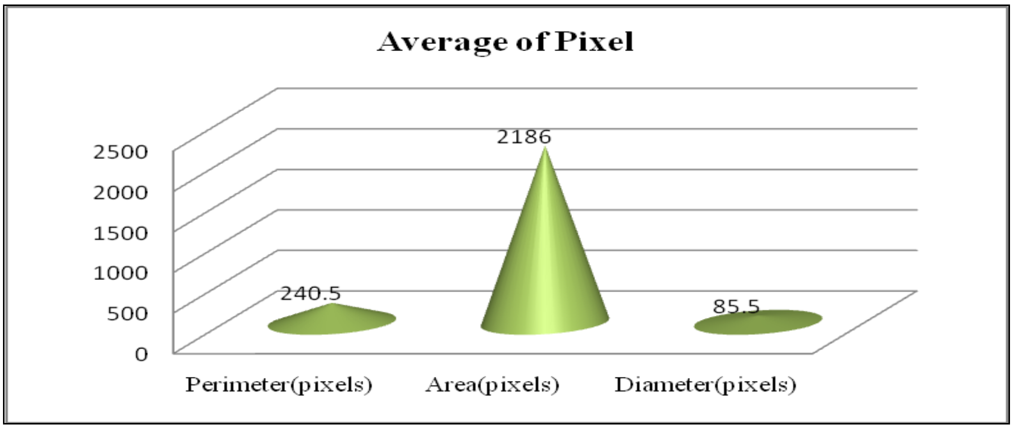


Fig 3. Result of pixel Density

VI. CONCLUSION

Early detection of breast cancer is of utmost importance, since only localized cancer is deemed to be treatable and curable, as opposed to metastasized cancer. Mammography is a widely used screening tool and is the gold standard for the early detection of breast cancer. However, many suspicious findings on mammograms are beginning. The most important mammographies signs of malignancy are masses and micro calcifications. Yet, the sensitivity of screening mammography is affected by the image quality and the radiologist's level of expertise

Moreover, since there is not a standard naming this method becomes an advantage in this aspect. Given that the images of MIAS database were considered unsuitable to be used as a test algorithm and even then achieved an accuracy rate of 85% for the distinction between left and right breast may be considered that the result was very positive. To emphasize the fact that the method of distinguishing between left and right breast is well developed it was obtained a accuracy rate of 96% for BreakHis repository.

REFERENCES

[1]. Mussarat Yasmin, Muhammad Sharif and Sajjad MohsinSurvey Paper on Diagnosis of Breast Cancer Using Image Processing Techniques, Research Journal of Recent Sciences,ISSN 2277-2502,Vol. 2(10), 88-98, October (2013).

[2]. Rajiv Parikh, A Survey on Computer Vision Based Diagnosis for Rkin Lesion Detection, International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 2, Issue 2, March 2013

[3]. D. Brown, I. Craw, J. Lewthwaite, A SOM based approach to skin detection with application in real time systems, BMVC01, 2001.

[4]. Albiol, A., Torres, L., Delp, E., Optimum color spaces for skin detection, In: Proceedings of the International Conference on Image Processing (ICIP). (2001) I: 122-124

[5]. H C Chen et al, Visible color difference-based quantitative evaluation of color segmentation. IEEE Proceeding, Vis image signal process vol.153 No.5 Oct 2006 pp 598-609.

[6]. S. Gundimada, L. Tao, and V. Asari, Face Detection Technique based on Intensity and Skin Color Distribution, *ICIP2004,* pp. 1413-1416, 2004.

[7]. Md.Amran Hossen Bhuiyan, Ibrahim Azad, Md.Kamal Uddin, "Image Processing for Skin Cancer Features Extraction", International Journal of Scientific & Engineering Research Volume 4, Issue 2, February-2013 1, ISSN 2229-5518.

[8]. F. Nachbar, W. Stolz, T. Merkle, A.B. Cognetta, T. Vogt, M. Landthaler, P.Bilck, O. Braun-Falco, and .Plewig, " The ABCD rule of dermatoscopy: High Prospective value in the diagnosis of doubtful melanocytic skin lesion," J.Amer.Acas.Dermatol., vol. 30, no.4,pp.551-559, Apr.1994.

[9]. R.J Friedman and D.S Riegel, "The clinical features of malignant melanoma," Dermatologic Clin., vol. 3, pp. 297-307, July 1982.

[10]. Hani K Al-Mohair, Abd Rahman bin Ramli, Elsadig A M. Shaiful J. Hashim, "Skin Detection in Luminance Images using Threshold Technique" Article. January 2007.

[11]. Richard Edson Eckert III, Spatially varying threshold models for the automated segmentation of radiodense tissue in digitized mammograms, Rowan University,Glassboro, New Jersey,2003.

[12]. Nafiza Saidin, Harsa Amylia Mat Sakim , Umi Kalthum Ngah and Ibrahim Lutfi Shuaib, Segmentation of Breast Regions in Mammogram Based on Density: A Review, IJCSI , Volume 6, Issue 4, July 2012.

[13]. Velkoff, V. A. and Adlakha, A., Womens' health in India. US Department of Commerce, Economics and Statistics Administration, Bureau of the Census, 1998.

[14]. Kushwah, V., The health status of women in India. Res. J. Chem. Environ. Sci., 2013, 1(3), 66–69.

[15]. Sridevi, T. R., Research evaluation of Indian Journal of Cancer: a bibliometric study. Res. J. Lib. Sci., 2014, 2(2), 1–5.

[16]. Subramanyam, K., Bibliometric studies of research collaboration: a review. J. Inf. Sci., 1983, 6, 33–38.

[17]. Parta, S. K. and Bhattacharya, P., Bibliometric study of cancer research in India. DESIDOC Bull. Inf. Technol., 2005, 25(2), 11– 18.

[18]. Bradford, S. C., Documentation, Public Affairs Press, Washington DC, 1950, pp. 106–121.

[19]. Egghe, L., A note on different Bradford multipliers. J. Am. Soc. Inf. Sci. Technol., 1990, 41(3), 204–209.

[20]. Egghe, L. and Rousseau, R., Introduction to Informetrics: Quantitative Methods in Library, Documentation and Information Science, Elsevier Science Publishers, 1990, p. 343.

[21]. Ortiz, A. P., Calo, W. A., Suárez-Balseiro, C., Maura-Sardo, M. and Suárez, E., Bibliometrsic assessment of cancer research in Puerto Rico, Salud. Publ., 2009, 25(4), 353–361.