Tumor Detection in Breast Ultrasound images

R. Vanithamani* and R. Dhivya**

Abstract: Breast ultrasound is becoming a popular screening modality for whole breast examination and is an important adjunct to mammography in breast cancer detection. As the breast ultrasound images are corrupted by speckle noise, it is challenging to develop a computer aided diagnosis system for detection of tumor in these images. The proposed approach uses a wavelet thresholding technique NeighShrink for reducing the speckle noise. Topographical watershed segmentation is incorporated and the features based on shape, texture, gradient and histogram are extracted and fed to AdaBoost classifier. To reduce the false positive and false negative detections the AdaBoost classifier is used in conjunction with binary logistic classifier. The proposed system is evaluated with 150 images (85-normal images, 65-abnormal images) and the performance is compared with neural network classifier. Experimental results reveal that the performance of the proposed system is better than the neural network classifier.

Index Terms: Ultrasound image; Speckle noise; Wavelet; NeighShrink; AdaBoost classifier.

1. INTRODUCTION

Early detection and treatment of breast cancer have been useful in reducing mortality rates. Mammography and Breast Ultrasound (BUS) are the two popular diagnostic modalities for the detection of breast tumor and for clinical examinations, mammography is mainly used. The drawbacks of mammography are high false positive rate, and it is not sensitive for women with dense breast tissue [1]. BUS is an adjunct imaging modality to mammography for detecting tumors in dense breasts [2] [3]. Clinicians prefer BUS, for initial clinical examinations for younger women, since it is radiation free. BUS can help to identify the difference between benign and malignant solid tumors, as well. The Computer-Aided Diagnosis (CAD) systems of BUS images have been proposed to accelerate the reviewing procedure and reduce oversight errors [4].

2. MATERIALS AND METHODS

Computer Aided Diagnosis (CAD) system involves four basic steps as illustrated in Fig. 1. BUS image suffers from the presence of granular pattern known as speckle, which deteriorates the resolution and overall quality of the image. Hence speckle suppression is an essential pre-processing step for improving the image quality. Segmentation of image divides the image into regions with similar properties. The feature extraction computes the features from an image or segmented region. These features will be used to

Figure 1: Steps in a CAD system

* Associate Professor, Department of Biomedical Instrumentation Engineering, Avinashilingam University for Women, Coimbatore, Email: vaniraj123@yahoo.co.in

** Assistant Professor, Department of Biomedical Instrumentation Engineering, Avinashilingam University for Women, Coimbatore, Email: dhivyaramasamy21@gmail.com
distinguish benign from malignant tumor in the classification stage. The classification is done by assigning data to one of the fixed number of possible classes.

2.1. Pre-Processing

Since speckle noise suppression is essential in ultrasound images, despeckling is included as a pre-processing step before segmentation and classification. In this paper wavelet thresholding technique is used as the pre-processing step and is given in Figure 2.

![Wavelet domain filter](image)

Figure 2: Wavelet domain filter

Noise reduction in wavelet domain is also known as wavelet thresholding. The Discrete Wavelet Transform (DWT) decomposes the image into approximation and detail subband coefficients. The detail subband coefficients are processed using NeighShrink. The wavelet domain image thresholding scheme NeighShrink [5] incorporates neighboring wavelet coefficients. The magnitude of the squared sum of all the wavelet coefficients within the neighborhood window of size 3x3 is taken into account for thresholding. Finally, Inverse Discrete Wavelet Transform (IDWT) is used to reconstruct the image.

2.2. Segmentation

Segmentation is an important step after pre-processing to extract information from the BUS image by dividing the image into region of interest and background [6]. The proposed method uses topographical watershed segmentation, since it produces a complete division of image into separated regions even if the contrast is poor and it is based on the concept of immersion.

2.3. Feature Extraction

After segmentation, the features are extracted to categorize the lesions into malignant or benign and several high precision features were suggested by researchers [7]. Since feature vectors affect the performance of the classification, an optimum feature set should have effective and discriminating features [8] [9]. The shape, texture, gradient based and histogram oriented descriptor features of BUS images are extracted and the features used in this study are listed below.

Morphological Features. Morphological features characterize the shape and size of an object of interest in the image. In this paper, morphological features such as area, perimeter and convex area are extracted.

GLCM features. Textural parameters are calculated from the Gray Level Co-occurrence Matrix (GLCM), of a pre-processed input image. In this study, Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Energy, Entropy, Homogeneity, Maximum probability, Sum of squares, Difference variance, and information measure of correlation features are extracted.

Tamura features. The six different Tamura features are coarseness, contrast, directionality, line-likeness, regularity and roughness.

MCHOG (Multi-Coordinate Histogram of Oriented Gradients) features. Among the gradient features, HOG features can depict the shape and Local Binary Pattern (LBP). LBP can reflect the texture information in a better way, many researchers coupled them together for object detection.
2.4. Classification
To classify the breast image as normal or cancerous, binary logistic regression classifier has been used. To reduce the false positive and false negative detections [10], AdaBoost classifier is used in conjunction with binary logistic classifier.

3. RESULTS AND DISCUSSION
The proposed algorithm is tested using MATLAB and for testing BUS images were obtained from the public ultrasound image database available at http://www.ultrasound-images.com.

3.1. Quality Metrics
The quality of a despeckled image is examined by the following standard image quality assessment metrics. The original image is represented by \( x(i, j) \) and the despeckled image is represented by \( \hat{x}(i, j) \).

*Peak Signal to Noise Ratio.* Peak Signal to Noise Ratio (PSNR) [11] is used to measure the difference between the original and despeckled images of size \( M \times N \), and is estimated using equation (1). It is expressed in decibel (dB).

\[
PSNR = 10 \log \frac{g^{2}_{\text{max}}}{MSE}
\]  

(1)

where \( g^{2}_{\text{max}} \) is the maximum intensity in the gray scale image and MSE is the Mean Squared Error and is given in equation (2).

\[
MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (x(i, j) - \hat{x}(i, j))^2
\]  

(2)

*Root Mean Square Error.* Root Mean Square Error (RMSE) [12] is the square root of the squared error averaged over \( M \times N \) window and is calculated using equation (3).

\[
RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (x(i, j) - \hat{x}(i, j))^2}
\]  

(3)

*Edge Preservation Index.* The edge preservation ability of the filter is assessed using Edge Preservation Index (EPI) [13] and is computed as in equation (4).

\[
EPI = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (\Delta x(i, j) - \bar{\Delta x})(\Delta \hat{x}(i, j) - \bar{\Delta \hat{x}})}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (\Delta x(i, j) - \bar{\Delta x})^2 (\Delta \hat{x}(i, j) - \bar{\Delta \hat{x}})^2}
\]  

(4)

where \( \Delta x(i, j) \) and \( \Delta \hat{x}(i, j) \) represents the edge images of original image \( x(i, j) \) and the despeckled image \( \hat{x}(i, j) \) respectively. The edge images are the high pass filtered versions of images \( x \) and \( \hat{x} \), obtained with a 3 \( \times \) 3 pixel standard approximation of the Laplacian operator. The \( \bar{\Delta x} \) and \( \bar{\Delta \hat{x}} \) are the mean intensities of \( \Delta x \) and \( \Delta \hat{x} \) respectively. If the edge is preserved well during despeckling process, the edge preservation index will be close to unity.

*Classification Accuracy.* Classification accuracy is calculated as the sum of correct classifications divided by the total number of classifications. The classification accuracy \( Ai \) is evaluated as in equation 5:

\[
Ai = \left( \frac{t}{N} \right) \times 100
\]  

(5)
where \( i \) depend on the number of samples correctly classified, \( t \) is the number of samples correctly classified, and \( N \) is the total number of samples.

### Table 1. PSNR, RMSE and EPI obtained for test image 1

<table>
<thead>
<tr>
<th>Filters</th>
<th>PSNR</th>
<th>RMSE</th>
<th>EPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speckled image</td>
<td>25.737</td>
<td>0.065</td>
<td>0.515</td>
</tr>
<tr>
<td>Anisotropic Diffusion</td>
<td>32.237</td>
<td>0.028</td>
<td>0.612</td>
</tr>
<tr>
<td>Soft thresholding</td>
<td>36.235</td>
<td>0.010</td>
<td>0.666</td>
</tr>
<tr>
<td>NeighShrink</td>
<td>37.812</td>
<td>0.008</td>
<td>0.865</td>
</tr>
</tbody>
</table>

For subjective evaluation, the output image is shown in Fig. 3. The quantitative results obtained in the pre-processing stage are listed in Table 1. The higher value of PSNR and lower values of RMSE indicates that the NeighShrink performs better than SRAD [14] and soft thresholding [15]. From the values of EPI, it is observed that the edge preservation ability of NeighShrink is better when compared to the other filters. Hence the image is pre-processed using the proposed hybrid method. To test the classifier accuracy a known set of 150 BUS images (85-normal, 65-abnormal) were used and the classification output images are shown in Figure 4 and Figure 5. The accuracy of the proposed classifier accuracy is compared with Neural Network and is given in Table 2. From the table it is observed that the accuracy of the proposed classifier is higher than that of the Neural Network Classifier.

### 4. CONCLUSIONS

In this paper, the image is pre-processed to reduce the speckle noise for feature preservation. For effective segmentation of lesions, topographical watershed algorithm is employed and features are extracted. The
extracted features are fed as input to the classifier. False positive and false negative detections have been reduced by using an adaptive boosting classifier in addition to the binary regression classifiers. It is observed from the results that the performance of the proposed classifier is encouraging.

References


