**IJESRT**

# INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY
## A REVIEW PAPER ON DIAGNOSIS OF BREAST CANCER USING DIGITAL MAMMOGRAM IMAGES

**Pavankumar M. Ahuja[*], Prof. R. D Ghongade ,Dr.D. G. Wakde**
*M.E  scholar, Department of Electronics and Telecomm, P.R Pote(Patil) COE & M,SGBAU, Amravati, India
Professor ,HOD ,Department of Electronics and Telecomm, P.R Pote(Patil)COE &M,SGBAU, Amravati, India
Principal, P.R.Patil College of Engineering & Technology, Amravati,India

## ABSTRACT
Breast cancer is presently a greatest threat existing among the women. Early detection of breast cancer helps to diagnose the disease and can enhance the life time of patient. The most common deformity that may signify breast cancer is masses. Digital mammogram images are the effective tool for its early detection. The interpretation of mammograms is not simple because of small differences in densities of different tissues within the image. Computer Aided Diagnosis (CAD) is the mainly preferable technique for the primary diagnosis of cancer using X-ray, CT, and mammogram or MRI images.

This paper makes a review on prior detection of breast cancer by making analysis of digital mammograms. It focus on initial processing which removes noise from the mammograms and then segmentation of image is done that lead to its partition and to helps to recognize the abnormalities that  can cause cancer.

**KEYWORDS**: Breast cancer, Mammogram- X-ray images of breast cancer, Computer Aided diagnosis (CAD); breast density.

## INTRODUCTION
In the medical terms, cancer is known as malignant neoplasm and simply it is an unregulated growth of cell of human body. It is a disease that form tumor, but all formed tumors are not necessarily cancerous. This in turn increases the complexity in the diagnosis of cancer. Currently, in the developed and less developed countries the breast cancer is become a serious problem. As per the data of national cancer institute 22% of new cases of breast cancer occurs every year and is considered to be second most common type of cancer in the world [7]. It is expected that in 2011 about 508.00 women died due to the disease breast cancer. Even though breast cancer is considered to be a disease of the developed world, approximately 50% of breast cancer and 58% of deaths take place in less developed country. Indonesia is found as the fourth country for breast cancer incidence, the data source was expected about 18.60 cases per year with 11.60 incidences rate. Early detection leads to solution of the problem of breast cancer, which can be achieved through mammography screening.

The Clinically suggested method for early detection of breast cancer is mammography [9]. The primary abnormality in the case of breast cancer is the mass that appears in the digital mammograms. Mammogram is medicinal means to distinguish the breast growth and gives better accuracy than clinical breast assessment. To detect early breast cancer, the sign of masses can be used. However, understanding of true digital mammograms can be crucial caused by the poor quality of images. In addition to that, it takes sufficient feature to discriminate the malignant or benign cases of cancer. Hence, the appropriate treatment and action could be done reduce the number of breast cancer patients. Figure1. Indicate images of digital mammograms with normal, benign, and malignant cases respectively.
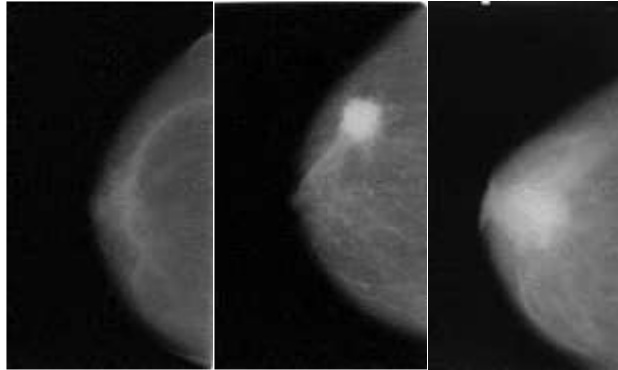
Figure1. Mammogram images of various cases of breast cancer.

The analysis of digital mammograms is done using the computer aided Diagnosis (CAD). It detects the presence of cancer in the breast. There are various forms of abnormality that may affect breast tissue. This Computer aided detection system make a use of available computer technologies for the detection of abnormalities that exist in mammograms, for example masses, calcifications and architectural distortion. It plays a significant role in the early hour's detection of breast cancer for radiologist and help to decrease the death rate among women in the world [16]. CAD reduces the difficulties occurring while reading the images for the radiologist. Image reading is become easier with computer algorithms. Hence in past some years, CAD system and associated methods have created the interest of both the radiologist and research scientists in it. By categories, in digital mammography CAD is distributed into Computer Aided Detection (CADe) which is used as a second reader then the final decision made by radiologist Computer Aided Diagnosis (CADx) which categorizes the detected region into normal, benign, and malignant cases. Here firstly the images are preprocessed to enhance their quality and to remove noise from them and then segmentation is done which is followed by feature extraction and selection. The purpose of feature selection is to select the suitable feature that is capable of distinguishing among the various cases. Finally the classification is done based to cancerous and non-cancerous region.

In this paper, enhancement of preprocessing and feature selection step is developed. The Preprocessing needs enhancing contrast and remove the noise as stated earlier, hence thus it is easier to segment the suspicious regions in segmentation step and then to classify accordingly.

## II LITERATURE SERVEY

Computer aided diagnosis is found as important means to detect and analyze the presence of breast cancer. In the preprocessing steps the analysis is done based on observation and mammograms obtained. A considerable amount of work is done in this field over the past 20 years. For the efficient CAD system many techniques have been developed.

B.Verma et al, [3] has formed a computer aided diagnosis system to determine a framework for computerize mammograms that give emphasis to neural-hereditary calculation characteristic determination technique and acquired precision was 85% on mammograms from Digital database for screening mammography (DDSM).

Domínguez and Nandi have proposed a way to detect masses in breast using mammograms in three steps. The first step improves contrast of image and equalizing the brightness across the image. In the second step the regions are targeted at multiple levels using the method of density slicing. Then In the third step, a ranking system is introduced to decide regions that are characterized as tumors.

Chan and Zheng et al[15], have proposed a detection algorithm in which segmentation technique is used with discrete wavelet transform and a random fied is selected with the help of multiresolutionmarkov algorithm. In addition to that, he also combined binary decision algorithm to identify suspicious areas from the selected features which are produced in the segmented part. Their study has carried out using MIAS database with 322 mammograms images along with total of 37 masses. The structure of masses was well defined, circumscribed or ill-defined masses

but not spiculated masses were included. The sensitivity of their proposed algorithm is estimated to be 97.3% with 3.92 averages of false positives per image.

A. Yadav, S. Singh and B. Singh, et al [4]. have compared the relationships of some novel and hybrid enhancement methods. This comparison is based on the basic observations (from the clinical point) and the performance estimation parameter (statistical-based) like CNR and PSNR. The techniques that have been evaluated are contrast stretching, mean filter, histogram equalization technique, median filter, hybrid technique, counter-let transform filter, Gabor filter and homomorphic filter. Their experimental results show that the enhancement algorithm can get the good quality of image.

For the purpose of feature selection technique, using an optimal set of feature, Wenfeng Han et al [18].identified the masses in digital mammogram. In his work, the region of interest known as ROI, firstly segmented by isocontour map method. Extraction of feature is based on textural and shape features. Then extracted features are selected by Correlation based Feature Selection (CFS) scheme and classified further to decide the masses or non-masses.  To train the dataset and compare with another method backpropagation neural network is used. The performance of each method was monitored by ROC curve.

 Roman W. Swiniarski and Andrzej Skowron [17] propose rough set method and principal component analysis (PCA). Here the basic concept in feature selection is selection is subset reduct. Feature extraction is founded on histogram in mammogram images. With MIAS database, this method is evaluated with 1024 x 1024 pixels image. The result was obtained 75% accuracy with backpropagation to identify the pattern of the selected feature. Extraction of feature is based on textural and shape features.

For early detection and to classify masses, William E. Polakowski et al.[5] offered a model based vision theory in digitized mammograms. His result involves five modules, out of five two of which carry out the detection and the reduction of FP's prior to classification. Wavelets methods are used here.The First module which is a focus of attention module, concentrate on a difference of Gaussians (DoG) filters tracked by threshold to select the regions of interest (ROI). The second module which is the index module calculates a fairly accurate mask of the mass in each ROI and decreases the number of false ROIs derived from the area, contrast and circularity of the masks. The training database is employed for optimization of an image. The ROI was given to prediction unit. Here comparison is done and features are extracted which make an ultimate conclusion that the tumor is malignant or benign. Inherent reductions in overall image cause reduction in contrast by this operation.

Firstly, the correlation between breast density and risk of developing breast cancer was studied by Wolfe [6]. He managed to organize mammographic images into different classes based on the density and the allocation of connective tissue: N1 fundamentally fat with few dense tissues), P1 (the dense elements reside in less than a quarter of the breast area), P2 (the dense elements inhabit over a quarter of the breast area) and DY (The area is essentially dense). As per the Wolfe study and analysis, merely the DY and P2 contain a high susceptibility to disease.

## III PROPOSED WORK

The analysis of mammograms is done using suitable techniques. The proposed approach is modeled into six parts image acquisition, enhancement or improvement, segmentation, feature extraction, fuzzy classification and diagnosis. The overall architecture is shown in Figure 2.and their  processes are described as follows.
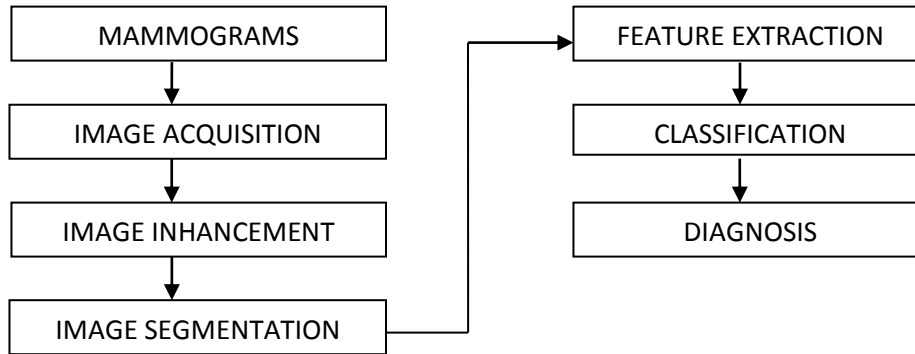
```
┌─────────────────────────┐        ┌─────────────────────────┐
│       MAMMOGRAMS        │        │    FEATURE EXTRACTION   │
└─────────────────────────┘        └─────────────────────────┘
            │                                  │
            ▼                                  ▼
┌─────────────────────────┐        ┌─────────────────────────┐
│    IMAGE ACQUISITION    │        │      CLASSIFICATION     │
└─────────────────────────┘        └─────────────────────────┘
            │                                  │
            ▼                                  ▼
┌─────────────────────────┐        ┌─────────────────────────┐
│    IMAGE INHANCEMENT    │        │        DIAGNOSIS        │
└─────────────────────────┘        └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    IMAGE SEGMENTATION   │────────┘
└─────────────────────────┘
```

Figure 2. Overall Architecture

*A. Image Acquisition*

The collection of images is done from the cancer hospital in the form of medical format which are of non-cancerous andcancerous patients. It also gives all needed tools fordiagnostically correct representation and processing of medical imaging data.

*B. Enhancements*

Enhancement practice improves the interpretability or observation of information in image for human being viewers. It leads to the contrast of images and provides better input for automated image processing technique. The contrast enhancement can limit in order to avoid the noise in an image. We make use histogram equalization (HE) because this technique is easy and efficient to advance the contrast of images by transforming the values in an intensity of an image.The HE preserving the input brightness of the image is essential to avoid the creation of non-existing artifacts in the output image. Enhancement can improve the performance.

*C. Segmentation*

There are various methods of image segmentation. In the process image segmentation, the CT image is divided into a set of region. The aim of segmentation in many tasks is for the regions to characterize meaningful areas of breast which consist of the nodules, spike and nodule boundaries.

Threshold binary image display the region of the breast extracted by using threshold value. The global threshold is computed for the enhanced image, in which it suppresses light structures related to image border.

$$T(x,y) = \begin{cases} 1 & f(x,y) > T \\ 0 & f(x,y) <= T \end{cases}$$

The single global threshold is selected for overall complete enhanced image $f(x,y)$. In which the intensity value ranges in between 0 and 1. To turn a gray-scale image into a binary image, the threshold value is used by means of thresholding method. The tissues are isolated by making use threshold and the artifacts attached to border are removed from threshold binary image.

In second step, the gradient magnitude is used as the segmentation function to display the nodule and nodule spike borders. Spikes in image, has the irregular structure, to recognize them gradient operator is used i.e. sobel and filter with the sobel mask with the replicate property there, values outside the bounds of the array are considered to be equal the nearest array border value.

$$G(x,y)= \sqrt{lx^2 + ly^2}$$

Where G(x,y)= output image

Lx=filtered co-ordinate in x direction

Ly=filtered co-ordinate in y direction

### D. Feature extraction

The feature extraction is especially essential and important step to extract region of interest (ROI). The nodule size, structure, nodule spike values and volume are considered as feature in proposed system. In feature extraction method we extract the feature from the images to discriminate the cancerous and non-cancerous region.

### E. Classification

The feature images calculated in feature extraction methods are used for the classification and diagnosis of cancerous and non-cancerous region or nodules. In it we categorize the nodule type by using fuzzy rules. Here nodule belongs to which one of the types such as normal, benign, malignant and advanced cancerous region or nodule.

## CONCLUSION

In this study, diagnosis of mammogram image based on CAD system is analyzed. Preprocessing, segmentation, feature extraction, feature selection, and classification are the steps to be undertaken. The proposed method focused on combination of feature extraction based on feature texture. The performance of CAD in classification step depends on the optimization of the feature selection. Correlation based feature selection (CFS) is used to choose the best feature. Mean, standard deviation, smoothness, angular second moment (ASM), entropy, and correlation are the best features that guarantee the improvement of classification with fewer features dimension. Here we are expecting to achieve the accuracy 96.66%, sensitivity 96.73%, specificity 97.35%, and ROC 96.6%. This finding is useful to help radiologist in detection of breast cancer using digital mammograms images. It is expected to contribute as material consideration in decision-making.

## REFERENCES

[1] Hanung Adi Nugroho,Faisal N, Indah Soesanti ,LinaChoridah "*Analysis of Computer Aided Diagnosis on Digital Mammogram Images",* 2014 IEEE International   Conference on Computer, Control, Informatics And its Application.

[2]  Professor. Kamalakannan J, Dr. P. Venkata Krishna, SITE, VIT University, Dr. M.Rajashekara Babu Kansagra  Deep Mukeshbhai SITE, VIT University"*Identificationof Abnoemality from Digital Mammo-Grams to Detect Breast cancer"* 2015 IEEE International Conference on Circuit, Powerand computing Technologies [ICCPCT].

[3]  B. Verma and P. Zhang,"A novel neural-   genetic algorithm  to find the most significant combination of featurein digital   mammograms", Applied Soft Computing,7, pp. 513-525, year, 2007.

[4] A. Yadav, B. Singh, and S. Singh, "Comparative analysis of different enhancement method on digital mammograms," 2011  2nd*Int. Conf. Comput. Commun.Technol.*, pp. 159–163, Sep.2011.

[5]  William E. Polakowski, Donald A. Cournoyer,  Steven K. Rogers, Mart in P. DeSimio, Dennis W. Ruck, Jeffrey W.  Hoffmeister. Computer-Aided Breast  Cancer Detect ion and  Diagnosis of MassesUsing the Difference of Gaussians and Derivative-Based Feature Saliency, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 16, NO. 6, DECEMBER 1997.

[6]  J. N. Wolfe, Cancer 37, 2486, 1976.

[7]  J Suckling et al (1994) "The Mammographic Image An lyses Society Digital Mammogram Database" Exe-Perta Medica. International Congress Series 1069 pp375-378

[8]  GLOBOCAN and International Agency for Research oncancer (IARC), "Cancer Incidence and Mortality     . .   worldwide",*http://www.iarc.fr/en/media/centre/iarcnews/2010/globocan*,*2008.php.2008.*

**[9]** Z. Luo, X. Wu, S. Guo, and B. Ye, "*Diagnosis of BreastcancerTumoronPCA and fuzzy support vector Machine classifer*" 2008,*forthInt. Conf. Nat Comput.*, pp. 363-367, 2008.

[10] C. H. Lee, "Screening mammography :proven benefit ,continued controversy," *Radiol.clin North Am..vol, 40,pp.395-407,2002*.

[11] M. P. Sampat, M. K. Markey, and A. C. Bovik, "Computer -aided detection and diagnosis in mammography," in *handbookof Image and Video Processing,*, A.C. Bovik, Ed., 2$^{nd}$ed. New York: Academic, 2005, pp. 1195–1217.

[12] MuhammedAnshad PY, S.S.KUMAR, Recent Methods for the Detection of Tumor using the Computer aideddiagnosis- A review, *2014 International Conference on control, instrumentation ,communicatio and Computational Technologies (ICCICCT)*.

[13] J. Tang, R. M. Rangayyan, J. Xu, I. El Naqa, , and Y. Yang, **"Computer-aided detection and diagnosisof breast cancer *with* mammography: recent advances.,"** *IEEE Trans Inf. Technol.,Biomed.*, vol. 13,2, pp. 236–51, Mar. 2009.

[14] World Health Organization (WHO), "Breast Cancer Burden".

[15] H.D. Cheng, X. Cai, X. Chen, L. Hu, X. Lou, "Computer-aided detectionand classification of microcalcific-ation mammograms: A survey" Pattern Recognit., 36, no.12, pp. 2967 – 2991,2003

[16] M. Meselhy Eltoukhy, I. Faye, and B. BelhaouariSamir,""A statistical based feature extraction method for Breast cancerdiagnosis in digital mammogram using multireresolutionrepresentation.,"*Comput. Biol.Med.,* vol. 42, 1, pp. 123–8, Jan. 2012.

[17] R. W. Swiniarski and A. Skowron, "Rough set methods in feature selection and recognition, ," *Pattern Recognit.Lett.*, vol. 24, no. 6, pp. 833–849, Mar. 2003.

[18] W. Han, J. Dong, Y. Guo M. Zhang, and J. Wang,,"identification of Masses in Digital Mammogram Using an optimal set of features" *2011IEEE 10th Int. Conf. Trust.Secur. Priv.Comput. Commun.*, pp.1763-1768.

## AUTHOR BIBLIOGRAPHY

| | |
|---|---|
|  | **1.Pavankumar M. Ahuja**received Bachelor of Engineering degree in Electronics and Telecommunication from SGB Amravati University & Pursuing Master of Engineering in Electronic and Telecommunication from P.R.Pote(Patil) ,College of Engineering and Management, Amravati. |
|  | **2.Prof. R .D Ghongade** has received M.E degree from SGB, Amravati University and is pursuing Ph.D. His area of specialization is Image Processing and signals and system and ANN. Currently he is working as Head of Dept., EXTC Dept.,inP.RPote (Patil), College of Engineering and Management, Amravati. |
|  | **3**. **Dr. D.G Wakde**has received M.Sc. in 1977 and Ph.D from Nagpur University in 1981. Currently he is Principal in P.R Patil College of Engineering and Technology,Amravati. He has Teaching experience of 27 years and research experience of 11 years as well as administration experience.He has published many papers in national and international journals. |