IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

Efficient Diagnosis of Breast Cancer Using Statistical Parameters

Sonal S. Naranje[1], Prof. Venkat N.Ghodke[2]

M .E Student, VLSI & Embedded System, Dhole Patil College of Engineering, Pune, India [1]

Asst. Professor, Department of E&TC Engineering, AISSMS Institute of Information technology, Pune, India [2]

Abstract: Breast cancer has become the leading cause of cancer deaths among women. To decrease the related mortality, disease must be treated early, but it is hard to detect and diagnose tumors at an early stage. Manual attempt have proven to be time taking and inefficient in many cases. Hence there is a need for efficient methods that diagnoses the cancerous cell without involvement of humans with high accuracy. This paper proposes an automated technique using artificial neural network as decision making tools in the field of breast cancer. Image Processing plays significant role in cancer detection when input data is in the form of images. Statistical parameter analysis of image is important in mammogram classification. Statistical parameters are extracted by using image processing The statistical parameter include entropy, mean, energy, correlation, texture, standard deviation, variance, MSE, PSNR. We have also include the comparative analysis of Statistical parameter using bar graph. This parameters will act as a inputs to ANN which will diagnose and give the result whether image is cancerous or non-cancerous.

Keywords: Artificial neural network, Image processing, Statistical parameter, ANN.

I. INTRODUCTION

In many developed countries breast cancer has become the leading cause of cancer deaths among women. According to recent research about 1 in 8 women in us develop breast cancer over the course of her life [1].early detection of breast cancer is needed. Breast image analysis can be performing in many ways .digital mammography is the most widely used method of breast cancer detection. Manual diagnosis done by oncologist does not give result with maximum accuracy. It is time consuming. This paper give the automatic detection of breast cancer using image processing techniques and artificial neural network .The mammogram images have taken from MIAS(mammography image analysis society). Image processing includes various techniques to make the digital mammogram image perfect for artificial neural network.

The input image undergoes through many processes which include pre-processing, image enhancement, noise removal, mass image, targeted image, segmentation, feature extraction. The statistical parameter is important step in mammogram classification. The best extracted feature is texture parameter, by which the abnormalities can be easily indentified. Texture is a method of capturing pattern in the image. Statistical parameters include texture, entropy, mean, standard deviation, energy, co-rrelation, variance, PSNR,MSE. This parameter will be given as input to classifier. There are different classifiers used for analysis of digital image applications. Artificial neural network is the leading classifier use nowadays. ANN is used for classification between cancerous and noncancerous image. The design and implementation of the proposed algorithm is done in matlab using advance image processing toolbox.

II. RELATED WORK

Prof Seema singh and Sushmita H[1] has presented An Efficient neural network based system for diagnosis of breast cancer in this author has implemented an efficient neural network for diagnosis of breast cancer .They have explained the supervised and unsupervised methods tested to develop the most efficient alternative for breast cancer diagnosis. They have used Back propagation algorithm. The comparision of different types of classifiers has been added.

B.M.Gayathri ,C.P.sumathi and t.santhanam [2] breast cancer diagnosis using machine Learning algorithms –a survey This paper summarizes the survey on breast cancer diagnosis using various machine learning algorithms and methods, which are used to improve the accuracy of predicting cancer. This survey can also help us to know about number of papers that are implemented to diagnose the breast cancer.

Chandra prasetyo utomo, Aan kardiana, Rika yuliwulandari [3] had published the paper , breast cancer diagnosis using artificial neural networks with extreme learning techniques in this research, author have implemented ann with extreme learning techniques for diagnosing breast cancer based on breast cancer wisconsin dataset. Results showed that extreme learning machine neural networks (ELM ANN) has better generalization classifier model than BP ANN.

Minavathi, Murali.S, M.S.Dinesh[4], presented reaserch on Classification of Mass in Breast Ultrasound Images using Image Processing Techniques. In the proposed method, ultrasound images are pre-processed using

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

Gaussian smoothing to remove additive noise and anisotropic diffusion filters to remove multiplicative noise (speckle noise). Active contour method has been used to extract a closed contour of filtered image which is the boundary of the spiculated mass. Spiculations which make breast mass unstructured or irregular are marked by measuring the angle of curvature of each pixel at the boundary of mass.

R.Nithya, B.Santhi[6] have presented the paper on comparative study on feature extraction method for breast cancer classification. This paper presents three different feature extraction methods they are intensity histogram, GLCM (grey level co-occurrence matrix) and intensity based features for classification of normal and abnormal patterns in mammogram.

This methods comes under texture measure. The results are proving that GLCM features based neural network is giving higher classification rate of 98%.a supervised classifier system based on neural network is used.

Ashmitha khaleel khan & Noufal P [9] have presented, the Wavelet based automatic lesion detection using improved active contour method. In this paper, we have proposed a method- improved active contour method that helps in the segmentation process for lesion Detection. In this method before segmentation preprocessing and Dwt of image in done.in this mean filter or average filter has been explain to improve the image quality for human viewers,and the wiener filter is used for noise Suppression. The adaptive median filter performs also mentioned about spatial processing used to determine which pixels in an image have been affected by impulse noise.

Chethan k, Dr. Krishna A N[10] presented Detection of breast masses in digital mammograms using multiple concentric layers.in this paper the implementation of multiple concentric layers (mcl) is applied to mammogram Segmentation; this technique has the automatic Detection of masses in digital mammograms. Perform the image to array converter then the image will be scale down . After this binary mask image are typically obtained by thresholding from the grayscale image. Image granulation step determines that group of pixels that are strongly connected in terms of spatial location and Intensity range.

In the above discussion on related work by many researchers ,methods of detection of breast cancer has been explained with different method, techniques and they have given the result with good accuracy ,but the processing is getting complicated as the concepts they have used is difficult. so to overcome this issue our paper has included the new concept of statistical parameter analysis where parameters are extracted for classification purpose .This parameters are sufficient to get maximum accuracy .It is also using artificial neural network classifier for diagnosis purpose. This paper gives the automatic method of detection of breast cancer.

III. METHODOLOGY

A. Input Dataset:
In this paper, we have used 51 Digital Mammogram X-ray Images Dataset (MIAS) .The dataset has 19 cancerous images and 32 non cancerous images. The biomedical images are normally damaged by number of noises, which affect the originality of image. The input image may be color images also. So the input image cannot be directly given to the system for the process.

B .Pre-processing:
Biomedical images are normally affected by various types of noises. Removal of these various noises without destroying the desired information is often a significant challenge. Pre-processing images commonly involves removing low frequency from background noise; this normalizing the intensity of the individual particle images. Digital Image preprocessing is the techniques of enhancing data images prior to computational processing. The output is said to been preprocessed from the input data. Preprocessing methods use a small neighborhood of a pixel to get a new brightness value in output image. Such preprocessing operation is called Filtration. The input image may be color image so first the image is converted to gray image. This gray image taken as input to proposed system, then noise will be removing. This image converted into enhanced image, which will be more accurate image to analyze in next step. The input image has been resized to ideal size for which the program has developed. Noise remove by using Median filter, this filtering method reduce 'salt and pepper' noise, remove the noise and preserve the edges. Direction of image is decided to any one side every time input image has taken as some images are of right side and some are of left side.

C. Enhanced Image
The pre-processed image has many undesirable white spots/pixels in the non-targeted area. The image contrast limit is increases or stretches so as to get the suspected object bright as compare to non targeted object by giving the pixel more intensity threshold value. The unwanted pixel in the undesired area has to be removed, and that part of image is mask .The mask creation of image is perform. Then target image is created by simply multiplying reference image and masked image and image is in size of eight bit.

D .Segmented Image
Segmentation is the process of changing a digital image into multiple segments that is set of pixel also refer as super pixel. Segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyse[5].The suspected objects are separated from background (lines and curves)objects and boundaries in image. It has been observed that after the segmentation process, the white shaded portions i.e. the abnormalities present in the breast region will appear.It include the background subtraction process. After that

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

advance thresholding is used. Gray threshold chooses the threshold to minimize the intraclass variance of the threshold black and white pixels.

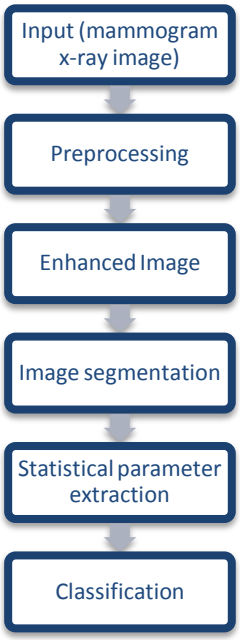

Fig 1: Flow diagram showing overall methodology

IV. STATISTICAL PARAMETER ANALYSIS

Statistical parameter is the best method to help radiologists to get more accurate diagnoses of breast cancer. Statistical parameter analysis is the advance method of detection of cancer. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant the input data will be transformed into a reduced representation set of features .Transforming the input data into the set of features is called features extraction which is nothing but the statistical parameters.
Statistical parameters are a method of capturing visual content of an image. The objective of feature extraction process is to represent raw image in its reduced form to facilitate decision making process such as pattern classification [7].

Texture parameter is extracted from the mammograms. This used to enhance the classification of result. Statistical parameter step is important step to get high classification rate. A set of features are extracted in order to allow a classifier to distinguish between normal and abnormal pattern. The abnormality can be identified on the basis of textural appearance. Extracted features are used in neural classifier to train it for the recognition of particular class either normal or abnormal. The ability of the classifier to assign the unknown object to the correct class is dependent on the extracted features. This will give maximum accurate result. The best feature extracted is Texture feature. We have calculated the following statistical parameter in the proposed system which will act as a input to neural network, sufficient for the classification purpose.

- Energy:

Energy represented the orderliness of a mammographic image. Energy is generally given by the mean squared of a mammographic image.

Gabor wavelet has been used for measurement of energy .Function for computing gabor features of a gray-scale image. Gabor wavelet function calculates gabor features. Mean-squared energy & mean Amplitude for each scale and orientation is returned.

$$\text{Energy}= \sum_{i=1}^{n} x(n)^2$$

- Textures:

It is one of the important features used for many applications. Texture features have been widely used in mammogram classification. The texture features are ability to distinguish between normal and abnormal pattern. Texture is an alteration and variation of surface of the image. In general, texture can be characterized as the space distribution of gray levels in a neighborhood.Texture feature has been proven to be useful in differentiating normal and abnormal pattern. A texture is a method of capturing pattern in the image. Gaussian filter is used for the texture measure. The mathematical expression is given as,

$$\text{Texture}=\int_{\infty}^{\infty} x(n)h(n-v)$$

- Standard deviation(sd):

The standard deviation is a parameter closely associated with the mean. These refer to the dispersion of values in a digital mammographic image around the mean value. The SD is,

$$sd=\sqrt{mean}$$

- Mean

The mean value gives the average intensity value of an image. Mammographic image that contain micro-calcification have a higher mean than those of normal images. Where mean calculated from the given image,

$$\text{Mean}=\frac{\text{number of pixel}}{\text{Total number of pixel}}$$

- Correlation

The relationship between things that happen or change together.

- Entropy

The amount of disorder in a mammographic image is called as entropy. The entropy value is high in micro Calcification. This is because the variation in intensity values in the image is high due to the presence of white calcification spots. Entropy is given as,

$$\text{Entropy}=\sum_{i=0}^{1} x(n)\log x(n)$$

- MSE (Mean Square error)

It is a error metric use to compare image compression quality. It is a cumulative squared error between compressed and original image.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

• PSNR(Peak signal to noise ratio)
It represent a measure of peak error.It is a error metric use to compare image compression quality.
PSNR=10*log10(MSE)

• Variance
It is the square of deviation of all the values in a set of image.

Variance=sd*sd

V. CLASSIFICATION

The project used a three layer artificial neural network. The proposed system uses the supervised technique for training the network. The schematic representation of neural network with 'n' inputs, 'm' hidden units and one output unit.

The main task of the classifier is to categorize the segmented image by considering the statistical parameters which is discussed in the previous section. The extracted features are considered as input to the neural classifier. A neural network is a set of connected input/output units in which each connection has a weight associated with it. The neural network trained by adjusting the weights so as to be able to predict the correct class. The network take the two important parameter, one is data feature and second is data class. Data features are the features which have been calculated in the previous section and data class are cancerous and noncancerous. The desired output was specified as 1 for cancerous and -1 for non cancerous.
The reason why ANN is selected for classification is it has good capacity of generalization, it is highly robust and work well with images .The classification process is divided into the training phase and the testing phase. In the training phase known data are given. In the testing phase, unknown data are given and the classification is performed using the classifier after training. The accuracy of the classification depends on the efficiency of the training.

VI. RESULT

The proposed system has taken the mammogram images as input image then the image goes through number of steps to get the desired output. The following fig shows the original image next to it is first half of image,then we taken second half of image to decide the direction of image.

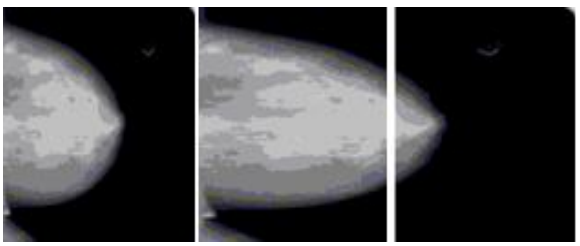

Fig 2: Original Image, First half of image, second half of image

The following figure shows the image at left side as direction of image is decided as left side then masked target image & mask image is calculated. The mask image is the one suspected region is calculated with respect to pixel intensity.

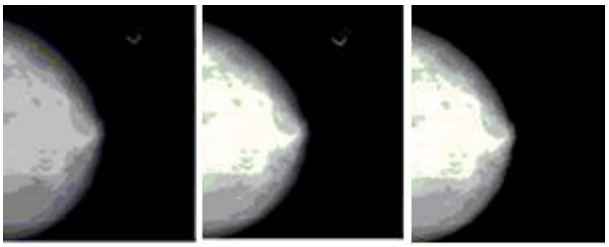

Fig 3: Imaged at left side, masked target image, Mask image

The following figure shows the enhanced image and the segmented image. The enhanced image is the pre-processed image with noise removed and its contrast limit is adjusted and in segmented image we get the region of interest which of use for detection of cancer.

Fig 4: Enhanced Image, Segmented image

The following figure shows the result output whether the input image is cancerous or non cancerous. The classification of cancer is done by ANN. Hence the proposed system has successfully detected the Breast Cancer.

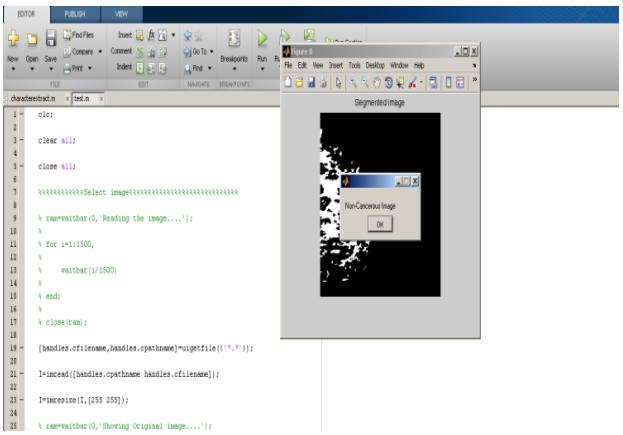

Fig 5: Screen showing the result output

We have also included the bar graph showing Comparative analysis of Statistical parameters which shows parameters on x axis. The parameters are energy, mean, texture, Standard deviation, entropy, correlation.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

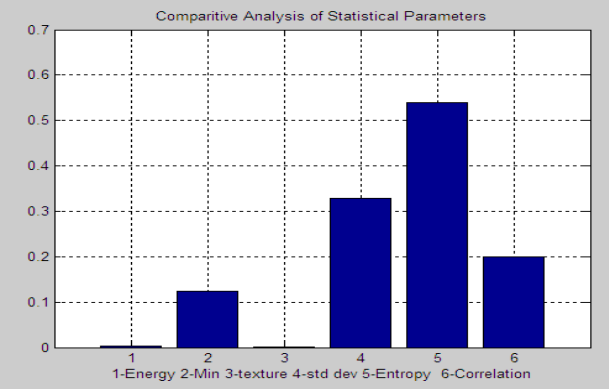

Fig 6: Comparative Analysis of Statistical Parameters

VII. CONCLUSION

Cancer is the most common disease in women in many countries. Early detection of cancer can reduce mortality rate. New technology is used for early detection. In this project first the noise removal of image takes place by the process of image enhancement includes pre-processing. Filters are use to enhance detail information of the image, and gives better sharpness. Segmented image ready for statistical parameter then the extracted feature will act as an input to Classifier ANN. The proposed system successfully achieved more than 90% classification accuracy, which is considered as a good result when compared with similar works in the same research field.

REFERENCES

1. Seema singh and Sushmita H, "An Efficient Neural network based system for diagnosis of Breast cancer", BMS Institude of Technology, India, IJCSIT, vol.5, no.3, pp.4354-4360, 2014,.
2. B.M.Gayathri , C.P.Sumathi and T.Santhanam, "Breast Cancer Diagnosis Using Machine Learning Algorithm –A Survay", SDNB Vaishnav College for Women, Chennai, India , IJDPS vol.4.no.3,May 2013.
3. Chandra Prasetyo Utomo, Aan Kardiana, Rika Yuliwulandari , "Breast Cancer Diagnosis using Artificial Neural Networks with Extreme Learning Techniques", YARSI University, Jakarta, Indonesia,IJARAI,vol.3,no.7.2014.
4. Minavathi, Murali.S, M.S.Dinesh, "Classification of Mass in Breast Ultrasound Images using Image Processing Techniques", Mysore University, India. IJOCA, vol.42.no.10, 2012.
5. Shahid Eqbal, Dr.M.AAnsari, A. Parlak, and H. Yasar, "Medical Image Feature Extraction for Computer Aided Diagnosis of Lung Cancer", GCET, IJARCSSE, vol.5, no. 6, pp. 193-197, June 2015.
6. R.NITHYA, B.SANTHI, "Comparative study on feature extraction method for breast cancer classification", School of Computing, SASTRA University, JATIT & LLS, vol.33.no.2. , 2005-2011.
7. S. Shanthi, V. Murali Bhaskaran, and M. Ghaffarpour, "Computer Aided System for Detection and Classification of Breast Cancer", Kongu Engineering College, Erode,1, Paavai College of Engineering, Namakkal, IJTCA,vol.2.no. 4pp.277-287,2012.
8. Dr. K. Usha Rani, "Parallel Approach for Diagnosis of Breast Cancer using Neural Network Technique", Sri Padmavathi Mahila isvavidyalayam, India.IJOCAvol.3, no.7, 2014.
9. Ashmitha khaleel khan & noufal p," Wavelet based automatic lesion detection using improved active contour method", Dept. Electronics & Communication, MES College of Engineering, Kuttippuram, Malappurum, Kerala,. IJERT, Vol. 3 Issue 6, June – 2014.
10. Said Jai-Andaloussi, Abderrahim Sekkaki, Gw´enol´e Quellec, Mathieu Lamard, Guy Cazuguel, Christian Roux"Mass Segmentation in Mammograms by Using Bidimensional Emperical Mode Decomposition BEMD" 35th Annual International Conference of the IEEE EMBS ,Osaka, Japan, 3 - 7 July, 2013.
11. Karthikeyan Ganesan, U. Rajendra Acharya, Chua Kuang Chua, Lim Choo Min, K. Thomas Abraham, and Kwan-Hoong Ng "Computer-Aided Breast Cancer Detection Using Mammograms: A Review" IEEE review in biomedical engineering, Vol. 6, 2013.

ACKNOWLEDGEMENT

I am deeply indebted to my Guide **Prof. Venkat N. Ghodke**, Assistant Professor, AISSMS'S Institute of Information Technology, Pune, who is considered one of the pioneers in the area of E&TC, for guiding me through the project. I thank him for his expertise in making me understand the problem, fruitful discussions on the subject, and equally for his critical comments, and his ever-smiling gestures, which helped me in getting through the project.

BIOGRAPHY

Venkat Ghodke received the B.E. degree in Electronics Engineering from Dr. B.A.M. University, Aurangabad, Maharashtra, India, in 1997, and the M.E degree in Electronics Engineering specialization with Digital System from Pune University, Maharashtra, India, in 2010. Currently He is an Assistant Professor in AISSMS'S Institute of Information Technology of Savitribai Phule University of Pune, India.His research interests include digital image processing and embedded system area .He had worked in various Institutes as UG and PG guide for Image and embedded system design related area. He had published books and also published papers in various International Journal.

Sonal S. Naranje received the B.E degree in Electronics Engineering from R.T.M.N.U, Nagpur, Maharshtra, India, in 2012, and pursuing M.E degree in VLSI & Embedded system from Savitribai Phule Pune University, Maharashtra, India. Her research interest include image processing and artificial neural network. She had published papers in International Journal.