

Automatic Detection and Classification of Microcalcification on Mammographic Images

¹Saranya R, ²M.Bharathi Ph.D., ³Showbana.R

¹PG Scholar, Bharath University, Chennai,

²Prof/Head, Bharath University, Chennai

³Assistant Professor, Avinashilingam University, Coimbatore.

Abstract: Mammography is a best technique for early detection of breast cancer. This paper deals with the detection of micro calcification on mammographic images. It appears as white speckles in images. Micro calcification is a tiny bit, which may not be clearly visible because of low contrast in the image. Our goal is to apply image processing technique to enhance the images and subsequently extract the features from the suspicious region with great accuracy and reliability to detect whether micro calcification is malignant or not.

Keywords: Micro calcification, malignant tumor

I. INTRODUCTION

Breast cancer is one of the leading causes of death in women over the age of 50. Among the large number of imaging modalities and detection tools, mammography is considered the most cost-effective method for detecting breast cancers at an early stage. It is a group of diseases in which cells in the tissues of the breast become abnormal and divide without order or control. These malignant cells form too much tissue and become a tumor. The tumor can grow into nearby tissue or cells break away and enter the bloodstream or lymphatic system eventually affecting other organs. The spread of cancer is called metastasis.

A cancer cell has characteristics that differentiate it from normal tissue cells with respect to: the cell outline, shape, structure of nucleus and most importantly, its ability to metastasize and infiltrate. When this happens in the breast, it is commonly termed as Breast Cancer.

1.2 COMPOSITION OF BREAST TISSUE

The sensitivity of mammography is inversely related to the tissue composition, or ratio of dense tissue to fatty tissue. The fatty breast tissue serves as a lucent background against which radiodense abnormalities can be identified, whereas normal tissue can obscure a mass. Therefore, the tissue composition is important for the referring physician to know because it correlates with the ability of mammography to detect lesions in the breast. The Three categories of tissue composition are

Fatty tissue: Less dense breasts are mainly made up of fat tissue. The whiteness on the mammogram is a combination of the glandular or ductal tissue of the breast and the fibrous supporting tissue. The grey is the fatty part of the breasts. The mammograms tends to look black or gray.

Fibro glandular tissue: Is a combination of glands and fibrous tissues. Fibro glandular breast tissue includes the milk ducts in the breast and the fibers that support the breast. Her mammogram will look very white or cloudy.

Dense tissue: Dense is usually a mathematical term meaning crowded or packed or a greater number of thing in a certain volume. Dense breast tissue is comprised of less fat and more glandular tissue. On a mammogram, dense tissue appears white and cancerous tumors appear white making it nearly impossible to "see" the tumor. Dense breasts contain more glandular and connective tissue. Breast cancer itself is made up of dense tissue. This means that on a mammogram, a tumor is harder to spot in dense tissue than in fatty tissue, because the tumor looks a lot like the tissue around.

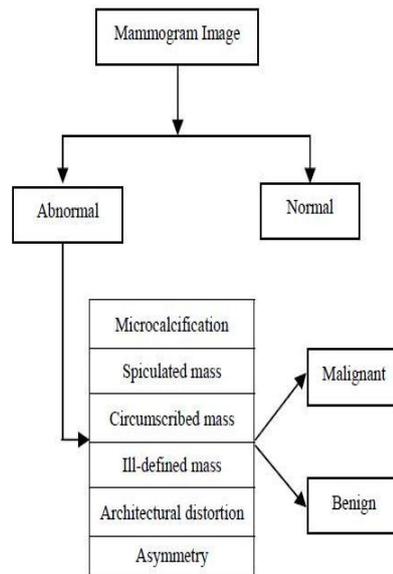


Fig.No.1: Block Diagram of Breast Cancer

1.3 SIGNS OF BREAST CANCER

Two most important signs of breast cancer that can be seen on a mammogram are

- Focal masses
- Micro Calcifications

Other signs are architectural distortions and Asymmetric breast tissue.

1.3.1 Mass

Masses in mammograms are compact areas that appear brighter than the tissue in which they are embedded because of higher attenuation of X-rays.

1.3.2 Calcifications

Breast calcifications are deposits of calcium that can be seen on a mammogram of the breast. There are two types

Macro calcifications: are bigger bits of calcium, and are not usually linked to breast cancer. They are coarse calcium deposits in the breast. They look like large white dots or dashes on a mammogram. They may be caused by calcium deposits in a cyst or in milk ducts as women get older.

Micro calcifications: are quite tiny bits of calcium, and may show up in clusters or in patterns (like circles or lines) and are associated with extra cell activity in breast tissue. Usually the extra cell growth is not cancerous, but sometimes tight clusters of micro calcifications can indicate early breast cancer. Scattered micro calcifications are usually a sign of benign breast tissue. So it is detected in early stages. Micro calcifications are actually tiny specks of mineral deposits (such as calcium) they can be distributed in various ways. Sometimes micro calcifications are found scattered throughout the breast tissue, and they often occur in clusters. Most of the time, micro calcification deposits are due to benign causes.

However, certain features and presentations of micro calcifications are more likely to be associated with malignant breast cancer. Microcalcifications appear grouped on a mammogram typically, at least 5 micro calcifications per square centimeter are required to be considered a cluster, but three suspicious micro calcifications could be enough to prompt a biopsy.

In theory, mammographically as a calcification is a concretion of varying size composed of calcium phosphates, mainly hydroxyapatite, or calcium oxalate crystals. Thus, the calcifications themselves are either benign nor malignant but are the result of cells that have undergone a benign or malignant transformation.

1.4 Micro calcifications according to form, size, density, and distribution.

Form, micro calcifications come in many shapes. They can be round, linear, coarse, granular (fine), monomorphic (all the same basic shape), or pleomorphic (many different shapes), or clustered.

Size, microcalcifications are described as either large or small, or when in clusters whether the sizes of micro

calcifications are homogeneous or not.

Density of micro calcifications may be high, low, or variable. And if in a cluster, the homogeneous or non-homogeneous nature of the density is noted.

Distribution refers to the overall placements of the micro calcifications within the breast image. Microcalcifications can be in single clusters or multifocal, unilateral or bilateral, diffuse, segmental, linear, or regional.

1.5 Best indicators for the evaluation of micro calcifications:

It can be said that as a general rule, when the micro calcifications are distributed in diffuse or bilateral arrangements in the acini, or with a round or punctate shape, or scattered in dense breast tissue, the situation is usually **benign**.

However, if the micro calcifications are in a branching or linear pattern and with irregular borders, or with variable density, or distributed in a segmental or haphazard way, this is highly suspicious of ductal carcinoma in situ, or **malignancy** in other words.

Sometimes the nature of micro calcifications can only be described as '**suspicious**' requiring a biopsy to find out more predictive information. When the distribution is linear and the form of the micro calcifications is either round, oval, punctate, or amorphous, this is suspicious.

II. MAMMOGRAPHY

Mammography is a specific type of imaging that uses a low-dose x-ray system to examine breasts. A mammography exam, called a mammogram, is used to aid in the early detection and diagnosis of breast diseases in women. Mammographic studies are performed for both diagnosis of breast disease and screening for cancer.

2.2 MAMMOGRAPHY PROJECTIONS

Mammography is the technique to detect breast cancer and it is based on the difference in absorption of X-rays between the various tissue components of the Breast such as fat, tumor tissue, and calcifications. The projection of the breast can be made from different angles. The two most common projections are

Medio-lateral oblique (side view taken at an angle):

The advantage of the Medio-lateral oblique projection is that almost the whole breast is visible, often including lymph nodes.

Cranio-caudal (top to bottom view): In Cranio-caudal view does not show the area close to the chest wall.

III. Overall Architecture Of The System

The basics behind this project are that, the whole project is done offline. The images are acquired and these acquired images are further processed in order to remove noises. The main problem countered often with the mammographic images is that noises will be more and also some valuable information's may get lost due to some processing.

3.1 PREPROCESSING

Preprocessing is used in an effort to reduce the effects of the normal anatomy of the breast. One approach for preprocessing is to fit a polynomial to the background and subtract the fitted background from the original image. Because microcalcifications contain relatively large amounts of high spatial frequency information, a high-pass filter would enhance the signals while reducing the background structure.

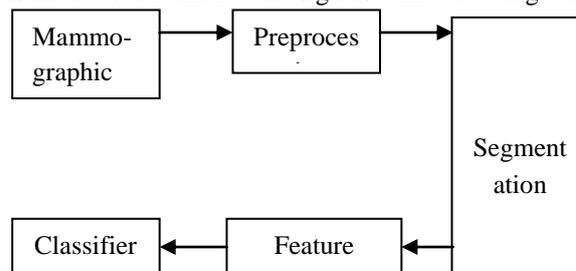


Fig.No.2: Architecture of the System

3.1.1 MEDIAN FILTER

In signal processing, it is often desirable to be able to perform some kind of noise reduction on an image or signal. The **median filter** is a nonlinear digital filtering technique, often used to remove noise and also preserves edges while removing noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (Ex.: edge detection on an image).

3.1.2 LOCALIZATION:

After filtering the image, Localization is the further step to proceed. Localization is done to find out the seed points in the images.

Algorithm:

Step 1: Read the image which has 1024*1024 pixels.

Step 2: Divide the image into 32*32 block.

Step 3: Find out the mean value for the entire block. And make it as threshold value.

Step 4: Find out the maximum intensity for each block.

Step 5: Check the following Conditions.

Condition 1: if maximum intensity < threshold
then Remove the block.

Condition 2: if maximum intensity > threshold
then Add the block.

Now, we can find the block where the microcalcifications are present.

Step 6: Apply z-score method to find where the seed point is located inside the block. $(x - m)/\text{standard deviation}$

Highest score value can be locate the seed point.

3.2 SEGMENTATION

Segmentation of tumor is an important prior step for further classification/ identification. Segmentation is an important step of CAD systems. Both automation and accuracy of segmentation is crucial. Automation of segmentation is important because it facilitates the complete automation of the CAD system. Accuracy of segmentation is important because many crucial features for discriminating benign and malignant lesions are based on the contour. After a lesion or lesions have been identified, they are segmented from the image. This essentially involves identifying the margin or border of the lesion(s). A pixel belongs to the lesion if the radial gradient is greater than a threshold value. In this way, all pixels that are a part of the lesion are identified. There are two different goals for the segmentation of microcalcifications. One is to obtain the locations of suspicious areas to assist radiologists for diagnose. The other is to classify the abnormalities of the breast into benign or malignant.

3.2.1 Region Growing Methods

The first region growing method was the seeded region growing method. This method takes a set of seeds as input along with the image. The seeds mark each of the objects to be segmented. The regions are iteratively grown by comparing all unallocated neighboring pixels to the regions. The difference between a pixel's intensity value and the region's mean δ is used as a measure of similarity. The pixel with the smallest difference measured this way is allocated to the respective region. This process continues until all pixels are allocated to a region. The main goal of segmentation is to partition an image into regions. Region-based segmentation is a technique for determining the region directly.

Algorithm

Input : Mammographic image

Step 1: Start

Step 2: Choose the seed pixel

Step 3: Check the neighboring pixels, add them to the region if they are similar to the seed.

Step 4: Repeat the previous step until no more pixels can be added.

Step 5: Stop

Output: Segmented image

3.3 FEATURES EXTRACTION

Features extraction step is the most critical in the context of microcalcifications segmentation and classification. In this work we propose the use of 10 different parameters in the preliminary step of the features extraction.

The parameters that we consider are subdivided into: geometrical features and textural features. The first ones are related to the shape of microcalcifications, to the distances among them, and to many statistical properties of these geometric features, and also to the shape of the cluster, represented by the convex polygon containing the microcalcifications (convex hull).

Instead, the second set of features deals with the interior of microcalcifications, taking into account the intensity variation within each microcalcifications and within the convex hull.

Finally, we also use more sophisticated texture features called Haralick parameters. A combination of three features, such as entropy, energy, and number of pixels, is the best combination to distinguish a benign microcalcification pattern from one that is malignant is used.

3.3.1 Need of feature extraction:

To reduce the number of false detections, features are extracted and are used to differentiate true calcifications from false detections. A large number of different features are being used by different investigators. The drawback of having a large number of features to choose from is that the selection of the optimum set of features is difficult to do, unless a very large number of images are available for feature selection. This is in addition to images needed for training and images needed for testing the technique.

3.3.2 TEXTURE FEATURE:

Features based on flow orientation in adaptive ribbons of pixels across the margins of microcalcifications are used to classify the regions detected as true regions or False-Positives (FPs). Those regions that were successfully segmented were further classified as benign or malignant disease by computing five texture features based on Gray-level Co-occurrence Matrices (GCMs).

Gray Level Image Texture Feature

Defined a set of image structural features are used for the classification of malignancy. Two categories of correlated gray-level image structure features are defined for classification of—difficult-to-diagnose cases. Surrounding Region Dependence Method (SRDM) is based on a second order histogram digitized mammograms. This approach is very efficient and effective for locating microcalcifications in the mammograms with various densities.

Definition of the contrast based on fuzzy homogeneity uses both local and global information and the contrast enhancement algorithm can enhance the main feature while suppress the noise.

The parameters can be easily adjusted to obtain different True-Positives (TP) and False-Positives (FP) rates that are useful matrix that is calculated from two surrounding regions of a pixel.

3.3.3 GEOMETRIC TEXTURE:

Shapes can be described by many aspects, they are:

Area: Actual number of pixels in the regions.

Perimeter: containing the distance around the boundary of each contiguous region in the image.

Major axis length: Scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region. This property is supported only for 2-D input label matrices.

Minor axis length: Scalar; the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region. This property is supported only for 2-D input label matrices.

Eccentricity: It is the measure of aspect ratio. It's ratio of length of major axis to minor axis. **Orientation:** the angle (in degrees ranging from -90 to 90 degrees) between the *x*-axis and the major axis of the ellipse that has the same second-moments as the region. This property is supported only for 2-D input label marices.

3.4 TRAINING USING ANN

An artificial neural network (ANN), often just called a "neural network (NN) as shown in the, is a mathematical model or computational model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation

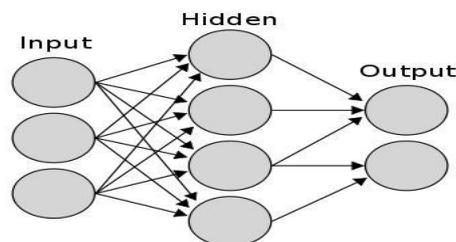


Fig. No. 6: Structure of Neural Network

3.5 LEARNING PARADIGMS

There are three major learning paradigms, each corresponding to a particular abstract learning task. These are supervised learning, unsupervised learning and reinforcement learning. Usually any given type of network architecture can be employed in any of those tasks.

3.5.1 SUPERVISED LEARNING

We wish to infer the mapping implied by the data; the cost function is related to the mismatch between our mapping and the data and it implicitly contains prior knowledge about the problem domain.

Tasks that fall within the paradigm of supervised learning are pattern recognition (also known as classification) and regression (also known as function approximation). The supervised learning paradigm is also applicable to sequential data (e.g., for speech and gesture recognition). This can be thought of as learning with a "teacher," in the form of a function that provides continuous feedback on the quality of solutions obtained thus far.

3.6 NEED FOR NEURAL NETWORK

Diagnosing of a disease is a major problem in medical network. For this certain basic difficulties have to be taken into account:

- The human probability to recognize patterns are low.
- For a valid diagnosis of a disease needs an experienced physician.

3.6.1 Back-Propagation Artificial Neural Network (BP-ANN)

A feed-forward back-propagation artificial neural network (BP-ANN) was used to predict the biopsy outcome from the mammographic findings and patient age. The BPANN was trained to minimize the sum-of-squares error using the back-propagation algorithm. The network had a single hidden layer of 12 neurons and each neuron in the network used a logistic activation function.

3.6.2 WHY WE GO FOR BACK PROPAGATION NETWORK?

Learning is accomplished by adjusting the weights in NN to solve a particular problem. For this type of network, the most common learning algorithm is called Back Propagation (BP). For each case. When comparing with other algorithms, this algorithm back propagates the error through the system and reduces error to obtain the desired output with minimum number of epochs in less time.

3.7 ARCHITECTURE OF FEEDFORWARD NETWORK

Feed forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of the neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range -1 to $+1$. On the other hand, if you want to constrain the outputs of a network (such as between 0 and 1), then the output layer should use a sigmoid transfer function (such as logsig).

3.8 TRAINING

Once the network weights and biases are initialized, the network is ready for training. The network can be trained for function approximation (non linear regression), pattern association, or pattern classification. The Training process requires a set of examples of proper network behavior –Network inputs p and target outputs t . During training the weights and biases of the network are iteratively adjusted to minimize the network performance function net.performFcn . The default performance function for feed forward networks is mean square error **mse** - the average squared error between the networks outputs and the target outputs t .

Algorithm:

Input unit: 15 neuron

Hidden unit: 10 neuron

Step 1: initialize weights, for each training pair compute the network output (forward)

Step 2: compute the following:

- Error for each layer (backward).
- learning rate for output layers:
- weight changes with new learning rate
- learning rate for hidden layers
- weight changes with new learning rate

- Update the weights
Step 3: Repeat step 1-3 while stopping condition is false.

IV. Conclusion

In this paper, we have examined various steps in detection of microcalcifications (1) the pre processing (2) segmentation algorithms, (3) features extraction, (4) classifiers, their performances were studied.. Although by now some progress has been achieved, there are still remaining challenges and directions for future research, such as, developing better preprocessing, enhancement and segmentation techniques; designing better feature extraction, selection and classification algorithms; integration of classifiers to reduce both false positives and false negatives; employing high resolution mammograms and investigating 3D mammograms.

In order to reduce false positive, several different types of features, sometimes, with clinical information should be used. As the complexity of algorithm increases, the time complexity of the CAD will also increase. Mammogram image analysis society database is standard test set but defining different standard test set (database) and better evaluation criteria are still very important. With some rigorous evaluations, and objective and fair comparison could determine the relative merit of competing algorithms and facilitate the development of better and robust systems.

References

- [1]. Y.Ireanus ,Digital mammogram segmentation and tumour detection using artificial neural networks, in international journal of soft computing ,2008, pp.112-119.
- [2]. J.Bozek,K.Delac,M.grgic, Computer aided detection and diagnosis of breast abnormalities in digital mammography in 50th international symposium sep2008
- [3]. A.V.Deshpande,S.P Narotee, A Region Growing Segmentation for detection of microcalcification in digitized mammogram, proceeding of international conference on cognition and Recognition.
- [4]. T.Nizialek,W.R.Brody, Segmentation algorithm for detecting microcalcification in mammogram ,IEEE Transactions on Information technology in Bio medicine,Jun1997,vol no.2,pp-141-149.
- [5]. S.Yang yu,L.Guan, A CAD system for the automatic detection of clustersd microcalcification in digitized mammogram films, IEEE Transactions on Medical Imaging ,vol.19,feb2000.
- [6]. J.Dengler,S.Behrens, Segmentation of microcalcification in mammograms, IEEE Transactions on medical imaging,Vol.12, Dec 2003
- [7]. M.mani, B.Revathy, Microcalcification detection in mammogram image using wavelet transform and neural network,, International journal of advanced scientific research and technology, issue 1, feb2012
- [8]. NCI[<http://www.cancer.gov/cancertopics/factsheet/support/prognosis-stats>]
- [9]. <http://www.cs.stir.ac.uk/~lss/31X7/BPlectures/>
- [10]. http://en.wikibooks.org/wiki/ArtificialNeural_Networks/ Error Correction _Learning