IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 12, Issue 11, November 2022, ||Series -I|| PP 38-43

Breast Cancer Tumor Detection Using Ann Classification In Image Processing Application

P Chitra

Corresponding author: P Chitra, St.Vincent Pallotti College, Bangalore-43.

Abstract: Breast cancer is the most common malignancy of women and is the second most common and leading cause of cancer deaths among them. At present, there are no effective ways to prevent breast cancer, because its cause is not yet fully known. Early detection is an effective way to diagnose and manage breast cancer can give a better chance of full recovery. This paper gives a clear idea of classification from the mammogram image to find cancer affected area which is a crucial step in breast cancer detection. The output of the classifier differentiates the normal, benign and malignant cases from applied digital mammographic images.
Keywords: ANN classification, breast cancer, malignant, mammogram image.

I. INTRODUCTION

The breast lesion act as the one of the significant problems now a day, the first line diagnosable solution by detecting all lesions of the breast early to help and improving prognosis of cancer and other health problem related to breast. More than one diagnostic modalities are used for breast cancer screening like computed tomography, resonance magnetic imaging and ultrasound. All of these modalities don't enough to detect all lesions of the breast in spite of their characteristics such as high sensitivity of soft tissue detection. This shortening of detecting lead to support the final diagnosis decision by other further examinations like biopsy and unfortunately that may increase the anxiety of the patients. The digital mammography is superior in detecting breast lesions than other diagnostic modalities. American cancer society recommended that MRI breast investigation should be mandatory for women with previous history of ovarian cancer or positive family history. Nowadays, a new method introduces to detect the beast lesions. This method is called digital mammography. Because its ability to link with image processing software [1].

Until now, there is no effective way to prevent the occurrence of breast cancer. Therefore, as it is well known, early detection is the first crucial step towards breast cancer diagnosis and treatment. In terms of medical diagnosis and screening techniques, X-ray mammography is currently the most common technique used in clinical practice due to its low cost and accessibility. Although screening mammography presents some limitations, such as low reliability on dense breast of young women or women who underwent a surgical intervention, it has been recommended as the most effective method for early detection of breast cancer as it provides high sensitivity on fatty breast and excellent performance on micro calcification detection [2]. As a result, a large number of mammograms need to be examined by a limited number of radiologists, resulting in misdiagnoses due to human errors by visual fatigue.

Most of the research has shown that women with family history of breast cancer have a higher hazard or getting disease. That is true whether the family history is on the mother's side or the father's. The death rate due to this cancer is reduced obviously due to the advanced screening programs. Premature detection of breast cancer increases the potentiality of survival rate where as postponed diagnosis considerably encounters the patient to a critical stage and occasionally results in death. Proper screening programs and diagnostic techniques that use digital mammogram to provide an image of the breast. These images, called mammograms, are used to find potential signs of breast cancer like tumors and abnormal changes in the skin [3].

The first major pointer of the malignant cancer is identified as masses. Masses are determined by the spaces recognized by lesions which can be pointed out by their structural formation and marginal property. The second pointer of calcification contains calcium sediments in the breast tissue. These are seen as small bright spots in mammogram descriptions. To classify cancer as benign or malign, the morphological dimension and features are to be evaluated. The third most common pointers of breast cancer are architectural distortions. They are accepted with the abnormal architecture [4].

Medical investigation on breast cancer is not original but absence of proper methods for early recognition is still a challenge [5]. With advancement in improving field, the contribution of information technology has presented a new dimension termed as Medical Image Processing. It has a distinct feature for investigating not only to cancer but also in other fields. By the use of image processing techniques, it has become easy to detect cancerous mass from an infected breast.

II. LITERATURE SURVEY

Classification is one of the most important and essential task in machine learning and data mining. About a lot of research has been conducted to apply data mining and machine learning on different medical datasets to classify Breast Cancer. Many of them show good classification accuracy.

In Lashkari [6], Gabor wavelets and ANN are used to classify normal and abnormal tissues which could increase the accuracy and save radiologist's time. Gabor wavelets transforms have a good attribute in image processing and computer vision. The result shows that this combination of neural networks has a good potential with 97% accuracy on unknown cases.

Vikas Chaurasia and Saurabh Pal compare the performance criterion of supervised learning classifiers; such as Naïve Bayes, SVM-RBF kernel, RBF neural networks, Decision trees (J48) and simple CART; to find the best classifier in breast cancer datasets. The experimental result shows that SVM-RBF kernel is more accurate than other classifiers; it scores accuracy of 96.84% in Wisconsin Breast Cancer (original) datasets [7].

Xin-Sheng Zhang (2014) proposed to classify and detect microcalcification clusters (MCs). They have formulated the classification problems as sparse feature learning based classification on behalf of the test samples with a set of training samples, which are also known as a "vocabulary" of visual parts. Visual information-rich vocabulary of training samples is manually built up from a set of samples, which include Mcs parts and no Mcs parts. With the prior ground truth of Mcs in mammograms, the sparse feature leaning is acquired by the regularized truth of Mcs in mammograms, the sparse feature learning is acquired by the sparse feature learning based Mcs classification algorithm using Twin Support Vector machine (TWSVM). To investigate its performance, the proposed method was applied to DDSM datasets and compared with SVMs with the same dataset. Experiments have shown that performance of the proposed method is more efficient or better than the state-of-art methods [8].

Liyang Wei (2009) propose a microcalcification classification scheme, assisted by content- based mammogram retrieval, for breast cancer diagnosis. We recently developed a machine learning approach for mammogram retrieval where the similarity measure between two lesion mammograms was modeled after expert observers.

In this work we investigate how to use retrieved similar cases as references to improve the performance of a numerical classifier. Our rationale is that by adaptively incorporating local proximity information into a classifier, it can help to improve its classification accuracy, thereby leading to an improved "second opinion" to radiologists. Our experimental results on a mammogram database demonstrate that the proposed retrieval-driven approach with an adaptive support vector machine (SVM) could improve the classification performance from 0.78 to 0.82 in terms of the area under the ROC curve [9].

Ayman A. AbuBaker 2015, proposes a method that can enhance the performance of Computer Aided Diagnosis (CAD) by automatically detecting and classifying the microcalcifications (MCs) in mammogram image accurately and efficiently using multi statistical filters and wavelet decomposition transform. The proposed method is divided to two main stages. In first stage, the potential MCs region (PMR) is detected based on visual characteristics of the MCs in the mammogram images. Then wavelet decomposition transform is implemented to classify the PMR to true positive and false positive regions based on extraction four wavelet features for the mammogram image. This novel method was found to be sensitive in detecting MCs in mammogram images by achieving a high true positive percentage of 98.1% and a low false positive rate 0.63 cluster/image for both MIAS and USF databases[10].

Malar et al 2012, this paper is to reveal the effectiveness of wavelet based tissue texture analysis for microcalcification detection in digitized mammograms using Extreme Learning Machine (ELM). Microcalcifications are tiny deposits of calcium in the breast tissue which are potential indicators for early detection of breast cancer. The dense nature of the breast tissue and the poor contrast of the mammogram image prohibit the effectiveness in identifying microcalcifications. Hence, a new approach to discriminate the microcalcifications from the normal tissue is done using wavelet features and is compared with different feature vectors extracted using Gray Level Spatial Dependence Matrix (GLSDM) and Gabor filter based techniques. A total of 120 Region of Interests (ROIs) extracted from 55 mammogram images of miniMias database, including normal and microcalcification images are used in the current research[11].

The network is trained with the above mentioned features and the results denote that ELM produces relatively better classification accuracy (94%) with a significant reduction in training time than the other artificial neural networks like Bayesnet classifier, Naivebayes classifier, and Support Vector Machine. ELM also avoids problems like local minima, improper learning rate, and over fitting.

Dheeba and WiselinJiji (2010) presented a new classification approach for detection of microcalcification clusters in digital mammograms. Two stages are used for detecting the micro calcification clusters. In the first stage, features are extracted to discriminate between textures representing clusters of microcalcifications and texture representing normal tissue. The original mammogram image was decomposed using wavelet decomposition and gabor features were extracted from the original image region of interest

(ROI). With these features individual microcalcification clusters were detected. In the second stage, the ability of these features in detecting micro calcification was done using Back Propagation Neural Network (BPNN). The proposed classification approach was applied to database of 322 dense mammographic images, originating the MIAS database. Results showed that the proposed BPNN approach gives a satisfactory detection performance [12].

III. RESEARCH METHODOLOGY

The proposed system is introducing a simple and efficient approach to detect the cancer region in mammogram images. Our approach also segments the cancer region on input mammogram image.

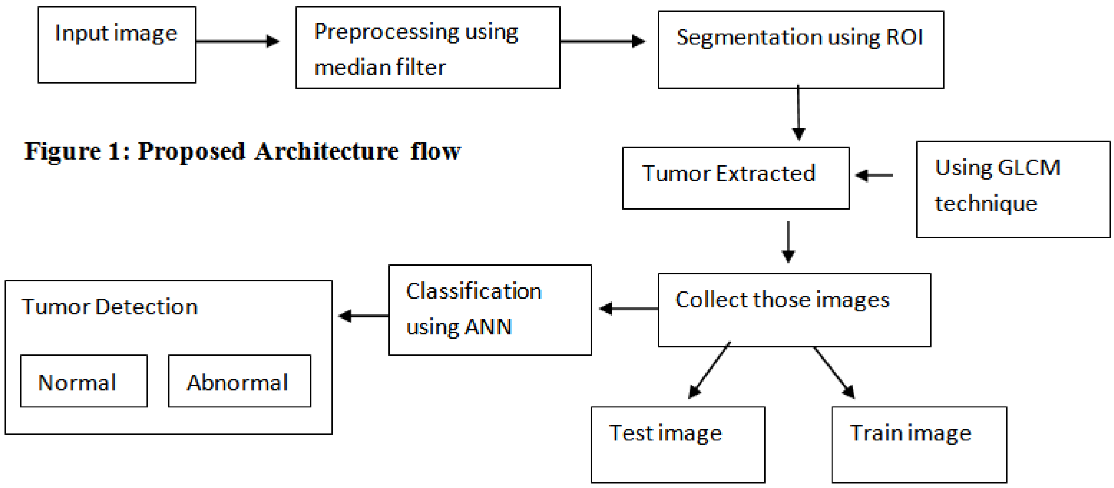

Figure 1: Proposed Architecture flow

Pre-processing using Median Filter

Mammography images are selected and converted into grayscale image of 2D matrix by the scale of 1024X1024. All the images of the database are of the same size, if image size differs from others then the image enchantment algorithm is applied to match the image resolution. These images are filtered through Noise removal algorithm. Then, they are filtered and adjusted to increase pixel intensity.

A median filter is a nonlinear filter is efficient in removing salt and pepper noise median tends to keep the sharpness of image edges while removing noise. The median channel is usually used to reduce disorder in an image, parallel to the mean channel. It is categorized as nonlinear channel used to save the edges of the image. For two images, median channel can be defined as in equation

$$median\ [S(y) + R(y)] \neq median\ S(y) + median\ R(y)$$

S(y) and R(y) are the functions of two images. These channels smooth the data as keeping the small and sharp details. The median is only the center value of all the values of the pixels in the neighborhood. Median filtering is extremely feasible at removing different kinds of disorders.

Segmentation using ROI

Segmentation divides image into its constituent regions or objects. Segmentation plays an important role in image analysis. The goal of segmentation is to isolate the regions of interest (ROI) depending on the problem and its characters. The approach is to partition an image based on abrupt changes in intensity, such as edges in an image and partitioning image into regions that are similar according to a set of predefined criteria. Segmentation is effectively performing the medical image processing for cancer application. A breast abnormality is commonly called ROI.

As per the review of existing report, it is identified that cancer disease is severe threat to human begins life. A mammogram contains two different regions: the exposed breast region and the unexposed air-background (non-breast) region. Background region in a mammogram usually appears as a black region, and it also contains high intensity parts such as bright rectangular labels, opaque markers, and artifacts (e.g. scratches). Breast regions can be partitioned into: Near-skin tissue region, Fatty region, Glandular regions, Hyper dense region.

A region of interest (ROI), is a selected subset of samples within a dataset identified for a particular purpose. The concept of a ROI is commonly used in many application areas. In medical imaging, the boundaries of a tumor may be defined on an image or in a volume, for the purpose of measuring its size. ROI extracted by entering coordinates X, Y and radius in pixels, according to data provided by the database for each abnormal mammogram image.

Feature extraction using GLCM:

Grey level co-occurrence matrix (GLCM) examines through statistical method by considering the spatial relationship of pixels. And the gray level co-occurrence matrix is also called as gray-level spatial dependence matrix. The pairs of pixels with specific values and in a specified spatial relationship occurring in image is calculated by GLCM, and then extracting statistical measures from this matrix. These statistics provide information about the texture of an image. The following table lists the statistics.

Statistic	**Description**
Contrast	It measures the local variation in the GLCM. $Contrast = \sum_{k=0}^{N-1} k^2 \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (f(x,y))$
Correlation	Measures the joint probability occurrence of the specified pixel pairs. $Correlation = \sum_{x=1}^{N-1} \sum_{y=1}^{N-1} (x-\mu)^2 f(x,y)$
Energy	Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment. $Energy = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)^2$
Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal $Homogeneity = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \{f(x,y)\}^2$

Classification using ANN

Artificial Neural Network (ANN) classifier is used in this work as it is a commonly used classifier for breast cancer classification. Neural Network composed of simple elements that are inspired by biological neuron operates in parallel. We train neural network to perform specific function by adjusting weights between elements. Neural network is trained to get desired output.

The network is adjusted based on the comparison with the output and the corresponding target until the network output matches the target. ANN classifier is based on two steps, i.e. training and testing. Classification accuracy depends on training. Once the features related to known Mammogram Images are extracted and selected, the features are used as inputs to train an ANN classifier to classify the images into benign or malignant. In testing/recognition the trained ANN is used to classify new images by comparing the extracted features with the features of the unknown sample Image of Mammogram.

Performance Analysis

The performance of the algorithm is evaluated using the measures like

1. Tissue Extraction - In this experiment, we choosen breast mammogram tissues (WM, GM and CSF). The results are shown in figures respectively. A quantitative performance evaluation is also given in Table. For a quantitative evaluation of the results, a pixel based evaluation approach is used. A comparison is made between pixels of the resulting image (Rt) and the ground truth (Rg). The following measures are used;

Dice index, Jaccard index, True Positive Fraction (TPF), False Negative Fraction (FNF), False Positive Fraction (FPF) and True Negative Fraction (TNF) [31– 33].

$$Dice = \frac{2\left|R_t \cap R_g\right|}{\left|R_t\right| + \left|R_g\right|}$$

2. $$Jaccard = \frac{R_t \cap R_g}{R_t \cup R_g}$$

3. $$TPF = \frac{R_t \cap R_g}{R_g}$$

$$FNF = \frac{R_g \cap R_t}{R_g}$$

$$FPF = \frac{R_t \cap R_g}{R_g}$$

$$TNF = 1 - \frac{R_t \cap R_g}{R_g}$$

5. Accuracy - Classification acc4uracy is the percentage of instances that are correctly classified by the model. It is calculated as the sum of correct classification divided by the total number of samples. It is given by the formula:

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN}$$

6. Specificity - This is a measure that is commonly used in two class problems where the focus is on a particular class. It is the proportion of the negative class that was predicted negative and it is also known as the true negative rate. Specificity formula:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

7. Recall / Sensitivity- is the measure of the ability of a classification model to select instances of certain class from the dataset. It is the proportion of actual positive which are predicted positive.

$$Sensitivity = \frac{TP}{TP + FN}$$

IV. CONCLUSION

This paper discussed with mammogram medical image for preprocessing techniques. In this image for detecting noises, that have detected noise also removed these noises from the above mammogram images by applying the various filtering techniques like Wiener filter, Median filtering, Gaussian filtering. Quality of image and accuracy's are the main factor of the research work. Quality of image evaluation and enhancements and also enhancements stages adopts less preprocessing technique which is established in cramer-rao approaches. This suggested techniques IBCR-SBI was effective on segmentations principle being the regions of passion foundations of featuring extraction obtains. By classification ,the parameter used for weighing are accuracy, specificity and sensitivity are calculated where sensitivity measures the percentage of truly predicted cancer class and specificity measures the percentage of truly predicted Benin /normal class .

The experiments and results show that weighted PSO performs better than other existing algorithms. It was observed that proposed weighted PSO outperforms the existing methods. Comparatively the approach Feature Extraction using naive bayes, SVM and Adaboost proposed weighted PSO has higher accuracy rate of 98.1. This experimental analysis will improve the accuracy of Mammogram Brest image. The results, which we have achieved, are more useful and they prove to be helpful for general medical practitioners to analyze the symptoms of the patients and in future more features can be considered and other data set can be used to increase the robustness of the system.

REFERENCES:

[1]. Yousif M.Y Abdallah, Sami Elgak, Hosam Zain, Mohammed Rafiq, Elabbas A. Ebaid, Alaeldein A. Elnaema," Breast cancer detection using image enhancement and segmentation algorithms", Biomedical Research, Vol. 29, Issue.20: Pg. 3732-3736, 2018.
[2]. R. Guzmán-Cabrera · J. R. Guzmán-Sepúlveda · M. Torres-Cisneros · D. A. May-Arrioja J. Ruiz-Pinales · O. G. Ibarra-Manzano · G. Aviña-Cervantes · A. González Parada, Digital Image Processing Technique for Breast Cancer Detection, Springer, International Journal Thermophys, Vol 34, Pg. 1519–1531, 2013.
[3]. Angayarkanni.N, Kumar.D and Arunachalam.G, "The Application of Image Processing Techniques for Detection and Classification of Cancerous Tissue in Digital Mammograms", journal of pharmaceutical sciences and research, Vol. 8(10), 1179-1183, 2016.
[4]. Dr.D.Devakumari, V. Punithavathi, Study of Breast Cancer Detection Methods using Image Processing with Data Mining Techniques, International Journal of Pure and Applied Mathematics, Volume 118 No. 18 2018, 2867-2873.
[5]. Baran, A., Kurrant, D., Zakaria, A., Fear, E., & LoVetri, J. (2014, July). Breast cancer imaging using microwave tomography with radar-derived prior information. In IEEE Radio Science Meeting (Joint with AP-S Symposium), 2014 USNC-URSI, pp. 259-259.
[6]. Lashkari, "Full automatic micro calcification detection in mammogramimages using artificial neural network and Gabor wavelets," in *Proceedings of the 6th Iranian Conference on Machine Vision and Image Processing (MVIP '10)*, Isfahan, Iran, October 2010.
[7]. V. Chaurasia and S. Pal, "Data Mining Techniques : To Predict and Resolve Breast Cancer Survivability," vol. 3, no. 1, pp. 10–22, 2014
[8]. Xin-Sheng Zhang, A New Approach for Clustered MCs Classification with Sparse Features Learning and TWSVM, Hindawi Publishing Corporation, the Scientific World Journal, Volume 2014.
[9]. Liyang Wei, Yongyi Yang, and Roberts M. Nishikawa, "Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis" NIH Public Access Pattern Recognit. 2009 June ; 42(6): 1126–1132
[10]. Ayman A. AbuBaker, Automatic Microcalcification Detection Using Wavelet Transform, International Journal of Computer Theory and Engineering, Vol. 7, No. 1, February 2015
[11]. E. Malar , A. Kandaswamy, D. Chakravarthy, A. Giri Dharan, A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine, Elsevier Computers in Biology and Medicine 42 (2012) 898–905.
[12]. Dheeba.J, Wiselin Jiji.G, Detection of Microcalcification Clusters in Mammograms using Neural Network, International Journal International Journalof Advanced Science and Technology of Advanced Science and Technology of Advanced Science and Technology Vol. 19 Vol. 19, June, 2010.