International Journal of Soft Computing 3 (2): 112-119, 2008
ISSN: 1816-9503

Digital Mammogram Segmentation and Tumour Detection Using Artificial Neural Networks

[1]Y. Ireaneus Anna Rejani and [2]S. Thamarai Selvi
[1]Anna University, Chennai, Tamilnadu, India
[2]Department of Information and Technology, MIT, Chennai, Tamilnadu, India

Abstract: This study presents an algorithm which aims to assist the radiologist towards fast detection and early diagnosis of breast cancer. It combines several image processing techniques like region splitting and region filling with the Discrete Wavelet Transform (DWT), artificial intelligence techniques and artificial neural networks for detection of masses in mammograms. The AI techniques include fractal dimension analysis, dogs and rabbits algorithm and others. The fractal dimension analysis is used to find the roughness value, which is used to locate the region suspicious for cancer in the mammogram. The dogs-and-rabbits algorithm initiates the clustering. Region splitting and filling are used to segment the suspicious region. The Back Propagation neural network is applied at the end to determine whether a given mammogram is suspicious for cancer. The algorithm is verified by using mammograms from Mammographic Image Analysis Society database.

Key words: Discrete wavelet transform, fractal dimension analysis, region splitting, region filling, back propagation neural network and tumor detection

INTRODUCTION

Each year around the world, millions of women develop new cases of breast cancer. Breast cancer is one of the leading causes of death among women. The death rate can be reduced if the cancer is detected at its early stage. Early diagnosis and treatment increases the chance of survival. Early breast cancer detection requires periodical reading of mammograms. Women over 40 years of age and those who have family history are recommended to take mammograms regularly for screening. Screening mammography has been recommended as the most effective method for early detection of breast cancer. At present mammogram readings are performed by radiologists and mammographers, who visually examine mammograms for the presence of deformities that can be interpreted as cancerous changes. Visually detecting these masses from mammogram is very difficult at the regions where a mass is overlapped with dense breast tissue. Screening mammography has been recommended as the most effective method for early detection of breast cancer. Manual readings may result in misdiagnosis due to human errors caused by visual fatigue. Computer aided diagnosis techniques are used to improve the diagnostic accuracy and efficiency of screening mammography. A radiologist can refer to an automated system for a second opinion.

The main aim of this research is detection of tumor from mammograms. We locate suspicious regions in the mammogram for more detailed examination by the attending physicians. There are several image processing methods proposed for detection of tumors in mammograms. In some cases the primary objective was to enhance the mammograms (Karssemeijer, 1997); in other cases, (Pisano and Shtern, 1994; Qian *et al*., 1998; Patricia *et al*., 1994) researchers have concentrated on identifying areas in mammograms that may contain cancerous changes. Steps have been taken (Petrick *et al*., 1996; Malfait and Roose, 1997), to fully automated mammogram analysis. Various technologies such as fractal analysis; multiresolution based image processing (Chen *et al*., 1997) Markov Random Field (MRF) (Li *et al*., 1995), have been used. In, a two stage adaptive Density Weighted Contrast Enhancement (DWCE) algorithm is explained for tumor detection from mammograms. Even though many algorithms are available for tumor detection the detection rate is still not high.

To improve the accuracy, we introduce an algorithm with image processing techniques and other technologies like artificial intelligence and artificial neural networks for detecting tumor in mammograms. The algorithm includes 3 major steps: Preprocessing, segmentation and classification. In the preprocessing step, the fractal dimension analysis technique is used to find the image roughness of the regions containing masses. The roughness of these regions is different from that of the normal tissue. In the second step the discrete wavelet transform based multiresolution Markov Random

Corresponding Author: Y. Ireaneus Anna Rejani, Anna University, Chennai, Tamilnadu, India

Field segmentation algorithm is used. The wavelet-based MMRF segmentation is very efficient in removing image noise introduced by veins and fibers and it generates clean images for further analysis. In this algorithm for better segmentation region splitting and filling is used. Features are generated from the segmented image and used as inputs to the back propagation neural network. The back propagation neural network is used to classify normal and suspicious areas.

MATERIALS AND METHODS

Mammographic data base: In this research we use mammograms from Mammographic Image Analysis Society (MIAS) MiniMammographic Data base. The database contains 322 mammograms including normal, mass, illdefined mass, architectural distoration, asymmetry and normal.

Discrete wavelet transform: Wavelet analysis represents a windowing technique with variable-sized regions. Wavelet analysis allows the use of long time intervals where more precise low-frequency information and shorter regions where high-frequency information are needed. Most natural images have smooth variations in color and fine details being represented as sharp edges. Smooth variations are given by low frequency component and sharp variations by high frequency. DWT separates these two to implement compression.

The sub band coding and the multiresolution analysis: In the discrete case, filters of different cutoff frequencies are used to analyze the signal at different scales. The signal is passed through a series of high pass filters to analyze the high frequencies and it is passed through a series of low pass filters to analyze the low frequencies. The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations and the scale is changed by up sampling and down sampling (sub sampling) operations. The procedure starts with passing this signal (sequence) through a half band digital low pass filter with impulse response h(n).

A half band low pass filter removes all frequencies that are above half of the highest frequency in the signal. This procedure can mathematically be expressed as

$$y[n] = \Sigma\, h[k] \times [2n-k] \qquad (1)$$

$$Y_{high}[k] = \Sigma x[n] g[2k-n] \qquad (2)$$

$$Y_{low}[k] = \Sigma x[n] h[2k-n] \qquad (3)$$

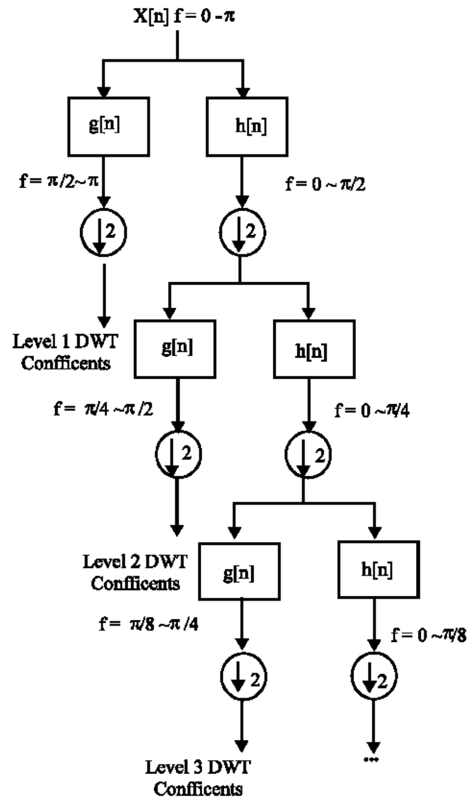

Fig. 1: Signal decomposition using DWT

Y_{high} [k] =Σ x[n] where y_{high}[k] and y_{low}[k] are the outputs of the high pass and low pass filters, respectively, after sub sampling by 2. Figure 1 illustrates this procedure, where x[n] is the original signal to be decomposed and h(n) and g[n] a low pass and high pass filters, respectively. The bandwidth of the signal every level is marked on the figure as f.

Discrete wavelet transforms splits the image into four different parts, the LL, LH, HL and HH. In this research the LL image is taken and then it is used for further processing because the low frequency components constitute the base of an image. 2D DWT is carried out for 3 levels.

Fractal analysis: The term "fractal" means break or fragment. Fractal is any pattern that reveals greater complexity as it is enlarged. For fractals, the counterparts of the familiar dimensions (0, 1, 2, 3) are known as fractal dimensions. The simplest variant of fractal dimension is the similarity dimension D. Applied to a point, a line, a square or a cube, D simply gives the number of ordinary dimensions needed to describe the object -0, 1, 2 and 3, respectively. The value of fractal dimension is between 1-2 for a curve and in between 2-3 for a surface. The larger the D is rougher is the surface.

The fractal window size is taken as 16×16 to find the fractal dimension. The mammograms are divided into blocks of 16×16 and the fractal dimension of each block is calculated. The fractal dimension D gives the roughness of a n× n block . The roughness of the pixel value is used to locate the timorous region in a given mammogram. The area of fractal surface is

$$Ar = K \times r^{2-D} \tag{4}$$

Ar = Surface area.
r = Ruled area.
K = Scaling constant.
D = Fractal dimension indicating the roughness in a given region.

Let the volume of an image region is u_r.
The gray value at position (I, j) is p(i, j) then

$$u_0(i, j) = p(i, j) \tag{5}$$

$$u_r = \max\{u_{r-1}(i, j) + 1, \max u_{r-1}(k, l)\} \tag{6}$$

The blanket volume is obtained by

$$Vr = \Sigma\, [u_r(i, j) - p(i, j)] \tag{7}$$

The blanket surface area at r is

$$Ar = Vr - Vr\text{-}1 \tag{8}$$

D is computed from

$$\log(Ar) = (2\text{-}D)\log(r) + K' \tag{9}$$

For all the subdivided blocks of a mammogram, the blocks that have a very smooth surface or a very rough surface are removed.

Dogs and rabbits algorithm: The prototypes are called dogs and the feature vectors are called rabbits. The dogs are in the field and a rabbit (feature vector) pops up. The dog (prototype) closest to the rabbit barks and chases the rabbit, but the other dogs see that another dog is closer and so they don't move much toward the rabbit.

The movement of the closest dog is determined by,

$$\bar{d}' = \bar{d} + \frac{2D}{(1+D)^f}(\bar{x} - \bar{d}) \tag{10}$$

And for the other dogs, they are determined by

$$\bar{d}' = \bar{d} + \frac{D}{A+D} \quad \frac{2D}{(1+D)^f}(\bar{x} - \bar{d}) \tag{11}$$

Where, $\bar{d}$ and $\bar{d}'$ are the previous and current positions of the dog respectively. D = ||x-d|| is the Euclidean distance between the current dog and rabbit and f≥1 is the fatigue ness of the dog which determines how much the current dog should move to the appeared data (rabbit). A>0 in the given equation determines the inhibition of the movement for the other dogs except the one closest to the rabbit. Using this clustering initialization, all data will be divided into k sets, extracted from region corresponding to the tumor .

Markov random fields: MRF is used to retrieve the original image from the observed data (original image plus noise). To achieve this, the MRF uses Maximum A Posterior (MAP) estimation technique. The MRF segments observed image such that the resultant image has the maximum probability to be original image. Let the observed image is Y. The lines of image pixel are represented by 1-D vector Y = { y_i }, here I = 1,2, 3,…… W×H. W and H are width and height of the image Y. After segmentation the result is labeled as X = {x_i}, I = {1, 2,….,W × H};. x_i = k decides the pixel at position i in the original image Y belongs to pixel k in the segmented image X. The index k å 1,2,…K is the index for the labels and K is the total number of distinct pixels. If $X_{|xi}$ denotes all labels of image except pixel x_i and X_{Ni} denotes the neighborhood pixels around pixel x_i then the MRF requires X to have following 2 properties.

$$P(X)>0 \text{ (positivity)} \tag{12}$$

$$P(x_i|X_{|xi})=P(x_i|X_{Ni}) \text{ (Markovianity)} \tag{13}$$

Neighborhood system and cliques are very important in MRF. In MRF a pixel x_i in a given image is related to other pixel through a neighborhood system N_i. A clique of a given neighborhood system is a subset of pixels within that neighborhood system. This system is adequate for expressing the relationships between pixels describing the suspicious masses. In this approach second order neighborhood system is used to avoid higher computational load.

Region splitting and filling: The image is subdivided into set of arbitrary, disjoint regions and then merging or splitting is done until the desired condition is achieved.

- Split the image into 4 quadrants, represented by the region R_i for which $P(R_i)$ = TRUE.

- Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)$ = TRUE.
- Stop when no further merging or splitting is possible.

Region filling is used to fill holes in the input binary image. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

Back propagation network: An ANS that is useful in recognition of complex patterns and in performing nontrivial mapping function is the back propagation network. It is a multi layer, feed forward network that uses the supervised mode of learning. It examines all the pixels in an image in parallel.

Generalized delta rule is the learning algorithm for the network.

Let an input vector $X_p = (X_1, X_2, \ldots X_n)t$ is applied to the input layer. The input units distribute the values to the hidden units

$$Net_j^h = \sum w_{ji}^h x\, I \quad (14)$$

The output of this node is $i_j = f_j^h(\text{net})$ (15)

The net input to the k^{th} output unit is

$$Net_k^o = \sum w_{kj}^o ij \quad (16)$$

The output of this node is $Ok = f_k^o(net_k^o)$ (17)

The transfer function used is log sigmoid function since it presents the best performance in the net training.

Architecture: The number of neurons in the input layer corresponds to the number of characters to be presented to the network. The number of neurons in the output layer corresponds to the number of classes for which it is being trained. The input layer consists of 10 input neurons, hidden layer 8 neurons and a single neuron in the output layer.

Training: The Back Propagation Algorithm was used. The interface made it possible for parameters as learning rate, moment, minimum error and variation in the architecture of the intermediate layers to be changed in the attempt of improving the performance of training (Fig. 2).

Classification: The tests were performed with training that reached with the minimum error.

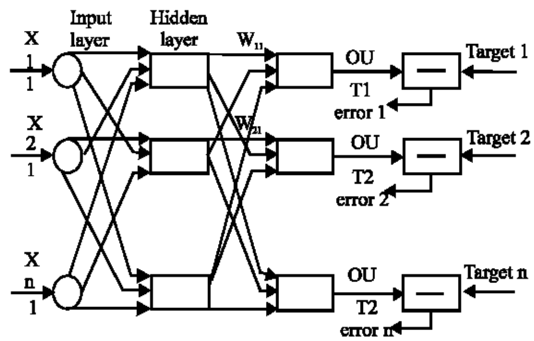

Fig. 2: Back propagation architecture

DEVELOPMENT OF THE ALGORITHM

Preprocessing

DWT compression: The 2D wavelet transform converts the image into 4 bands of data labeled as LL, HL, LH and HH. The LL band can be decomposed once again in same manner, thereby producing even more sub bands (Fig. 3). This can be done up to any level, thereby resulting in a pyramidal decomposition as shown (Fig. 3)

The steps involved in DWT compression are:

- The digital mammogram was taken as the input.
- First level decomposition takes place on application of DWT to the image.
- First level decomposed image was again decomposed for the second level and the resulting LL sub band was further divided into 4 sub bands.

LL sub band image was used as the input to the next step in this algorithm. It contains the low frequency components, which has more information. In this process the original 1024×1024 digital mammogram is converted into 128×128 without the loss of information.

Fractal analysis: To compute the roughness of each and every pixel in the image:

- The LL sub band of the decomposed image was taken as the input for fractal analysis.
- This image was divided into blocks of size 16×16.
- Blanket's method was used to find the roughness of each pixel of each block.
- Fifty mammograms were used for fractal analysis and the threshold value for roughness was estimated. Roughness value is between 2 and 3 for a surface.

The roughness value calculated lies between 2.45 and 2.7 for the tumor region. The region that has the

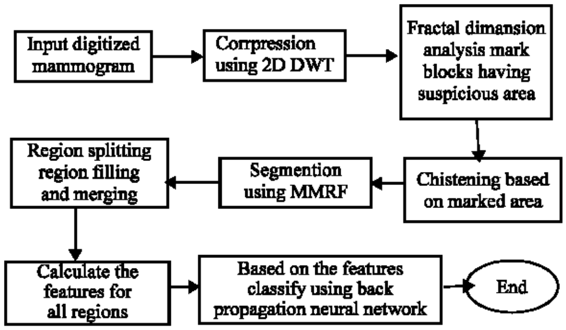

Fig. 3: Flow chart of tumor detection algorithm

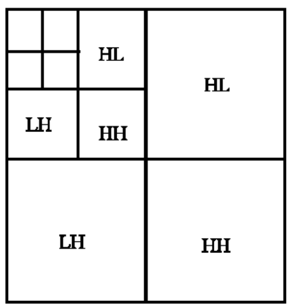

Fig. 4: Pyramidal decomposition of image

roughness values in this range was taken for subsequent analysis and the regions below and above this region are discarded.

Segmentation

Dogs-and-rabbits clustering algorithm: Before the image is segmented, it has to be initialized. For this the dogs-and-rabbits clustering algorithm was used. The steps involved are:

- Fractally analyzed image was taken as the input.
- K data points were taken randomly and they were named dogs.
- Again some random data points were selected and they were named as rabbit.
- Distance between the rabbit and the dogs were calculated and the shortest distance was found out.
- Based on this distance, every dog was moved towards the rabbit.
- The above steps are repeated until the convergence criteria had reached.

MMRF segmentation: MMRF segmentation is initiated using the dogs-and-rabbits clustering algorithm. Here the cliques and neighborhood concept was used and the clustering process was carriedout, until the finest resolution level had been reached. In this research, second order neighborhood system is used. Thus the region for further analysis is considerably reduced and the noise in the image was removed to certain extent.

Region splitting and filling: Splitting and filling is done to group the similar pixels. For this, first the output of MMRF segmentation was divided into blocks. Then each block was considered and checked for any dissimilarity within them. If any dissimilarity is detected, those regions are again subdivided. If there is any similarity between adjacent regions, they are merged. This process is continued until there was no further splitting or merging possible.

Tumor classification: Neural network is used for classification. It is trained with the back propagation algorithm. Figure 2 illustrates a Back Propagation Network (BPN).

The network identifies the affected and unaffected mammograms and they are classified.

The learning rate varied from 0.2 and 0.7 learning rate is around 0.3 and the minimum error is 0.099. The regions that are suspicious for tumor are located based on the criteria as follows:

Mean gradient within current region: This parameter measures the average gradient of each pixel in the region:

$$mwg = 1/N \Sigma g_k$$

Where N equals the total number of pixels within the given region and g_k is the gradient at each pixel k.

Mean Gradient of region boundary: This parameter indicates the sharpness of the region boundary.

$$mg = 1/N' \Sigma g_{k'}$$

Where N' equals the total number of pixels on the boundary of the given region and g_k' is the gradient along the boundary of the given region.

Gray value variance: var measures the smoothness of the given region

$$Var = (1/N \Sigma (X(i, j)- mean(X))^{1/2}$$
$$Mean(X) = 1/N \Sigma X(i, j)$$

Where X(i, j) is the gray level of each pixel within region A and N is the total number of pixels in the given region.

Edge distance variance-edu: Edu measures the shape of the region and its rotational symmetry

$$edv = (1/N\ \Sigma\ (d_k-d_x)^2)/d_x$$

Where d_k represents the distance from pixel k on the edge to the center of the region and d_x is the mean value of all edge distances.

Area-A: This parameter is the total number of pixels within a certain extracted region.

Compactness-cmp: It reflects the shape of the given region.

Cmp = Area of the given region/Area of the smallest rectangle circumscribe the given region.

The feature parameters of a given region are input to the neural network classifier. The Back propagation network is trained and it is used for testing the mammograms for normal and tumor cases.

RESULTS AND DISCUSSION

The algorithm was tested using the digital mammograms obtained from Mammographic Image Analysis Society (MIAS) database. The original MIAS Database (digitized at 50 micron pixel edge) has been reduced to 200 micron pixel edge and clipped/padded so that every image is 1024×1024 pixels. The outputs obtained at various stages of this technique is shown in the following sections.

This algorithm is verified using fifty digital mammograms. During the experiment, A 3 level DWT decomposed image had been chosen 3 level decomposition reduces an image of size 1024×1024-128×128. Wavelet decomposition is carried out for 3 level so that over smoothing is avoided. For fractal analysis a block size of 16×16 was used Fig. (5a). The fractal dimension for the suspicious area obtained by us is between 2.45 and 2.7. Figure 5(b) shows the 3 level DWT filtered image. Figure 5(c) shows the Three level 2D wavelet approximated image.

Figure 5(d) shows the Clustered Output, Fig. 5(e) shows MMRF output, Fig. 5(f) shows theTrained output and Fig. 5(g) shows the Tumour affected region.

The performance of the Back Propagation classification can be given as in Table 1.

Table 1: Performance of back propagation classifier

Features	TP	FP	TN	FN
features extracted	0.78	0.15	0.81	0.17
+Roughness	0.83	0.11	0.78	0.08

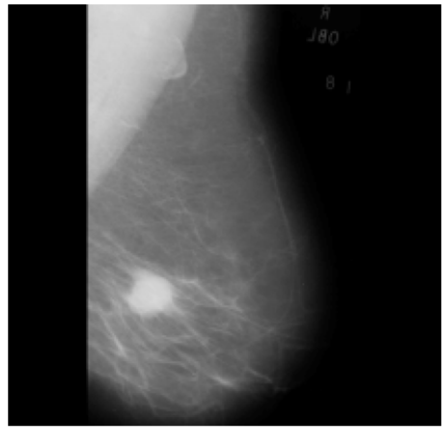

Fig. 5: Original image

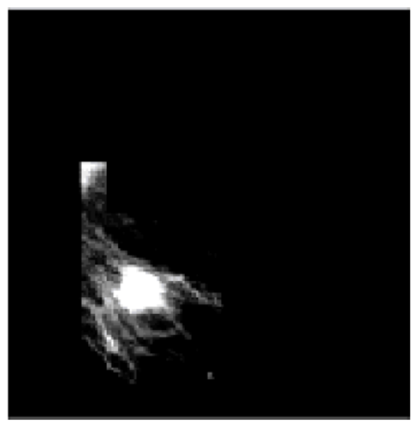

Fig. 5(a): Fractal analysis output

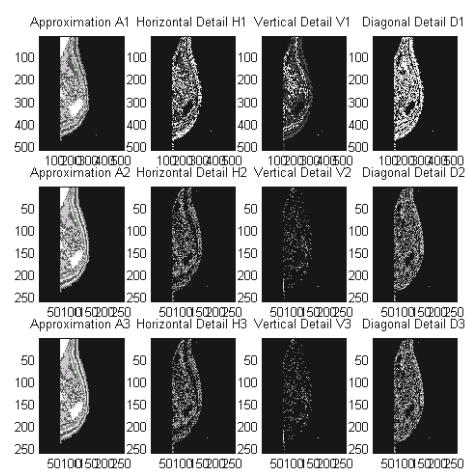

Fig. 5(b): 3 level 2D wavelet decomposition

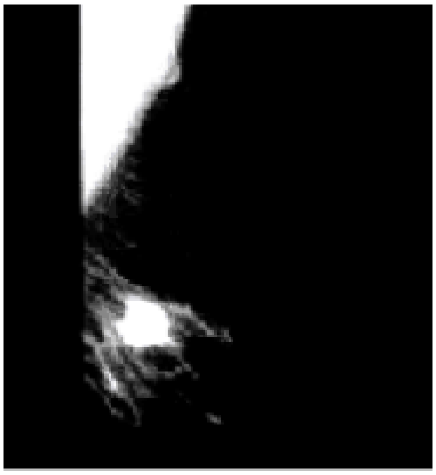

Fig. 5(c): 3 level 2D wavelet approximated image

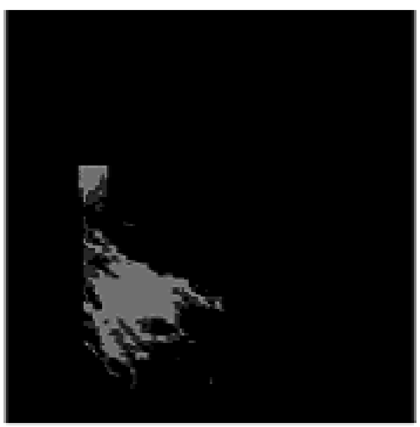

Fig. 5(d): Clustered output

Fig. 5(e): MMRF output

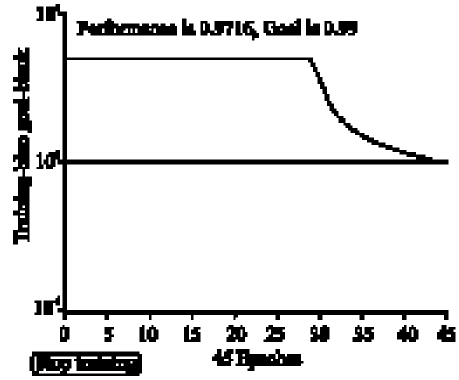

Fig. 5(f): Trained output

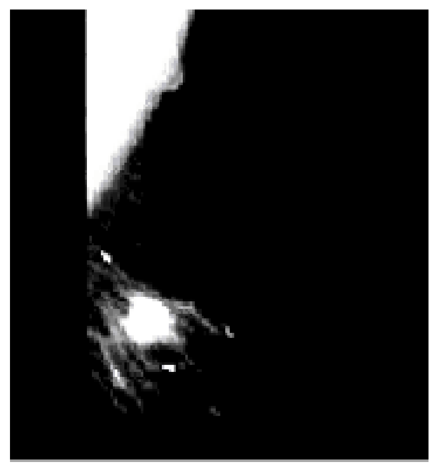

Fig 5(g): Tumour affected region

CONCLUSION

In this study, a technique for detection of tumors in digital mammograms is developed. This method includes preprocessing stage in which the fractal dimension of the mammogram is calculated. Here each pixel is considered for the measurement so that very small sized tumor can also be identified. Initially segmentation is started by the dogs-and-rabbits clustering algorithm while the final segmentation results are obtained from DWT-based MMRF. For better segmentation results region splitting and filling techniques are used. A back propagation network is used for classification. The proposed method is more effective in breast cancer detection environment and it gives an accuracy of 92%. It will assist the radiologists in a better way for their diagnosis of early breast cancer.

ACKNOWLEDGEMENT

The authors are thankful to The International Cancer Research Institute, Neyoor,Tamilnadu for their help.

REFERENCES

Chen, C.H. and G.G. Lee, 1997. On digital mammog-ramsegmentation and Microcalcification detection using multiresolution wavelet analysis. Graphical Models Image Processing, 59: 349-364.

Karssemeijer, N., 1997. Computer-Assisted Reading of Mammograms. Eur. Radiol., 7: 743-748.

Li, S.Z., M. Kallergi, L.P. Clarke, V.K. Jain and R.A. Clark, 1995. Markov random field for tumor detection in digital mammograph. IEEE. Trans. Med. Imag., 14: 567-576.

Lei Zheng and Andrew K. Chan, 2001. An artificial intelligent algorithm for tumor detection in screening mammogram. IEEE. Trans. Med. Imaging, 20: 7.

Malfait, M. and D. Roose, 1997. Wavelet-based image denoising using a Markov random field a priori model. IEEE Trans. Image Processing, 6: 549-565.

Pisano, E. and F. Shtern, 1994. Image Processing and Computer-aided Diagnosis in Digital Mammography in State in The Art of Digital Mammographic Image Analysis. World Scientific Publishing, 7: 280-291.

Patricia McKenzie and Michael Alder, 1994. Unsupervised learning: the dog rabbit strategy. Proc. IEEE. Int. Conf. Neural Networks, Vol. 2.

Petrick, N., H.P. Chan and D. Wei, 1996. An adaptive density weighted contrast enhancement filtermammographic breastmassden. IEEE. Trans. Med. Imag., 15: 59-67.

Qian, W., L. Li, L.P. Clarke, F. Mao and R.A. Clark, 1998. Adaptive CAD modules for mass detection in digital mammography. In: Proc. 20th Ann. Int. Conf. IEEE. Eng. Med. Biol. Soc., 20: 1013-1016.

Richardson, W.B., Applying Wavelets to Mammograms. IEEE. Trans. Eng. Med. Biol., pp: 551-560.