

# Detection and Diagnosis of Breast Tumors using Digital Mammogram Images: A Fuzzy Approach

Pavankumar M. Ahuja<sup>1</sup>, R. D. Ghongade<sup>2</sup>, Dr. D. G. Wakde<sup>3</sup>

<sup>1</sup>P. G. Scholar, EXTC Dept. P.R. Pote (Patil) College of Eng. And Management, Amravati, Maharashtra, 444601, India

<sup>2</sup>Professor and Head of Dept., EXTC Dept. P.R. Pote (Patil) College of Eng. and Management, Amravati, Maharashtra, 444601, India

<sup>3</sup> Principal, P.R. Patil College of Engineering & Technology, Amravati, Maharashtra, 444601, India

**Abstract:** Breast cancer is one of the major causes of death among women. Tiny clusters of micro calcifications appear as collection of white spots on mammograms show an early warning of breast cancer. Primary prevention seems impossible because the causes of this disease are still remaining unknown. An improvement of early diagnostic techniques is important for women's quality of life. Mammography is the key test use for screening and pre-diagnosis. Early detection performed on X-ray mammography is the key to improve breast cancer prognosis. In order to increase radiologist's diagnostic performance, several computer-aided diagnosis (CAD) methods have been developed to advance the detection of primary signatures of this disease: masses and micro calcifications. Masses are space-occupying lesions, described by their shapes, structure, boundaries, and denseness properties. A benign neoplasm is smoothly marginated having regular shape, whereas a malignancy is characterized by an indistinct and irregular border that becomes more speculated with time. Because of the slight differences in X-ray attenuation between masses and benign glandular tissue, they appear with low contrast and often very blurred. Micro calcifications are small deposits of calcium that appear as small bright spots in the mammogram. This paper presents a staging approach on mammography images using gradient magnitude Sobel filter for cancer tumor mass segmentation. After Enhancement, the first step of the cancer signs detection should be a segmentation procedure can distinguish masses and micro calcifications from background tissue followed by fuzzy classifier. The proposed technique shows good results.

**Keywords:** Breast cancer, Mammogram- X-ray images of breast cancer, Computer Aided diagnosis (CAD), Masses, Shapes.

## 1. Introduction

Breast cancer represents a main health problem worldwide being a major cause of death amongst women. This disease became a commonest cancer among women. If the cancer can be detected early, the options of treatment and the chances of total revival will increase.

In the medical terms, cancer is recognized as malignant neoplasm and merely 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. In turn this 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]. Even though breast cancer is considered to be a disease of the urban 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 occurrence, the data source was expected about 18.60 cases per year with 11.60 incidence rate. Early detection leads to solution the problem of breast cancer, which could be achieved through mammography screening. The clinically recommended scheme for the 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 therapeutic means to distinguish the breast growth and gives better accuracy than clinical breast evaluation. To identify early breast cancer, the sign of masses can be used. However, understanding of true

digital mammograms can be vital caused by the poor quality of images.

The Figure 1 shows the typical image of breast tumor.



**Figure 1:** Typical image of tumor

The examination 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 deformity 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 discovery 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. 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 as shape and structure. Finally the classification is done.

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

## 2. Related Work

To detect and analyze the presence of breast tumor computer aided diagnosis is found an important means. Radiologist carefully makes the observation of digital mammograms. Here lesions and micro calcification are important. 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 and still the research is going on.

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).

EDDIE et al, "Breast Imaging : A Survey", April 10 2011:- This paper describes the different types of modalities used for screening and detection of breast cancer, such as, Mammography used as one of the most effective and popular technique used for uncovering of cancer which are at low risk. Another technique named Ultrasound and MRI is a method used when the tolerant cancer is at high risk and also when the condition of cancer is dense which cannot be detected in the mammogram. Likewise some more different methodologies are used to identify cancer such as: - Breast thermography (is a technique in which cancerous and noncancerous cells have higher metabolic rate resulting in the growth of new blood cells provided that nutrients to the cancer cells for faster growth. Thus the temperature covering the surrounding area of non-cancerous and cancerous breast tissue is different when compared and comes to be higher than normal tissue), PET(Positron Emission tomography is a nuclear medicine imaging technique which is used to produce 3D images), Scintimammography (this technique used the radio isotropic to view the lesions of breast).

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 multiresolution markov algorithm. In addition to that, he also collective 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 evaluate are contrast stretching, histogram equalization technique, hybrid technique, median filter, counter-let transform filter, mean filter, Gabor filter and homomorphic filter. Their experimental result shows that the enhancement algorithm can get the good quality of image.

For the function of feature selection technique, using an optimal set of characteristic, Wenfeng Han et al [18] has 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 extract features are selected by Correlation based Feature Selection (CFS) scheme and classified further to decide the masses or non-masses.

Hilary Alto et al, "content Based Retrieval and Analysis of Mammographic masses", April-June 2005:-This paper explains the trial investigated upon patient's mammogram on their shape, edge sharpness and texture features. The experiment was conducted upon 57 regions in which 20 are malignant and 37 are benign or normal. They used 3 shape features like: - compactness, fractional concavity, and speculation, 14 textural features and 4 edge sharpness features were computed to detect accumulation regions. Then out of these features some features are selected and using them classification is done.

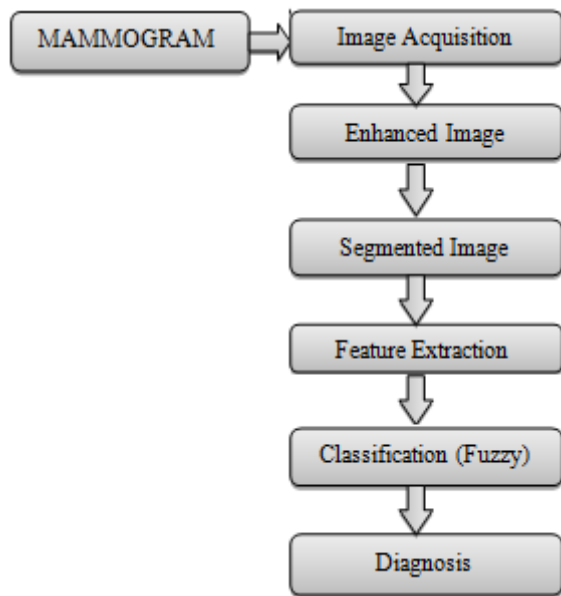
Roman W. Swiniarski and Andrzej Skowron [17] have proposed rough set method and principal component analysis (PCA). Here the basic concept in feature selection is subset reduct. Feature extraction is founded on histogram in mammogram images. It used MIAS dataset. The result was obtained 75% accuracy with back propagation to identify the prototype 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 categorization. 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.

### 3. Our Methodology

We are performing various image processing operations on mammograms. The analysis of mammograms is done using suitable techniques. Our approach is modeled into six parts image acquisition, enhancement or improvement, segmentation, feature extraction, fuzzy classification and diagnosis. The complete architecture is shown in Figure 2. and their processes are described as follows.



**Figure 2:** Implemented Architecture

#### 3.1 Image Acquisition

We have collected 80 mammograms of cancerous and non-cancerous patients from radiologist. These images are given as input to our designed system which are in DICOM format (.dcm) which is a standard format in medical system. It also gives all needed tools for diagnostically correct representation and processing of medical imaging data.

#### 3.2. Image 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 employed histogram equalization (HE) because this technique is easy and resourceful to advance the contrast of images by transforming the values in an intensity of an image. The HE preserves brightness of image.

#### 3.3. Image Segmentation

There are various methods of image segmentation. In the process image segmentation, the mammogram image be divided into a set of region. The aim of segmentation in numerous 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 extract by using threshold value. The global threshold is computed for the enhanced image, in which it suppresses light structures related to image border.

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

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 formation, to recognize them gradient operator is used i.e. sobel and filter with the sobel mask with the imitate property there, values outside the bounds of the array are considered to be equal the nearest array border value.

$$G(x, y) = \sqrt{I_x^2 + I_y^2}$$

Where  $G(x, y)$  = output image

$I_x$  = filtered co-ordinate in x direction

$I_y$  = filtered co-ordinate in y direction

#### 3.4. Feature Extraction

The feature extraction is especially essential and important step to extract region of interest (ROI). Here we are extracting the nodule size, shape or structure, nodule spike and are considered as feature in this system. Based on size of tumor we are staging the tumor. In feature extraction method we extract the feature from the images to discriminate the cancerous and non-cancerous region.

#### 3.5. 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 means of fuzzy rules. We are extracting out the shape and structure of the nodule from segmented image and further based on the size of the nodule we are utilizing T component of TNM classification. We classify the tumor into one of from the different stages. We are staging them based on the information obtained from American Joint Committee on Cancer. These stages are determined from size of nodule. We are using fuzzy classifier.

This staging is done based on size of tumor given below in table 1.

**Table 1**


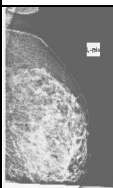
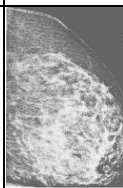
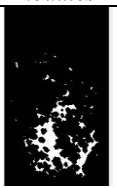

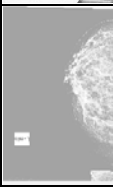
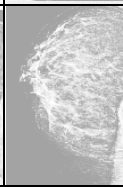



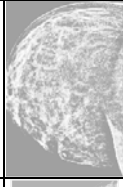



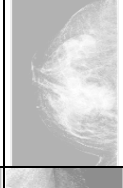


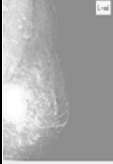
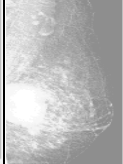

Stage of Tumor	Size of nodule in mm
T1 (T1mi, T1a, T1b, T1c)	T1mi - 0-1mm T1a - 2-5 mm T1b - 6-10 mm T1c - 11-20mm
T2	21-50 mm
T3	>50mm

**Fuzzy Logic:** Fuzzy logic is a control system methodology that lends itself to implementation in systems ranging from easy, small, embedded systems, networked, multi-channel Personal Computer or workstation-based data acquisition and control systems. Fuzzy rule could be applied for building a classifier, a model used for prediction, or it can be applied to form a choice support system.

#### 4. Performance Measurement

In our Implementation we are taking mammogram in DICOM image with extension .dcm as input. After performing preprocessing steps we are extracting nodules based on their regular and irregular shapes. Regular shapes generally come under benign category. We collected 80 images from radiologist and made their diagnosis based on our approach. Some of them are shown below in table 1 at different instance of our approach.

**Table 2**

Original Image (Dicom)	Enhanced Image	Segmented Image	Image with extracted nodules	Area of Main nodule (mm)
				230mm
				8mm
				1mm
				195mm
				48mm

#### 5. Results

For testing purposes, we collected mammograms of 80 patients. We have extracted the nodules from images and also determined the number of nodules in mammogram including small to big size and also total Mean value of nodules. Further based on the size of nodule, we classified them in

three stages as shown in table 1. The experimental results with respect to above five images is given in below table 3.

**Table 3**

Image no.	Size of main nodule	Mean Value	Classification (Staging)
1.	230 mm	0.0575993	T3
2.	8mm	0.00380899	T1b
3.	1mm	0.00043653	T1mi
4.	195mm	0.0481018	T3
5.	48mm	0.0131076	T2

#### 6. Conclusion and Future Scope

Breast cancer is one of the major causes of death among women. So early diagnosis through regular screening and timely treatment has been shown to put off cancer. In this study, diagnosis of mammogram image based on CAD system is analyzed here. Preprocessing, segmentation, feature extraction, feature selection, and classification are the steps performed. In this paper, we have presented a novel approach to identify the presence of breast cancer mass in mammograms. Our work utilizes basic image processing operation and gradient magnitude sobel filter to detect the surrounding and shape of nodule and based on that their shapes are identified and further classified using fuzzy classifier.

It is a new approach using sobel filter with sobel mask, using this we have successfully detected the breast cancer masses in mammograms. The grades indicate that this system can facilitate the doctor to detect breast cancer in the early stage of diagnosis process.

For future research, this method can be extended to include parameters like the number of invaded axillary nodes, calcifications, disease extent, etc. and so improve the prognostic risk judgment.

#### 7. Acknowledgement

We are grateful to Matoshri Radiology Clinic, Amravati for providing Mammograms in DICOM format and incalculably, our management for outspreading their support in providing us foundation and allowing us to use them in the successful completion of our research paper.

#### References

- [1] Hanung Adi Nugroho, Faisal N, Indah Soesanti, Lina Choridah "Analysis of Computer Aided Diagnosis on Digital Mammogram Images", 2014 IEEE International Conference on Computer, Control, Informatics And it's Application.
- [2] Professor. Kamalakannan J, Dr. P. Venkata Krishna, SI TE, VIT University, Dr M.Rajashekara Babu Kansagra Deep Mukeshbhai SITE, VIT University" Identification of Abnormality from Digital Mammogram to Detect Breast cancer", 2015 IEEE International conference on Circuit, Power and Computing Technologies[ICCPCT].



- [3] B. Verma and P. Zhang, "A novel neural-genetic algorithm to find the most significant combination of feature In digital mammogram", *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 2<sup>nd</sup> Int. Conf. Comput. Commun. Tech. pp. 159–163, Sep. 2011.
- [5] W. E. Polakowski, D. A. Cournoyer, S. K. Rogers, M.P DeSimio, D. W. Ruck, J.W. Hoffmeister, and R. A Raines "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative based feature saliency," *IEEE Trans. Medical Image.*, vol.16, no.6, pp. 811–819, Dec. 1997, 10.1109/42.650877.
- [6] J. N. Wolfe, *Cancer* 37, 2486, 1976.
- [7] J Suckling et al (1994) "The Mammographic Image Analyses Society Digital Mammogram Database" *Experta Medica. International Congress Series* 1069 pp.375-378
- [8] GLOBOCAN and International Agency for Research On cancer (IARC). "Cancer Incidence and Mortality Worldwide," 2008.php.2008.  
<http://www.iarc.fr/en/media/centre/iarcnews/2010/globocan>
- [9] Z.Luo, X.Wu, S. Guo, and B. Ye, "Diagnosis of Breast Cancer Tumor on PCA and fuzzy support vector machine classifier" 2008, forth Int. Conf. Nat Comput., pp. 363–367, 2008.
- [10] C.H.L "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 *handbook of image and video processing*, A.C.Bovik, Ed., 2<sup>nd</sup> edtn. New York: Academic, 2005, pp.1195-1217.
- [12] Muhammed Anshad PY, S.S.KUMAR, Recent method For the Detection of Tumor using the Computer aided Diagnosis- A review, 2014 International Conference on Control, instrumentation, communication and Computational Technologies (ICCICCT).
- [13] J.Tang, R. M. Rangayyan, J. Xu, I. El Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography :Recent Advances", *IEEE Trans Inf. Technol., Biomed.* vol.13, no. 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 detection and classification of micro calcification in mammograms: a survey", *Pattern Recognition*.
- [16] M. Meselhy Eltoukhy, I. Faye, and B. Belhaouari Samir "A statistical based feature extraction method for the Breast cancer diagnosis in digital mammogram using Multi resolution representation.," *Comput. Biol. Med.* Vol.42, no. 1, pp. 123–8, Jan. 2012.
- [17] R. W. Swiniarski and A. Skowron, "Rough set method 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 Optimization Set of Features," 2011 IEEE 10th Int.conf. Trust.

### Author Profile



**Pavankumar M. Ahuja** received Bachelor of Engineering in Electronics and Telecomm. from SGB Amravati University and Pursuing Master of Engg. from P.R.Pote (Patil) COEM, Amravati.



**Prof. R.D Ghongade** currently working as Head of Dept, EXTC Department, in P.R.Pote (Patil), College of Engg. And Management, Amravati. His area of Specialization is Signals and System, Image Processing and ANN.



**Dr. D.G Wakde** has received M.Sc in 1977 and his Ph.D from Nagpur University. Now, He is working as Principal in P.R Patil College of Engineering and Technology. He has Teaching as well as Research Experience.