Novel Margin Features for Mammographic Mass Classification

Ali Bagheri-Khaligh, Ali Zarghami, and M. T. Manzuri-Shalmani
Department of Computer Engineering
Sharif University of Technology
Tehran, Iran
{abagheri, zarghami}@ce.sharif.edu, manzuri@sharif.edu

Abstract— Computer-Aided Diagnosis (CAD) systems are widely used for detection of various kinds of abnormalities in mammography images. Masses are one type of these abnormalities which are mostly characterized by their margin and shape. For classification of masses proper features are needed to be extracted. However, the number of well-known features for describing margin is much fewer than geometrical, shape, and textural ones. In addition, most of the existing margin features are highly dependent on segmentation accuracy. In this work, new features for describing margin of masses are presented which can handle inaccuracies in segmentation. These features are obtained from a set of waveforms by wavelet analysis among each of them. For each of these waveforms an edge probability distribution is computed. Then, features are extracted from these probability distributions. Although these features are called margin features, they are highly related to the texture of the mass. For experimentation DDSM dataset was used and our simulations show the great performance of these features in classification of masses.

Index Terms— Mammography; mass classification; margin; feature extraction; computer-aided diagnosis.

I. INTRODUCTION

Cancer is one of the major causes of death in the world. Breast cancer is the most common type of cancer and the second cause of death by cancer among women in United States [1]. Diagnosis of breast cancer in early stages increases the chance of full recovery. The most effective method for detection of breast cancer before it becomes clinically intense is mammography [2]. Nowadays, Computer-Aided Diagnosis (CAD) systems have a great function in detecting various kinds of abnormalities in mammography images [3].

According to Breast Imaging Reporting and Data System (BI-RADS) [4] a mass is: “a space-occupying lesion seen in two different projections.” Masses usually characterized by their shape and margin [5]. Different shapes of masses can be categorized as round, oval, lobular, and irregular while irregular masses have the most probability of being malignant. Mass margins are categorized as well-defined, obscured, micro-lobular, ill-defined, and speculated. Most of the benign masses have well-defined margins [6] and as the same order they have been listed, the probability of malignancy increases.

There are two common approaches based on feature extraction for classification of masses. Some extract various kinds of features from Region of Interest (ROI) which contains the mass, such as [7] and the others extract features from a segmented mass such as [8]. One of the first things that radiologists do in confronting masses is to examine mass margin [9]. However the number of existing margin features for characterizing a mass is far fewer than geometrical, shape, or textural ones. This might be because of the fact that the qualification of these features is highly dependent on the accuracy of segmentation. In [10, 11] some features related to margin are presented.

In this work efficacious features which are not dependent on the accuracy of segmentation are introduced. Although it is proper to have a accurate segmentation, but our features can handle inaccuracies as well because we extract features from a set of waveforms which can capture abruptness inside and outside of the mass. First, preprocessing and segmentation are explained on the images. Then, constructing waveforms and extracting features will be explained in detail. Fig. 1 shows the diagram of the proposed method.

II. PROPOSED METHOD

In this section each step of the diagram shown in Fig. 1 is explained in detail. The image enhancement step which is presented here can be used in most of the CAD systems working on mammograms. For segmentation a relatively new method based on level set was used. After these steps, margin feature extraction will be explained.

A. Preprocessing and Segmentation

Before analyzing margin a set of preprocessing steps are needed. Currently, there are some methods for automatic mass detection and extracting corresponding Region Of Interest (ROI) in mammograms [12, 13]. However, for avoiding complexity and propagation of detection error we extracted these ROIs manually. From this point forward, all steps are accomplished on the ROI. In order to eliminate artifacts such as bright spots, artifact removal which is presented in [14] is used. This method eliminates isolated peaks in ROI histogram. Then, for making ROI smoother, median filter of size 3-by-3 is applied.

Since mammograms usually have low contrast, an
Fig. 1. General diagram of the proposed scheme for mass margin feature extraction

Enhancement is essential in most of the mammography image processing techniques. We exploited power law transformation for contrast enhancement as

\[ I_{\gamma} = J^{\frac{1}{\gamma}} \]

where \( \gamma \) was set as 2. \( I \) and \( J \) are the images after and before enhancing, respectively. In the next step, the edges are enhanced using an edge-enhancing anisotropic diffusion filter [15, 16]. This filter constructs diffusion tensor \( D \) by computing its eigenvalues. Let \( \mathbf{u}_{\sigma} \) and \( \mathbf{v}_{\sigma} \) be eigenvectors of \( D \), then they have to be such that \( \mathbf{u}_{\sigma} \perp \nabla \mathbf{u}_{\sigma} \) and \( \mathbf{v}_{\sigma} \parallel \nabla \mathbf{u}_{\sigma} \), where \( \mathbf{u}_{\sigma} \) is the input image smoothed by a Gaussian filter with standard deviation \( \sigma \). We set \( \sigma \) as 0.3 because we have already smoothed the image by a median filter. Edge-enhancing diffusion filter chooses eigenvalues of \( D \) as

\[ \lambda_1 = g(s) \lambda_{\sigma}, \quad \lambda_2 = 1 \]

where \( \lambda_{\sigma} \) is the largest eigenvalue of the structure tensor \( J_{\rho} \) which is smoothed version of (a Gaussian filter with standard deviation \( \rho \) ) of \( \nabla \mathbf{u}_{\sigma} (\nabla \mathbf{u}_{\sigma})^T \). \( \rho \) was also set as 3. Fig. 2-a shows a mass before preprocessing and the result of contrast enhancement and diffusion filter is shown in Fig 2-b.

Segmentation is a very challenging task for masses because of their ambiguous margins. A method based on level set were used for segmentation [17], the interesting part of this method is using a sparse representation technique named Morphological Component Analysis (MCA) [12, 18] to obtain a shape constraint initialization. Fig 3-a and 3-b are smooth and texture components of Fig 2-b, obtained by MCA. Followed by evolving of shape constraint curve [19], nine feature maps along with vector valued Chan-Vese [20] were used to compute final segmentation result as

\[ \frac{\partial \phi}{\partial t} = \delta (\mu \text{div}(\nabla \phi) - g_{\delta} \sum_{i=1}^{N} \omega_i \lambda_i^c (u_{i+} - c_i^c)^2 - \sum_{i=1}^{N} \omega_i \lambda_i^c (u_{i-} - c_i^-)^2) \]

where \( \omega_i \) is the weight of \( i \text{th} \) feature map. \( \phi \) is the level set function, \( c_i^c \) and \( c_i^- \) are the average of \( u_{i+} \), outside and inside of the boundary, respectively. \( g_{\delta} \) is a distance function based on Gaussian distribution to make sure that the final curve is in annular zone which is also presented in [17]. In the end, final segmentation result is refined. This refinement was added to the original segmentation method in order to prevent twisted margins and intersecting waveforms which will be explained later. This refinement is accomplished on zero level set contour with \( a=0 \) in [21]. Shape constraint and final segmentation result of Fig. 2-b are shown in Fig. 3-c and 3-d, respectively.

B. Margin Analysis

When the segmented boundary is available, analysis of margin can be carried out. First, a set of waveforms should be extracted. The idea of radial lines (not exactly the waveforms presented here) was also presented in [22] but for
a different purpose. In [22] no segmentation was accomplished and radial scan lines were used to determine minimum and maximum radius of a mass. Let \( \partial \Omega \) be the margin of mass. The path of the margin is traversed and waveforms are placed every 5 pixels (margin is downsampled at the rate 5). \( Q_1 \) is the coordinate of the center of gravity of the shape and is computed as

\[
Q_1 = \frac{1}{\Omega} \sum_{i,j \in \Omega} (i, j),
\]

where \( \Omega \) indicates region of the shape. If \( Q_2 \) is the pixel which is selected while traversing margin, then waveform \( l \) will have the same direction as the line \( Q_1 Q_2 \). The length of \( l \) is computed by

\[
|l| = 2^{\left\lfloor \log_2 r \right\rfloor + 1},
\]

where \( r \) is the length of \( Q_1 Q_2 \). Thus, the length of \( l \) is a power of two and appropriate for wavelet transformation. So far, length and direction of \( l \) have been determined, now it should be located such that its middle element is \( Q_2 \). Fig. 4 shows an example of two possible waveforms \( l_1 \) and \( l_2 \), where \( Q_1 \) and \( Q_2 \) are their corresponding pixels on the margin. A real example of constructing waveforms is also shown in Fig. 5 where the rate at which the margin was downsampled, was set as 10 to capture a visually good representation. Note that as mentioned earlier, the margin was refined so that it gets closer to a convex shape.

The waveforms are the basis for the margin analysis of masses in the proposed scheme. From this point on, we focus on analysis of these signals to obtain information about margin. For this purpose, one dimensional wavelet transformation was applied to each of them using haar filter. Let \( n \) be the number of waveforms and \( l_i \) be the \( i \)th waveform. After \( L \) level decomposition, wavelet coefficients \( \{d_1, d_L, d_{L-1}, \ldots, d_1\} \) are obtained. \( L \) is computed as

\[
L = \log_2 |l| - 2,
\]

where \( |l| \) is the length of the \( i \)th waveform. We only used the detail coefficients and because of ambiguous texture of mammograms \( d_1 \) was dropped. The values of each level are normalized separately to be between 0 and 1.

Using wavelet coefficients, edge probability distribution is computed for each \( l_i \). It is worth to mention that we are not looking for an edge, instead the probability distribution of edges is computed. Let \( D \) be the length of \( d_i \), then the Probability Mass Function (PMF) of edges has also \( D \) elements. The edge probability \( P_j \) (\( j \)th element of \( P_i \)) is
Fig. 6. Schematic model of computing edge probabilities of each waveform

computed as

$$P_i = S_i / \sum_{j=1}^{L} S_j$$  \hspace{1cm} (6)

$$S_j = \sum_{i=2}^{L} w_i d_i^{(j,i)}$$ \hspace{1cm} (7)

$$u(t, j) = \left[ \frac{j}{2^{i-2}} \right]$$ \hspace{1cm} (8)

where $d_i$ is the $i$th element of the $j$th level of details in the decomposition. $w_i$ is the weight which is assigned to the $j$th level of details and computed as

$$w_i = t / \sum_{i=2}^{L} t$$ \hspace{1cm} (9)

The lowest weight is assigned to the second level and the highest weight is assigned to the $L$th level of decomposition, and $\sum_{i=2}^{L} t = 1$. The reason of choosing these weights is that energy in higher levels is more than lower ones in the decomposition of each waveform. In addition, the effect of noise is also more sensible in lower levels. For more clarification of Eq. (6-8), a schematic example is shown in Fig. 6 where $L=4$ and $D=8$. Each $P'$ is computed by multiplying corresponding elements by their weights and then summing in the direction of arrows. Finally, for the sake of the constraint $\sum_j P' = 1$, elements of $P$ are normalized to their sum which is stated in Eq. (6).

Before ending discussion about computing PMF, it is worth to mention that we computed a discrete probability distribution. One can go further and compute a distribution using techniques such as Parzen windows and K-Nearest Neighbors (KNN) [23]. For our purpose this simple discrete distribution is adequate, but with the proper kernels and increasing number of bins more qualified features can be obtained.

### C. Feature Extraction

As mentioned earlier, there are two common approaches for mass classification: 1) ROI-based 2) segmentation-based. In the latter case accurate geometrical and shape features can be extracted [24]. For each distribution function $P_i$ which was obtained from $l_i$, following parameters are computed

- **Kurtosis**
  $$K_i = \frac{(1 / D) \sum_{j=1}^{D} (P_i - \bar{P})^4}{((1 / D) \sum_{j=1}^{D} P_i - \bar{P})^2} - 3.$$ \hspace{1cm} (10)

- **Entropy**
  $$E_i = -\sum_{j=1}^{D} P_i \log P_i.$$ \hspace{1cm} (11)

- **Index of the maximum probability**
  $$R_i = (\arg \max_j P_j) - \frac{D}{2}.$$ \hspace{1cm} (12)

Kurtosis measures how much peaked is a probability distribution and for well-defined margins it gets higher values. In general, well-defined margins have an abrupt transition along a waveform. It is also expected that well-defined margins have lower entropy. Index of the maximum probability is another measure which is shifted to be zero on the margin, positive for outside, and negative for inside of the margin. This index is used to capture variations of most probable edge places among the margin. Each of the above vectors has $n$ elements ($n$ is the number of waveforms). The following features can be obtained from these vectors

- Mean and standard deviation of $K$.
- Mean and standard deviation of $E$.
- Total variation of $R$: $\sum |R'_j|$.

The features are not dependent on segmentation accuracy, because they are obtained along waveforms and if the boundary changes by some deflation or inflation, they will be the same.

### III. EXPERIMENTAL RESULTS

In order to test margin features the Digital Database for Screening Mammography (DDSM) [25] was used. Mammograms which were used were digitized by Lumisys scanner with spatial resolution of 50 $\mu$m and depth of 12-bit per pixel. Each case has an associated document which contains the expert radiologists’ opinions about the abnormalities (if there is any) and their characteristics. For masses an approximate boundary, type of shape and margin, pathology, and some other information are included.
Before any classification task, it is worth to say that $K$, $E$, and $R$ can be used directly by radiologists to examine the margin. Fig. 7 shows these signals for two kinds of mass margins. The right column pertains to a well-defined (circumscribed) margin and the left column is for an ill-defined margin. It can be seen that in a well-defined mass margin $R$ is smoother, $E$ has lower values, and $K$ has higher values. We expected that the edge probability distribution in each waveform of a well-defined mass was peakier than an ill-defined one; higher kurtosis values and lower entropy values in Fig. 7 show it. The higher variations of $R$ in an ill-defined mass margin show erratic boundaries of this kind of margin which is exactly the definition of this type. In order to show the discrimination that margin features make between different types of margins, a two-class classification task for two types of margins was accomplished using 3-nearest neighbor classifier. We used 25 cases for each of well-defined and ill-defined classes, 10 cases of each class were used for training phase and the other cases were used for test phase. The classification was repeated 5 times and each time training cases were selected randomly. Table I shows the result of this classification task. Recall, Precision, and accuracy for each class were computed by

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (13)
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (14)
\]

\[
\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (15)
\]

As it is shown in Table I, these two different types of margin with minimum and maximum probability of malignancy were separated properly.

We also tested our features in classification of masses as benign and malignant. First, the 25 features for texture and shape in [8] were extracted and Support Vector Machine (SVM) classifier [23] with polynomial kernel (degree of 3) and soft margin were used for classification. Leave-one-out cross-validation was used to find optimal value of the soft margin parameter. Then, we added our 5 features and the classification task was accomplished again using 30 features. An aggregate of 100 mammograms including 50 benign and 50 malignant cases were used for evaluation. For training phase 20 cases of each class were selected randomly while the other ones were used for test. The task was repeated 5 times. Table II shows the result of classification tasks with and without using margin features. As it can be seen adding margin features improve the accuracy significantly. Note that here our purpose is not to obtain the best possible accuracy; instead, our emphasis is on the improvement of it by using margin features.

### IV. CONCLUSION

In this paper a set of margin features for masses were introduced which have the least dependency on segmentation accuracy. The number of existing margin features for masses is much fewer than shape and texture ones while the margin type is one of the most important characteristics that radiologists diagnose based on it. Margin features presented in this work, obtained from a set of waveforms and it is possible to extract even more qualified features from them.
These features showed a great improvement of accuracy in classification of masses.

REFERENCES


