

# DWT based Feature Extraction for Classification of Untreated MRI Mammogram of Breast Cells and Normal Cells

Sushma S.  
JSS University  
Mysore

Balasubramanian S.  
JSS University  
Mysore

Latha K. C.  
JSS University  
Mysore

## ABSTRACT

A standout amongst the most effective strategies for bosom malignancy early discovery is mammography. Another strategy for identification and arrangement of miniaturized scale calcifications is displayed. It should be possible in four phases: in the first place, pre processing stage manages clamour expulsion, and standardized the picture. Second stage, K-Means bunching (KMC) is utilized for division and pectoral muscle extraction utilizing territory figuring lastly smaller scale calcifications identification. Third stage comprises of two dimensional discrete wavelet changes are separated from the discovery of miniaturized scale calcifications. And after that, nine measurable components are figured from the LL band of wavelet change.

## Keywords

DWT (Discrete Wavelet Transform), K-nearest neighbor, mean, standard deviation, MRI Mammogram

## 1. INTRODUCTION

Screening strategy is completed on the premise of mammograms, X-beam pictures are utilized to uncover bumps in the bosom. DWT is utilized to ascertain the zone furthermore demonstrating the disease influenced territories. Breast growth is a sort of disease starting from bosom tissue, most usually from inward coating of drain pipes or the lobules that supply the pipe with milks. The attributes of the growth decide the treatment, which may incorporate surgery, medicine (hormonal therapy and chemotherapy) radiation and immunotherapy. Forecast and Survival rates for breast cancer vary greatly depending on the malignancy sort, stages, treatment, and geological area of the patient. [1]

### 1.1 Symptoms of Breast Cancer:

New bumps or a thickening in the bosom or under the arm, Nipple delicacy, release, or physical changes, Skin aggravation or changes, for example, puckers, dimples, terror factor, or new wrinkles, warm, red, swollen bosoms with a rash taking after the skin of an orange, Pain in the bosom, no noticeable or clear indications (asymptomatic) [2] are the side effects of bosom growth in various patients.

### 1.2 Risk elements of bosom malignancy:

Disease can be distinguished in various numbers of ways. The nearness of certain kind of tumors can be distinguished in various ways, which incorporates the nearness of certain sign and side effects, screening test or restorative imaging. Other potential hazard considers include: absence of childbearing or absence of bosom encouraging more elevated amounts of certain dietary examples, and heftiness. The two most ordinarily utilized as a part of screening strategies, physical examination is utilized as a part of bosoms tissues by a social

insurance supplier and mammography. This can offer a suitable probability that a protuberance is growth furthermore distinguishes some different sores, for example, a basic blister. Numerous sorts of growth could be counteracted by lessening smoking in men, eating more foods grown from the ground. [3]

## 2. PROPOSED METHODOLOGY

The early identification and finding of bosom malignancy builds the treatment alternatives and lessens the deadly rate in ladies. The Discrete Wavelet Transform is utilized to get the power estimations of tumor cells. The order of evaluated mean and difference of the given database is done utilizing Multilayer preceptor.

Given a breast MRI image, the first step enhances the image, the second step segments the bosom tumor image and in the third step post processing using morphological operations and windowing technique takes place. As a result of these steps, we get a final breast tumor detected image.

### 2.1 Preprocessing

Preprocessing of bosom MR image is the first step in our proposed technique. Preprocessing of an image is done to reduce the noise and to enhance the bosom MR image for further processing. The purpose of these steps is basically to improve the image and the image quality and to get more surety and ease in detecting the tumor.

### 2.2. Overview of the Convolution DWT

The Multi-Resolution Analysis (MRA) capacity and time-scale area qualities of the Discrete Wavelet Transform (DWT) have built up it as an intense instrument for various applications, for example, flag examination, picture pressure and numerical investigation, as expressed by Mallet [4]. This has driven various research gatherings to create calculations and equipment models to execute the DWT. DWTs are in effect progressively utilized for picture coding. This is because of the way that the DWT underpins highlights, similar to dynamic picture transmission, simplicity of compacted picture control, district of enthusiasm coding and so forth., as talked about by Gonzalez.

The innate points of interest of the Discrete Wavelet Transform over different changes, similar to the DCT, DST and DHT make it reasonable for JPEG2000 pressure gauges, as proposed by Christopoulos et al [5]. The multi-determination highlight of the wavelets conquers the blocking curios issues in the DCT, as proposed by Mallet (1989). The convolution DWT utilizes Finite Impulse Response (FIR) channel banks for actualizing sub-band decay. In the traditional convolution strategy for DWT, a couple of Finite Impulse Response channels (FIR) is connected in parallel, to infer high pass and low-pass channel coefficients.

DWT is a technique used to extract features of each image from bosom MRI, which extracts maximum highlighting pixels present in images to progress results. Fourier series analysis chooses sinusoids as the basis function whereas wavelet analysis is based on a decomposition of a signal using an orthonormal (typically, although not necessarily) family of basic functions and it is an interesting and relatively recent tool. Sinusoids are useful in analyzing periodic and time-invariant phenomena, while wavelets are well suited for the analysis of transient, time-varying signals because its energy concentrated in time. Furthermore, in spatial domain, DWT analysis also gives the best performance in detecting discontinuities or abrupt changes in signals.

### 3. FEATURE EXTRACTION

#### 3.1 Breast Density:

Area of breast is calculated by DWT. These also helps to showing the portions which affected by tumors. The Region of Interest (ROI) of the affected tumors cells are identified by using 3-level decomposition using DWT. Area is calculated by intensity values of the pixels. [6]

#### 3.2 Intensity based features:

Intensity based features are first order statistics depends only on individual pixel values. The pixel intensities are Mammogram proposed method simplest available feature useful for pattern recognition. The intensity and its variation inside the mammograms can be measured by features like mean and standard deviation. [7]

#### 3.3 Mean value

The mean value gives the average intensity value of an image. Here i, j indicates the row and column of the given image P (i, j) is the cell denoted by the row and the column of the image of size MxN.

$$\mu = \frac{1}{M \times N} \sum_{i=1}^m \sum_{j=1}^n p(i, j)$$

#### 3.4 Standard deviation:

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

$$\text{Standard Deviation} = \sqrt{\text{variance}}$$

### 4. CLASSIFICATION

In this section, an approach for breast cancer diagnosis with cluster k nearest neighbor is explained. The systems starts by extracted features using wavelet transform, then energy and variance are utilized to select the most relevant features and finally those selected feature are used to train the classifier.

Wavelet transform is being utilized in this system to extract coefficients that represent the image. Wavelets decompose an image and represent them in different multiscales. Figure 1 shows an example of wavelet mutliscale decomposition.

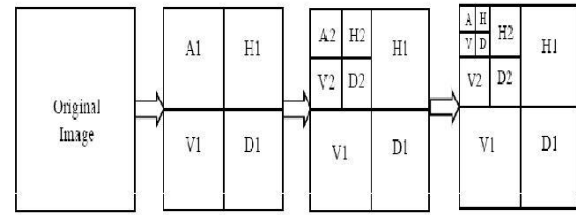


Figure 1. Multiscale decomposition with wavelet transform

The method for feature extraction using wavelets consists of decomposing a set of images. The obtained coefficients are used to construct a matrix  $K \times N$  where K is the image number and N the coefficients number. After that, it is necessary to decrease the number of coefficients since most of them will affect the performance of the classifier negatively. This is due to the redundancy in the information contained as well as the irrelevant information. The energy and variance are utilized in this paper to select the relevant coefficient.

#### 4.1 k nearest neighbor classifier(K-NN)

This algorithm calculates the distance  $d(x_{i,n}, x_{i,m})$  for all  $x_i C_i$ , and then starts to cluster each class to

$N_i$  sub-classes, where  $card(C_i) = N_i$ . Then each two closet data will put in the same sub-class

$$C_{i,l} = \{x_{i,n}, x_{i,m}\},$$

$$\text{where: } \min_{n, m} d(x_{i,n}, x_{i,m})$$

The other data will be gathered at separated sub-class

$$C_{i,j} = \{x_{i,j}, j \in \{1, \dots, N_i\}\} \cup \{n_0, m_0\}$$

Now the index  $n_j$  and  $m_j$  be

$$\text{Min } d(x_{i,n_j}, x_{i,m_j})$$

$$\text{for } n_j, m_j \in \{1, \dots, N_i\}$$

The sub-class  $C_{i,r}$  will split into two other sub-classes if the  $x_{i,n_l}$  and  $x_{i,m_l}$  belong to it. But if they belong to different sub-classes, they will be put in that classes based on the classes  $card$ . The iteration will stopped after K-subclasses were obtained and the initial K-vector will become the means of each sub-classes. Figure 2 will show that each class  $C_i$  will be divided into number of sub-classes  $C_{i,j}$  represented by the mean  $\mu_{i,j}$  of the data.

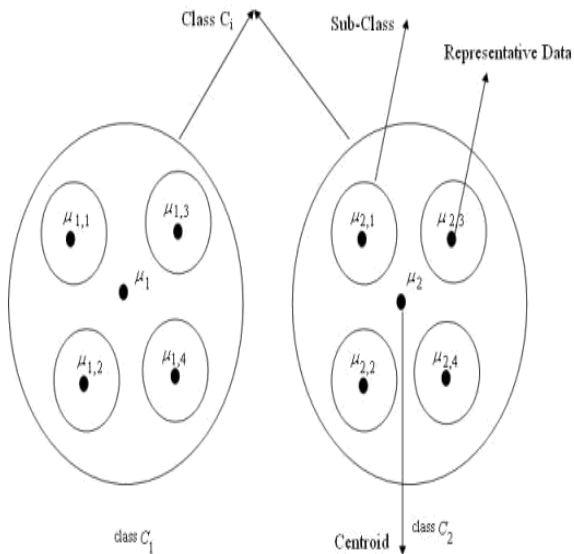


Figure 2. Classes, sub-class, representative data

## 5. EXPERIMENT RESULTS

During the experiments 10 training images were used, which was divided into two size classes: namely Cancer mammography image and Normal mammography image.

From each image classes K means Segmentation is applied. Further for each training set the DWT based features were calculated.

The obtained dimensional feature vectors are classified using K-Nearest Neighbor Classifier. Further the Accuracy is measured for different values of K. The algorithms were coded in MATLAB, because this system is computationally is rather fast, and the code generation is very simple.

FIGURE 3 indicates the classification result for the normal mammography image.

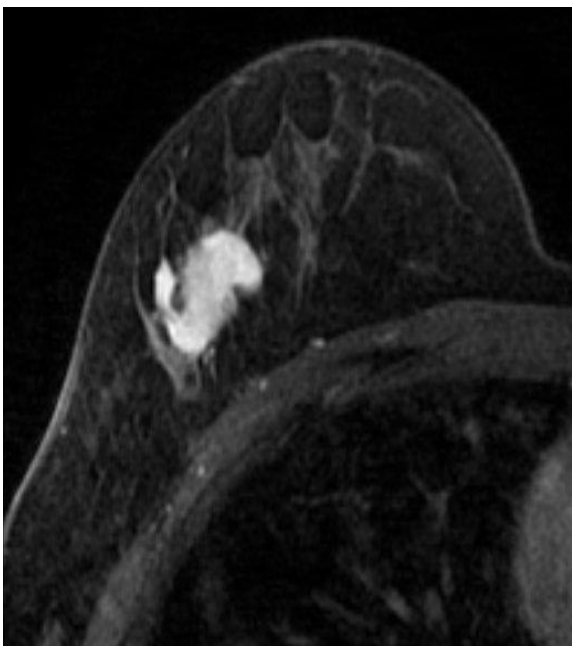


Figure 3. Normal mammography image (output1)

FIGURE 4 indicates the Recognition Performance for different Values of K (Since we are using K-NN Algorithm for Classification).

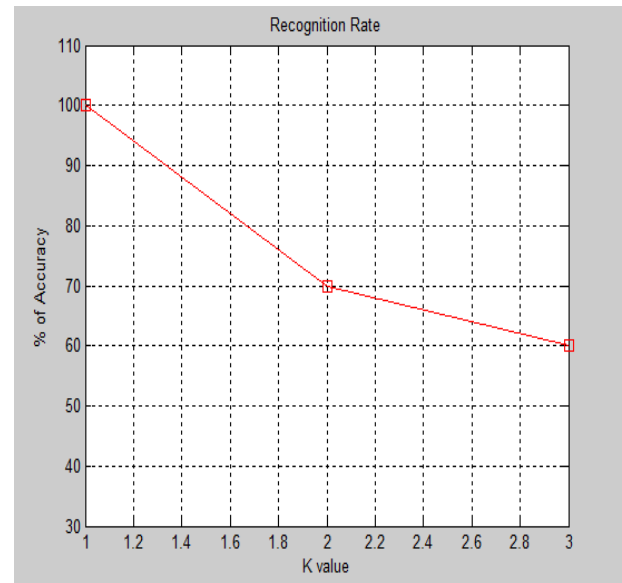


Figure 4. Performance for different Values of K)

FIGURE 5 indicates the classification result for the abnormal mammography image.

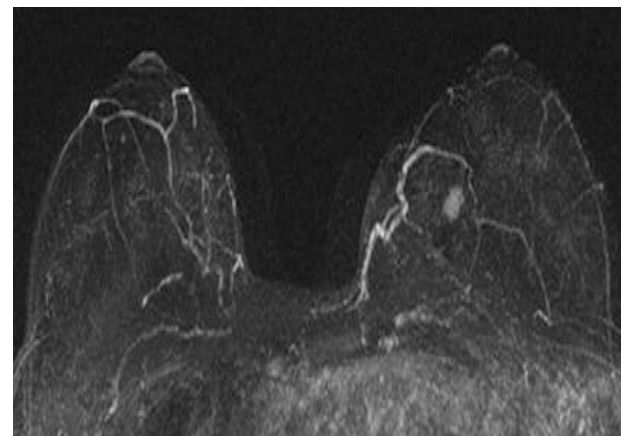


Figure 5. Abnormal mammography image

## 6. CONCLUSION

A new method for detection and classification of micro calcifications was presented. It was done in four stages: first, preprocessing stage deals with noise removal, and normalized the image. Second stage, K-means clustering is used for segmentation and pectoral muscle extraction using area calculation and finally micro calcifications detection. Third stage consists of two-dimensional discrete wavelet transforms are extracted from the detection of micro calcifications. Then, nine statistical features are calculated from the LL band of wavelet transform. It shows that as the K Value Increases the Accuracy Decreases.

## 7. REFERENCES

- [1] Jernal A, Bray F,Center M M,Ferlay J,Ward E,Forman D.Global cancer statistic.CA cancer J clin.vol 61,2011
- [2] <http://en.wikipedia.org/wiki/cancer>
- [3] [www.cancerquest.org/breast-cancer-risks.html](http://www.cancerquest.org/breast-cancer-risks.html)

- [4] Prastawa, M., Bullitt, E., Gerig, G.: ‘Simulation of brain tumors in MRI for evaluation of segmentation efficacy’, *Med. Image Anal.*, 2009, 13, pp. 294–311
- [5] Padma, A., Sukanesh, R.: ‘Automatic diagnosis of abnormal tumor region from brain computed tomography images using wavelet based statistical texture features’, *Int. J. Comput. Sci., Eng. Inf. Technol. (IJCSEIT)*, 2011.
- [6] E.A. Rashed, I.A. Ismail, S.I. Zaki, Multiresolution mammogram analysis in multilevel decomposition, *Pattern Recognition Letters* 28, 2007, pp. 286–292.
- [7] S. Liu, C.F. Babbs, E.J. Delp, Multiresolution detection of spiculated lesions in digital mammograms, *IEEE Transactions on Image Processing* 10 (6) , 2001, pp. 874–884.