AD_____

Award Number: DAMD17-00-1-0245

TITLE: Robust Detection of Masses in Digitized Mammograms

PRINCIPAL INVESTIGATOR: Lihua Li, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, Florida 33620

REPORT DATE: June 2002

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2002	Annual Summary (15 May 01 - 14 May 02)

4. TITLE AND SUBTITLE
Robust Detection of Masses in Digitized Mammograms

5. FUNDING NUMBERS
DAMD17-00-1-0245

6. AUTHOR(S)
Lihua Li, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of South Florida
Tampa, Florida 33620

E-Mail: lilh@moffitt.usf.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

This project is to develop a robust computer aided diagnosis (CAD) system for mass detection with high sensitivity and specificity in digitized mammograms. The research scope in past year is to improve and optimize detection performance and classification generalizability. Several major progresses have been made including (1). A novel graph-based algorithm was proposed to segment stellate masses in mammograms by separating the adjacent regions while keeping the spiculation of masses. It is helpful for the improvement of stellate mass and distortion detection. (2). A hybrid "hard"-"soft" classification method was proposed, where the "hard" decision classifier is cascaded with a "soft" decision classification with the objective to reduce false-positives (FPs) in the cases with multiple FPs retained after the "hard" decision classification. It has a much better performance and generalizability of classification. (3). A training database was generated for fine tuning the parameters of CAD system. An FROC curve of CAD mass detection using training database was obtained.It is expected that these processing will be very helpful in improving the robustness of the detection system.

14. SUBJECT TERMS	15. NUMBER OF PAGES
breast cancer, mass, detection, CAD, mammography, robust, adaptive, segmentation, classification	41
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

This project is to develop a robust computer aided diagnosis (CAD) system for mass detection with high sensitivity and specificity in digitized mammograms. As listed in the Statement of Work, the research scope in the second year of project is to improve and optimize detection performance and classification generalizability.

BODY

***Objective 1:* to enhance the detection performance of stellate mass/distortion**

Accomplishments:

A major problem in detecting stellate mass is that the mass region may be extended to surrounding tissue area in mass segmentation due to the similarity of pixel intensity. Region growing is widely used in mass segmentation. Some methods attempt to utilize shape constraints in conventional region growing to regularize the segmented partitions. But they didn't take the local shape of spiculation and the connected area between adjacent objects into account. In order to improve the detection performance of stellate mass/distortion, we present a graph-based region growing method for segmenting masses in digital mammograms. In the proposed algorithm, the procedure of region growing is represented as a growing tree whose root is the selected seed. Active leaves, which have the ability to grow, in the connection area between adjacent regions are deleted to stop growing, then separating the adjacent regions while keeping the spiculation of masses. A complete procedure of graph-based region growing algorithm is described in the attached paper. In summary, it has following steps:

(1). Select a seed for region growing and define the seed as an active leaf in T_0.

(2). At iteration i, every neighboring pixels (x, y) of active leaves in T_{i-1} are checked. If they are not in R_{i-1} and satisfy $|I(x,y) - M_{i-1}| < T_{\text{int}}$, they are supposed to be in R_i.

(3). Compute the number of the supposed active leaves $|la|$ and dead leaves.

(4). Label the active leaves groups (G s) and calculate the size of active groups $|G|$.

(5). If $|G| < T_a \times (|la| + |ld|)$, compute the number of the neighboring dead leaves $|lsd|$ of the active groups. If $|lsd| > T_h \times |G|$, the active group is a hole (H).

(6). Compute PE for supposed active leaves in H. If $PE \geq T_{ext}$, the supposed active leaves in the H are eliminated from R_i.

(7). Continue the procedure, the algorithm stops when no active leaves are found.

***Objective 2:* to improve the performance and generalizability of classification.**

Accomplishments:

The major difficulties of classification result from the great similarity in appearance between mass and dense normal tissue, and the great variation in feature distribution of different masses. From the classification perspective, the former requires more elaborate features to be extracted for classification while the later means the classifier structure should be more flexible. In this research, we focus on the exploration of new classification strategy with less effort on new feature design.

1. Feature extraction

Seven features are used in this work for FP reduction. Due to the page limit, they are simply listed as follows. More detailed definition can be found in the attached publications.

Morphological features: (i) Area; (ii) Circularity; (iii) Normalized deviation of radial length.

Intensity features of region: (i) Intensity variation; (ii) Mean intensity difference.

Intensity features of boundary: (i) Mean gradient of region boundary; (ii) Mean intensity difference along region boundary.

2. Hybrid classification

There are several types of classifiers used in discrimination of masses from normal tissue regions, such as decision tree, Bayes classifier, neural networks, linear discriminant analysis (LDA), and quadratic classifiers. A common characteristics of current FP reduction schemes is that a "hard" discriminant criteria is developed by training to evaluate each segmented suspicious region. From the standpoint of region-based classification, it is reasonable, and the task of classifier design is to find an optimal discriminant hyperplane in feature space. However, the segmented FPs can not be reduced efficiently by a single "hard" decision classifier. With an analysis of "hard" decision-making classification results, it can be found that the distribution of FPs is quite different from case to case, generally less FPs for fatty breast and more FPs for dense breast in which more mass-like regions are segmented. In other words, from the image-based standpoint, the efficiency of classification strategy by means of "hard" decision criteria is questionable. The proposed hybrid classification method takes two different decision-making strategies as described below, where the current "hard" decision classifier is cascaded with a "soft" decision classification with the objective to reduce FPs in the cases with multiple FPs retained after the "hard" decision classification.

(1) "Hard" classification using a modified fuzzy decision tree method

The segmented regions are first classified using a modified fuzzy decision tree (MFDT) method. It is a modification of our previously developed fuzzy binary decision tree (FBDT) method, in which the classification decision at each tree node is made based on the region feature as well as region size. The design of MFDT structure is based on an observation that the feature distributions are statistically different for different mass size.

(2) "Soft" classification using committee decision method

The "hard" classification method is region-oriented. The thresholds at the decision nodes of MFDT are selected by training to ensure a high detection sensitivity. However, at the mean time, several FPs in each image may be produced especially for that with non-uniform high density tissue. The "soft" classification strategy proposed is to reduce FPs by an image-based analysis for selecting the best ones among the pre-classified candidate regions as the "real" masses. It is

performed by a committee decision method based on a simple premise that most features of a mass should be top in individual feature ranking among all the candidate suspicious regions in a single mammogram. The committee consists of six "experts", each provides a ranking list of the suspicious regions with a likelihood to be a mass using a single feature except the feature "Area". A region is decided by the committee to be a possible mass if it is ranked among top K_0 ($K_0 \leq M$) by more than C_0 "experts". Here M is the candidate region number.
A very detailed description of the proposed classification method and the comparative study of false-positive reduction can be found in the attached papers.

Objective 3: systematic optimization of the CAD method

Accomplishments:

A training database containing 30 normal and 47 abnormal mammograms with totally 70 masses was generated for fine tuning the parameters in each CAD module in mass detection system. An FROC curve of detection on training database was obtained (see the attached paper). The operating point of system is chosen to be at sensitivity TP=93% and false positive rate FP=3.1 per image. By analyzing the five masses missed in training at the operating point, two of them are due to extremely small size (<4 mm) and lower contrast (<4.0); another two of them are on the boundary of breast area; one is due to its great deviation of shape feature.

KEY RESEARCH ACCOMPLISHMENTS

1. A novel graph-based algorithm was proposed to segment stellate masses in mammograms by separating the adjacent regions while keeping the spiculation of masses. It is helpful for the improvement of stellate mass and distortion detection.
2. A hybrid "hard"-"soft" classification method was proposed, where the "hard" decision classifier is cascaded with a "soft" decision classification with the objective to reduce false-positives (FPs) in the cases with multiple FPs retained after the "hard" decision classification. It has a much better performance and generalizability of classification.
3. A training database was generated for fine tuning the parameters of CAD system. An FROC curve of CAD mass detection using training database was obtained.

REPORTABLE OUTCOMES

1. Presentation and/or proceedings paper

(a) Yong Chu, Lihua Li, and Robert A. Clark, "Graph-based Region Growing for Mass Segmentation in Digital Mammography," Proc. of SPIE Medical Imaging, 1999.

(b) Lihua Li and Robert A. Clark, "Hybrid Classification Method for False-Positive Reduction in CAD for Mass Detection," Proc. of 5th Intern. Workshop on Digital Mammography, 2000.

(c) Lihua Li, W. Qian, L.P. Clarke, R.A. Clark, J. Thomas, "Improving Mass Detection by Adaptive and Multi-Scale Processing Methods in Digitized Mammograms," Proc. of SPIE Medical Imaging, 1999.

2. Journal paper

(a) L. Li, Y. Zheng, L. Zhang, R.A. Clark, "False-positive reduction in CAD mass detection using a competitive strategy," Medical Physics, 28(2), Feb. 2001.

3. Fundings Applied

(a) "New CAD Strategies for Early Detection of Breast Cancer", a proposal submitted to U.S. ARMY Medical Research and Material Command, BCRP01-IDEA

CONCLUSIONS

The great variation of characteristics of mammograms and masses hinders us in developing a high detection performance and more generalizable CAD system. The typical variations between different mammograms result either from the imaging process (such as film exposure, film label), digitization process (such as spatial / intensity resolution, response function to optical density), or most importantly the inherent breast tissue characteristics. The variations of masses include its size, contrast, shape, location, intensity pattern and its relation to the surrounding tissues. The research work taken in second year of this project is directed to address these problems by developing and optimizing advanced segmentation and classification methods. The results demonstrated their effectiveness.

False-positive reduction in CAD mass detection using a competitive classification strategy

Lihua Li,[a] Yang Zheng, Lei Zhang, and Robert A. Clark
Department of Radiology, College of Medicine, and the H. Lee Moffitt Cancer Center and Research Institute at the University of South Florida, Tampa, Florida 33612

(Received 5 May 2000; accepted for publication 22 November 2000)

High false-positive (FP) rate remains to be one of the major problems to be solved in CAD study because too many false-positively cued signals will potentially degrade the performance of detecting true-positive regions and increase the call-back rate in CAD environment. In this paper, we proposed a novel classification method for FP reduction, where the conventional "hard" decision classifier is cascaded with a "soft" decision classification with the objective to reduce false-positives in the cases with multiple FPs retained after the "hard" decision classification. The "soft" classification takes a competitive classification strategy in which only the "best" ones are selected from the pre-classified suspicious regions as the true mass in each case. A neural network structure is designed to implement the proposed competitive classification. Comparative studies of FP reduction on a database of 79 images by a "hard" decision classification and a combined "hard"–"soft" classification method demonstrated the efficiency of the proposed classification strategy. For example, for the high FP sub-database which has only 31.7% of total images but accounts for 63.5% of whole FPs generated in single "hard" classification, the FPs can be reduced for 56% (from 8.36 to 3.72 per image) by using the proposed method at the cost of 1% TP loss (from 69% to 68%) in whole database, while it can only be reduced for 27% (from 8.36 to 6.08 per image) by simply increasing the threshold of "hard" classifier with a cost of TP loss as high as 14% (from 69% to 55%). On the average in whole database, the FP reduction by hybrid "hard"–"soft" classification is 1.58 per image as compared to 1.11 by "hard" classification at the TP costs described above. Because the cases with high dense tissue are of higher risk of cancer incidence and false-negative detection in mammogram screening, and usually generate more FPs in CAD detection, the method proposed in this paper will be very helpful in improving the performance of early detection of breast cancer with CAD.
[DOI: 10.1118/1.1344203]

Key words: CAD, false-positive, classification, mass detection, mammography

I. INTRODUCTION

The computer-aided diagnosis (CAD) method has been proposed as a "second opinion" or a "pre-reader" strategy to improve mammogram interpretation. The methods reported have included the detection and classification of either microcalcification clusters or masses.[1–3] Compared to microcalcification cluster detection, mass detection clearly poses a much more difficult problem and usually generates more false-positive signals in detection because masses are often: (a) of varying characteristics in size, shape, and density; (b) exhibit poor image contrast in breast with high density tissues; (c) have a similar characteristic to the nonuniform high dense tissue background; and (d) are highly connected to the parenchymal tissue, particularly for spiculated masses.

Numerous investigators have addressed mass detection in the past several years by following a general approach consisting of three major steps: preprocessing, localization/feature extraction, and classification. Most of the proposed researches have been directed to the first two CAD modules and the feature selection part in classification module. In preprocessing, Lai *et al.*[4] proposed a modified median filtering method to remove background noise before using template matching to detect circumscribed masses. Yin *et al.*[5] used both linear and nonlinear bilateral subtraction to enhance possible masses. Petrick *et al.* proposed a two-stage adaptive density-weighted contrast enhancement (DWCE) filtering technique to enhance objects in mammogram.[6] Polakowski *et al.*[7] used a difference of Gaussians (DOG) filter to highlight suspicious regions in the mammogram. We developed a wavelet-based mass detection system where a multi-orientation transform and a multiresolution wavelet transform are used as preprocessing for improving mass segmentation and feature extraction.[8–10] The feature extraction involves either region based and/or pixel based methods, where the regions could be a manually extracted region-of-interest (ROI)[11] or an automatically segmented one.[6,8,12] The most commonly used features in mass detection include textures derived from the gray level dependence matrix,[11] texture energy obtained from the output of Law's filters,[13] features derived from pixel orientation,[14,15] and density and morphological features.[4,6,16] Due to the complexity of mammograms and the similarity of mammographic characteristics between mass and dense tissue, classification faces a great challenge in reducing false-positives. Research efforts have

mostly been focused on feature selection[17,18] and the application of widely used pattern classifiers in computer vision.[11]

Although the mass detection methods proposed to date have met varying levels of success, the high false-positive rate remains to be one of the major problems to be solved in CAD study because too many false-positively cued signals will degrade the performance of detecting true-positive regions and increase the call-back rate in the CAD environment. In this paper, we first presented a brief overview of the FP problem in CAD. A new approach to classification was then developed with the objective to reduce the FPs in cases with high FP number.

II. FALSE-POSITIVE REDUCTION

With analysis of the CAD mechanism, it can be found that the direct cause of FPs is in the stage of locating suspicious areas in which the number of located suspicious regions is usually much higher than the expected true mass number in order to minimize the false-negative detection. Classification is designed to discriminate the mass from the normal tissue region among the located suspicious regions. The criteria for success are therefore two-fold: the final detected region number should be minimized and the probability of inclusion of the true mass regions should be maximized. A common approach to reduce the FP rate is to perform a further feature analysis of the segmented/located regions and to use a classifier to discriminate masses from the normal tissue region. However, due to the fact that (1) mass and the normal dense tissue region have a great similarity in mammographic appearance, and (2) there is a great variation in feature distribution of different masses, the classification is recognized as a very difficult task in CAD. From the classification perspective, the former requires more elaborate features to be extracted for classification while the latter means the classifier structure should be more flexible.

Many features have been designed and/or tested for FP reduction, ranging from intensity domain, and morphological domain, to textural domain. Different feature ranking and selection techniques such as genetic algorithm were applied to optimize the feature selection under various performance criteria.[17] Most of the typical pattern classifiers have been explored for discrimination of masses from normal tissue regions, such as decision tree,[15] Bayes classifier,[16] *k*-nearest-neighbors (kNN),[14] neural networks,[19,20] linear discriminant analysis (LDA),[11] and quadratic classifiers.[21] A common characteristics of current FP reduction schemes is that a "hard" discriminant criteria is developed by training to evaluate each segmented/located suspicious region. From the standpoint of region-based classification, it is reasonable. The task of classifier design is to find an optimal discriminant hyperplane in feature space. However, because the classifier is trained by regions and the criteria is fixed for all images, the segmented FPs cannot be reduced effectively by a single "hard" decision classifier especially for high-dense cases. With an analysis of "hard" decision-making classification results, it can be found that the distribution of FPs is quite different from case to case. Usually there are less FPs for fatty breast and more FPs for dense breast in which more masslike regions are segmented. In other words, from the image-based standpoint, the efficiency of classification strategy by means of "hard" decision criteria is questionable. Figure 1 shows some detection examples with a "hard" decision classifier used in our CAD system as described later. Figure 1(a) is a fatty case with one mass as indicated. It is detected with one FP in the detection output. Figure 1(b) is a case with mixed fatty-dense tissue. The mass is detected but there are 11 FPs generated. Figure 1(c) is another case with

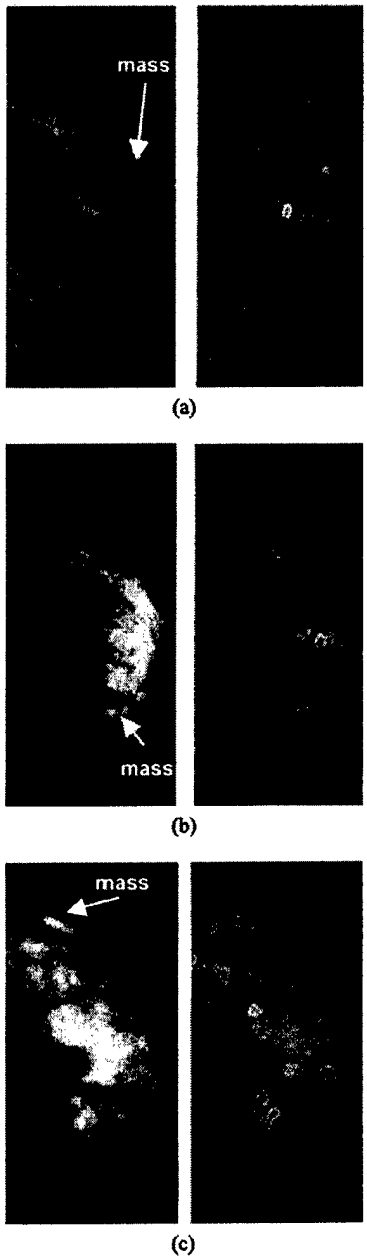


FIG. 1. Some detection examples with a "hard" decision classifier (a) mostly fatty breast; (b) mixed fatty-dense breast; (c) mostly dense breast.

high-dense tissue. It produced a similar result to that of Fig. 1(b).

Because the dense breast usually has a higher risk of cancer incidence and a higher false-negative detection in mammogram screening, the reduction of FP in cases with multiple FPs in CAD detection is of great significance. This paper addresses this FP reduction problem. As described later, it focuses on the exploration of new classification strategy with less effort on a new feature design.

III. HYBRID CLASSIFICATION METHOD

A. Feature extraction

Seven features are used in this proposed FP reduction method. They are similar to those we reported before[8] except that: (a) one more mixed boundary-intensity feature is added; and a (b) the calculation of some features is modified with more reasonable definitions. For the convenience of description, they are listed in the following subsections.

1. Morphological features

(a) Area: The number of pixels in the extracted region.

(b) Circularity (circ): It provides information about the gross shape of the mass but contributes nothing specific about the fine detail of the mass boundary. It is defined as

$$\text{circ}=\frac{P^2}{S}, \tag{1}$$

where P is the perimeter and S is the area of the extracted region.

(c) Normalized deviation of radial length (drl): The normalized deviation of radial length is calculated using

$$drl=\sqrt{\frac{1}{N_b}\sum_{k=1}^{N_b}\left(\frac{r_k-m_r}{m_r}\right)^2}, \tag{2}$$

where r_k is the radial length from the centroid (x_0,y_0) to the kth pixel (x_k,y_k) on the boundary of the extracted region, m_r is the mean radial length, and N_b is the total number of pixels on the boundary of the region.

It is a measure of how the boundary changes in a microscopic way. Normalization is done due to the fact that the edge roughness increases with the size of mass.

2. Intensity features of region

(a) Intensity variation (iv): This is a measure of the smoothness of the pixel intensity in extracted region and defined as

$$iv=\sqrt{\frac{1}{N_a}\sum_{(i,j)\in A}(I(i,j)-m_a)^2}, \tag{3}$$

where N_a is the total number of pixels in segmented region A, and m_a is the mean intensity value of region A;

$$m_a=\frac{1}{N_a}\sum_{(i,j)\in A} I(i,j). \tag{4}$$

(b) Mean intensity difference (mid): This parameter measures the intensity difference between the extracted region and its surrounding area

$$\text{mid}=\frac{1}{N_a}\sum_{(i,j)\in A} I(i,j)-\frac{1}{N_a}\sum_{(i,j)\in A_s} I(i,j), \tag{5}$$

where N_a is the total number of pixels in segmented region A, and N_s is the total number of pixels in region A_s surrounding region A. It was obtained by morphological dilation operation on region A with N_s to be approximately equal to N_a.

3. Intensity features of boundary

(a) Mean gradient of region boundary (mg): This parameter measures the edge contrast of the extracted region:

$$mg=\frac{1}{N_b}\sum_{k=1}^{N_b} g_k, \tag{6}$$

where N_b is the total number of pixels on the boundary of extracted region, and g_k is the edge gradient value of the k-th boundary pixel. The edge gradient is calculated by the Sobel operator.

It is a measure of the intensity change in perpendicular direction along the edge of region in microscopic way.

(b) Mean intensity difference along region boundary (mid-b): This parameter measures the mean intensity difference between inside and outside the extracted region along its boundary

$$\text{mid-}b=\frac{1}{N_{\text{in}}}\sum_{(i,j)\in A_{\text{in}}} I(i,j)-\frac{1}{N_{\text{out}}}\sum_{(i,j)\in A_{\text{out}}} I(i,j), \tag{7}$$

where N_{in} is the total number of pixels in inside region A_{in} and N_{out} is the total number of pixels in outside region A_{out} surrounding region A. They are obtained by one erosion and one dilation operation on region A, respectively, with an eight-connection structuring element.

It is a measure of intensity change along the edge of region in an average meaning.

B. Hybrid classification

The hybrid classification is a cascaded classification structure. Each segmented region is first pre-classified by a "hard" classifier. Then the pre-selected regions in each image take an image-based competition. The candidates are voted by a committee to select the best ones to be "true" masses.

1. "Hard" classification using a fuzzy decision tree method

The segmented regions are first classified using a fuzzy decision tree method. It is a modification of our previously developed fuzzy binary decision tree (FBDT) method.[8,22] Its decision tree structure is depicted in Fig. 2, in which the classification decision at each tree node is made based on the region feature as well as region size. The design of decision tree structure is based on an observation that the feature dis-

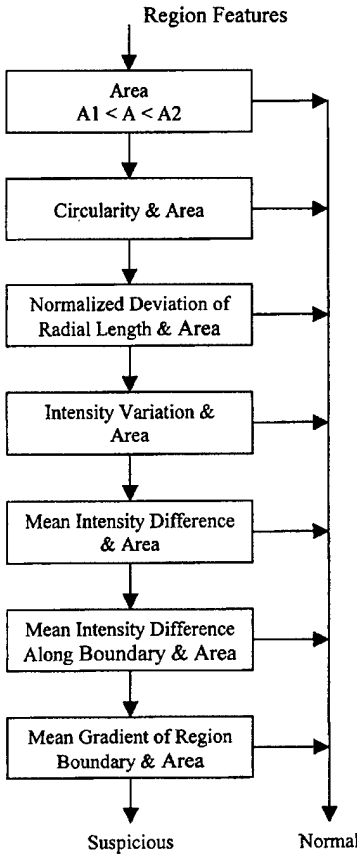


FIG. 2. Flow chart of the decision tree for "hard" classification.

tributions are statistically different for different mass size. The fuzzy membership functions at different decision nodes is similar to that described in our previously reported work.[8,22]

2. "Soft" classification using committee decision method

(a) Principle. The "hard" classification method is region-oriented. The thresholds of the decision nodes in the fuzzy decision tree are selected by training to be small enough to ensure a high detection sensitivity. However, at the mean time, several FPs in each image may be produced, especially for that with nonuniform high density tissue. The "soft" classification strategy proposed is to reduce FPs by an image-based analysis for selecting the best ones among the pre-classified candidate regions as the "true" masses. It is performed by a committee decision method based on a simple premise that most features of a mass should be top in individual feature ranking among all the candidate suspicious regions in a single mammogram.[23]

The committee consists of N "experts," each provides a ranking list of the suspicious regions with a likelihood to be a mass using a single feature. A region is decided by the committee to be a possible mass if it is ranked among top $K_0(K_0 \leq M)$ by more than C_0 "experts," where M is the number of suspicious regions in each image. Mathematically, it can be described as follows: For a region i among M candidate regions, its ranking value by feature j is

$$r_i^{(j)} = k, \quad k \in [1, M] \tag{8}$$

if feature j of region i is at the order k among M regions.

The committee score of region i is defined as

$$C_i = \sum_{j=1}^{N} \cup(r_i^{(j)}), \tag{9}$$

where

$$\cup(r_i^{(j)}) = \begin{cases} 1 & k \leq K_0 \\ 0 & k > K_0 \end{cases}. \tag{10}$$

K_0 is an order threshold, and N is the "expert" number.

Region i is classified to be a mass if $C_i \geq C_0$, i.e., more than C_0 "experts" vote region i for mass class provided they can select K_0 regions from all M regions as candidate masses. Please notice that the final "mass" number decided by committee is not fixed from case to case. It depends on following factors:

(1) the number of candidate regions and their distribution of features;
(2) the number of regions each "expert" can vote to be "mass," i.e., the order threshold K_0;
(3) the threshold of committee score C_0, i.e., the minimum number of favoring "expert."

The maximum number of "mass" in each image is

$$MX = \max\left\{ \text{floor}\left[\frac{k_0 * N}{C_0} \right], M \right\}, \tag{11}$$

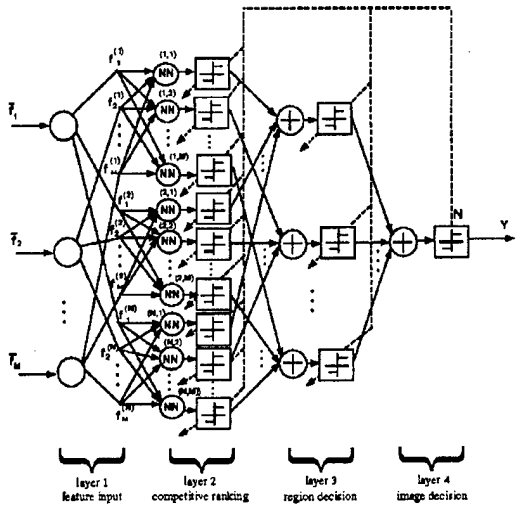
where floor[*] is the operation of taking the smallest integer value.

(b) Neural network implementation. The "soft" classification algorithm described above is implemented using an adaptive competitive classification neural network (CCNN). The structure of CCNN is depicted in Fig. 3, where (b) is a sub-network of the nodes in layer 2. It is a feed-forward four-layer neural network with fixed weights and adaptive neuron bias. All weights of each connection are units except those shown in the sub-network (b). The four layers are described as follows.

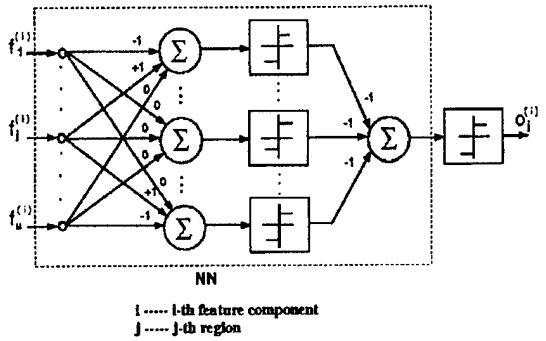
Layer 1: *Feature input.* Each region has a feature vector. Their components are feed into next layer individually so that they can be evaluated together with the corresponding component from other regions in the subsequent competitive ranking;

Layer 2: *Competitive ranking.* It is used to determine if the region is ranked in top K_0 of all the input regions based on individual feature competition, where K_0 is the bias/threshold value of bi-polar step function of neuron. For each region, there is N ranking outputs.

Layer 3: *Region decision.* The outputs of individual feature ranking are input to layer 3 for region classification. The neuron of this layer has a step-function with a threshold C_0, which is a decision threshold.



(a)



(b)

FIG. 3. Adaptive competitive classification neural network (CCNN) for "soft" classification. (a) CCNN structure; (b) the sub-neural network used in (a).

Layer 4: *Image decision*. Layer 4 is an optional layer. It is designed to be a decision-feedback controller to limit the maximum "mass" number per image, which is set to be the threshold of output neuron. It is performed by controlling the neuron thresholds in layer 2 (K_0) and layer 3 (C_0).

IV. RESULTS

A comparative evaluation of false-positive reduction by means of a "hard" decision classifier and a combined "hard"–"soft" decision classifier was performed using a testing database consisted of 30 normal and 49 abnormal mammograms with totally 74 masses. This database is visually divided into three categories with different density, including 18 dense, 21 fatty, and 40 mixed mammograms. The classifiers work at their operating points as developed in our second generation CAD mass detection system design.[24] The electronic truth file was formed for each abnormal mammogram where the masses were labeled by an expert mammographer based on visual criteria and biopsy results. All the features except feature "Area" listed in Sec. III are used in "soft" classification, i.e., there are six "experts" in the decision-making committee. The feature "Area" was not included simply because the size of suspicious region is not a good feature for ranking in "soft" classification. In other words, the bigger or smaller of region size does mean a higher likelihood to be a mass. The threshold of committee score C_0 is 3, i.e., at least half of the committee members have to vote it for a suspicious region to be a "mass." The order threshold K_0 is selected automatically with an objective constraint of the final detected mass number less than 6. With an initial value $K_0=5$, if there are more than six regions detected as masses, the threshold value K_0 will be reduced recursively until it meets the constraint.

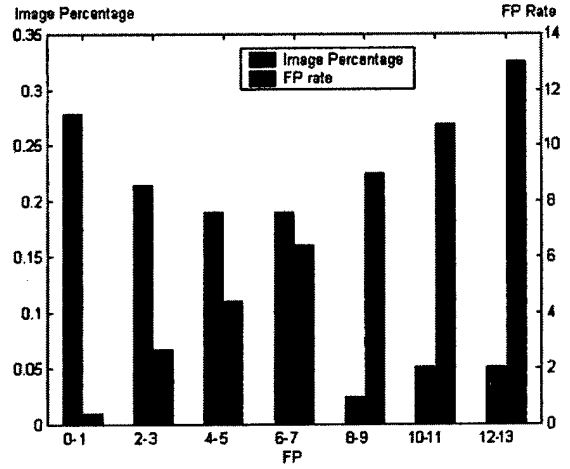


FIG. 4. Images and their corresponding average false-positive rate distributions versus false-positive signal numbers after initial "hard" classification.

The evaluation was taken in three steps. The first step is to get a set of detection results by our CAD detection system with only a decision tree based "hard" classification, as shown in Fig. 2. They are used as the reference of comparative studies of the proposed classification method. Figure 4 shows the distributions of image and their group-average false-positive rate versus false-positive signal numbers on the image in a detection by our CAD mass detection system on the testing database used in this study. Figure 5 is the corresponding accumulative distributions of FPs versus two different image percentages. It is observed that the generated FP number is quite different from case to case. Some cases do not have any FPs but some cases can have more than ten FPs. For different density categories as listed earlier, the FP rates are 7.67, 2.24, and 3.60 per image for dense, fatty, and mixed mammograms, respectively. One-third cases account for almost two-thirds of all FPs in the detection.

The second step is to evaluate the FP reduction performance with "hard" classification by increasing the threshold value of classifier. Figure 6 are the distributions of images and their corresponding average FP rate. Compared to Fig. 4, the percentage of cases with lower FP is increased, which means some FPs are removed in the cases with higher FP

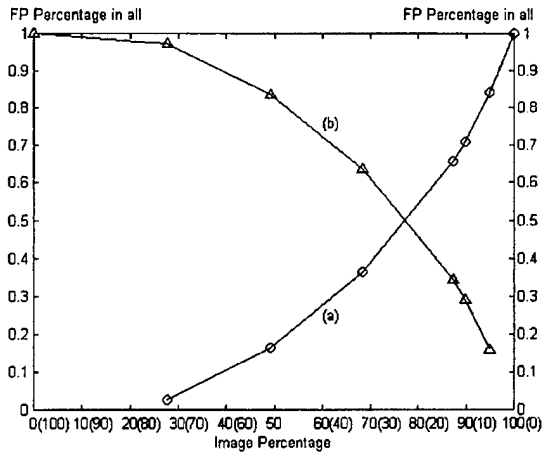


FIG. 5. Accumulative distribution of FPs versus two different image percentages. (a) beginning from the high FP image group, (b) beginning from the low FP image group.

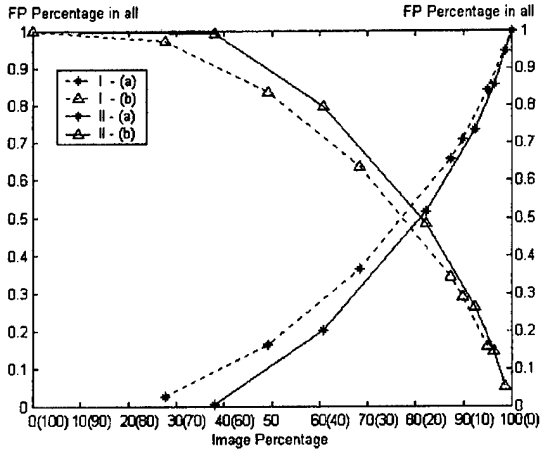


FIG. 7. Accumulative distributions of FPs versus two different image percentages before (I) and after (II) using "hard" classification with a higher threshold. (a) beginning from low FP image group, (b) beginning from high image group.

number. Figure 7 is the corresponding accumulative distributions of FPs versus two different image percentages. For convenience of comparison, the curves of Fig. 5 are also displayed. It is observed that the accumulative FP percentage-in-all beginning from high FP image group is consistently higher than that before FP reduction, while the accumulative FP percentage-in-all beginning from low FP image group is consistently lower, which means the FP distribution among different cases is even more uneven after FP reduction by classification with higher threshold. The 17.7% cases with higher FP numbers generated almost half of the all FPs. Figure 8 shows the distribution of FP rate and its FP reduction of each image group with different numbers of FPs. It is observed that although the overall average FP rate is reduced (from 4.16 to 3.05 per image in whole database), the average reduction in images with higher FP incidence is only a little bit more than that in lower FP images, and the distribution of FPs remains to be very uneven. Furthermore this very limited FP reduction is obtained at a big cost of 14% loss of detection sensitivity as shown in the FROC curve (see Fig. 14). Figure 9 shows the FP reduction results of the three cases shown in Fig. 1. It is noticed that: (1) FPs are reduced in all the three cases, but the true masses in two cases are also removed; (2) the cases with high number of FPs remain to have many FPs. An FROC analysis of the "hard" classification performance at different thresholds was taken. The FROC curve (shown in Fig. 14) together with the "hard"–"soft" classification analysis.

The third step of evaluation is taken on the proposed hybrid classification scheme in which the "hard" classifier is cascaded with a novel "soft" classification as described in Sec. III. Figure 10 shows the distributions of images and their corresponding average false-positive rate versus false-

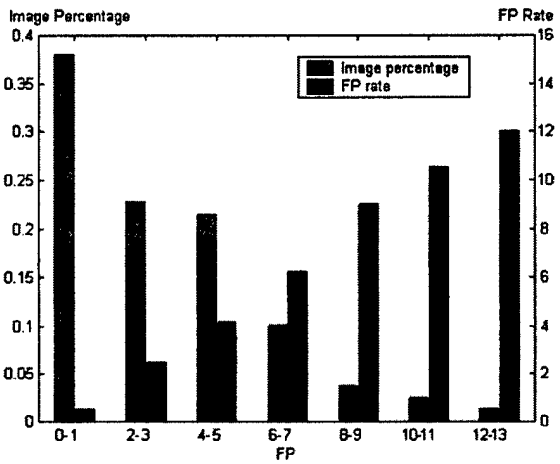


FIG. 6. Images and their corresponding average false-positive rate distributions after FP reduction with a loss of 14% TP rate by using "hard" classification with a higher threshold.

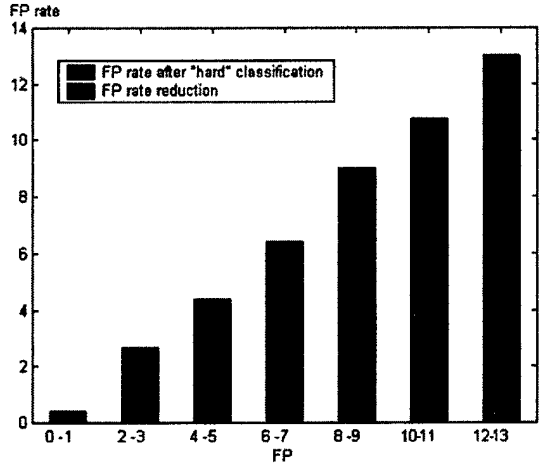


FIG. 8. Distribution of false-positive rate and its FP reduction in different image groups at a loss of 14% TP rate by increasing threshold of "hard" classification.

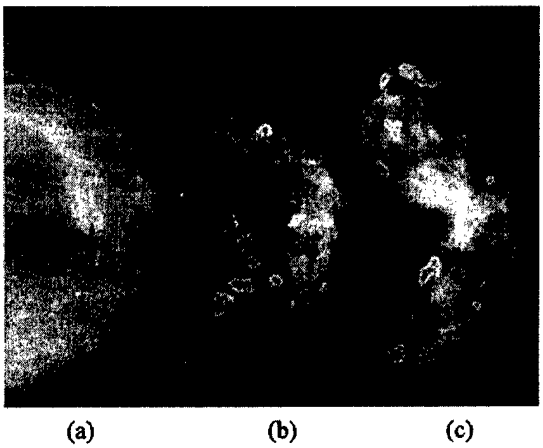


FIG. 9. FP reduction results of the three cases shown in Fig. 1 by increasing the threshold of "hard" classifier.

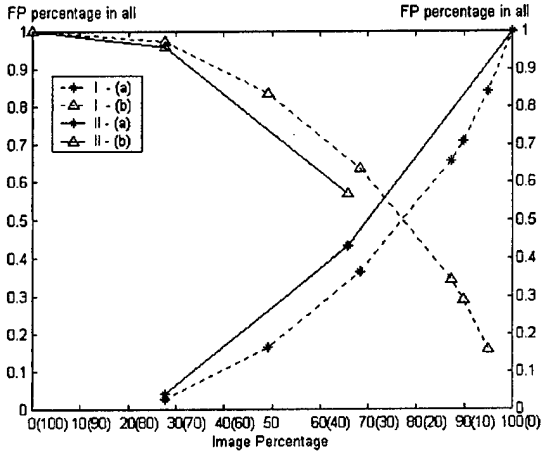


FIG. 11. Accumulative distributions of FPs versus two different image percentages before (I) and after (II), using "soft" classification. (a) beginning from low FP image group, (b) beginning from high FP image group.

positive signal number in image after "soft" classification. The corresponding accumulative distributions are shown in Fig. 11, where the curves of Fig. 7 are also displayed. A significant change of image distribution is observed, where: (1) no image has an output with more than five FPs; (2) the accumulative FP percentage-in-all beginning from higher (lower) FP image group is consistently lower (higher) by the new FP reduction method. For a comparison with the result of single "hard" classification, a distribution of FP rate and its FP reduction in each image group is shown in Fig. 12. It is observed that, compared to the FP reduction scheme with increased threshold "hard" classification, a much bigger false-positive reduction is achieved in the images with more FP detection signals. The final average detection FP rate in whole database is reduced to 2.58 with only 1% loss of detection sensitivity. Figure 13 shows the FP reduction results of the same cases as shown in Fig. 1 and Fig. 9. The FPs in high FP cases is significantly reduced while the true masses are well kept. The distribution of FP rate in different image group is more uniform, which means most of the images generate a similar number of FPs except for the one with very few FPs. A FROC analysis of the "hard"–"soft" classification performance was taken at different thresholds as that in the FROC analysis of "hard" classification. The FROC curves are shown in Fig. 14. The comparison between two FROC curves shows the effect of "soft" classification strategy. All these results demonstrate that "soft"-decision classification can effectively remove the FPs in higher FP incidence cases with less impact on the images with less FPs for securing a high sensitivity.

V. DISCUSSION

The great variation of mammographic characteristics of mammograms and the great similarity between mass and the dense normal tissue regions hinder us in developing a high detection performance and more generalizable CAD system. A major problem in current CAD mass detection systems is

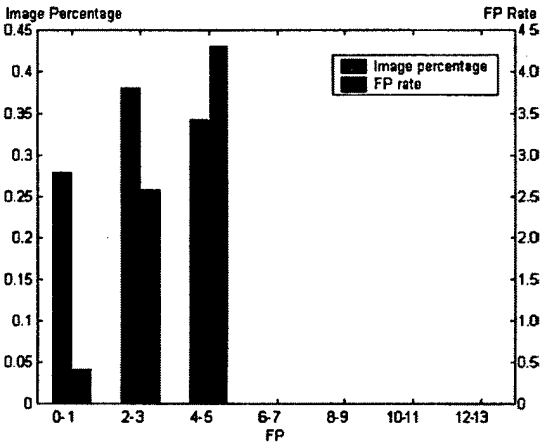


FIG. 10. Images and their corresponding average false-positive rate distributions after FP reduction with a loss of 1% TP rate by using "soft" classification.

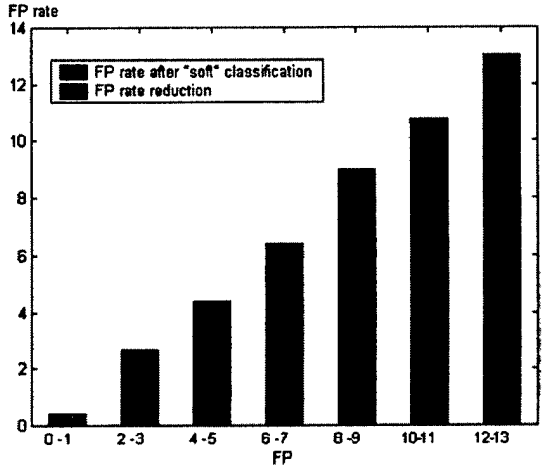


FIG. 12. Distribution of false-positive rate and its FP reduction at a loss of 1% TP rate by "soft" classification.

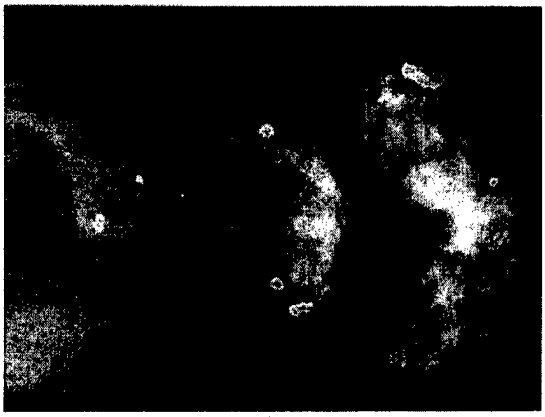


FIG. 13. FP reduction results of the three cases shown in Fig. 1 by "soft" classification.

that they usually have a low specificity in order to get a high sensitivity. Some studies have indicated that the FP signals may degrade the performance of detecting true-positive regions in cued CAD environment. In past years, a lot of efforts have been directed to classification module aiming at FP reduction. However, most of them focused on the study of application of existing classifiers (such as LDA and NN) and the optimal feature selection. In this work, we address for the first time the classification strategy in FP reduction. The basic motivation originated from the observation that: (1) the FP signals generated in current classification schemes have a very uneven distribution, specifically less FPs in fatty mammogram and much more in high dense case; (2) the FP problem may have a bigger impact on cancer screening with CAD of dense cases, because the dense breast usually has a higher risk of cancer incidence and higher false-negative detection.

In addition to the inherent difficulty because of the similarity of mammographic characteristics between dense tissue

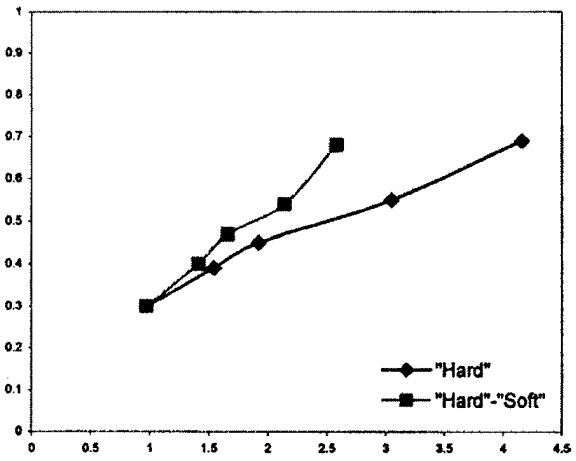


FIG. 14. FROC analysis of the effect of competitive classification strategy on false-positive reduction.

region and mass, a major problem in current CAD detection is the classification takes a "hard" classification criteria. The purpose of this work was to develop a new classification scheme by introducing a "soft" decision strategy in FP reduction. This new decision-making method provides a flexible criteria in each mammogram. The classification is image-based as opposed to region-based in "hard" classification. The evaluation results have demonstrated its efficiency. Compared to the conventional (region-based) classification, a higher FP reduction can be achieved with less loss of detection sensitivity. What should be pointed out here is this research is directed to the analysis of new classification strategy, and all the evaluations were taken on only one-scale mode in the detection system we developed.[24] Therefore the sensitivity listed in Fig. 14 is relatively low. It is believed that this new classification strategy can also be successfully applied to FP reduction in microcalcification cluster (MCC) detection.

VI. CONCLUSION

High false-positive rate is one of the major problems to be solved in a CAD study because too many false-positively cued signals will potentially degrade the performance of detecting true-positive regions and increase the call-back rate in CAD environment. A very common observation of current CAD detection is the distribution of FPs is quite different from case to case, usually less FPs for fatty breast and much more for dense breast. A direct cause of this deficiency is the existing FP reduction classifiers take "hard" decision criteria. The method proposed in this paper is a combination of "hard" and "soft" classification, where the "soft" classification takes a competitive classification strategy and only the "best" ones are selected from the pre-classified suspicious regions as the "true" mass in each case. The evaluation results demonstrate its efficiency both in reducing the FPs in high incidence cases and in the overall FP reduction. However, it should be pointed out that an optimal classification can only be made by "optimal" features with a corresponding "optimal" classifier. In order to have a better FP reduction performance, further studies will be needed including (a) optimal feature selection for hybrid classification, (b) optimized combination of "hard" and "soft" classifiers.

ACKNOWLEDGMENTS

This work is supported by a grant from Navy Medical Research and Development (USUHS) and a Career Development Award from the USAMRMC (DAMD 17-00-1).

[a]Correspondence: Lihua Li, Ph. D., Department of Radiology, College of Medicine, University of South Florida, 12901 Bruce B. Downs Blvd., Box 17, Tampa, FL 33612-4799, Tel: (813) 979-6718, Fax: (813) 979-6724, lihua@splinter.rad.usf.edu

[1]M. L. Giger, "Computer-aided diagnosis," in *RSNA Syllabus: A Categorical Course in Physics, Technical Aspects of Breast Imaging*, edited by A. G. Haus and M. J. Yaffe (RSNA, Oak Brook, 1993), pp. 283–298.

[2]R. M. Nishikawa, M. L. Giger, K. Doi, C. E. Metz, F. F. Yin, C. J. Vyborny, and R. A. Schmidt, "Effect of case selection on the performance of computer-aided detection schemes," Med. Phys. **21**, 265–269 (1994).

[3] Z. Huo, M. L. Giger, C. J. Vyborny, U. Bick, P. Lu, D. E. Wolverton, and R. A. Schmidt, "Analysis of spiculation in the comuterized classification of mammographic masses," Med. Phys. **22**, 1569–1579 (1995).
[4] S. M. Lai, X. Li, and W. F. Bischof, "On techniques for detecting circumscribed masses in mammograms," IEEE Trans. Med. Imaging **8**, 337–386 (1989).
[5] F. F. Yin, M. L. Giger, K. Doi, C. E. Metz, C. J. Vyborny, and R. A. Schmidt, "Computer-aided detection of masses in digital mammograms: Analysis of bilateral subtraction images," Med. Phys. **18**, 955–963 (1991).
[6] N. Petrick, H. P. Chan, B. Sahiner, and D. Wei, "An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection," IEEE Trans. Med. Imaging **15**, 59–67 (1996).
[7] W. E. Polakowski *et al.*, "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency," IEEE Trans. Med. Imaging **16**, 811–819 (1997).
[8] L. Li, W. Qian, and L. P. Clarke, "Digital mammography: computer-assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms," Academic Radiology **4**, 724–731 (1997).
[9] L. Li, W. Qian, and L. P. Clarke, "Wavelet transform for directional feature extraction in medical imaging," IEEE Proc. of Int. Conf. on Image Processing. Oct. 1997.
[10] L. Li, W. Qian, and L. P. Clarke, "Image feature extraction for mass detection in digital mammography: effects of wavelet analysis," Proc. of SPIE Medical Imaging, **3338** (1998) [also appear in Med. Phys. **25**, 402–408 (1999)].
[11] H. P. Chan, D. Wei, M. A. Helvie, B. Sahiner, D. D. Adler, M. M. Goodsitt, and N. Petrick, "Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space," Phys. Med. Biol. **40**, 857–876 (1995).
[12] B. Zheng, Y.-H. Chang, and D. Gur, "Computer-aided detection of masses in digitized mammograms using image segmentation and a multilayer topographic feature analysis," Academic Radiology **2**, 959–966 (1995).
[13] R. Gupta and P. E. Undrill, "The use of texture analysis to delineate suspicious masses in mammography," Phys. Med. Biol. **40**, 835–855 (1995).
[14] N. Karssemeijer and G. M. te Brake, "Detection of stellate distortion in mammograms," IEEE Trans. Med. Imaging **15**, 611–619 (1996).
[15] W. P. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screen for spiculated lesions," Radiology **191**, 331–337 (1994).
[16] D. Brzakovic, X. M. Luo, and P. Brzakovic, "An approach to automated detection of tumors in mammograms," IEEE Trans. Med. Imaging **9**, 233–241 (1990).
[17] B. Sahiner, H. P. Chan, D. Wei, N. Petrick, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue," Med. Phys. **23**, 1671–1684 (1996).
[18] B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Design of a high-sensitivity classifier based on a genetic algorithm: application to computer-aided diagnosis," Phys. Med. Biol. **43**, 2853–2871 (1998).
[19] B. Sahiner, H. P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," IEEE Trans. Med. Imaging **15**, 598–610 (1995).
[20] S. B. Lo *et al.*, "Detection of mammographic masses using sector features with a multiple circular path neural network," Proc. SPIE Med. Imaging **3338** (1998).
[21] K. S. Woods and K. Bowyer, "Computer detection of stellate lesions," *Digital Mammography*, edited by A. G. Gale *et al.* (Elsevier Science B.V., 1994).
[22] H. Li, M. Kallergi, L. P. Clarke, V. K. Jain, and R. A. Clark, "Markov random field for tumor detection in digital mammography," IEEE Trans. Med. Imaging **14**, 565–576 (1995).
[23] D. Black, *The Theory of Committees and Elections* (Cambridge University Press, 1963).
[24] L. Li, W. Qian, L. P. Clarke, R. A. Clark, and J. Thomas, "Improving mass detection by adaptive and multi-scale processing in digitized mammograms," Proc. SPIE Med. Imaging **3661** (1999).

Graph-based Region Growing for Mass Segmentation in Digital Mammography

Yong Chu [*a,b], Lihua Li [**b], and Robert A. Clark [b]
[a] Department of Computer Science and Engineering
[b] Department of Radiology, College of Medicine
H. Lee Moffitt Cancer Center & Research Institute
University of South Florida

ABSTRACT

Mass segmentation is a vital step in CAD mass detection and classification. A challenge for mass segmentation in mammograms is that masses may contact with some surrounding tissues, which have the similar intensity. In this paper, a novel graph-based algorithm has been proposed to segment masses in mammograms. In the proposed algorithm, the procedure of region growing is represented as a growing tree whose root is the selected seed. Active leaves, which have the ability to grow, in the connection area between adjacent regions are deleted to stop growing, then separating the adjacent regions while keeping the spiculation of masses, which is a primary sign of malignancy for masses. The new constrained segmentation was tested with 20 cases in USF moffitt mammography database against the conventional region growing algorithm. The segmented mass regions were evaluated in terms of the overlap area with annotations made by the radiologist. We found that the new graph-based segmentation more closely match radiologists' outlines of these masses.

Keyword: Digital mammography, Mass segmentation, Region growing, Graph-based segmentation

1. INTRODUCTION

Mammography is important in early detection of breast cancer. Masses contain important signs of breast cancer and are hard to detect as they often occur in dense glandular tissue. Numerous segmentation methods have been developed to segment masses in mammograms. A two-stage adaptive density-weighted contrast enhancement segmentation method is used by Petrick et al.[1] to detect masses. Li et al.[2] used a segmentation method based on Markov random fields to segment regions based on texture information. Multiresolution wavelet analysis techniques are developed by Chen and Lee[3] to detect mass edges. Li et al.[4] presented a statistical model supported approach for enhanced segmentation and extraction of suspicious mass area from mammographic images.

Region growing is also popularly used for mass segmentation. Some methods attempt to utilize shape constraints in conventional region growing to regularize the segmented partitions. Matthew A. Kupinski[5] proposed a new algorithm to add additional a priori information into the creation of the mass partitions. The original image is

* chuy@moffitt.usf.edu; phone 1 813 972-8400-6318; fax 1 813 979-6724; Department of Radiology, College of Medicine, H. Lee Moffitt Cancer Center & Research Institute at University of South Florida, 12901 Bruce B. Downs Blvd., MDC 17, Tampa, FL, USA 33612;
** lilh@moffitt.usf.edu; phone 1 813 979-6718; fax 1 813 979-6724; Department of Radiology, College of Medicine, H. Lee Moffitt Cancer Center & Research Institute at University of South Florida, 12901 Bruce B. Downs Blvd., MDC 17, Tampa, FL, USA 33612;

Medical Imaging 2002: Image Processing, Milan Sonka, J. Michael Fitzpatrick, Editors, Proceedings of SPIE Vol. 4684 (2002) © 2002 SPIE · 1605-7422/02/$15.00

multiplied by a function, called the constraint function that suppressed distant pixel values because masses are tend to be compact. An isotropic Gaussian function centered on the seed point location with a fixed variance is chosen as the constraint function. More circular regions are generated when this method is used because it is possible that spiculations are more likely to be eliminated when the image is multiplied by Gaussian function. Huo et al.[6] extracted mass regions using region growing methods and calculated two features, circularity Circ() and size Size(), for every grown regions. The final partition is chosen by analyzing these functions and determining transition points or jumps in the features. But if a transition point cannot be found, the segmentation algorithm fails to return a final partition. Both methods didn't take the local shape of speculations and the connected area between adjacent objects into account.

In this paper, we present a graph-based region growing method for segmenting masses in digital mammograms. In the proposed algorithm, the procedure of region growing is represented as a growing tree whose root is the selected seed. Active leaves, which have the ability to grow, in the connection area between adjacent regions are deleted to stop growing, then separating the adjacent regions while keeping the spiculation of masses. The contents of the paper are organized as follows. Section 2 presented a graph theory description of region growing. A new graph-based region growing method for mass segmentation is proposed in Section 3. Section 4 describes the results of the mass segmentation experiments. Finally we have a conclusion of this study in Section 5.

2. GRAPH THEORY DESCRIPTION OF REGION GROWING

2.1. Region growing

Region growing (RG)[7] constructs a region by starting from a user-provided pixel called seed. At each iteration, the algorithm finds the boundary pixels of the region that participates in the growing procedure. For each of these boundary pixels, the neighborhood is checked for pixels that are not assigned to this region. If the intensity of checked pixel is close to the mean intensity of the region under consideration, then the pixel is assigned to this region. The algorithm stops when no pixels to be classified to the region can be found.

Given an image, we define the set of coordinates in this image as:

$$IM = \{(x, y): x = 1,2,\cdots,m \quad and \quad y = 1,2,\cdots,n\}. \tag{1}$$

The function describing the pixel gray levels of the image is given by $I(x, y)$ where $(x, y) \in IM$. Assume region growing stops at iteration N. For conventional region growing segmentation, the segmented region R_i at iteration i: $i = 1,\cdots,N$ can be represented as

$$R_i = \{(x, y): |I(x, y) - M_{i-1}| < T_{\text{int}} \quad and \quad NBR_{i-1}(x, y) = 1\}. \tag{2}$$

Where M_{i-1} is the mean intensity of R_{i-1} at iteration $i-1$, T_{int} is a gray-level homogeneous threshold. $NBR_{i-1}(x, y)$ is defined as

$$NBR_{i-1}(x, y) = \begin{cases} 1 & (x-1, y), (x+1, y), (x, y-1) \text{ or } (x, y+1) \in R_{i-1} \\ 0 & \text{All } (x-1, y), (x+1, y), (x, y-1), (x, y+1) \notin R_{i-1} \end{cases} \tag{3}$$

Region growing is a local thresholding process, which utilizes only the gray-level information around the seed point.

2.2. Some concepts about directed graph

In the proposed method, region growing is represented as a growing tree whose root is the selected seed. A rooted tree is a directed tree. For the sake of convenience in method description, some concepts about directed graph[8] are introduced first.

Directed graph: A directed graph D is a pair of sets (V, A) where V is nonempty, and A is a set of ordered pairs of elements of V. The elements of V are called the vertices of D and the elements of A are called the directed edges or arcs.

Indegree and outdegree: In a directed graph, at any vertex there are some number of arcs directed toward, or coming into, that vertex and some number of arcs directed away from, or coming out of, that vertex. The number of arcs coming into a vertex v is called the indegree of v. The number of arcs coming out of a vertex v is called the outdegree of v.

Rooted tree: A rooted tree T is a directed graph that has precisely one vertex r that has indegree 0. The vertex is called the root of the rooted tree. Those vertices with outdegree 0 are called leaves.

Parent vertex and child vertex: If there is an arc from vertex v to vertex w, we say that w is a child vertex of v, v is the parent vertex of w.

2.3. Growing tree for RG

In the proposed method, region growing is represented as a growing tree whose root is the selected seed. The vertices correspond to the pixels in the segmented region. The segmented region R_i at iteration i is represented as the rooted tree T_i. The children of a vertex will be within the 4-neighborhood of it. The 4-neighborhood $N_4(v)$ of a vertex $v=(x,y)$ consists of the four pixels whose positions in the image differ from v in only one coordinate by ± 1, shown in Figure 1(b).

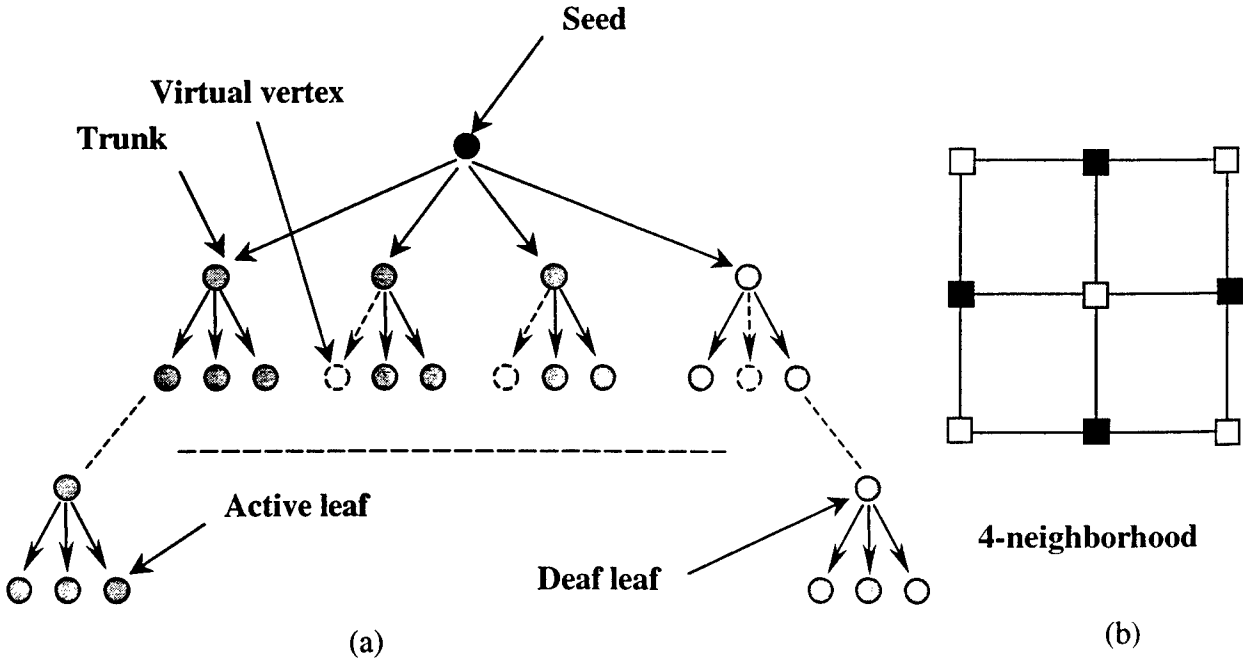


Figure 1 (a) Iteration i of region growing as a tree T_i (b) The 4-neighborhood of a vertex

We need to introduce some new definitions for growing tree used in the proposed method, shown in Figure 1(a).

(1).**Active vertex** (va): All vertices, which are in T_i but not in T_{i-1}, are called active vertices in T_i. They correspond to the pixels, which are assigned to R_i but not R_{i-1}. The root is the only active vertex in T_0.

(2). **Real child vertex**: For any active vertex va in T_{i-1}, if there is a pixel in $N_4(va)$ assigned to R_i but not R_{i-1}, this pixel called a real child vertex of va in T_i with one real arc coming into it. This pixel is also a real vertex (vr) in T_i. All real vertices in T_{i-1} are real vertices in T_i. All active vertices T_i are the real child vertices of active vertices in T_{i-1}.

(3).**Virtual child vertex**: For any active vertex va in T_{i-1}, if there is a pixel in $N_4(va)$ assigned to R_{i-1} except its parent vertex, this pixel is viewed as a virtual child vertex of va in T_i with one virtual arc coming into it.
(4). **Indegree, outdegree and vtdegree**: we only consider real vertices. The number of real arcs coming into a real vertex vr in T_i is called the indegree of vr in T_i. The number of real arcs coming out of a real vertex vr in T_i is called the outdegree of vr in T_i. The number of virtual arcs coming out of a real vertex vr in T_i is called the vtdegree of vr in T_i.
(5). **Active leaf** (la): Active vertices in T_i are called active leaves in T_i, which are real vertices in T_i but not in T_{i-1}. The sum of vtdegree and outdegree of an active leaf in T_i is 0. Active leaves in T_i are child vertices of active leaves in T_{i-1}.
(6). **Dead leaf** (ld): Dead leaves in T_i are real vertices in T_{i-1} with the sum of vtdegree and outdegree smaller than 3.
(7). **Trunk**: Trunks in T_i are real vertices in T_{i-1} with the sum of vtdegree and outdegree equal to 3.

Based on the definitions above, region growing constructs a rooted growing tree in which the root is a user-provided seed. At iteration i, the segmented region is represented as T_i. For any vertex except the root, there are one parent vertex and at most three child vertices (no parent vertex and at most four child vertices for the root). The active leaves and dead leaves in T_i constitute the boundary pixels of the boundary pixels of the region R_i. Only the active leaves contribute to the growing of the tree T_i to the tree T_{i+1}. For each of these active leaves, the neighborhood is checked for the vertices that are not in T_i. If the intensity of checked pixel is close to the mean intensity of region under consideration, then the pixel is assigned to T_{i+1}, which are the active leaves of T_{i+1}. The algorithm stops when no active leaves can be found.

3. GRAPH-BASED RG FOR MASS SEGMENTATION

In mammography, masses may contact with surrounding tissues, which have the similar intensity. The new method discussed in this paper attempts to solve this problem by computing the probability of extension to other tissues of every pixel for every iteration in region growing. The pixels with high probability of extension are not allowed to grow, and therefore separating the adjacent tissues from masses while keeping spiculations.

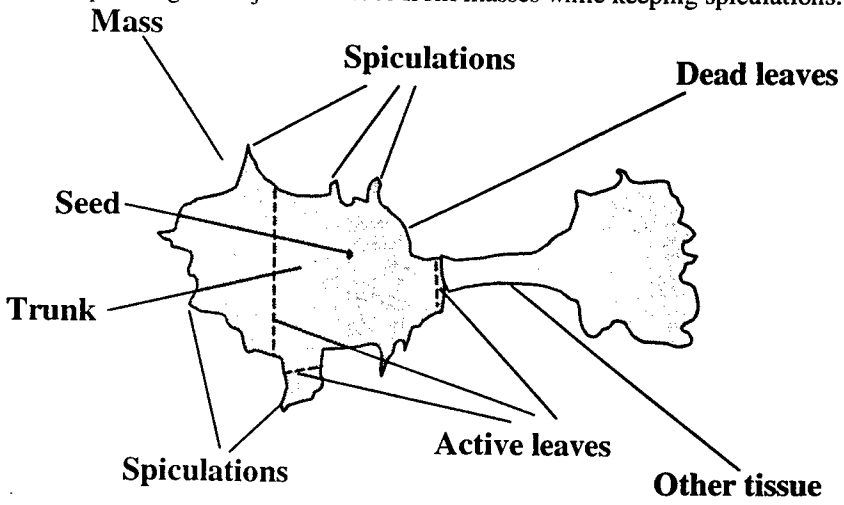


Figure 2 Mass segmentation at iteration i in region growing

3.1. The probability of extension into other tissue

All active leaves at iteration i (real vertices in T_i but not in T_{i-1}) are the child vertices of the active leaves of T_{i-1}. It means that only active leaves have the ability to grow at each iteration, so the probability of extension into other tissues (PE) for each trunk and dead leaf is 0.

At iteration i, all neighboring pixel of active leaves in T_{i-1} are checked. If they are not in R_{i-1} and satisfy $|I(x,y)-M_{i-1}|<T_{\text{int}}$, they are supposed to be in R_i. The number of the supposed active leaves and dead leaves are computed. Then active leaves (la s) are labeled into active groups (G s). Active groups are the finite (8-) connected set of points, that are, discrete curves. The size of an active group ($|G|$) is defined as the num of points in G. From Fig.2, we can see that a spiculation or a bridge between the mass and other tissue, at iteration i, can be characterized as a small set of the active leaves surrounded by dead leaves. An active group with small $|G|$, which is surrounded by dead leaves, is called a hole (H) in the proposed method. We can say that the size of H ($|H|$) is decreasing if a hole represents a spiculation, but not for a bridge between the mass and other tissue. Let a hole grow forward 2n iterations, the mean size of the hole in first n iterations is given by M_{fHL}, and the mean size of the hole in the second n iteration is given by M_{sHL}. The PE of each pixel in this hole can be given by $\frac{M_{sHL}}{M_{sHL}+M_{fHL}}$. Let us summarize it as following.

$$PE_i(x,y)=\begin{cases}0 & (x,y)\text{is a dead leaf or a trunk at iteration i}\\ 0 & (x,y)\text{is an active leaf and (x, y)}\notin\text{ hole at iteration i}\\ \dfrac{M_{sHL}}{M_{sHL}+M_{fHL}} & (x,y)\text{is an active leaf and (x, y)}\notin\text{ hole at iteration i}\end{cases}. \quad (4)$$

3.2. Graph-based RG in Mass segmentation

The conventional RG method makes use of the fact that the intensities of pixels inside the mass are close to each other, but it does not take shape into account in growing. Irregular segmentation results can arise in conventional region growing. The graph-based region growing algorithm adds constrained condition to prevent mass region from effusing into other tissues while keeping the spiculations of the mass. Mathematically, the new algorithm is defined as

$$R_i=\{(x,y):|I(x,y)-M_{i-1}|<T_{\text{int}} \quad and \quad NBR_{i-1}(x,y)=1 \quad and \quad PE_i(x,y)<T_{ext}\}. \quad (5)$$

where PE is the probability of extension into other tissues, T_{ext} is the threshold of PE. The pixels, which have the probabilities greater than T_{ext}, will be deleted from the segmented region.

A complete procedure of graph-based region growing algorithm is described below:

(1). Select a seed for region growing and define the seed as an active leaf in T_0.

(2). At iteration i, every neighboring pixels (x,y) of active leaves in T_{i-1} are checked. If they are not in R_{i-1} and satisfy $|I(x,y)-M_{i-1}|<T_{\text{int}}$, they are supposed to be in R_i.

(3). Compute the number of the supposed active leaves $|la|$ and dead leaves.

(4). Label the active leaves groups (G s) and calculate the size of active groups $|G|$.

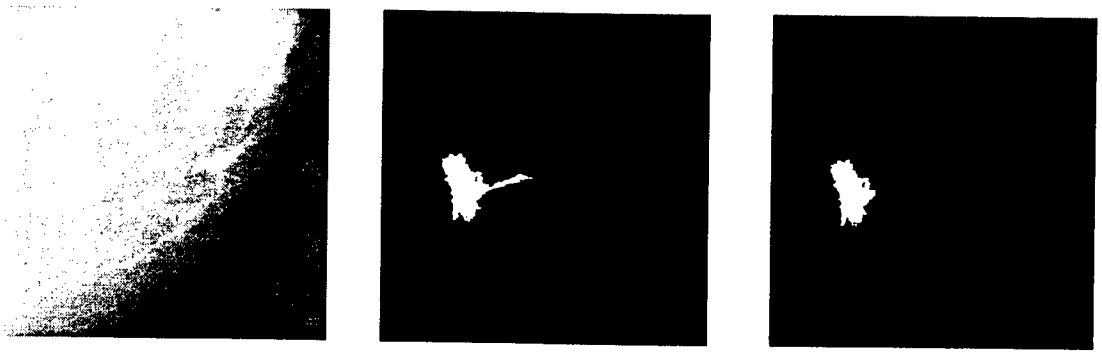
(5). If $|G| < T_a \times (|la| + |ld|)$, compute the number of the neighboring dead leaves $|lsd|$ of the active groups. If $|lsd| > T_h \times |G|$, the active group is a hole (H).

(6). Compute PE for supposed active leaves in H. If $PE \geq T_{ext}$, the supposed active leaves in the H are eliminated from R_i.

(7). Continue the procedure, the algorithm stops when no active leaves are found.

4. RESULTS AND DISCUSSION

Segmentation results for a mass connected with surrounding tissues are shown in Figure 3. Conventional region growing extend into the background and thus a connected tissue is included in the segmented region. The graph based region growing does not get confused by the tissue connected with the mass and correctly segments this mass while keeping spiculations. From figure 3, we can see that the graph based region growing method segment the mass better than conventional region growing.



(a) (b) (c)

Figure 3. Segmentation results for (a) a mass contacted with surrounding tissues using (b) graph-based region growing (c) conventional region growing

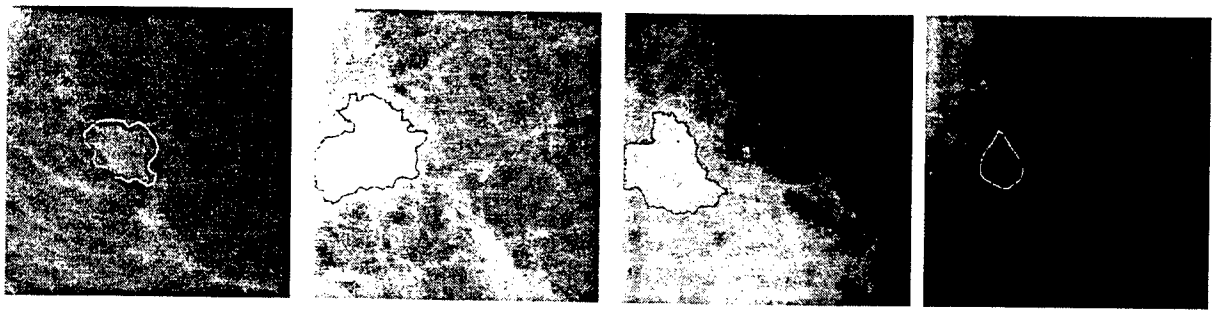


Figure 4. Comparison results of segmentation. Black boundaries denote the graph-based segmentation. White boundaries denote the radiologist-segmented results.

In order to examine the performance of the proposed method, the segmentation results of conventional region growing and the constrained region growing are compared against radiologist's outlines of the masses as shown in Figure 4. The screening database of masses with a total of 20 visible masses was employed.

T_{int} and T_{ext} in (5) is unknown. T_{int} is determined by the mean intensity of the surrounding pixels of the seed(M_{seed}). T_{int} is empirically chosen as $0.025 \times M_{seed}$ in our experiment. A value of 0.5 is chosen to be T_{ext}. For each mass, the seed point was selected within the radiologist's outline. Once the mass was segmented, an overlap O was calculated:

$$O = \frac{area(S \cap T)}{area(S \cup T)}. \tag{6}$$

where S is the segmented mass and T is the radiologist's hand-drawn segmentation result. The value of O is between 0 (no overlap) and 1 (exact overlap). To evaluate the segmentation result, we need to choose a threshold for an "adequate" segmentation. If O is greater than a certain value, the mass is considered to be correctly segmented. Figure 5 shows a plot of the fraction of masses correctly segmented at various overlap threshold levels. The graph-based region growing algorithm substantially outperformed the conventional region growing. At an overlap threshold of 0.3, conventional region growing correctly identified 50% of the lesions in our database, while the graph-based segmentation methods correctly segmented 90% of the masses.

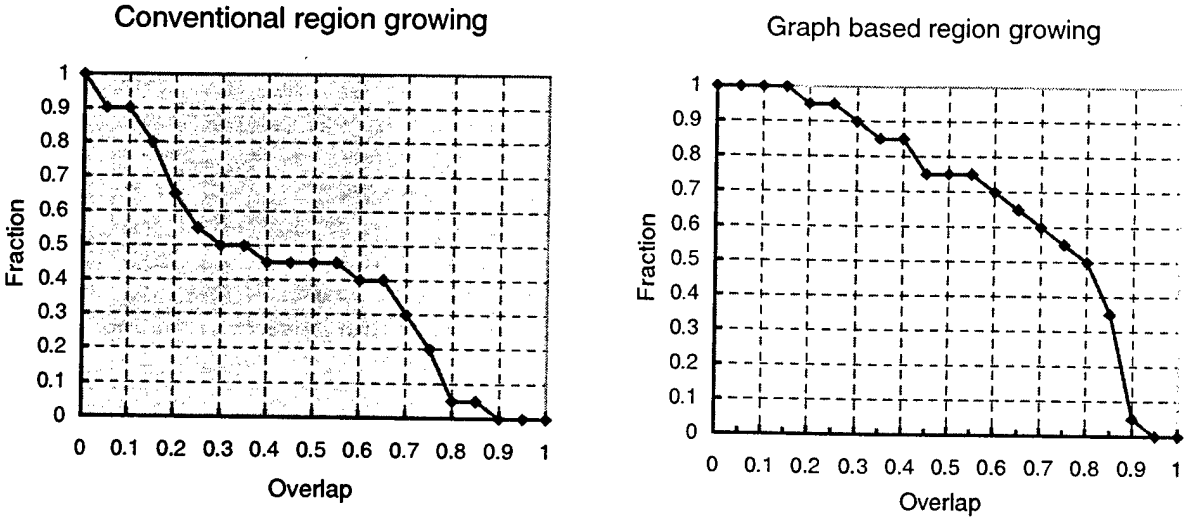


Figure 5. Results for the methods using the overlap evaluation. The left figure shows the results of conventional region growing. The right figure shows the results for the graph-based region growing.

5. CONCLUSION

We have developed a new method of seeded mass segmentation for use in digital mammography. This new method substantially outperforms conventional region growing segmentation. At an overlap threshold of 0.3, conventional region growing correctly identified 50% of the lesions in our database, while the graph-based segmentation methods correctly segmented 90% of the masses. This method can be potentially applied to the discrimination between benign and malignant masses, because the spiculation feature can be effectively preserved in segmentation. However, there are still some rooms for improvement. For example, this method does not work well yet for large mass. Extensive testing and evaluation is needed with this method.

ACKNOWLEDGEMENTS

This work is supported by The Naval Medical Research and Development Command, funds managed by The Henry M. Jackson Foundation for the Advancement of Military Medicine (#600-0675299-169) and a Career Development Award from the USAMRMC (DAMD 170010245).

REFERENCE

1. N. Petrick, H.P. Chan, B. Sahiner, and D. Wei, "An Adaptive Density-Weighted Contrast Enhancement Filter for Mammographic Breast Mass Detection," *IEEE Transaction on Medical Imaging*, **vol. 15, no. 1**, pp. 59-67, February 1996.
2. H.D. Li, M. Kallergi, L.P. Clarke, V.K. Jian, and R.A. Clark, "Markov random field for tumor detection in digital mammography," *IEEE Transaction on Medical Imaging*, **vol. 14, no. 3.**, pp. 565-576, September 1995.
3. Chen CH and Lee GG, "On digital mammogram segmentation and microcalcification detection using Multiresolution wavelet analysis," *Graphical Models and Image Processing*, **59(5)**, pp. 349-364, 1997
4. H. Li., Y. Wang, K.J.R. Liu, S.C.B. Lo, and M.T. Freedman, "Computerized Radiographic Mass Detection-Part I: Lesion Site Selection by Morphological Enhancement and Contextual Segmentation," *IEEE Trans. Medical Imaging*, **vol. 20, no. 4**, April 2001
5 Matthew A. Kupinshi and Maryellen L. Giger, "Automated Seeded Lesion Segmentation on Digital Mammograms," *IEEE Transaction on Medical Imaging*, **vol. 17, no. 4**, pp. 510-517, August 1998.
6 Zhimin Huo, Maryellen L. Giger, "Analysis of Spiculation in the Computerized Classification of Mammographic Masses," *Medical Physics*, **vol. 22, no. 10**, pp. 1569-1579, 1995.
7 Jonathan Gross and Jay Yellen, *Graph Theory and its applications*, pp. 1-46, CRC Press, USA,1999.
8 Kenneth R. Castleman, *Digital Image Processing*, pp. 468-470, Prentice-Hall International Inc., 1996.

Hybrid Classification Method for False-Positive Reduction in CAD for Mass Detection

LIHUA LI
ROBERT A. CLARK
Department of Radiology, College of Medicine
H. Lee Moffitt Cancer Center and Research Institute
University of South Florida, Tampa, Florida

INTRODUCTION

Computer-aided diagnosis (CAD) has been proposed as a "second opinion" or a "pre-reader" strategy to improve mammogram interpretation (Giger 1993). Methods reported have included the detection and classification of either microcalcification clusters or masses. Compared to microcalcification cluster detection, mass detection clearly poses a much more difficult problem and usually generates more false-positive (FP) signals in detection because masses are often (a) of varying characteristics in size, shape, and density; (b) exhibit poor image contrast in breast with high density tissue; (c) have a similar characteristics to the nonuniform dense tissue background; and (d) highly connected to the parenchymal tissue, particularly for spiculated mass.

Numerous investigators have addressed mass detection and classification in the past several years (Karssemeijer et al. 1996, Li et al. 1999, Petrick et al. 1996, Yin et al. 1991, Zheng et al. 1995). The detection process usually consisted of three major steps: preprocessing, localization, and classification. Most of the proposed research has been directed to the first two CAD modules and the feature selection part in classification module. Although the methods proposed to date have met varying levels of success, the high false-positive rate remains to be one of the major problems to be solved in CAD study because too many false-positively cued signals will significantly degrade the performance of detecting true-positive regions in the CAD environment. In this paper, we propose a new approach to classification. Its basic strategy is to improve the classification performance in cases with more false-positive signals by using "soft" decision making.

HYBRID CLASSIFICATION METHOD

The major difficulties of classification result from the great similarity in appearance between masses and dense normal tissue, and the great variation in feature distribution of different masses. From the classification perspective, the former requires more elaborate features to be extracted for classification while the latter means the classifier structure should be more flexible. Here we focus on the exploration of new classification strategy with less effort on new feature design.

Feature Extraction

Seven features are used in this work for FP reduction. They are similar to that we reported before except that (a) one more mixed boundary-intensity feature is added, and (b) the calculation of some features is modified with more reasonable definition (Li et al. 1997). Due to page limitations, they are simply listed as follows:

Morphological features: (i) Area; (ii) Circularity; (iii) Normalized deviation of radial length.

Intensity features of region: (i) Intensity variation; (ii) Mean intensity difference.

Intensity features of boundary: (i) Mean gradient of region boundary; (ii) Mean intensity difference along region boundary.

Hybrid Classification

There are several types of classifiers used in discrimination of masses from normal tissue regions, such as decision tree (Kegelmeyer et al. 1994), Bayes classifier (Brzakovic et al. 1990), neural networks (Sahiner et al. 1995, Lo et al. 1998), linear discriminant analysis (LDA) (Chan et al. 1995), and quadratic classifiers (Woods et al. 1994). A common characteristics of current FP reduction schemes is that a "hard" discriminant criteria is developed by training to evaluate each segmented suspicious region. From the standpoint of region-based classification, it is reasonable, and the task of classifier design is to find an optimal discriminant hyperplane in feature space. However, the segmented FPs cannot be reduced efficiently by a single "hard" decision classifier. With an analysis of "hard" decision-making classification results, it can be found that the distribution of FPs is quite different from case to case, generally less FPs for fatty breast and more FPs for dense breast in which more mass-like regions are segmented. In other words, from the image-based standpoint, the efficiency of classification strategy by means of "hard" decision criteria is questionable. The proposed hybrid classification method takes two different decision-making strategies as described below, where the current "hard" decision classifier is cascaded with a "soft" decision classification with the objective to reduce FPs in the cases with multiple FPs retained after the "hard" decision classification.

"Hard" Classification Using a Modified Fuzzy Decision Tree Method

The segmented regions are first classified using a modified fuzzy decision tree (MFDT) method. It is a modification of our previously developed fuzzy binary decision tree (FBDT) method (Li et al. 1997), in which the classification decision at each tree node is made based on the region feature as well as region size. The design of MFDT structure is based on an observation that the feature distributions are statistically different for different mass size.

"Soft" Classification Using Committee Decision Method

The "hard" classification method is region-oriented. The thresholds at the decision nodes of MFDT are selected by training to ensure a high detection sensitivity.

However several FPs in each image may be produced especially for that with non-uniform high density tissue. The "soft" classification strategy proposed is to reduce FPs by an image-based analysis for selecting the best ones among the pre-classified candidate regions as the "real" masses. It is performed by a committee decision method based on a simple premise that most features of a mass should be top in individual feature ranking among all the candidate suspicious regions in a single mammogram (Black 1963).

The committee consists of six "experts"; each provides a ranking list of the suspicious regions with a likelihood to be a mass using a single feature except the feature "Area." A region is decided by the committee to be a possible mass if it is ranked among top K_0 ($K_0 \leq M$) by more than C_0 "experts."

Mathematically, it can be described as follows:

For a region i among M candidate regions, its ranking value by feature j is

$$r_i^{(j)} = k,\ k \in [1, M] \tag{1}$$

if feature j of region i is at the order k among M regions.

The committee score of region i is defined as

$$C_i = \sum_{j=1}^{N} \cup\left(r_i^{(j)}\right) \tag{2}$$

where

$$\cup\left(r_i^{(j)}\right) = \begin{cases} 1 & k \leq K_0 \\ 0 & k > K_0 \end{cases} \tag{3}$$

and K_0 is an order threshold, N is the "expert" number.

The region i is classified to be a mass if $C_i \geq C_0$, i.e. more than C_0 "experts" vote region i for mass class provided they can select K_0 regions from all M regions as candidate masses. The maximum number of "mass" in each image is

$$MX = \max\left\{\left[\frac{k_0 * N}{C_0}\right], M\right\} \tag{4}$$

where [*] is the operation of taking maximum integer value.

EVALUATION

A comparative evaluation of false-positive reduction by means of a "hard" decision classifier and a hybrid "hard"-"soft" decision classifier was performed using a database consisting of 30 normal and 49 abnormal mammograms with a total of 74 masses. The electronic truth file was formed for each abnormal mammogram where the masses were labeled by an expert mammographer based on visual criteria and biopsy results.

The evaluation was taken in three steps. First, a detection result was obtained by our CAD detection system with only "hard" classification working at its trained operating point. Figure 1 shows the distributions of image and their group-average false positive rate versus false-positive signal numbers on the image. Its corresponding accumulative distributions of FPs versus two different image percentages is shown in figure 2. It is observed that the generated FP number is quite different from case to case. Some cases do not have any FPs in detection but some cases generated more than 10 FPs per image. One-third of the cases account for almost two-thirds of all FPs in detection. The second step is to evaluate the FP reduction with "hard" classification by increasing the threshold value of classifier. Figure 3 shows the distributions of image and their corresponding average FP rate. Compared to figure 1, the percentage of images with lower FP is increased, which means some FPs are removed in the cases with higher FP number. Figure 4 shows the distribution of FP rate and its FP reduction of each image group with different numbers of FPs. It is observed that although the overall average FP rate is reduced (from 4.16 to 3.08 per image in whole database) and the average reduction in images with higher FP incidence is more than that in lower FP images, the distribution of FPs remains very uneven. The 17.7% cases with higher FP numbers (more than 5 FPs) generated almost half of all the FPs. Furthermore, this very limited FP reduction is obtained at a large cost of 14% loss of detection sensitivity. The third step of evaluation is taken on the proposed hybrid classification scheme in which the "hard" classifier is cascaded with a novel "soft" classification. Figure 5 shows the distributions of images and their corresponding average false-positive rate versus false-positive signal number in each of the images after "soft" classification. A significant change of image distribution is observed, where no image has an output with more than 5 FPs. For a comparison with the result of single "hard" classification, a distribution of FP rate and its FP reduction in each image group is shown in figure 6. It is observed that, compared to the FP reduction scheme with increased threshold

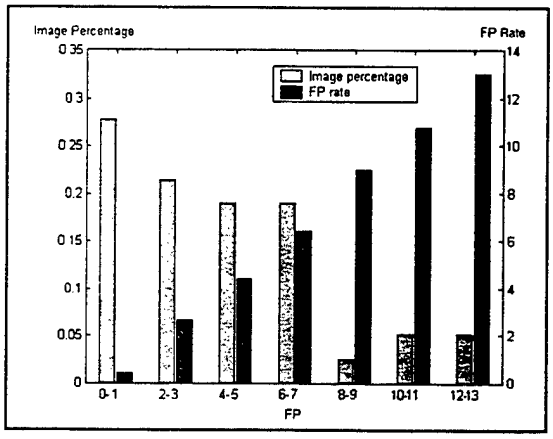


Figure 1. Images and their corresponding average false positive rate distributions versus false-positive signal numbers after initial "hard" classification.

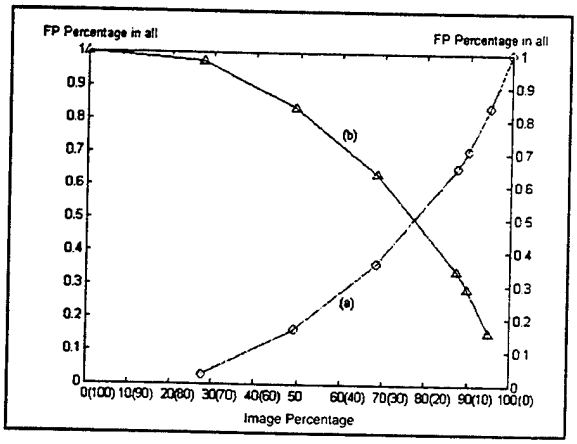


Figure 2. Accumulative distribution of FPs versus two different image percentages. (a) Beginning from the low FP image group, (b) Beginning from the high FP image group.

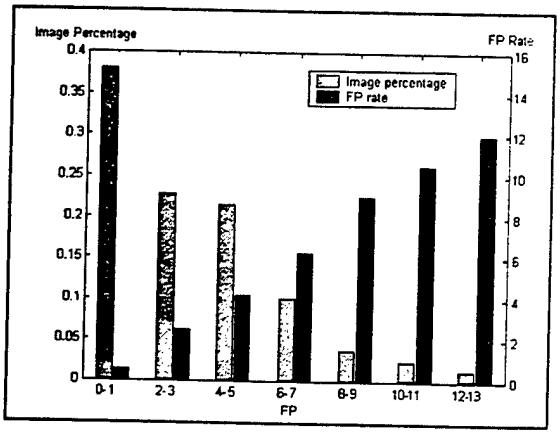


Figure 3. Images and their corresponding average false positive rate distributions after FP reduction with a loss of 14% TP rate by using "hard" classification with increased threshold.

"hard" classification, a much bigger false-positive reduction is achieved in the images with more FP detection signals. The final average detection FP rate in the whole database is reduced to 2.24 with only 1% loss of detection sensitivity. The distribution of FP rate in different image groups is more uniform, which means most of the images generate a similar number of FPs except for the ones with very few FPs. This result demonstrates that "soft"-decision classification can effectively remove the FPs in higher FP incidence cases with less impact on the images with less FPs for securing a high sensitivity.

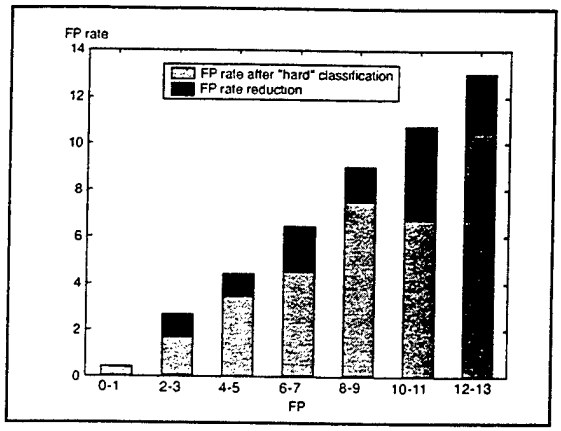


Figure 4. Distribution of false positive rate and its FP reduction at a loss of 14% TP rate by threshold-increased "hard" classification.

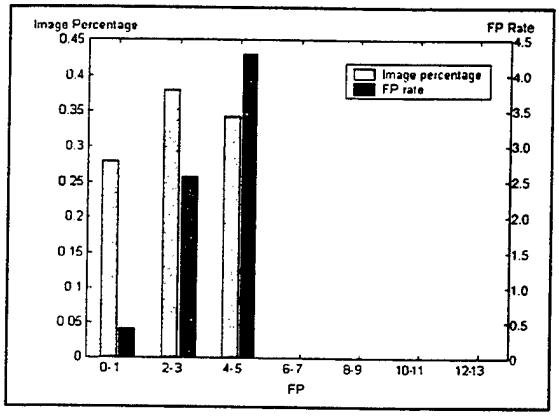


Figure 5. Images and their corresponding average false positive rate distributions after FP reduction with a loss of 1% TP rate by using "soft" classification with increased threshold.

CONCLUSION

This paper proposes a new classification scheme by introducing a "soft" decision strategy in FP reduction. This new decision-making method provides a flexible criteria in each mammogram. The classification is image-based as opposed to region-based in "hard" classification. The evaluation results have demonstrated its efficiency. Compared to the conventional (region-based) classification, a higher FP reduction can be achieved with less loss of detection sensitivity. However, it should be pointed out that an optimal classification can only be made with "optimal" features and "optimal" classifiers. To have a better FP reduction performance, further studies will be taken including (a) optimal feature selection for hybrid classification, and (b) optimized combination of "hard" and "soft" classifiers.

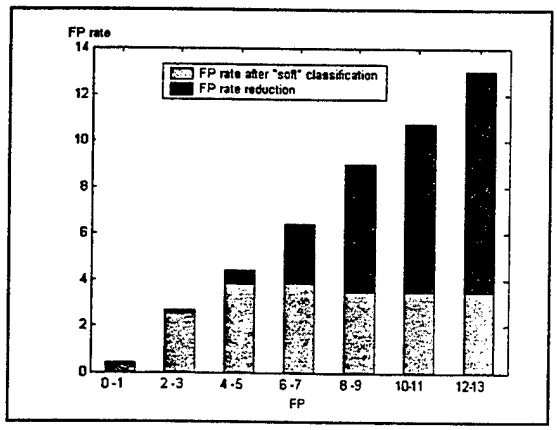


Figure 6. Distribution of false positive rate and its FP reduction at a loss of 1% TP rate by "soft" classification.

REFERENCES

Black, D., *The Theory of Committees and Elections.* Cambridge: The Cambridge University Press, 1963.

Brzakovic, D., X. M. Luo, and P. Brzakovic (1990). "An approach to automated detection of tumors in mammograms." *IEEE Trans. Med. Imag.* 9: 233–241.

Chan, H. P., D. Wei, M. A. Helvie, B. Sahiner, D. D. Adler, M. M. Goodsitt, N. Petrick (1995). "Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space." *Phys. Med. Biol.* 40: 857–876.

Giger, M. L. "Computer-Aided Diagnosis" in Radiological Society of North America (RSNA) Syllabus: *A Categorical Course in Physics, Technical Aspects of Breast Imaging.* A. G. Haus and M. J. Yaffe (eds.). Reston, Virginia: RSNA, pp. 283–298, 1993.

Karssemeijer, N., and G. M. te Brake (1996). "Detection of stellate distortion in mammograms." *IEEE Trans. Med. Imag.* 15(5): no page nos. given.

Kegelmeyer, W. P., J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper (1994). "Computer-aided mammographic screen for spiculated lesions." *Radiol.* 191: 331–337.

Li, L., W. Qian, and L. P. Clarke (1997). "Digital mammography: Computer-assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms." *Acad. Radiol.* 4: 724–731.

Li, L., W. Qian, L. P. Clarke, R. A. Clark, and J. Thomas (1999). :Improving mass detection by adaptive and multi-scale processing in digitized mammograms." *SPIE Med. Imag.*

Lo, S. B., H. Li, A. Hasegawa, Y. J. Wang, et al. (1998). "Detection of mammographic masses using sector features with a multiple circular path neural network." *Proc. SPIE Med. Imag.*

Petrick, N., H. P. Chan, B. Sahiner, and D. Wei (1996). "An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection." *IEEE Trans. Med. Imag.* 15(1): no page nos. given.

Sahiner, B., H. P Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt (1995). "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images." *IEEE Trans. Med. Imag.* 15(5): no page nos. given.

Woods, K., and K. Bowyer. "Computer Detection of Stellate Lesions" in *Digital Mammography*, vol. 1069 of Excerpta Medica International Congress Series. A. Gale, S. Astley, D. Dance, A. Cairns (eds.), Proceedings of the 2nd International Workshop on Digital Mammography, York, England, June 1994. Amsterdam, The Netherlands: Elsevier Science BV, 1994.

Yin, F. F., M. L. Giger, K. Doi, C. E. Metz, C. J. Vyborny, R. A. Schmidt (1991). "Computer-aided detection of masses in digital mammograms: analysis of bilateral subtraction images." *Med. Phys.* 18: 955–963.

Zheng, B., Y.-H. Chang, D. Gur (1995) "Computer-aided detection of masses in digitized mammograms using image segmentation and a multi-layer topographic feature analysis." *Acad. Radiol.* 2: 959–966.

Improving mass detection by adaptive and multi-scale processing in digitized mammograms

Lihua Li, Wei Qian, Laurence P. Clarke, Robert A. Clark, Captain Jerry Thomas *

Department of Radiology, College of Medicine and the H. Lee Moffitt Cancer Center
and Research Institute at the University of South Florid
* Uniformed Services University of the Health Sciences, Bethesda, Maryland, Washington DC

ABSTRACT

A new CAD mass detection system was developed using adaptive and multi-scale processing methods for improving detection sensitivity / specificity, and its robustness to the variations in mammograms. The major techniques developed in system design include (a) image standardization by applying a series of preprocessing to remove extrinsic signal, extract breast area, and normalize the image intensity; (b) multi-mode processing by decomposing image features using directional wavelet transform (DWT) and non-linear multi-scale representation using anisotropic diffusion (AD); (c) adaptive processing in image segmentation using localized adaptive thresholding (LAT) and adaptive clustering (AC); and (d) combined "hard"-"soft" classification by using a modified fuzzy decision tree and committee decision-making method. Evaluations and comparisons were taken with a training dataset containing 30 normal and 47 abnormal mammograms with totally 70 masses, and an independent testing dataset consisting of 100 normal images, 39 images with 48 minimal cancers and 25 images with 25 benign masses. A high detection performance of sensitivity TP=93% with false positive rate FP=3.1 per image and a good generalizability with TP=80% and FP=2.0 per image are obtained.

Keywords: Mass Detection, Adaptive, Multi-Scale Processing, Classification

1. INTRODUCTION

Breast cancer is the second leading cause of cancer death among American women. The American Cancer Society projects 181,600 new cases and 44,190 deaths from breast cancer in 1997 [1]. There is considerable evidence that early diagnosis and treatment of breast cancer can significantly increase chances of survival. Of all screening methods currently available, mammography is the most reliable and has demonstrated its benefit in the early detection of breast cancer with mortality reduction of 20-40% [2]. However, because mammogram interpretation is performed by radiologists by visual examination of the films for the presence of abnormalities that may be malignant, the shortage of radiologists and the large volume of mammogram to be analyzed, most of which are normal, make such readings labor intensive, cost ineffective, and often inaccurate. Similarly there is significant inter- and intra-variability in reading mammogram. Studies indicate that, of the breast cancers that are visible in retrospective studies, 10 to 30% are missed during mammographic interpretation, and 40% of the missed cancers appear as masses on the mammogram [3][4]. The missed lesions can be caused by a number of factors, but a significant percentage have been attributed to subjective or varying decision criteria , distraction by other image features, or simple oversight. In an attempt to reduce the cost and increase the effectiveness of mammography, alternative techniques and systems have been developed to improve mammographic imaging and interpretation, among which computer-aided diagnostic (CAD) method is a low cost and efficient approach [5]. There are several examples in the literature of increasing lesion detection through the use of CAD methods [6][7]. However, to date,

there is few practical CAD system for clinical use because (a) they have not achieved sufficient performance (sensitivity and specificity); (b) they have poor reproducibility and adaptivity, i.e. have a great variation in performance for different mammograms at different time and/or sites; (c) their required processing is usually not cost-effective. This paper addresses the first two problems described above in mass detection aiming at improving its detection performance and robustness by making the CAD method adaptive and more generalizable.

2. DETECTION METHOD

The new mass detection system is a modification of our previously developed wavelet based detection method by using some novel techniques such as adaptive, nonlinear multiscale processing and hybrid classification methods. Figure 1 shows the schematic diagram of the proposed mass detection system. It is a modular structure and explained as follows.

2.1. Preprocessing

The images used in CAD algorithm design and testing are obtained mostly by scanning clinical screen film mammograms with digitizers. Due to the difference in film resources and digitizers, and the variation of imaging procedure and characteristics of digitizer, the appearance of digitized images may have a great difference. To reduce the influence of non-breast signals on mass detection, the digitized mammograms are "standardized" by a series of preprocessing of breast area extraction, removal of Be-Be mark, noise suppression and image intensity normalization.

2.2. Feature Decomposition Using Directional Wavelet Transform

Different types of masses tend to be of different image appearance and characteristics. For most of masses, they are primarily related to the region feature. However, for architectural distortion and spiculated masses, their important signs are reflected as directional feature as well as their regional features. Although, for a systematic detection purpose, it will be a great help to utilize both types of features, each may be a "trouble" for another if they are not dealt with appropriately. For example, spiculation is an important feature for detection of stellate mass and classification of malignancy, but it frequently results in the connection of central mass area to surrounding tissue area in mass segmentation. More often is the case when vessels overlap on mass in the projected image, which complicated the segmentation of suspicious regions. Therefore, it is necessary to decompose the mammographic image into a region-based image and line-based image so that each can be processed more efficiently. Directional wavelet transform proposed in our previous work is used here to decompose the mammogram feature [8][9]. For an input mammographic image, two output feature images are obtained: one is a directional texture image, another is the smoothed version of original image.

2.3. Multi-Scale Representation of Mammogram

A multi-scale representation of a signal is an ordered set of derived signals intended to represent the original signal at different levels of scale. A major reason for multi-scale representation is to explicitly represent the multi-scale aspect of the data, and to suppress and remove unnecessary and disturbing details such that the later stage processing tasks can be simplified [10]. Due to fact that the appearance of masses in mammograms has a great variation in size, shape, intensity contrast and intensity variation inside masses, multi-scale description of mammograms can hopefully be effective in revealing the features of different kind of masses at different scale levels, which provides the basis for further multi-mode processing. Anisotropic diffusion (AD) is a "semantically meaningful" multi-scale description method [11], and is used for mammogram processing in this work.

The multi-scale representation of image can be considered to be a series of "smoothing" operation, where each "smoothing" operation can be formulated as a diffusive process,

$$\frac{\partial}{\partial t} I(x,y,t) = \Delta\left(c(x,y,t)\right) \nabla I(x,y,t) \qquad (1)$$

with initial condition $I(x,y,0)=I_0(x,y)$ be the original image. The diffusion strength is controlled by C(x,y,t) where (x,y) is the coordinates of image, t is the scale-space parameter. Two typical functions are [11]:

$$C_1(x,y,t) = \exp\left(-\left(\frac{|\nabla I(x,y,t)|}{K}\right)^2\right) \tag{2}$$

$$C_2(x,y,t) = \frac{1}{1+\left(\frac{|\nabla I(x,y,t)|}{K}\right)^{1+\alpha}} \qquad \alpha>0 \tag{3}$$

where parameter K is chosen according to noise level and edge strength.

To generate different scale representations using AD, we need to determine the diffusion function C(*), parameter K in C(*), and the iteration number of diffusion process. Because the flow function with C_1 has a better manageable "band-pass" and edge preservation property, $C_1(x,y,t)$ is used as the diffusion function in this work. The parameters K and t are selected empirically, larger for the representation of mammographic image with large mass and vice versa. In this work, two scale images are generated using the AD method, one is of small scale features which is suitable for detection of small to medium size masses; the second is directed to reveal large scale feature and suppress small regions which will be of a help in detection of large-size masses. Fig. 2 shows the 3-D images of region-of-interest (ROI) of a mammogram containing a small and large mass respectively. Their small scale and large scale representations are shown for comparison in Fig. 2. It is demonstrated that different masses can be effectively enhanced with different scale representations.

2.4. Adaptive Segmentation

A region-based detection approach is taken in this method where all the suspicious regions are first separated from fat and glandular normal tissues using image segmentation techniques.

Step 1: Localized adaptive thresholding

Since mass is generally radiographically denser than surrounding tissue, the locally bright spot of appropriate size is extracted by an initial segmentation using a localized adaptive thresholding method. For each pixel I(x,y) in breast area, a decision is made to classify it into a potential mass pixel class (class 1) or a normal pixel class (class 2) by the following rule:

$$I(x,y) \in \text{Suspicious} \quad \text{if } I(x,y) > TH(x,y) \tag{4}$$
$$I(x,y) \in \text{Normal} \quad \text{if } I(x,y) \le TH(x,y) \tag{5}$$

where TH(x,y) is an adaptive threshold value calculated by

$$TH(x,y) = M(x,y) + \alpha\,(\,I_{max}(x,y) - I_{min}(x,y)\,) \tag{6}$$

where M(x,y) is an average of the pixel intensity in small window around pixel I(x,y); $I_{max}(x,y)$ and $I_{min}(x,y)$ are the maximum and minimum intensity value in large window as illustrated in Fig. 3; α is a thresholding bias coefficient and is chosen empirically.

To overcome the difficulty of segmentation of masses due to a large variation in size and contrast, two-mode segmentation is performed on the output of multi-scale representation, one is specifically tuned to small-medium size mass (mode S), the other is for medium-large mass segmentation (mode L).

Step 2. Relaxation of segmentation using adaptive clustering

Localized adaptive thresholding provides an efficient approach to separate the potential abnormal pixels especially for low contrast regions. However, because the localized adaptive thresholding is a pixel-based operation, the segmented pixels are usually not well grouped. With localized adaptive thresholding as its initial segmentation, a relaxation process by adaptive clustering is used to refine the segmentation [12], in which the pixel classification is updated by incorporating local spatial context constraints in the conventional gray-level clustering. It is an iterative process:

(a) with the initial segmentation, calculate an estimate of confidence in the segmentation at each pixel;
(b) for each pixel, modify the segmentation and the confidence estimate based on the pixels in local region;
(c) repeat (b) until the segmentation is completed.

The confidence estimate of a pixel I(x,y) to be a suspicious one (class 1) and normal one (class 2) at i-th iteration can be described respectively as:

$$class1: \quad p_1^{(i)} = \exp\{-(I(x,y) - m_1^{(i)}(x,y))^2 + \beta N_1^{(i)}(x,y)\} \tag{7}$$

$$Class2: \quad p_2^{(i)} = \exp\{-(I(x,y) - m_2^{(i)}(x,y))^2 + \beta N_2^{(i)}(x,y)\} \tag{8}$$

where

$$m_1^{(i)} = \frac{1}{n_1} \sum_{(k,l)} I(k,l) \qquad (k,l) \in B_s(x,y) \cap I(k,l) \in Class\ 1 \tag{9}$$

$$m_2^{(i)} = \frac{1}{n_2} \sum_{(k,l)} I(k,l) \qquad (k,l) \in B_s(x,y) \cap I(k,l) \in Class\ 2 \tag{10}$$

$N_1^{(i)}(x,y)$ and $N_2^{(i)}(x,y)$ are the number of pixels belonging to Class 1 and Class 2 in the 8-connection neighborhood of I(x,y) respectively. $B_s(x,y)$ is a pre-defined neighbor area and decreasing with iteration. The constraint coefficient is $\beta = 2\sigma^2$, where σ^2 is an estimate of noise variance.

The update of pixel label of I(x,y) is performed according to following rule:

I(x,y) updated to be Class 1, if $p_1^{(i)} > p_2^{(i)}$ (11)
I(x,y) updated to be Class 2, if $p_1^{(i)} < p_2^{(i)}$ (12)

The relaxation process continues until a criteria is met, such as the number of changed pixels less than a threshold T_1, and / or the iteration number of relaxation greater than a threshold T_2.

2.5. False-Positive Reduction

A common approach to reduce the FP rate is to perform a further feature analysis of the segmented regions and use a classifier to discriminate masses from normal tissue region. The major difficulties of classification result from the great similarity in appearance between mass and dense normal tissue, and the great variation in feature distribution of different masses. From the classification perspective, the former requires more elaborate features to be extracted for classification while the later means the classifier structure should be more flexible. In this paper, we focus more on the exploration of new classification scheme with less effort on new features design.

Seven features are used in this work for FP reduction including Area, Circularity, Normalized deviation of radial length, Intensity variation, Mean intensity difference, Mean gradient of region boundary, Mean intensity difference along region boundary. They are similar to that we reported before [9] except that (a) one more mixed boundary-intensity feature is added; (b) the calculation of some features is modified with more reasonable

definition.

There are several types of classifiers used in discrimination of masses from normal tissue regions, such as decision tree [13], neural networks [14], linear discriminant analysis (LDA) [15]. A common characteristics of these FP reduction schemes is that a "hard" discriminant criteria is developed by training to evaluate each segmented suspicious region. From the standpoint of region-based classification, it is reasonable, and the task of classifier design is to find an optimal discriminant hyperplane in feature space. However, due to the fact that (1) the mammogram characteristics is usually quite different from case to case; (2) there is a great variation of feature among masses; (3) there is a great similarity between mass and normal tissue regions, the segmented FPs can not be reduced efficiently by a single "hard" decision classifier.

A hybrid classification method is proposed as follows, where the segmented regions are first classified using a modified fuzzy decision tree (MFDT) method. It is a modification of our previously developed fuzzy binary decision tree (FBDT) method [9]. The "hard" decision classifier is then cascaded with a "soft" classification with the objective to reduce FPs in the cases with multiple FPs retained after the "hard" decision classification by an image-based analysis for selecting the best ones among the pre-classified candidate regions as the "real" masses. It is performed by a committee decision method based on a simple premise that most of features of a mass should be top in individual feature ranking among all the candidate suspicious regions in a single mammogram [16].

2.6. Combination of Results

There is a great difference of the detected suspicious regions in size and shape from two processing modes. Generally, the small scale processing mode produces more and small detected regions while the large scale processing mode produces fewer and large detected regions. In order to get a single detection output, the detection results from different processing modes have to be combined. In this work, following scheme is proposed for results combination: Taking the detection results from large-scale mode as the major result, using small-scale detection as a reference. For a detected region from small-scale mode, if there is a detection region from large scale mode having a common detection area, we will take the later as the output while discard the former one. On the other hand, for a detection region from small-scale mode, if there is not a detection region from large-scale modes having a common detection area, we will take it as the detected region for output.

3. DATABASES AND RESULTS

Two image datasets are generated for mass detection system design and testing. The training dataset contains 30 normal and 47 abnormal mammograms with totally 70 masses. It has the same cases as before for the training of last version of detection algorithm, where the mammograms were digitized by an ImageClear R3000 digitizer (DBA System Inc., Melbourne, FL), so that a comparison of detection performance on training database can be made. The testing dataset was generated independently, consisting of 100 normal images, 39 images with 48 minimal cancers and 25 images with 25 benign masses. It can only be accessed for final testing of algorithm.

The mass detection system is first evaluated with the training database. Fig.4 is a detection FROC curve. The operating point of system is chosen to be at sensitivity TP=93% and false positive rate FP=3.1 per image as indicated in Fig. 4. It is then tested independently using the testing database. A good generalization performance was obtained with TP=80% and FP=2.o per image. Two representative mammograms and their detection results are shown in Fig. 5. By analyzing the five masses missed in training at the detection operation point, two of them are due to extremely small size (≤4 mm) and lower contrast (≤4.0); another two of them are on the boundary of breast area where the intensity value decreases dramatically caused by the reduction in thickness at the margin of the compressed breast; one is due to its great deviation of shape feature. Among the testing outputs, a higher detection sensitivity is obtained for benign masses (88%) than that for minimal cancers (75%) at a similar false-positive rate. Again two of the major causes of missing in mass detection are mass located on boundary and lower contrast. In addition to that, a high intensity MCC inside mass is a great interference for mass detection because it results in a great deviation of intensity-related feature values; the masses with too big size could not be detected in testing even though it is not a significant issue from the point of view of CAD.

4. DISCUSSION AND CONCLUSION

The great variation of characteristics of mammograms and masses hinders us in developing a high detection performance and more generalizable CAD system. The typical variations between different mammograms result from either the imaging process (such as film exposure, film label), digitization process (such as spatial / intensity resolution, response function to optical density), and most importantly the inherent breast tissue characteristics. The variations of masses include its size, contrast, shape, location, intensity pattern and its relation to the surrounding tissues. The CAD mass detection method described in this paper proposed a systematic approach to address these problems including (a) image standardization strategy by applying a series of preprocessing to remove extrinsic signal, extract breast area, and normalize the image intensity; (b) multi-mode processing strategy by decomposing image features and multi-scale representation; (c) adaptive processing strategy in localized image segmentation; and (d) combined "hard"-"soft" decision making strategy by using a modified fuzzy decision tree and competitive classification neural network. Compared to the CAD mass detection method we developed earlier, a great improvement is made both in detection performance and generalizability.

REFERENCES

[1] S.L. Parker, T. Tong, S. Bolden and P.A. Wingo. Cancer statistics, 1997. CA Cancer J Clin, 47: 5-27, 1997.

[2] Kerlikowske, D. Grady, S. Rubin, et al., " Efficacy of screen mammography," JAMA 1995; 273(2): 149-154.

[3] R.E. Bird, T.W. Wallace, and B.C. Yankaskas, "Analysis of cancers missed at screning mammography," Radiology, vol.184, Sept. 1992.

[4] M.G. Wallis, M.T. Walsh, and J.R. Lee, "A review of false negative mammography in a symptomatic population," Clinical Radiology, vol.44, 1991.

[5] D.D. Adler, R.L. Wahl, "New methods for imaging the breast: techniques, findings and potential," AJR, vol.164, pp.19-30, 1995.

[6] R.A. Schmidt and R.M. Nishikawa, "Clinical use of digital mammography: The present and the prospects," Journal of Digital Imaging, vol8, no.1, Feb. 1995.

[7] M. L.Giger, " Computer-aided diagnosis ," RSNA Syllabus: A Categorical Course in Physics, Technical Aspects of Breast Imaging, eds. AG Haus, MJ Yaffe, 1993; 283-298.

[8] L. Li, W. Qian, L.P. Clarke, "Wavelet transform for diretional feature extraction in medical imaging," IEEE Proc. of Int. Conf. on Image Processing, Oct. 1997.

[9] L. Li, W. Qian, L.P. Clarke, "Digital mammography: computer-assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms," Academic Radiology 1997; 4:724-731.

[10] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, 1994.

[11] P. Perona, J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.12, July, 1990.

[12] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," IEEE Trans. on Signal Processing, vol.40, no.2, April, 1992.

[13] W. P. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs and M. L. Nipper , " Computer-aided mammographic screen for spiculated lesions ," Radiology, vol.191, pp.331-337, 1994.

[14] B. Sahiner, H. P Chan, N. Petrick, D. Wei, M. A. Helvie, D.D. Adler, and M.M. Goodsitt, "Classification of mass and normal breast tissue: A convolution beural network classifier with spatial domain and texture images," IEEE Trans. Medical Imaging, Vol. 15, No. 5, Oct. 1995.

[15] H. P. Chan, D. Wei, M. A. Helvie, B. Sahiner, D. D. Adler, M. M. Goodsitt, N. Petrick , " Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space ," Phys. Med. Biol., vol.40, pp.857-876, 1995.

[16] D. Black, *The Theory of Committees and Elections*, The Cambridge University Press, 1963.

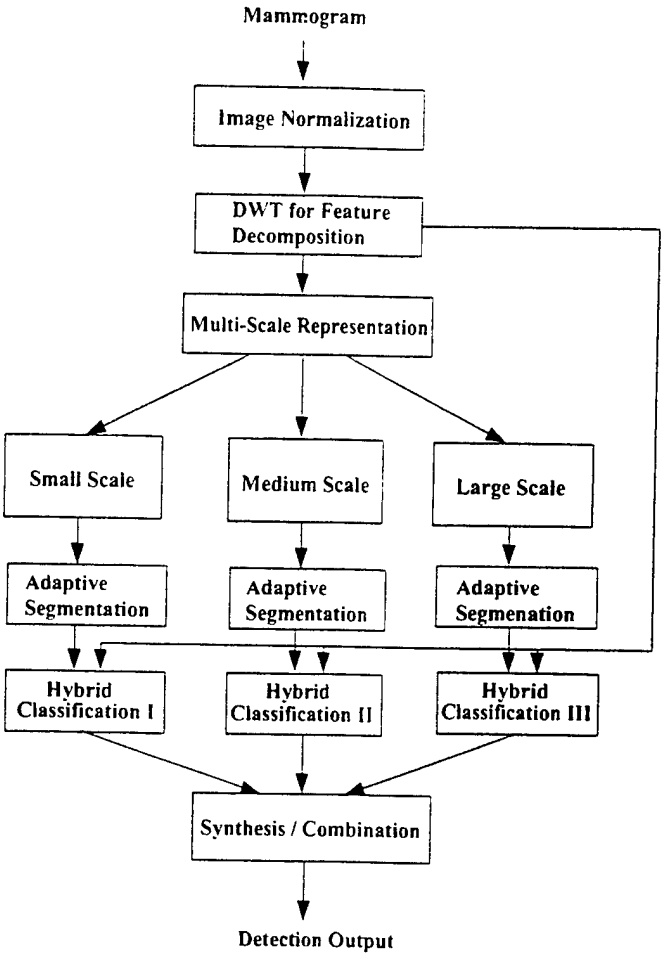


Fig. 1 A schematic diagram of the proposed mass detection system

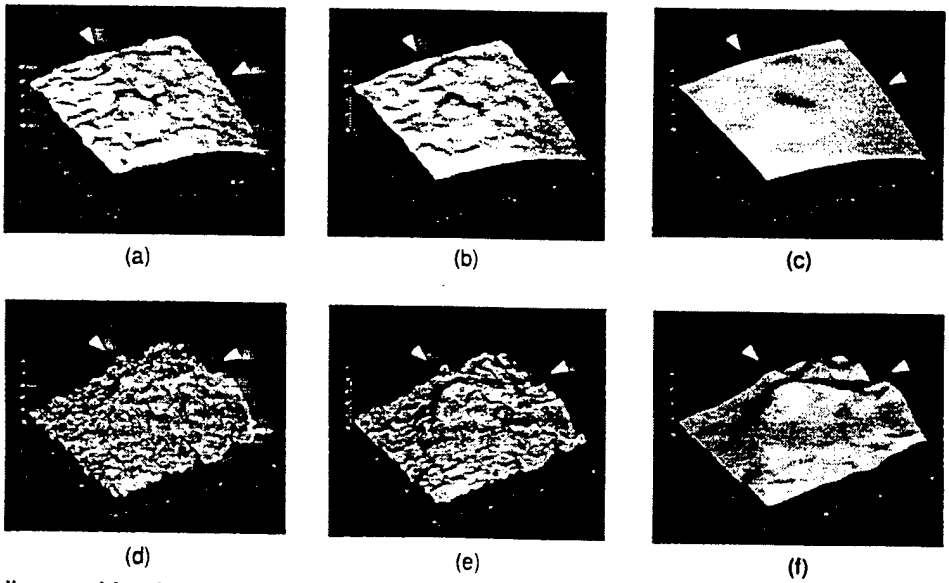


Fig. 2 Nonlinear multi-scale representation of two mammographic ROIs. (a) is a 3-D image of the ROI with a small mass (as the arrows indicated), (b) and (c) are its representations at two different scales; (d) is a 3-D image of the ROI with a large mass (as the arrows indicated), its representations at the same scales as that of (b) and (c) are shown in (e) and (f) respectively.

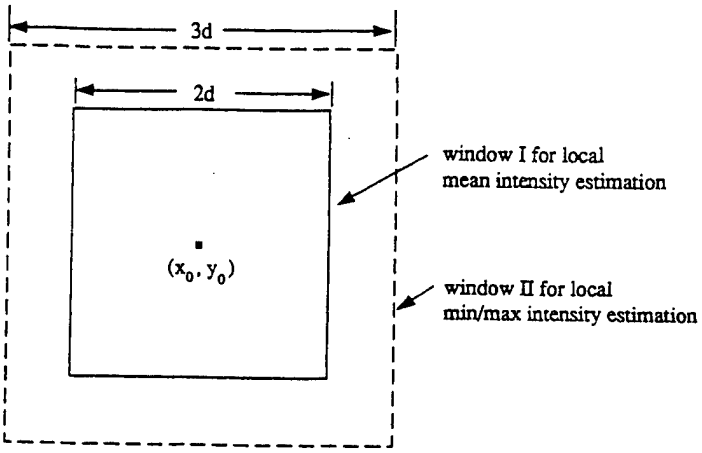


Fig. 3 Windows for localized adaptive thresholding.

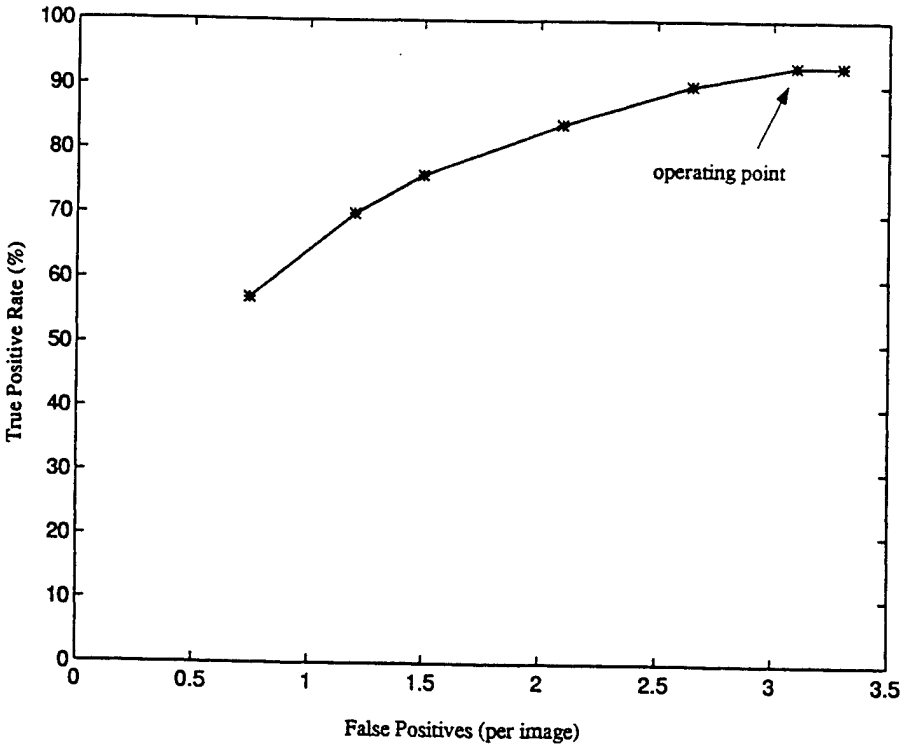


Fig. 4 FROC curves of the true positive detection rate versus the number of false positive findings per image.

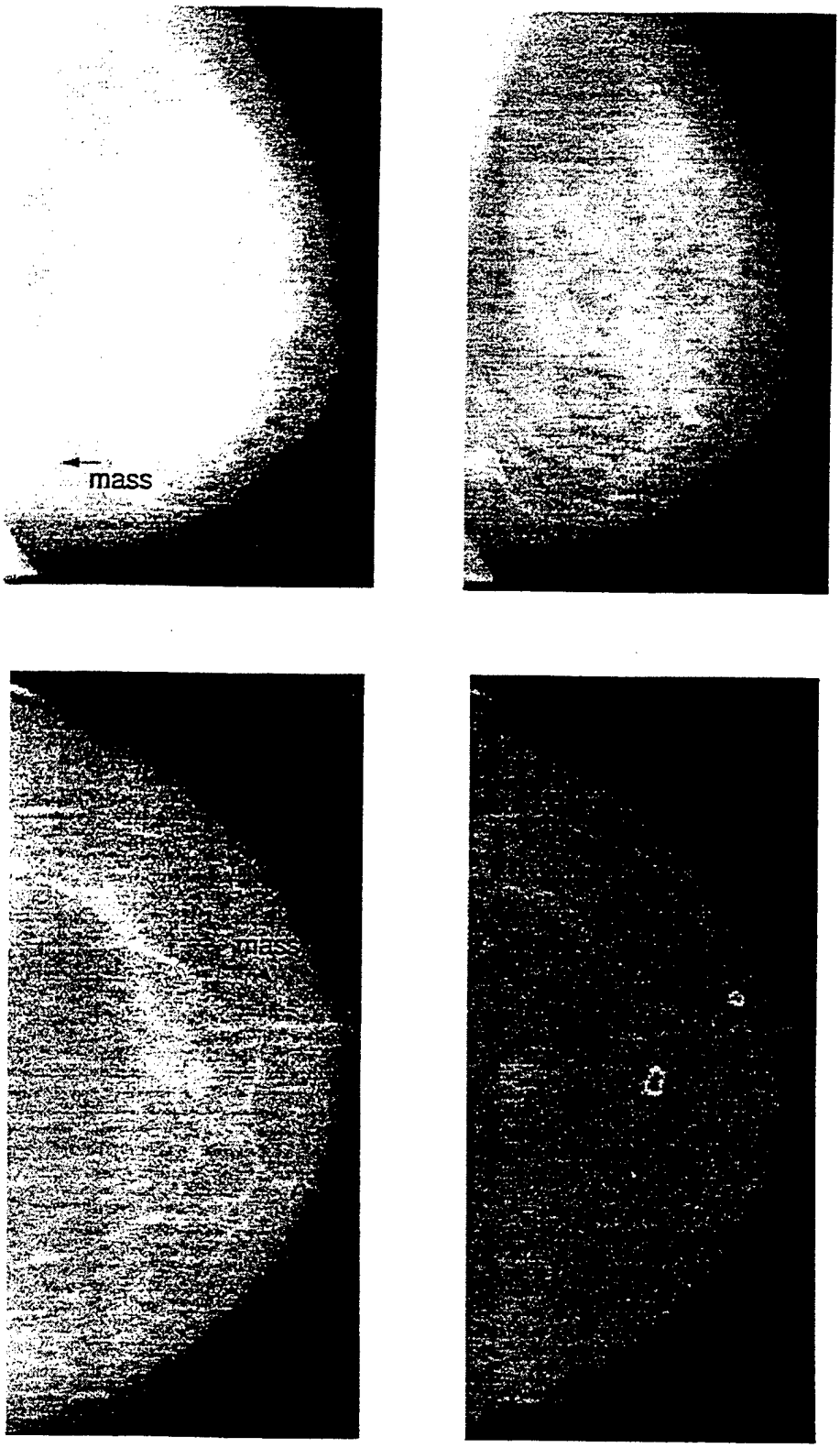


Fig. 5 Two representative mammographic images (preprocessed) and their detection results.