Classification Of Micro Calcification In Digital Mammogram Using Tetrolet Transform

ISSN 2319-9725

P. Indra
Assistant Professor, Government College of Engineering,
Salem, India.

Abstract: In this study an efficient method is presented to classify the microcalcification severity in digital mammogram images. The leading cause of cancer mortality among women arises due to breast cancer. Digital mammogram plays an important role for cancer diagnosis. The early sign of breast cancer is the appearance of microcalcifications clusters on mammogram images. In this proposed method Tetrolet transform is utilized as feature extraction technique. Tetrolet transform also named as adaptive Haar transform, in which features are extracted from the Tetrolet decomposed mammogram. The extracted features are fed as input to the classifier. The classification of microcalcification clusters into benign or malignant is done by k nearest neighbor classifier (KNN). The proposed Tetrolet based classification of microcalcification approach achieves satisfactory performance than the conventional Haar transform.

Keywords: Tetrolet transform, microcalcification, mammogram, benign, malignant, nearest neighbor classifier, energy features.
1. Introduction:

Computer aided detection (CAD) diagnosis system is implemented for classification of microcalcification (MC) on digital mammogram images based on support vector machine (SVM) in Dheeba and selvi (2011), which constructs a set of hyper planes in a vast dimensional space. Law’s texture energy is extracted as feature from the Region of interest (ROI) mammogram image. The extracted feature is given as input to the SVM classifier for MC classification. Veldkamp et al (1996) described importance of segmentation on classification of microcalcifications in digital mammography. Contrast of microcalcifications can be used to classify benign and malignant types. Mean and maximum contrast, with and without correction for microcalcification size are used for different measures for contrast. For classification the k-Nearest-Neighbor method is used and for testing the “leave-one-out-method”. From this analysis the segmentation strongly influences classification.

Geetha et al (2008) presented a novel approach that is New Particle Swarm Optimization for Feature Selection and Classification of Microcalcifications in Mammograms. The Spatial Gray Level Dependence Method (SGLDM) is used for feature extraction. The selected features are fed to a three-layer Back propagation Network hybrid with New Particle Swarm Optimization (BPN) for classification. And the Receiver Operating Characteristic (ROC) analysis is performed to evaluate the performance of the feature selection methods with their classification results.

Weidong et al (2007) described a novel computer-aided diagnosis method for detection and classification of microcalcifications (MCs) based on discrete wavelet transform (DWT) and adaptive neuro fuzzy inference system (ANFIS). DWT is used to extract the high-frequency signal of the images, and thresholding with hysteresis was applied to locate the suspicious MCs. Then, filling dilation was applied to segment those desired regions. During the detection, ANFIS is used to adjust the parameters, making the CAD algorithm more adaptive. Finally, the suspicious MCs are classified with multilayer perception.

A novel classification of microcalcification approach is implemented in (Ren et al, 2013).optimized decision making is establish for accurate and effective mammogram classification, where classification of microcalcification is achieved by improved support vector machine (SVM) scheme. Eltoukhy MM et al., (2010), presented a Texture analysis based on curvelet transform for the classification of mammogram tissues. The most discriminative texture features of regions of interest are extracted. Then, a nearest neighbor
classifier based on Euclidian distance is constructed. It consists of two steps, detecting the abnormalities and then classifies the abnormalities into benign or malignant tumors.

M.Suganthi and M.Madheswaran (2009) presented a computer aided decision support system for an automated diagnosis and classification of breast tumor using mammogram. It differentiates two breast diseases namely benign masses and malignant tumors. From the preprocessed mammogram image, texture and shape features are extracted. The optimal features can be extracted by using a feature selection scheme based on the multi objectives genetic algorithm. A system based on fuzzy-neural and feature extraction techniques for detecting and diagnosing microcalcifications patterns in digital mammograms is developed by Brijesh Verma and John Zakos (2001). The back propagation technique is used for classification of features into benign or malignant

2. Tetrolet Transform:

A new adaptive Haar wavelet transform, called Tetrolet transform provides efficient image representation. Tetrolet is wavelet based efficient and effective transform, whose supports tetrominoes which are shapes are made by connecting four equal-sized squares. The corresponding filter bank algorithm is simple but enormously effective. In every level of the filter bank algorithm the low-pass image is divided into $4 \times 4$ blocks. Then in each block a local Tetrolet basis is determined, which is adapted to the image geometry in this block. In discrete wavelet transformation (DWT), horizontal and vertical directions only preferred, which fails to achieve optimal results with images that contain geometric structures in other directions. To improve the treatment of orientated geometric image structures, Tetrolet transform is introduced by Krommweh, J. (2010).

In the Haar filter bank, the low-pass filter and the high-pass filters are just given by the averaging sum and the averaging differences of each four pixel values which are arranged in a $2 \times 2$ square. Then we can determine the low-pass part $a^1 = \{a^1(i,j)\}_{i,j=0}^{N-1}$ with

$$ a^1[i, j] = \sum_{(i', j') \in I_{i,j}} a [i', j'] $$

(1)
as well as the three high-pass parts for \( l=1,2,3 \) \( w^l_i = (w^l_i[j,k])_{j,k=0}^{N-1} \) with

\[
w^l_i[j,k] = \sum_{(i',j') \in I_{i,j}} \mathcal{L}[i',j'][i',j']
\]

(2)

where the coefficients \( \in [l,m], l,m = 0,\ldots,3 \) are entries from the Haar wavelet transform matrix

\[
W := (e[l,m])^3_{l,m=0} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}
\]

(3)

The decomposition algorithm of Tetrolet transform is processed by four ways, which are as follows.

a. Divide the image into 4 × 4 blocks.

b. Find the sparsest Tetrolet representation in each block.

c. Rearrange the low- and high-pass coefficients of each block into a 2 × 2 block.

d. Store the Tetrolet coefficients (high-pass part).

e. Apply step 1 to 4 to the low-pass image.

Input image is considered as \( a^0 = (a^l_{i,j})_{i,j=0}^{N-1} \) with \( N = 2^J \), \( J \in \mathbb{N} \) then \( J-1 \) levels can be applied.

In the \( r \)-th level, \( r \leq J-1 \), the following computations are performed.

**Step 1:** Divide the low pass image \( a^{r-1} \) into blocks \( Q_{i,j} \) of size 4x4, \( i,j = 0,\ldots, N^{4^{-r}}-1 \).

**Step 2:** In each block of \( Q_{i,j} \) there are 117 admissible tetromino coverings are considered i.e. \( c=1,\ldots,117 \). For each tiling \( c \) Haar wavelet transform is applied to the four tetromino subsets \( I^{(c)}_s \), \( s=0,1,2,3 \). In this way, four low-pass coefficients and 12 Tetrolet coefficients are obtained for each tiling of \( c \). More precisely, in \( Q_{i,j} \) we compute analogously to (1) and (2) the pixel averages for every admissible tetromino configuration \( c=1,\ldots,117 \) by

\[
a^{r,(c)} = (a^{r,(c)}[s])_{s=0}^3
\]
\[ a^{r,(c)}[s] = \sum_{(m,n) \in I^c_{\ell}} \epsilon[0,L(m,n)] a^{r-1}[m,n], \tag{4} \]

as well as the three high pass parts \( l = 1, 2, 3 \)

\[ w^r_{l,c} = (w^r_{l,c})^\ast = \sum_{(m,n) \in I^c_{\ell}} \epsilon[l,L(m,n)] a^{r-1}[m,n], \tag{5} \]

where the coefficients \( \epsilon[l,L(m,n)] \) are given in (3) and where \( L \) is the bijective mapping, which relating the four index pairs \( (m,n) \) of \( I^c_{\ell} \) with the values 0, 1, 2, and 3 in descending order. That means, by the one-dimensional indexing \( J(m,n) \) the smallest index is identified with the value 0, while the largest with 3.

Then the covering \( c^* \) is chosen such that the \( l^{-1} \)-norm of the 12 Tetrolet coefficients becomes minimal

\[ c^* = \arg \min_c \sum_{l=1}^{3} \left| \sum_{l=1}^{3} w^r_{l,c} \right| = \arg \min_c \sum_{l=1}^{3} \sum_{s=0}^{3} \left| w^r_{l,c} [s] \right|. \tag{6} \]

Hence, for every block \( Q_{i,j} \) we get an optimal Tetrolet decomposition \( [a^{r,(c^*)},W^{r,(c^*)}_1,W^{r,(c^*)}_2,W^{r,(c^*)}_3] \). By doing this, the local structure of the image block is adapted. The best covering \( c^* \) is a covering whose tetrominoes do not intersect an important structure like an edge in the image \( a^{r-1} \). Because the Tetrolet coefficients become as minimal as possible a sparse image representation will be obtained. For each block \( Q_{i,j} \) the covering \( c^* \) is stored such that has been chosen, since this information is necessary for reconstruction. If the optimal covering is not unique, then the tiling \( c^* \) has taken that already chosen most frequently in the previous blocks. Thus, the coding of the used coverings becomes cheaper.

**Step 3:** In order to be able to apply further levels of the Tetrolet decomposition algorithm, the entries of the vectors \( a^{r,(c^*)} \) are arranged and \( W^{r,(c^*)}_l \) into \( 2 \times 2 \) matrices using a reshape function \( R \),

\[ a^r_{[Q_{i,j}]} = R(a^{r,(c^*)}) = \begin{pmatrix} a^{r,(c^*)}[0] & a^{r,(c^*)}[2] \\ a^{r,(c^*)}[1] & a^{r,(c^*)}[3] \end{pmatrix}. \tag{7} \]
and in the same way $w^l_q \left| Q_{i,j} \right. = R(w^l_j (c^s))$ for $l = 1, 2, 3$. For an efficient representation in the next level, a suitable arrangement of the low-pass values is essential. That means, the order of labeling the tetrominoes of $c^s$ in each block $Q_{i,j}$ by $s = 1, 2, 3$ is very important. The labeling should be done in a way, such that the geometry of the tiling is suitably mapped to

$$\left( \begin{array}{ccc} 0 & 2 \\ 1 & 3 \end{array} \right).$$

Therefore the four shapes of the chosen partition $c^s$ are labeled by comparing with the square case.

Among the 24 possibilities to label the four tetrominoes, the numbering with the highest correlation with the Haar partition is preferred. Tetrominoes with corresponding low-pass image block is shown in Figure 1. For comparison of different labeling of the four tetrominoes, the number of deviations from the Haar wavelet tiling is computed, i.e. the number of small squares in a block $Q_{i,j}$ is counted, where the label differs from the label of these squares in the Haar wavelet tiling the order with minimal deviations is applied. This optimal order needs not to be unique. The labeling in Figure 1(a) would lead to a distorted low-pass image, while 1(b) shows a reasonable order.

![Figure 1: Example of labeling tetrominoes with corresponding low-pass image block. (a) Bad order (ten deviations from square case), (b) best order (eight deviations).](image)

**Step 4:** After finding a sparse representation in every block $Q_{i,j}$ for $i, j = 0, \ldots, \frac{N}{4^l} - 1$, the low-pass matrix is stored.

$$a^l = \left( a^l_q \right)_{i,j=0}^{\frac{N}{4^l} - 1}$$

and the high pass matrices
\[ w^r_j = \left( w^r_{j[l]} \right)_{l,j=0}^{N-1}, \quad (9) \]

\[ I = 1, 2, 3, \) replacing the low pass image \( a^{r-1} \) by the matrix

\[
\begin{pmatrix}
  a^r & w^r_2 \\
  w^r_1 & w^r_3 \\
\end{pmatrix}
\] \quad (10)

After a suitable number of decomposition steps, a shrinkage procedure to the Tetrolet coefficients in order to get a sparse image representation is applied.

\[ S_\lambda(x) = \begin{cases} 
 x, & |x| \geq \lambda, \\
 0, & |x| < \lambda, 
\end{cases} \quad (11) \]

For the reconstruction of the image, low-pass coefficients are needed from the coarsest level and the Tetrolet coefficients as usual. Additionally, the information about the respective covering in each level and block is necessary.

### 3. Proposed Method:

The proposed classification of microcalcification system is integrated by two stages, which are feature extraction and classification stage. The block diagram of proposed classification system is shown in figure 2.
3.1. Feature Extraction Stage:

The first stage of the proposed classification of microcalcification system is feature extraction, in which the input data is transformed into reduced form of features. Digital mammogram image is given as an input of feature extraction stage. At first, the input image is decomposed by Tetrolet transform in predefined level of decomposition. The decomposed coefficient contains lower and higher frequency sub bands, in which the higher sub bands are considered for further feature extraction process. In this feature extraction process, mean, standard deviation and variance of all higher sub band taken as a feature. All the extracted features are stored in the database for classification. This process repeated for all training images in the proposed classification system.

3.2. Classification Stage:
In this proposed classification stage KNN classifier is used for classification. At the start of the classification process, the unknown image is decomposed by Tetrolet transform in predefined level of decomposition. The decomposed coefficient contains lower and higher frequency sub bands, in which the higher sub band of unknown image is given to feature extraction process. In this process, mean, standard deviation and variance are evaluated and taken account into the feature vector for classification. The extracted unknown feature vector and stored feature database is fed to the KNN classifier, where unknown image is classified into benign or malignant based on minimum distance between feature space unknown and feature database. To achieve maximum classification accuracy Euclidean distance measure is used in KNN classifier.

4. Result And Discussion:

The experimental result of the proposed mammogram classification system is evaluated by MIAS database images. From the MIAS database 12 benign and 13 malignant images, totally 25 microcalcification images are used for the result evaluation. Generally there are two abnormalities are available in MIAS database, which are mass and microcalcification. The proposed approach concentrates the microcalcification image classification. The ROI of the mammogram image is used for the result analysis, the size of the ROI image is 256x256. Figure 3 shows the microcalcification images of MIAS database.

![Figure 3: Microcalcification images of MIAS database (a) benign and (b) malignant.](image-url)
Table 1: Classification accuracy of proposed microcalcification system using Tetrolet transform

<table>
<thead>
<tr>
<th>Level of decomposition</th>
<th>Classification accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>1</td>
<td>91.67</td>
</tr>
<tr>
<td>2</td>
<td>91.67</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2: Classification accuracy of microcalcification system using conventional Haar transform

<table>
<thead>
<tr>
<th>Level of decomposition</th>
<th>Classification accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>83.33</td>
</tr>
<tr>
<td>3</td>
<td>83.33</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>91.67</td>
</tr>
</tbody>
</table>

The performance of proposed classification of microcalcification system evaluated in terms of classification accuracy. Table 1 shows the classification accuracy of proposed microcalcification system using Tetrolet transform. Table 2 shows the classification accuracy of microcalcification system using conventional Haar transform. It is clearly observed from table 1 and table 2 the maximum average classification accuracy 100% is achieved at level of 5.
decomposition 4 in the proposed Tetrolet transform and 92.31% is achieved in conventional haar transform. Figure 4 shows graphical representation of average classification accuracy of proposed method and conventional haar transform. From that result the proposed microcalcification system achieve better classification accuracy.

![Figure 4: Average classification accuracy of proposed method](image)

5. Conclusion:

In this study classification of microcalcification system is proposed in digital mammogram images. The proposed approach is employed based on feature extraction system with KNN classifier. In order to obtain the feature set, the given mammogram image is decomposed by tetrolet transform. Hence, higher and lower tetrolet transformed coefficients are attained from tetrolet decomposition, in which higher band coefficients alone considered for further process. Features such as mean, standard deviation and variance are evaluated and taken account into the features form higher band tetrolet coefficients. The extracted features are given into the KNN classifier, where unknown mammogram image is classified into benign or malignant by measuring minimum distance between unknown feature vector and stored feature set. The proposed tetrolet based mammogram classification system is compared with state of the conventional haar based approach and 100 % classification accuracy is achieved.
References:


