STATISTICAL DETECTION OF BREAST CANCER BY MAMMOGRAM IMAGE

RAJINI GK[1]*, SHAIK NASEERA[2], SARAVANAN M[1]*

[1]Department of Instrumentation, School of Electrical Engineering, VIT University, Tamil Nadu, India. [2]Department of Software Systems, School of Computing Science and Engineering, VIT University, Tamil Nadu, India. Email: rajini.gk@vit.ac.in/m.saravanan2014@vit.ac.in

ABSTRACT

Objective: To create awareness about the breast cancer which has become one of the most common diseases among women that leads to death if not recognized at early stage.

Methods: The technique of acquiring breast image is called mammography and is a diagnostic and screening tool to detect cancer. A cascade algorithm based on these statistical parameters is implemented on these mammogram images to segregate normal, benign, and malignant diseases.

Results: Statistical features - such as mean, median, standard deviation, perimeter, and skewness - were extracted from mammogram images to describe their intensity and nature of distribution using ImageJ.

Conclusion: A noninvasive technique which includes statistical features to determine and classify normal, benign, and malignant images are identified.

Keywords: Breast cancer, Benign, Malignant, Mammogram image, ImageJ.

INTRODUCTION

The statistical data reveal that the death due to cancer is 21.5/100,000 women a year. These rates are age-adjusted and based on 2009-2013 cases are categorized based absolute risk, lifetime risk, and relative risk. The first symptom that is identified for breast cancer is typically a lump that feels different from the rest of the breast tissue. The earliest breast cancers are detected by a mammogram lumps found in lymph nodes located in the armpits which indicate breast cancer. Indications of breast cancer may include a lump in the breast, a change in breast shape, dimpling of the skin, fluid coming from the nipple, or a red scaly patch of skin [1]. To efficiently treat or inhibit this disease, an automatic detection based on mammogram image monitoring is significantly required.

Review shows that breast cancer can be detected through DNA damage recognition and chromatin remodeling [2]. Indications of breast cancer other than a lump may include thickening different from the other breast tissue, one breast becoming larger or smaller, a nipple changing position or shape or becoming inverted, skin puckering or dimpling, a rash on or around a nipple, discharge from nipples, constant pain in part of the breast or armpit, and swelling beneath the armpit or around the collarbone. Pain is an unreliable tool in determining the presence or absence of breast cancer but may be indicative of other breast health issues. The performance of nana medicine in treating cancer has been focused and emphasized [3].

System implementation

The mammogram image data for normal, cancer and benign are obtained from digital database for screening mammography (DDSM) database. These images are processed with the following few steps and are shown in Fig. 1. The features extraction is done for a mammogram image such as the mean, variance, standard deviation, and skewness are obtained. These parameters are used to detect cancer-related issue subsequently, classification of cancer-related images are done based on the range of prescribed values of the parameters.

Preprocessing

The following procedure is carried out to implement the preprocessing for the images such as 8 bit (grayscale), 16 bit (grayscale), 32 bit-float (grayscale), and RGB color images. If the images do not pertain to the above-mentioned formats, it is converted to specific formats using ImageJ plug in tool [4]. The following few steps are implemented as mentioned below.

Image sharpening

Image sharpening has been done to reduce the influence of noise present in an image, and some of the sharpening techniques are high boost filtering and high-frequency emphasis is implemented [5]. This procedure is carried out to increase contrast and to highlight edges and fine details present in an image.

Enhance contrast

This process changes the image intensity value distribution to cover a wide range of gray levels. The following steps are used to improve low contrast images using either normalization and histogram stretching or histogram equalization.

Normalization

Normalization is the process that automatically modifies the image that recalculates the pixel values of an image, so the range is equal to the maximum range for the data type, or 0-1.0 for float representation of images. The maximum range is 0-255 for 8-bit images and 0-65535 for 16-bit images.

Equalize histogram

Histogram equalization attempts to achieve a flat histogram - so all possible pixel intensities are equally represented in the image as mentioned. This technique is based on the histogram and statistics obtained from the neighborhood around each pixel [6] as shown in Fig. 2.

Subtract background

Background subtraction is a popular method to detect an object as a foreground by segmenting and it attempts to provide more valuable information for further processing. This method is used to subtract background image from cancer-related tissues in benign and cancer images as shown in Fig. 3.

Feature extraction

Feature extraction aims to detect useful features that help to identify important data and to eliminate irrelevant details present in an image.
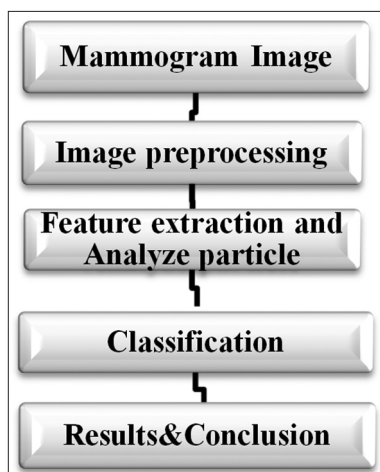
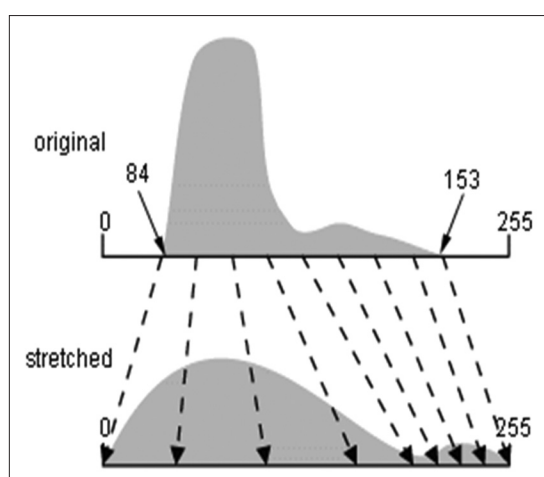Fig. 1: Flow chart of the algorithm

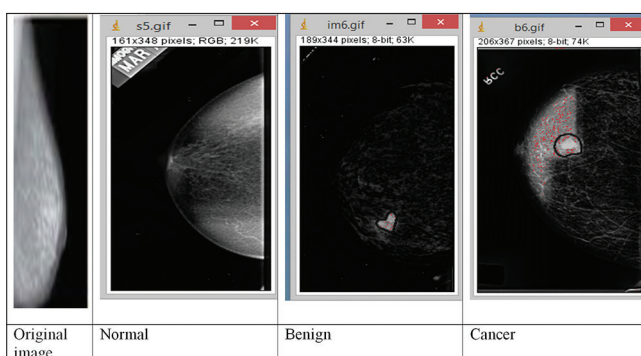Fig. 2: Represents the histogram equalization

Fig. 3: Subtract background images

Fig. 4: Edge detected images
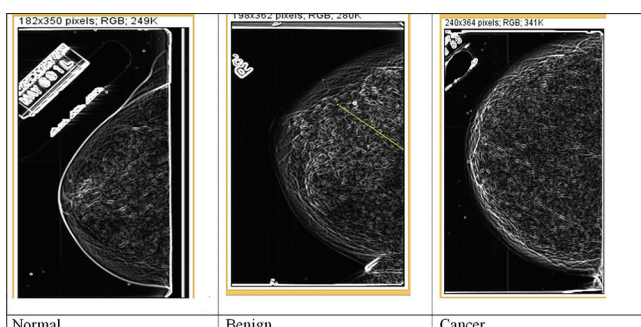
Table 1: Represents mathematical expressions for statistical parameters

Features	Expression
Mean	$\bar{X} = \dfrac{\sum x}{n}$
Median	$M = x_{\frac{(n-1)}{2}}$ for n odd $M = \dfrac{1}{2}\left[x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right]$ for n even
Perimeter	The length of the outside boundary of the selection
Standard deviation	$SD = \sqrt{\dfrac{\sum (x-\bar{x})^2}{n}}$
Skewness	The third order moment about the mean

Table 2: Statistical data for normal, benign, and cancer images

Features	Normal	Benign	Cancer
Mean	22-80	65-79	81-160
Median	25-80	95	111-150
Perimeter	55-140	112	143-225
Standard deviation	00-07	6.5-7.9	08-15
Skewness	−0.3-0.3	0.4	2.1-11

The features include various parameters such as areas average values topological functions, edges, and statistical parameters [7].

Find edges

Edge in an image represents a boundary across which the intensity level of pixel changes precipitously. Most of the shape information of an image is enclosed in edges [8]. Hence, first we detect these edges in an image and using image filters and then by enhancing those areas of image which contains edges as shown in Fig. 4.

The statistical parameters namely mean, median, perimeter, standard deviation, and skewness are obtained from mammogram images, and their mathematical expressions are tabulated in Table 1.

RESULTS

The results were arrived for the database DDSM using mammogram images [9]. The extracted features for normal and benign and cancer are compared and tabulated in Table 2. From the table, it is evident that the statistical parameters for normal mammogram are lower than the cancer mammogram images, and the features are moderate for the benign and higher range of these parameters are observed in cancer images.

CONCLUSION

Automatic detection of mammogram image is very essential in medical practice which leads to early detection of cancer disorder and might aid to contribute to reduced mortality. In this study, the use of image processing techniques for extraction of features from the mammogram image has been presented using ImageJ. The main advantage of this study is that by acquiring statistical parameters of a mammogram image, the morphological differences between several types of cancer-related issues are emphasized, and the extracted features exhibit the differences more predominantly.

REFERENCES

1. Available from: https://en.wikipedia.org/wiki/Breast Cancer.
2. Sabitha K, Ahmad, K, Thangarajan R. Computational analysis of mutations in really interesting new gene finger domain and BRCA1 c terminus domain of breast cancer susceptibility gene. Asian J Pharm Clin Res 2016;9(3):96-102.
3. Yeole MP, Dhole SN, Kulkarni NS. Peptide Nano medicine in cancer

treatment. Asian J Pharm Clin Res 2013;6 Suppl 1:29-32.
4. Available from: http://www.epfl.ch.
5. Rafael GC, Woods RE. Digital Image Processing. 3rd ed. New York: Pearson Education Inc., Publishing, Prentice Hall; 2008.
6. Wang Q, Ward RK. Fast image/video contrast enhancement based on weighted threshold histogram equalization. IEEE Trans Consum Electron 2007;53(2):757.
7. Hall EL, Kruger RP, Dwyer SJ 3rd, McLaren RW. A survey of pre-processing and feature extraction techniques for radiographic images. IEEE Trans Comput 1971;20:1032-44.
8. Rajini GK, Reddy GR. Performance evaluation of edge detection techniques for biomedical imaging. J Technol 2014;4:128-42.
9. Available from: http://www.marathon.csee.usf.edu/Mammography/Database.