Classification of breast masses and normal tissues in digital tomosynthesis mammography

Jun Wei*, Heang-Ping Chan, Yiheng Zhang, Berkman Sahiner, Chuan Zhou, Jun Ge, Yi-Ta Wu, Lubomir M. Hadjiiski

Department of Radiology, The University of Michigan, Ann Arbor, MI 48109

ABSTRACT

Digital tomosynthesis mammography (DTM) can provide quasi-3D structural information of the breast by reconstructing the breast volume from projection views (PV) acquired in a limited angular range. Our purpose is to design an effective classifier to distinguish breast masses from normal tissues in DTMs. A data set of 100 DTM cases collected with a GE first generation prototype DTM system at the Massachusetts General Hospital was used. We reconstructed the DTMs using a simultaneous algebraic reconstruction technique (SART). Mass candidates were identified by 3D gradient field analysis. Three approaches to distinguish breast masses from normal tissues were evaluated. In the 3D approach, we extracted morphological and run-length statistics texture features from DTM slices as input to a linear discriminant analysis (LDA) classifier. In the 2D approach, the raw input PVs were first preprocessed with a Laplacian pyramid multi-resolution enhancement scheme. A mass candidate was then forward-projected to the preprocessed PVs in order to determine the corresponding regions of interest (ROIs). Spatial gray-level dependence (SGLD) texture features were extracted from each ROI and averaged over 11 PVs. An LDA classifier was designed to distinguish the masses from normal tissues. In the combined approach, the LDA scores from the 3D and 2D approaches were averaged to generate a mass likelihood score for each candidate. The $A_z$ values were 0.87±0.02, 0.86±0.02, and 0.91±0.02 for the 3D, 2D, and combined approaches, respectively. The difference between the $A_z$ values of the 3D and 2D approaches did not achieve statistical significance. The performance of the combined approach was significantly ($p<0.05$) better than either the 3D or 2D approach alone. The combined classifier will be useful for false-positive reduction in computerized mass detection in DTM.

Keywords: Digital Tomosynthesis Mammography (DTM), Simultaneous Algebraic Reconstruction Technique (SART), Breast mass, Receiver operating characteristic (ROC) analysis

1. INTRODUCTION

Breast cancer is one of the leading causes of cancer mortality among women\textsuperscript{1}. Although mammography is a powerful screening tool for detecting breast cancer, the sensitivity of cancer detection is often limited by the presence of the overlapping dense fibroglandular tissue in the breast. Digital tomosynthesis mammography (DTM) is one of the new breast imaging modalities that have potential to improve the detection and diagnosis of breast cancer. In DTM, a serial of low-dose projection view images (PVs) are acquired when the x-ray source is rotated about the fulcrum over a limited angular range. DTM slices reconstructed from the serial of PVs can provide quasi-3D structural information of the breast which may improve the conspicuity of breast cancer and also reduce the effects of the overlapping dense tissues.

In conventional mammography, computer-aided detection system (CADe) can increase the breast cancer detection rate by radiologists both in the laboratory and in clinical practice\textsuperscript{2,3}. We are developing a CADe system to assist radiologists in detecting breast masses in DTMs. In this preliminary study, our purpose is to design an effective classifier to distinguish breast masses from normal tissues for false positive (FP) reduction in computerized mass detection in DTMs.

* jywei@umich.edu, phone: 734-647-8553, Med-Inn Building C478, 1500 E. Medical Center Dr., Ann Arbor, MI 48109-5842
Figure 1. Histogram of the mass sizes measured on the central slices of the reconstructed DTMs. Mass sizes are measured as the longest dimension of the mass by an experienced MQSA radiologist. The size of the masses in this data set ranged from 5.5 mm to 43.4 mm with a mean size of 17.4±7.3 mm.

Figure 2. An example of a spiculated mass in DTMs. (a) PV at 0 degree, (b) the DTM slice approximately intersecting the center of the mass, (c) ROI image on PV at 0 degree, and (d) ROI image on the DTM slice in (b).
2. MATERIALS AND METHODS

2.1 Materials

DTMs in this study were collected with Institutional Review Board (IRB) approval in the Breast Imaging Research Laboratory at the Massachusetts General Hospital with a GE first generation prototype DTM system. The DTM system has a flat panel detector with a pixel size of 0.1 mm x 0.1 mm. It acquired 11 PVs in 5-degree increments over a total tomosynthesis angle of 50-degree. The total dose for the 11 PVs was designed to be less than 1.5 times that of a single conventional mammogram. Only mediolateral oblique view (MLO) DTM is collected for each case. The data set we used in this study contained 100 cases. Each case had one single breast DTM which was reconstructed by a simultaneous algebraic reconstruction technique (SART) developed in previous study\(^8\). SART method iteratively updates the imaged volume using each PV image sequentially. The number of reconstructed slices in this data set ranged from 45 to 90 depending on the compressed breast thickness. An experienced MQSA radiologist identified the location of the masses in the reconstructed DTM slices. There were 69 malignant and 31 benign masses in this data set including some large-area architectural distortions. Figures 1 showed the histograms of mass sizes measured on the central slices. The mass size ranged from 5.5 to 43.4 mm with a mean size of 17.4±7.3 mm on the central DTM slices. Figure 2 is an example of a DTM with a spiculated mass; the PV at 0 degree and the SART reconstructed slice approximately intersecting the center of the mass are shown.

2.2 Methods

2.2.1 CADe System Overview

![Block diagram of the CADe system for mass detection on DTMs.](image)

Figure 3. Block diagram of the CADe system for mass detection on DTMs.

Our CADe system consists of five processing steps: 1) segmentation of breast region on each reconstructed slice, 2) prescreening of mass candidates, 3) identification of suspicious objects, 4) extraction of morphological and texture features, and 5) classification between the normal tissue and the mass regions by using LDA classifiers. The block diagram for the CADe system is shown in Figure 3.
We have described a prototype CADe system for detection of masses in the reconstructed DTM volumes previously.\(^9\) For each reconstructed slice, the CADe system first segments the breast region from the background by using a breast boundary detection algorithm. The rest of the processes is only applied to the segmented breast region. For the pre-screening stage, we designed a 3D filter using the information of gradient field directions to identify mass candidates. After this enhancement filtering, the local maxima within the breast region were identified as the mass candidates on each DTM volume. Volumes of interest (VOI) were then identified in the breast with the center of a VOI placed at each of the mass candidates. The suspicious structure in each identified location was extracted by a 3D region growing method.

During the feature analysis and classification step, we evaluated three approaches to distinguish breast masses from normal tissues. In the first approach, referred to as 3D approach, we extracted morphological and run-length statistics texture features from the DTM slices. A linear discriminant analysis (LDA) classifier with stepwise feature selection was then trained to merge the useful features. In the second approach, referred to as 2D approach, the raw input PVs were first preprocessed with a Laplacian pyramid multiresolution enhancement scheme. A mass candidate was then forward-projected onto the each of the 11 PVs in order to determine the corresponding regions of interest (ROIs) on the preprocessed PVs. Spatial gray-level dependence (SGLD) texture features were extracted from each ROI and averaged over the 11 PVs. An LDA classifier with stepwise feature selection was designed to distinguish the masses from normal tissues. In the third approach, referred to as the combined approach, the LDA scores of a detected object, which represented the likelihood of the object being a mass, from the first and second approaches were merged to generate a combined mass likelihood score for each mass candidate. In this study, a simple averaging method was used to merge the information.

### 2.2.2 Training and test CADe system

The leave-one-case out re-sampling method was used for training and testing the FP classifiers. During training, feature selection with stepwise LDA was employed to obtain the best feature subset and reduce the dimensionality of the feature space to design an effective classifier. Briefly, for a data set of \(n\) cases, all mass candidates from a test case are left out while the other \((n-1)\) cases are used for selection of predictor variables from the feature pool and estimation of the LDA classifier weights in each leave-one-case-out cycle. Since the appropriate threshold values for feature entry, feature elimination, and tolerance of feature correlation were unknown, we used an automated simplex optimization method to search for the best combination of thresholds in the parameter space. The trained LDA classifier was then applied to the mass candidates in the left-out case to obtain the mass likelihood score of each candidate. This process was performed with each of the \(n\) cases left out in turn so that all objects would be assigned a mass likelihood score at the completion of the \(n\) cycles.

### 2.2.3 Evaluation methods

In order to compare the performance of the three feature analysis and classification methods, we employed the receiver operating characteristic (ROC) analysis. The ROCKIT curve fitting software and statistical significance test for ROC analysis developed by Metz et al.\(^{10}\) is used in this study. The discriminant scores from the three different approaches were input as the decision variable in the ROCKIT program, which fits a binormal ROC curve based on maximum likelihood estimation.

### 3. EXPERIMENTAL RESULTS

Figure 4 shows the test ROC curves for classification of breast masses and normal tissues on DTMs. The areas under the ROC curves \((A_z)\) were 0.87±0.02, 0.86±0.02 and 0.91±0.02 for the 3D, 2D, and combined approaches, respectively. The difference between the \(A_z\) values of the 3D and 2D approaches did not achieve statistical significance \((p=0.50)\). The performance of the combined approach was significantly better than either the 3D or the 2D approach \((p=0.04\) and \(p=0.03\), respectively). Table 1 summarized the ROC results.
Figure 4. The test ROC curves for classification of breast masses from normal tissues on digital tomosynthesis mammography (DTM). The $A_z$ values are 0.87±0.02, 0.86±0.02 and 0.91±0.02 for the 3D, 2D and combined approaches, respectively.

Table 1. Estimation of the statistical significance in the difference between the ROC performances of three feature analysis and classification approaches. A: 3D approach, B: 2D approach, C: combined approach.

<table>
<thead>
<tr>
<th>Az Value</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.87</td>
</tr>
<tr>
<td>B</td>
<td>0.86</td>
</tr>
<tr>
<td>C</td>
<td>0.91</td>
</tr>
</tbody>
</table>

A vs B = 0.5
B vs C < 0.05
C vs A < 0.05

4. DISCUSSION AND CONCLUSIONS

In mammography, 3D anatomical structures are projected onto a 2D image plane, and the overlapping fibroglandular breast tissues reduce the visibility of breast cancer. DTM is one of the promising methods that provide quasi-3D structural information and may improve the sensitivity and specificity of breast cancer detection. In this study, we compared three methods for classification of breast masses and normal tissues in DTMs. The forward projection for extraction of 2D image information from the PVs is a new method for identifying corresponding ROIs for the detected objects. The combined approach which fuses the information on both the 2D projection views and 3D reconstructed slices is a promising approach to improving the differentiation of breast masses from normal tissues on DTMs. It is expected that the classifier will be useful for FP reduction in a computerized mass detection system.

ACKNOWLEDGMENTS

This work is supported by USPHS grants CA120234 and CA95153, U. S. Army Medical Research and Materiel Command grants DAMD 17-02-1-0214, W81XWH-04-1-0475 and W81XXH-07-1-0324. The content of this paper does not necessarily reflect the position of the funding agencies and no official endorsement of any equipment and product of any companies mentioned should be inferred. The authors are grateful to Charles E. Metz, Ph.D., for the LABROC and ROCKIT programs.
REFERENCES