Classification of Breast Masses Using Color Doppler Flow Imaging

Yingtao Zhang¹, H. D. Cheng¹,2, Yan-Hui Guo¹, Jiawei Tian³, Jianhua Huang¹, Bo Liu¹, Yanxin Su³

¹School of Computer Science and Technology, Harbin Institute of Technology. Harbin 150001. China
²Department of Computer Science, Utah State University. Utah . U.S.A
³Department of Ultrasound, Second Affiliated Hospital of Harbin Medical University. Harbin 150086. China

Abstract

Color Doppler flow imaging takes a great value in diagnosing and classifying benign and malignant breast tumors. However, previous computer aided diagnosis (CAD) systems have not analyzed blood flow information. In this paper, we present a new approach to classify benign or malignant masses based on two dimensional (2D) ultrasound plus Color Doppler flow imaging in CAD system. In the proposed method, geometric features, textural features and blood flow features are extracted automatically. The experimental results show that the proposed system can improve true positive (TP) and decrease false positive (FP) detection rates greatly, and it will useful for breast cancer control.

Keywords: Breast cancer, Color Doppler Flow Imaging (CDFI), blood flow features, support vector machine (SVM).

1. introduction

Breast cancer is the most prevalent cancer among women [¹]. Ultrasound examination is non-invasive, painless, real time and safe for the patient [²]. Thus, there is an increased interest in the use of ultrasound in breast cancer diagnosis [³].

In medical ultrasound image, flow data is usually recorded with B-mode data. Tumor vascularity particularly has played an important role in promoting cancer growth, invasion, and metastasis [⁴]. Most malignant tumors have copious blood supply. The formation of new blood vessels is the vital for rapid growth of solid malignant tumors. Tumors with volume of 1-2mm³ or more have to acquire a vascular supply in order to grow [⁵]. Doppler spectral analysis is a valuable indicator for distinguishing benign and malignant breast neoplasm.

In this paper, Color Doppler flows are employed to differentiate the benign and malignant lesions. A novel breast mass classification system is proposed, whose block diagram is illustrated as Figure 1.

A detailed description of this paper is organized as follows. In Section 2, the procedure of preprocessing is discussed,
which including mass segmentation as described in (Huang et al. 2008)\textsuperscript{[6]}, for segmenting the suspicious areas from 2D ultrasonography and blood detection from color Doppler flow imaging. In Section 3, we analysis and extract of three kinds of features: geometric features, textural features and blood flow features. In Section 4, we briefly introduce the procedure of classification using SVM. Finally, the results and the conclusions are summarized in Section 5.

1. Preprocessing

Here, we briefly introduce the preprocessing methods used in our CAD system. The preprocessing consists of two stages. The first stage is segmentation of candidate lesions from 2D ultrasonography; and the second one is detection of blood flow information from color Doppler sonography.

The method of breast ultrasound image segmentation is based on homogeneity histogram. Texture and edge features are used to compute homogeneity. Both global and local information is considered, which has not been achieved by previous algorithms. The image is divided into homogeneity subset and the non-homogeneity subset according to the threshold computed from the maximum entropy principle. The two subsets are segmented separately thereafter. Figure 2 illustrates an example of mass segmentation.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig1.png}
\caption{The diagram of the proposed system}
\end{figure}

Doppler color flow systems assign a given color to the direction of flow; red is flow toward, and blue is flow away from the transducer, as shown in Fig 3. Therefore it is very easy to detect the blood flow.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2.png}
\caption{An example of mass segmentation.}
\end{figure}
3. Feature Extraction

A key stage of CAD (computer-aided diagnosis) systems is feature analysis and extraction. Many useful features of masses have been employed in previous CAD systems \[7\]. In this study, we have utilized several blood flow features, which reflect the blood-supplying characteristic of the masses. The typical features by proposed system are listed in table1.

<table>
<thead>
<tr>
<th>Feature Sub-Space</th>
<th>Features</th>
</tr>
</thead>
</table>
|                   | Geometric Features | 1. compactness  
|                   | 2. border roughness  
|                   | 3. aspect ratio |
|                   | Textural Features | 1. Gray Asymmetry  
|                   | 2. Grads Asymmetry  
|                   | 3. Energy  
|                   | 4. Gray Mean  
|                   | 5. Grads Mean  
|                   | 6. Gray Variance  
|                   | 7. Grads Variance  
|                   | 8. Correlation  
|                   | 9. Gray Entropy  
|                   | 10. Grads Entropy  
|                   | 11. Intertia |

In color Doppler, progressively increasing velocities are encoded in varying hues of either red or blue. The more dull the hue, the slower the velocity. The brighter the hue, the faster the relative velocity is. High velocities away from the transducer will appear as lighter shades of blue, and higher velocities toward the transducer will be represented by lighter shades of red, or even yellow. Low velocity flow will be represented by darker shades of these colors.

The flow spectrum features include peak systolic velocity \( V_{\text{max}} \), diastolic terminal velocity \( V_{\text{min}} \), mean velocity \( M_{\text{V}} \), resistant index \( R_{\text{I}} \) and pulsatile index \( P_{\text{I}} \). It is proved that \( V_{\text{max}}, P_{\text{I}}, R_{\text{I}} \) and \( M_{\text{V}} \) are significantly higher in the malignant lesions than those in the benign ones. Their definitions are given as follows.

1) Peak systolic velocity

\[
V_{\text{max}} = C \max_{k} \left( \frac{\sum_{i=1}^{N} H_{(u,v)}(k) - H_{(u,v)}(k-1)}{\Delta t} \right)
\]  

Where \( R \) is the suspected region of blood flow in CDFI video, and \( H_{(u,v)}(k) \) is hue value of \((u,v)\) in the frame \( k \). Generally, the velocity of human artery is below 0.22m/s. So we set a coefficient \( C \). \( N \) is the number of frames in one heart period, and \( \Delta t \) is delay between two frames.

2) Diastolic terminal velocity

\[
V_{\text{min}} = C \min_{k} \left( \frac{\sum_{i=1}^{N} H_{(u,v)}(k) - H_{(u,v)}(k-1)}{\Delta t} \right)
\]  

3) Mean velocity

\[
M_{\text{V}} = C \sum_{k \in R} \left( \frac{\sum_{i=1}^{N} (H_{(u,v)}(k) - H_{(u,v)}(k-1))}{\Delta t} \right)
\]  

4) Resistant index

\[
R_{\text{I}} = \frac{V_{\text{max}} - V_{\text{min}}}{V_{\text{max}}}
\]  

5) Pulsatile index

\[
P_{\text{I}} = \frac{V_{\text{max}} - V_{\text{min}}}{M_{\text{V}}}
\]
4. Classification

SVM has been shown to provide higher performance than traditional learning machines [8]. In order to overcome the small sample size problem, we choose SVM and use the leave one-out method to evaluate the performance of classification. Validation set targets was used in training and model selection. The final choice of model, including the choice of kernel, use of a bias term, use of weighting in training and model selection and the choice of model selection criterion were all determined by minimizing the leave-one-out balanced error rate.

5. Results and Discussions

In the experiment, all of the ultrasound images and CDFI videos were acquired with an ultrasonic scanner ATL 3000 unit (GE, VIVID7), using a 7.5-14MHz linear probe, and captured directly from the video signals. The data consists of 65 cases, including 29 malignant solid masses and 36 benign solid masses. The natures of the lesions were all confirmed by pathology (either with surgical excision or with US-guided percutaneous core-needle biopsy). One case consists of at least three B-mode images and one Doppler video.

The results are expressed in terms of four parameters, True Positive (TP); False Positive (FP); True Negative (TN); and False Negative (FN). Table 2 shows the results of this research, and the comparisons with radiologist assessments.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Radiologist</th>
<th>Proposed system</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>FP</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>TN</td>
<td>25</td>
<td>36</td>
</tr>
<tr>
<td>FN</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Accuracy</td>
<td>75.4%</td>
<td>93.8%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>88.9%</td>
<td>92.6%</td>
</tr>
<tr>
<td>Specificity</td>
<td>65.8%</td>
<td>94.7%</td>
</tr>
</tbody>
</table>

In the future, the proposed system needs more test cases to prove its reliability and accuracy.

6. Acknowledgement

The work was supported, in part, by Natural Scientific Research Innovation Foundation in Harbin Institute of Technology, Project HIT.NSRIF.2008.48, and Natural Science Foundation of China NO.60873142 and No. 30670546.

Reference


[8] C. Burges, "A Tutorial on support vector machines for Pattern Recognition ", Data Mining