J Digit Imaging (2014) 27:649–660
DOI 10.1007/s10278-014-9681-4

Computerized Breast Mass Detection Using Multi-Scale Hessian-Based Analysis for Dynamic Contrast-Enhanced MRI

Yan-Hao Huang · Yeun-Chung Chang · Chiun-Sheng Huang · Jeon-Hor Chen · Ruey-Feng Chang

Published online: 1 April 2014

Abstract This study aimed to investigate a computer-aided system for detecting breast masses using dynamic contrast-enhanced magnetic resonance imaging for clinical use. Detection performance of the system was analyzed on 61 biopsy-confirmed lesions (21 benign and 40 malignant lesions) in 34 women. The breast region was determined using the demons deformable algorithm. After the suspicious tissues were identified by kinetic feature (area under the curve) and the fuzzy c-means clustering method, all breast masses were detected based on the rotation-invariant and multi-scale blob characteristics. Subsequently, the masses were further distinguished from other detected non-tumor regions (false positives). Free-response operating characteristics (FROC) curve and detection rate were used to evaluate the detection performance. Using the combined features, including blob, enhancement, morphologic, and texture features with 10-fold cross validation, the mass detection rate was 100 % (61/61) with 15.15 false positives per case and 91.80 % (56/61) with 4.56 false positives per case. In conclusion, the proposed computer-aided detection system can help radiologists reduce inter-observer variability and the cost associated with detection of suspicious lesions from a large number of images. Our results illustrated that breast masses can be efficiently detected and that enhancement and morphologic characteristics were useful for reducing non-tumor regions.

Keywords Breast · Magnetic resonance imaging · Detection · Morphologic · Hessian

Y.-H. Huang · R.-F. Chang (✉)
Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan 10617, Republic of China
e-mail: rfchang@csie.ntu.edu.tw

Y.-C. Chang (✉)
Department of Medical Imaging, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan, Republic of China
e-mail: ycc5566@ntu.edu.tw

C.-S. Huang
Department of Surgery, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan

J.-H. Chen
Tu & Yuen Center for Functional Onco-Imaging, Department of Radiological Sciences, University of California, Irvine, CA, USA

J.-H. Chen
Department of Radiology, E-Da Hospital and I-Shou University, Kaohsiung, Taiwan

R.-F. Chang
Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

Introduction

Breast cancer is one of the most common cancers in women [1], and treatment of breast cancer in an earlier stage will reduce the mortality rate [2]. If a tumor is detected in its earlier stages, the chance of successful treatment is higher with proper diagnosis and treatment [2]. Some algorithm on mammography [3] and breast ultrasound (US) [4] have been used for breast cancer detection and diagnosis. However, some limitations might occur using these computer-aided algorithms in different imaging techniques. In mammography, dense fibroglandular tissues can obscure some breast cancers [5], resulting in missed lesions. Mogatadakala et al. [6] proposed a method using order statistic features from multiresolution decompositions of energy-normalized subregions to discriminate surrounding normal and tumor regions in two-dimensional (2-D) US images. However, these algorithms are difficult to be applied to a large three-dimensional (3-D) volume data because the algorithm and feature are

designed for 2-D images. Improper US machine settings and scanner operation error may increase the risk of missed lesions.

Compared with US and mammography image, time-intensity curve obtained from dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) can increase the diagnostic accuracy if morphological appearance is indeterminate [7, 8]. Breast magnetic resonance imaging has a moderate specificity but high sensitivity [9, 10]. The detection of early breast cancer in high-risk populations in the preoperative cancer assessment becomes increasingly important [11]. In MRI study, some breast lesion can be identified from maximum intensity projection (MIP) images if manual rotation in different projections is performed [7] while some suspicious lesions are more likely missed in 2-D MIP examination. Because breast lesion is inherent 3-D, 2-D characterization using MIP might be unable to extract the accurate 3-D lesion information due to overlapping between the enhancing fibroglandular tissue and lesion. In viewing of the only 2-D MIP images, the depth of the lesion perpendicular to plane might be lost. Due to the large number of thin-section 3-D DCE-MRI images obtained from the current state-of-the-art MR scanners, manual tumor identification by radiologists is inefficient and subjective to inter-observer variability. To assist in characterizing inherent 3-D lesion and reducing the inter-observer variability, a computer-aided detection (CADe) system for 3-D breast DCE-MRI image data to detect tumor is proposed for clinical practice. In this study, a 3-D multi-scale tumor detection system is developed for detecting not only suspicious-enhanced tissues by time-intensity curve but also whole 3-D mass-like lesions using morphological and textural characteristics.

Materials

Patients Enrollment

This study is approved by the Institutional Review Board, and informed consent is waived for our retrospective study. In our experiment, the dynamic contrast-enhanced magnetic resonance images were obtained by radiologists from August 2006 to August 2009 and consisted of 61 biopsy-proved lesions (21 benign and 40 malignant lesions, size range from 5 to 60 mm, mean 19.3±12.8 mm) in 34 women (age range from 36 to 77 years, mean 51.0± 10.4 years) to evaluate the performance of the proposed detection system for breast MRI. The benign lesions included 15 fibrocystic changes, 4 papillomas, and 2 fibroadenomas. The malignant lesions included 35 invasive ductal carcinomas (IDC), 3 ductal carcinoma in situ (DCIS), and 2 invasive lobular carcinomas.

MRI Study

DCE-MRI is a medical imaging technique that traces changes in contrast enhancement of internal tissues, such as breast or brain tissue. In our experiment, DCE-MRI images were acquired using a 1.5T superconductive MR scanner (Signa Excite HD, GE Healthcare, Wauwatosa, WI, USA) with a dedicated eight-channel breast coil from patients in the prone position. The dynamic study was performed with the following parameters: pulse sequence VIBRANT (Volume Imaging for BReast AssessmeNT), fat-suppressed 3-D fast spoiled gradient echo (FSGR), repetition time/echo time/inversion time (TR/TE/TI)=3.5/1.7/14 ms, flip angle 12°, matrix 256×160, image size 256×256 pixels, slice thickness 2.5 mm without gap, acquisition 0.75, and field of view 24×24 to 30×30 cm. There were 56 slices obtained for each acquisition of dynamic study with 14 cm coverage in craniocaudal distance There were 35 total acquisitions, each lasting 15 s. Intravenous injection of a contrast agent (0.5 mmol/ml, gadodiamide; Omniscan, GE Healthcare AS, Oslo, Norway or gadopentetate dimeglumine; Magnevist, Bayer Healthcare Pharmaceuticals, Montville, NJ, USA) was performed with antecubital venous access through a plastic cannula (20G), with a bolus injection (4 ml/s) at the beginning of fifth acquisition.

Method

The proposed detection algorithm includes four stages. First, the breast region was segmented by template-based segmentation [12]. Based on the area under the curve (AUC) [13], the fuzzy c-means (FCM) clustering method [14] was used to identify suspicious tissue regions. Mass candidates with different sizes were then detected from these suspicious regions using multi-scale detection. The high frequency of false positives (FPs) could lead to unnecessary biopsies. The FPs are further reduced from candidates by blob, enhancement, texture, morphologic, and combined features.

Breast Region Segmentation

Breast MRI images include not only the breast region but also additional chest regions, such as the lung and the heart. Therefore, the other chest or thorax should be excluded to determine the breast region of interest for tumor detection. However, the shape of the breast and thorax regions is difficult to automatically define due to variation in breast shape and density patterns [12]. Thus, manual inspection of each data is needed to avoid degrading segmentation accuracy from such variation. In the template-based segmentation algorithm [12], the outline between the breast and thorax is

manually delineated on one 2-D slice, producing a template to initialize the segmentation outline of other slices. In this study, the manual outline was delineated in the middle transverse slice of the AUC image (Fig. 1). The AUC image was generated from the accumulation of contrast enhancement on time-intensity curve. The intensity of each pixel in AUC image was composed of the area under the curve and similar with the area-under-curve (AUC) color maps in previous study [13]. This outline could exclude most of thorax, chest body region, except for the axillary region. To exclude the axillary region, the thoracic spine and the two breast margins of the bilateral pectoralis muscles were manually identified to define the bilateral margins on the template by a radiologist who had more than 10 years of experience of interpreting breast MR (Yeun-Chung Chang) (Fig. 1). The v-shape cut [12] formed by these three landmarks can exclude the axillary region (Fig. 1). The breast region was determined as the region below the outline inside the v-shape cut (Fig. 1). To determine the breast region in adjacent slices, the outline and landmarks of adjacent slices were used. To evaluate the deformation of chest regions between two adjacent slices, the thorax was defined with a manual bounding box (Fig. 1) to register the chest region inside the same bounding box in the adjacent slice using the demons deformable algorithm [12, 15]. When registering to neighboring slices based on the computed deformation from two adjacent chest regions, the outline and the landmarks were simultaneously deformed and used to extract the breast region of the neighboring slice. Based on the high similarity between the two adjacent slices, the procedure above could be propagated to each two adjacent slices from middle transverse slice to the cranial margin and caudal margin. Figure 2a shows the sequential segmentation results.

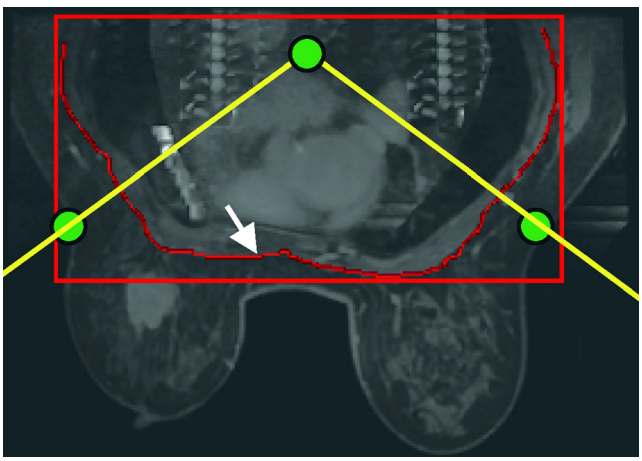

Fig. 1 The *solid arrow* illustrates the manual outline on the AUC image. Three landmarks were connected to depict the v-shape cut formed by the two oblique lines

Finally, the breast region from each slice was segmented using the boundary curve and the v-shape cut (Fig. 1). The suspicious tissues were detected from the segmented breast regions for further mass detection.

Suspicious Tissue Detection

Because diseased tissues usually show more intense enhancement compared to normal tissues [16], the pixels with more enhancement are initially considered as potentially suspected lesions. In this study, the AUC feature computed by the accumulation of contrast enhancement on time-intensity curve [13, 17] was to compute the enhancement degree. The voxels with larger AUCs in the first half of the AUC cumulative histogram [17] computed from the segmented breast region are considered as enhanced tissues. Furthermore, the enhanced tissues were divided into four clusters [14] using FCM clustering method [15, 24], and the average AUC of each cluster was computed. The voxels inside the two clusters with higher average AUC were identified as suspicious (Fig. 2b, c).

However, the suspicious tissues include some enhanced normal tissues (e.g., fibroglandular tissue, skin, and vessels). Besides, if the tumor is connected with other non-tumor tissues, the real tumor can be obscured within an over-segmented region, as shown in Fig. 3. The over-segmented regions, including tumor tissues and normal tissues, are difficult to diagnose due to their poor shape and synthetic composition. Hence, the tumor-like tissues were further extracted from suspicious tissues using the Hessian method based on the morphologic characteristics.

Breast Mass Detection

The Hessian method, which has been applied to detect several kinds of geometrical structures [18–20], was used in our study to detect breast masses [21]. The Hessian method calculate the blob-like degree of each suspicious voxel based on the second derivatives along the three-dimensional directions by convoluting the AUC image. The analysis of eigenvalues [18] from the Hessian matrix were defined by

$$H = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \quad (1)$$

where the elements inside the Hessian matrix H are the partial second derivatives of the image [18], computed by

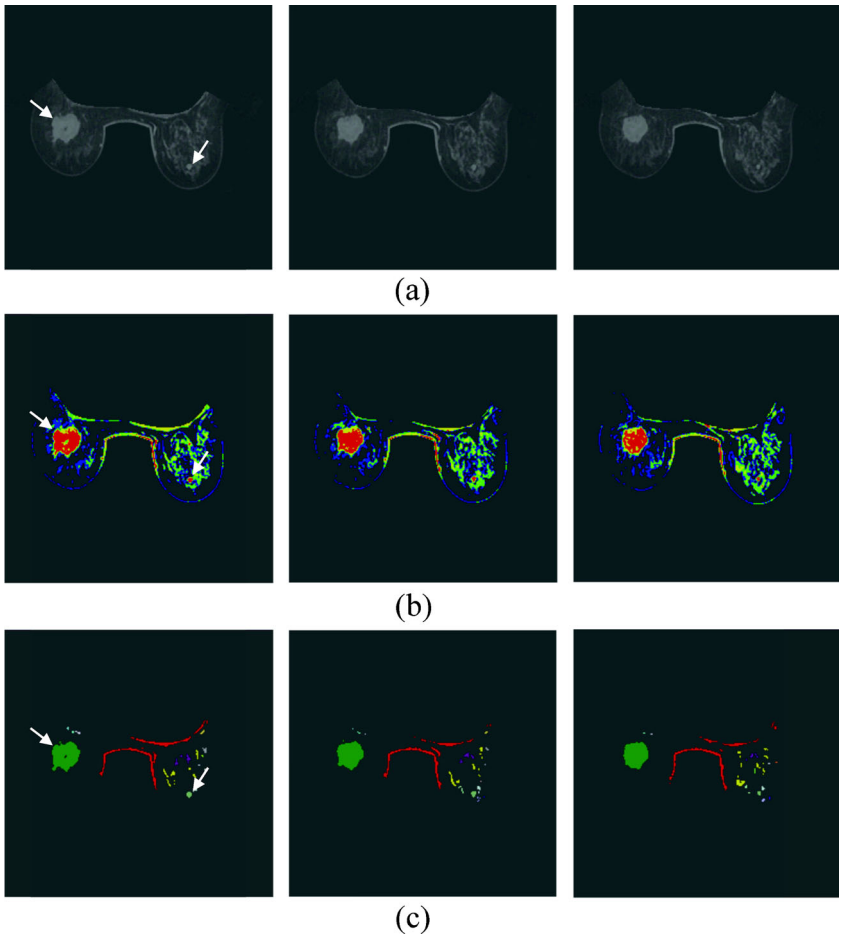

Fig. 2 The breast segmentation of three continue slices of 3.5 cm IDC and 0.5 cm fibrocystic changes, indicated using *white arrows* on the AUC image. **a** AUC image. **b** The four FCM clusters results are represented by *different colors*. Non-breast regions and tissues without significant enhancement are *colored black*. **c** The connected regions from the two clusters with higher average AUC. The different connected regions are colored by *different colors*

convoluting the AUC image $I(x,y,z)$ with second-order derivatives of 3-D Gaussian kernels $G(x,y,z,\sigma)$ [19]. The kernels scale σ at each voxel is defined as the radius of the largest blob-like structure centered at this pixel. The eigenvalues of H are defined as λ_1, λ_2 and λ_3 ($|\lambda_1|<|\lambda_2|<|\lambda_3|$) and the blob-like degree of each suspicious voxel is computed by $O_\sigma(\lambda)$ defined as

$$O_\sigma(\lambda) = \left(1-e^{\frac{-R_A^2}{2\alpha^2}}\right)e^{\frac{-R_B^2}{2\beta^2}}\left(1-e^{\frac{-S^2}{2c^2}}\right) \tag{2}$$

$$R_A = \frac{|\lambda_2|}{|\lambda_3|} \tag{3}$$

$$R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2\lambda_3|}} \tag{4}$$

$$S = \sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2} \tag{5}$$

where α, β, and c are user-defined and fixed to 0.5, 0.5, and 20, respectively [18]. If a region is more blob-like or mass-like, the blob-like degrees $O_\sigma(\lambda)$ will be higher. By means of blob-like degrees, we can distinguish mass from tubular and plate-like region. The maximum matching response [18], $O_\sigma(\lambda)$, of each suspicious voxel evaluated the multi-scale blob-like degree at a range of spatial scales σ of the Gaussian kernel to detect the mass with different sizes, defined by

$$O(\lambda) = \max_{\sigma\in[\sigma_{\min},\sigma_{\max}]} O_\sigma(\lambda) \tag{6}$$

where the kernels scale $\sigma_{\min}$ and $\sigma_{\max}$ are used to control the minimum and the maximum radius of the blob-like structures. Because skin and vessels are not in blob shapes, these enhanced tissues should be of lower blob-like degrees and excluded. The degree was firstly normalized to a range between 0 and 1, and the voxels with the degree lower than 0.2 were eliminated.

After detecting the mass voxels from the suspicious tissues, the voxels were grouped into regions using the 3-D connected component algorithm [22, 23] (Fig. 4). Then, the mass-like likelihood of each connected region was evaluated to differentiate the mass from other non-tumor regions.

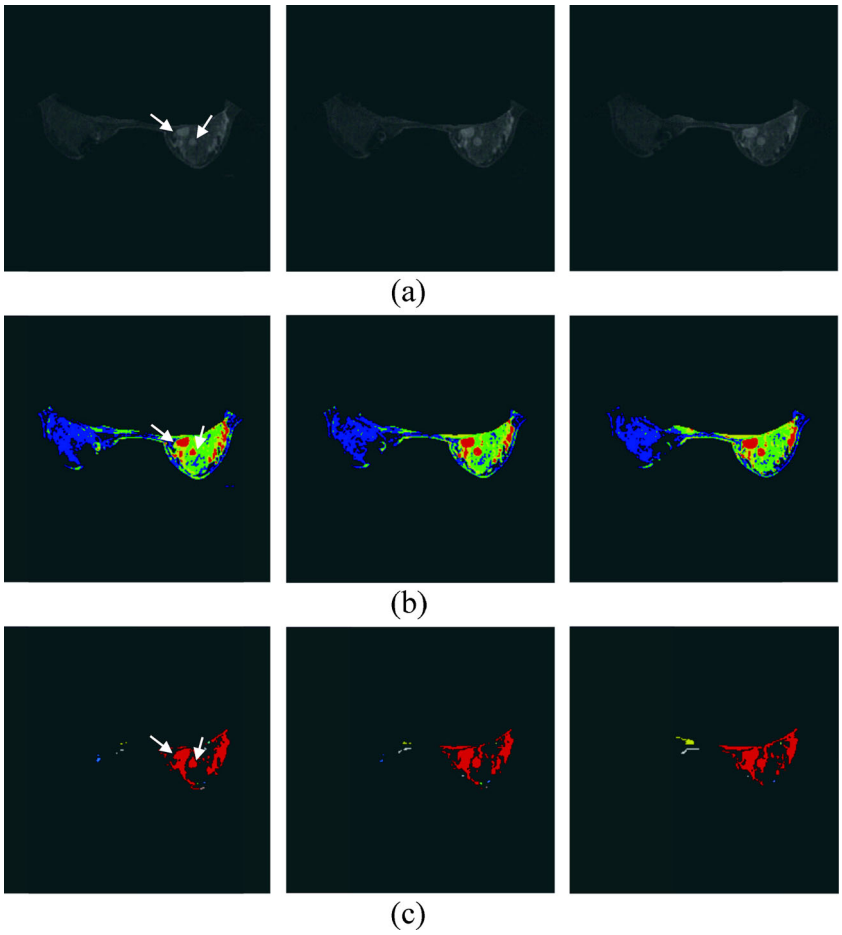

Fig. 3 The 1.4 cm IDC and 0.7 cm fibrocystic changes, indicated by *white arrows*. **a** AUC image. **b** The four FCM clusters results are represented by *different colors*. **c** The connected regions from the two clusters with higher average AUC. The different connected regions are colored by *different colors*

Mass Selection from Candidates

Some false positives were detected as candidates. If all candidates were considered real mass and estimate the tumor-like likelihoods, the processing might be time-consuming and the performance of the CADe system was poor. Before evaluating the likelihoods of these candidates, most false positives should be filtered out to avoid unnecessary evaluation. Because the size

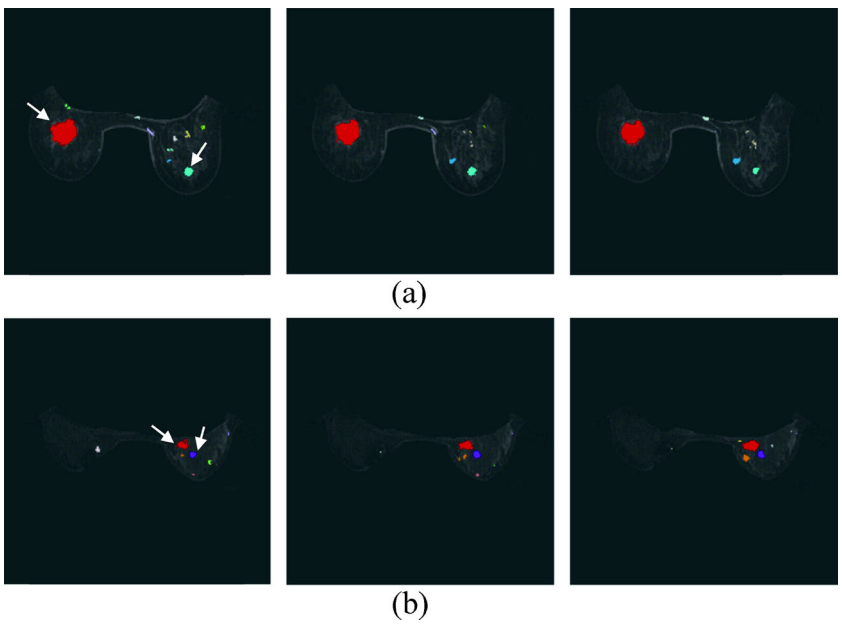

Fig. 4 Breast mass candidates from **a** Fig. 2 and **b** Fig. 3. The different connected regions are displayed by *different colors*. The tumor is indicated using *white arrow*

of most detected FPs was small, the candidates with smaller volumes [24] were considered the scattered FPs and reduced. Hence, only the 50 largest detected regions per case were considered as mass candidates for further tumor selection.

A probability value, called likelihood, was evaluated for predicting the mass-like likelihood. For evaluating the mass-like likelihood of each candidate, each mass candidate was evaluated for blob, enhancement, texture [25], morphologic [26], and combined features, including previous four feature sets. If the likelihood of a candidate was higher than the threshold value, TH_{reg}, this candidate was considered to be a mass.

Mass Selection by Blob and Enhancement Features

To select the masses from the candidates, the blob-like degree of all voxels inside the candidate was averaged as the blob-like degree of candidate. On the other hand, the enhancement degree of each candidate was computed by averaging the AUC characteristics of all the voxels inside the candidate. To select the relatively higher blob-like and enhancement degrees for each case, the rank features were used to sort the 2°, respectively, in descending order. That is, if a candidate is assigned with a smaller rank value, it is comparatively more blob-like or enhanced than the other candidates. Therefore, beside the 2° (B_d and E_d), two rank values (B_r and E_r) of the blob-like and enhancement degrees were applied (Table 1).

Table 1 Definition of the features (blob, enhancement, morphological, and texture) for mass selection

Features		Description
Category	Symbol	
Blob features	B_r	Blob-like degree
	B_d	Blob-like rank
Enhancement features	E_r	Enhancement degree
	E_d	Enhancement rank
Morphological features	C_1	Conventional compactness
	C_2	New compactness
	R	Radius
	R_s	Spiculation
	ER_{axis}	Axis ratio
	$ER_{surface}$	Surface ratio
	V	Volume
Texture features	G_1	Energy
	G_2	Entropy
	G_3	Correlation
	G_4	Inverse difference moment
	G_5	Inertia
	G_6	Cluster shade
	G_7	Cluster prominence
	G_8	Haralick's correlation

Mass Selection by Morphologic Features

Morphological features have been widely used to characterize breast tumors [26, 27]. In our study, seven morphologic features were used to measure the mass-like likelihood of candidates, including two compactness features (C_1 and C_2), radius (R), spiculation (R_s) [26], surface ratio ($ER_{surface}$), axis ratio (ER_{axis}) [26], and volume (V) features computed by the number of tumor voxels and image resolution (Table 1).

Mass Selection by Texture Features

The texture features computed from the gray level co-occurrence matrix (GLCM) of the first post-contrast acquisition measured not only the intensity distribution but also how correlated spatially a voxel is to its neighbors. GLCM features have been applied for medical image processing, including breast diagnosis [26] and classification [28]. In our study, eight GLCM characteristics [25] were applied to measure the mass-like likelihood of candidates for discrimination between tumors and normal candidates. These GLCM features included energy (G_1), entropy (G_2), correlation (G_3), inverse difference moment (G_4), inertia (G_5), cluster shade (G_6), cluster prominence (G_7), and Haralik's correlation (G_8) (Table 1).

Statistical Analysis

To compare the ability to distinguish between mass and FP, statistical analysis was used to analyze features. Each feature was analyzed for normal distribution by the Kolmogorov-Smirnov test [29]. If the p value of Kolmogorov-Smirnov test was greater than or equal to 0.05, it indicated that the feature was normally distributed and the mean and standard deviation of the feature could be further analyzed. For these features, the differences between the masses and FPs were tested by Student's t test [29]. Otherwise, the median of the feature was given, and the differences between masses and FPs were tested by the Mann-Whitney U test [29]. If the p value of Student's t test or Mann-Whitney U test was less than 0.05, the features of the masses were considered significantly different from the FP features.

The combined features were selected using the backward elimination algorithm [29]. Two performance indices, detection rate, and FPs per case were used to evaluate the mass detection performance. The detection rate is defined as the number of the selected candidates divided by the number of all true masses confirmed by a radiologist. A higher detection rate means fewer missed masses. FPs per case indicates the number of detected non-tumor regions for each case. A smaller number of FPs indicate that fewer wrongly detected regions needed to be further identified by a radiologist. Using the proposed mass selection features, the mass-like likelihood of each selected region was evaluated by logistic regression [29].

The 10-fold method [30] was adopted for cross validation. All masses were partitioned into 10 groups that were approximately equal in size. In each step of the validation, all candidates in one group were considered as the testing data and the candidates in other nine groups were training data used to train the logistic regression model. The mass-like likelihood of each candidate in the testing data can be predicted by the trained regression model. The procedure was repeated until the likelihoods of all candidates were evaluated. If the evaluated likelihood of the mass region was greater than or equal to the threshold value, TH_{reg}, it was considered to be a mass. TH_{reg} increased at 0.1 increments from 0.0 to 1.0 to generate the free-response operating characteristics (FROC) [31] curves.

Experiments and Results

Statistical Analysis for Mass Selection

Not all features were normally distributed according to the Kolmogorov-Smirnov test; therefore, the median value and p value from the Mann-Whitney U test were used to analyze the mass selection (Table 2). Except for one texture feature, G_4, all blob, enhancement, and morphologic features were significant different between masses and FPs ($p<0.05$) (Table 2).

Mass Selection by Different Features

After applying the proposed detection method, all masses were included among the candidates. Based on logistic regression, the tumor-like likelihood of each candidate was evaluated for the proposed selection features. In addition to the blob, enhancement, morphologic, and texture features, the combined features were selected from the all features with the backward elimination algorithm to select the masses (Table 3). Based on these feature sets, the detection performance was analyzed by the detection rate and the FPs per case.

Mass Selection

The 50 largest regions from each patient detected by the Hessian method were considered to be mass candidates. The likelihood of each candidate was evaluated by blob, enhancement, morphologic, texture, and combined features (Table 3). The relation between the detection rate and FPs per case was analyzed using different likelihood thresholds (Table 4). The threshold, TH_{reg}, was set at 0.1 and 0.5 to separate masses from FPs based on the combined features. For all masses, the detection rate was 100 % (61/61) with 15.15 FPs and 91.80 % (56/61) with 4.56 FPs based on the combined features and TH_{reg} set at 0.1 and 0.5, respectively. For benign masses, the detection rate was 100 % (21/21) with 16.27 FPs and 76.19 % (16/21) with 4.80 FPs. For malignant masses, the detection rate was 100 % (40/40) with 15.27 FPs per case and 100 % (40/40) with 4.68 FPs per case. All masses were detected while TH_{reg} set at 0.1. But 5 benign masses were missed using

Table 2 Comparison among different features for selecting masses from candidates

Feature categories	Features	Type	Median	p value
Blob features	B_r	Mass	11	<0.001*
		FP	54	
	B_d	Mass	150	<0.001*
		FP	136	
Enhancement features	E_r	Mass	5	<0.001*
		FP	48	
	E_d	Mass	599	<0.001*
		FP	425	
Morphological features	C_1	Mass	21.80	<0.001*
		FP	29.56	
	C_2	Mass	70.45	<0.001*
		FP	54.19	
	R	Mass	4.84	<0.001*
		FP	3.12	
	R_s	Mass	1.37	<0.001*
		FP	1.21	
	$ER_{surface}$	Mass	1.61	<0.001*
		FP	1.39	
	ER_{axis}	Mass	0.61	<0.001*
		FP	0.47	
	V	Mass	539.03	<0.001*
		FP	107.67	
Texture features	G_1	Mass	0.36	0.036*
		FP	0.50	
	G_2	Mass	1.78	<0.001*
		FP	1.15	
	G_3	Mass	0.28	<0.001*
		FP	<0.001	
	G_4	Mass	0.83	0.45
		FP	0.83	
	G_5	Mass	0.35	<0.001*
		FP	0.27	
	G_6	Mass	−0.03	<0.001*
		FP	<0.001	
	G_7	Mass	0.55	<0.001*
		FP	0.07	
	G_8	Mass	7.59×10^6	<0.001*
		FP	6.15×10^6	

*$p<0.05$ is considered statistically significant

The median value and p value of Mann-Whitney U test are used to compare each feature between mass and FP

The blob features are the blob-like degree and its rank value (B_d and B_r). The enhancement features are the enhancement features ant it rank value (E_d and E_r). Morphological features are two compactness (C_1 and C_2), radius (R), spiculation (R_s), surface ratio ($ER_{surface}$), axis ratio (ER_{axis}), and volume feature (V). Texture features are energy (G_1), entropy (G_2), correlation (G_3), inverse difference moment (G_4), inertia (G_5), cluster shade (G_6), cluster prominence (G_7), and Haralik's Correlation (G_8)

Table 3 Each feature set of the combined features used for mass selection

Blob features	Enhancement features	Morphologic features	Texture features
B_r	E_r, E_d	C_1, C_2, R, R_s, $ER_{surface}$, V	G_1, G_3, G_5, G_6, G_7, G_8

TH_{reg} set at 0.5. These five masses (three fibrocystic changes and two papillomas) are considered as more easily missed by our system.

Detection Performance Analysis with FROC Curves

To compare the detection performances of the proposed CADe system using different feature sets, the FROC curve was used to analyze the relationship between the detection rate and the FPs per case (Fig. 5). The combined features were most useful to select masses, which are with higher detection rate and lower FPs. FROC curves for all, benign, and malignant masses using the combined features were shown in Fig. 6. Moreover, the detection performance estimated by the jackknife alternative of FROC-1 (JAFROC) figure of merit (FOM) [32, 33] was applied to analyze the FROC curve of different feature sets statistically (Table 5). The FOM describe whether the mass-like likelihood of tumor is higher than that of FPs. A larger FOM score indicates that the feature set could select mass well. The result reveals that the combined features are most useful to select breast masses and significantly better than other feature sets ($p<0.05$).

Detection Performance Analysis According to Different Size Groups

Because small tumors are easily missed [8, 34], the relationship between the detection rate and the tumor size was further analyzed. The detection rate for different size groups is listed in Table 6. Among the five missed masses using TH_{reg} 0.5, four masses were smaller than 1 cm and one mass was between 1 and 2 cm (Table 6).

Discussion

Early tumor detection and proper treatment is important for increasing cancer survival rate [35]. Because MRI is the most sensitive technology for screening high-risk women for breast cancer [36], it has been widely applied by radiologists [8, 34]. Because the amount of three-dimensional (3-D) imaging data with current state-of-the-art MRI techniques is tremendous, the increasing capability of computer assistance was not only to avoid missing tumors but also to reduce interpretation time. To the best of our knowledge, however, multi-scale blob-based detection algorithms have not been applied to detect breast masses for 3-D DCE-MRI. In this study, we proposed a robust algorithm of 3-D tumor detection, including breast segmentation, mass detection, and mass selection to detect differential cancers from everything else in the breast.

In previous MRI detection studies [8, 34], malignant tumors were manually detected from MRI maximum intensity projection plane of enhanced MRI image sets. Nevertheless, the inter-observer inconsistency and variation are inevitable, and performance depends on the experiences of interpreting radiologists. To overcome these issues, automatic detection of

Table 4 The detection rate and FPs per case for mass detection based on combined feature selection

	All cases		Benign cases		Malignant cases	
	Detection rate	FPs per case	Detection rate	FPs per case	Detection rate	FPs per case
Th=0.0	100.00 %	43.74	100.00 %	46.27	100.00 %	43.30
Th=0.1	100.00 %	15.15	100.00 %	16.27	100.00 %	15.27
Th=0.2	96.72 %	10.41	90.48 %	11.00	100.00 %	10.57
Th=0.3	91.80 %	8.00	76.19 %	8.60	100.00 %	8.07
Th=0.4	91.80 %	6.18	76.19 %	6.80	100.00 %	6.10
Th=0.5	91.80 %	4.56	76.19 %	4.80	100.00 %	4.43
Th=0.6	81.97 %	3.53	52.38 %	3.67	97.50 %	3.43
Th=0.7	78.69 %	2.68	42.86 %	3.00	97.50 %	2.67
Th=0.8	73.77 %	1.88	33.33 %	2.13	95.00 %	1.90
Th=0.9	62.30 %	1.12	19.05 %	1.20	85.00 %	1.20
Th=1.0	0.00 %	0.00	0.00 %	0.00	0.00 %	0.00

If the likelihood of detected candidates estimated by the combined features is larger than the logistic regression threshold value, Th, it is selected as a mass. If a case included benign and malignant masses, the case is considered as both benign and malignant case. There are 15 benign cases and 30 malignant cases

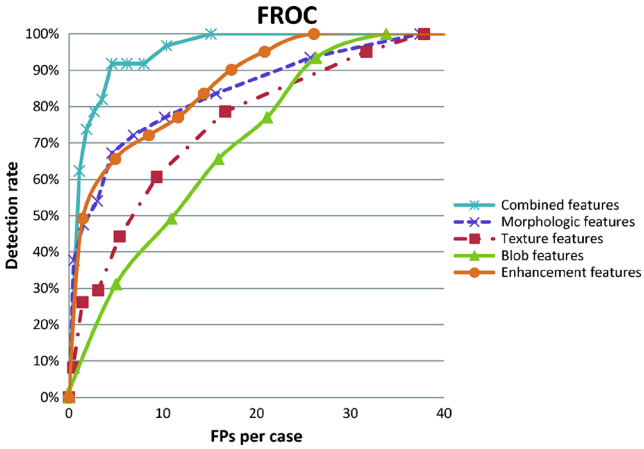

Fig. 5 The free-response receiver operating characteristic (FROC) curve of the proposed system for all cases using different feature set

benign and malignant breast masses using computer assistance was proposed for 3-D DCE-MRI data in this study. Renz's study [37] proposed a detection algorithm to detect potential tumors using a hierarchical 3-D Gaussian pyramid to outline regions with high local intensity similarities. However, only enhanced contrast characteristics were used to evaluate tissues. Some normal tissues, such as fibroglandular, skin, and vessels, were also detected to increase the number of FPs. As a result, two malignant lesions were missed and the detection performance was 95.74 % with 17 FPs. In our study, not only enhancement but also blob features characterized the tumor tissues to improve the detection performance.

The whole breast region must be determined before detecting masses. In previous study [38], the border was identified manually by radiologists. However, manual segmentation was usually inconsistent and caused inter-observer variability. A straight line method [39] to segment breast cannot segment the breast tissue near the axilla. In contrast with the previous segmentation methods [38, 39], demons segmentation technique preserves the most of breast region because the outline can deform to segment breast [12, 15]. The semiautomatic breast segmentation initialized by three manual landmarks in this study can segment the precise breast region compared to previous studies. Although the manual landmarks might result in little inter-observer variability, it can improve the accuracy of breast segmentation for clinical practice.

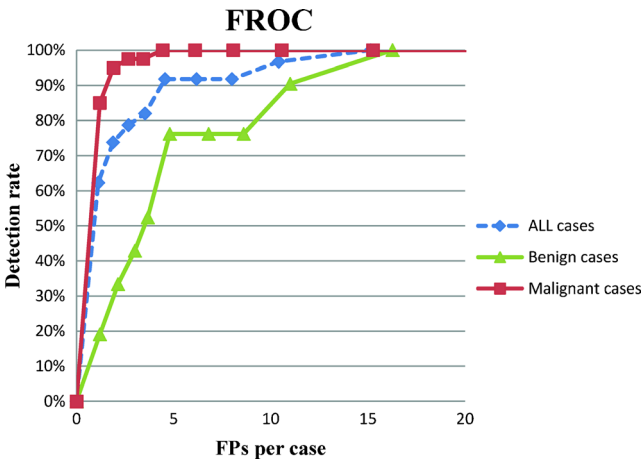

Fig. 6 The free-response receiver operating characteristic (FROC) curve of the proposed system for all cases, benign cases, and malignant cases using combined features

All suspicious tissues from the segmented breast were identified using the FCM clustering method based on AUC feature. However, suspicious tissues identified only using the enhancement characteristic included many flat or lengthy non-tumor tissues. The blob characteristics [40] of masses were applied to further identify the masses from other normal tissues. After mass detection stage, we apply some features, such as the blob, enhancement, texture, morphologic, and combined features, to select the masses from the other normal candidates. Based on our experiment (Table 2), most features show significant difference between masses and FPs ($p<0.05$), except for one texture feature, inverse difference moment (G_4). In the previous study, the morphologic features [27] of breast MRI are important criteria to characterize breast masses and enhancement-related [13] features are useful to interpret tissues on DCE-MRI. According to the median and p value, the different masses and FPs were more significant in blob, enhancement, and morphological features than texture features. The experiment results indicated the high importance of 3-D morphological and enhancement for tumor detection. The detection performance using the five categories of features is listed in Fig. 5. The combined features performed better than individual features, with higher detection rates and fewer FPs per case. Table 4 and Fig. 6 show that the malignant masses were more easily selected than benign masses. The detection rate of malignant tumors is 100 %, higher than that of benign tumors at 76.19 %. To maintain a high sensitivity for detecting breast cancer, the malignant tumors should not be reduced and at the mass selection process. TH_{reg} values set at 0.5 as a suitable threshold to detect all malignant tumors in our experiment, compared to TH_{reg} values as 0.1 (Table 4). For the combined features, the detection performances were 100 % (61/61) and 91.80 % (56/61), respectively. Five benign masses were missed using a TH_{reg} of 0.5. The JAROC-1 FOM of different feature sets is listed in Table 5. The combined features and morphological features were more capable of detecting breast mass, which are with higher ROM values, 0.45 and 0.36, respectively. The p value for FOM between different pairs of feature sets indicated that the combined features can significantly improve the detection performance from other individual feature set. The morphological features are most used in clinical practice. The differences between morphological and texture features, as well as blob features, are also statically significant.

Table 5 FOM values of five feature sets for mass selection

Combined features	Morphological features	Texture features	Blob features	Enhancement features
$0.45^{\circledcirc\dagger\Phi\parallel}$	$0.36^{\Theta\ddagger}$	$0.18^{\$}$	$0.09^{\#}$	0.31

All p values of FOM between different pairs of feature set

$^{\circledcirc}p<0.05$ while comparison between combined features and morphological features; $^{\dagger}p<0.05$ while comparison between combined features and enhancement features; $^{\Phi}p<0.05$ while comparison between combined features and texture features; $^{\parallel}p<0.05$ while comparison between combined features and blob features; $^{\Theta}p<0.05$ while comparison between morphological features and texture features; $^{\ddagger}p<0.05$ while comparison between morphological features and blob features; $^{\$}p<0.05$ while comparison between texture features and enhancement features; $^{\#}p<0.05$ while comparison between blob features and enhancement features

The p value of all the features for different types of detected regions and the five missed masses were analyzed using a TH_{reg} of 0.5. For the clinically used feature, enhancement degree (E_d), the missed masses were more similar with the FPs than malignant and benign masses ($p>0.05$). The enhancement degrees (E_d) of the malignant and benign masses (753.78±367.22 and 619.38±209.63) were significantly larger than that of the missed masses (400.00±119.26) ($p=0.018$ and 0.032). The tumor volumes (V) of malignant masses (2765.03±3347.59 mm^3) were significantly larger than the missed masses (265.37±140.55 mm^3) ($p=0.003$). Like previous studies [7, 34], masses which are relatively small or have a slight absence of perceptible enhancement could be missed. The miss-detected benign tumors were difficult to be detected using only enhancement and volume features. The morphologic features, C_1, C_2, R, R_s, $ER_{surface}$, ER_{axis}, and V computed from benign tumors were similar to those from FPs ($p>0.05$). The result indicates that morphologic characteristics between the miss-detected masses and benign tumors, as well as normal tissues, are not significantly different. Therefore, they are difficult to be distinguishable only based on the morphologic characteristics. Because all texture features were larger than 0.05, single texture feature was not useful for mass detection. The detection rate analyzed according to different tumor size groups is listed using TH_{reg} of 0.5 (Table 6). The malignant masses were detected correctly among all size groups. The missed masses were benign and smaller in size. Prior study [41] with a large cohort of MRI-guided needle localization shows that lesions smaller than 5 mm have low likelihood of being cancer.

Table 6 The detection rate for different size groups

Tumor size	All masses	Benign masses	Malignant masses
≦1 cm	77.78 % (14/18)	71.43 % (10/14)	100 % (4/4)
>1 cm and ≦2 cm	95.45 % (21/22)	83.33 % (5/6)	100 % (16/16)
>2 cm	100 % (21/21)	100 % (1/1)	100 % (20/20)

Some limitations in this study could be considered further. The voxels with blob-like degrees higher than the 0.2 obtained by experiments were considered to be masses in our study. However, the threshold to determine whether the voxel is a mass tissue can affect the detection performance. Because of poor spatial resolution (image size 256×256 pixels and slice thickness 2.5 mm), tumor tissues might seem to connect to enhanced normal tissues on images. Hence, these tumor tissues can be evaluated as lower blob-like degrees and easily filtered out by an automatically defined threshold. Therefore, an automatic algorithm to determine the threshold of blob-like degree on a higher resolution image will be important for future work involving more precise detection of blob-like tissues. Because the missed benign masses contain only small numbers of voxels, a few enhanced voxels can make analysis difficult. The use of higher resolution images might overcome this problem. On the other hand, kinetic features usually analyzed the malignancy likelihood of benign and malignant tumors [17], but the study about using time-intensity curve to distinguish tumor from non-tumor tissues were rare. In addition to the blob, morphological, and texture features, the pharmacokinetic might be incorporated to improve our system in the future.

Conclusion

In the proposed breast DCE-MRI computer-aided detection system, whole breast region is determined using the modified demons deformable algorithm. Masses are detected using multi-scale Hessian method. The mass detection rates are 100 % (61/61) with 15.15 false positives per case and 91.80 % (56/61) with 4.56 false positives per case. To date, this is the first study using a multi-scale blob-based tumor detection algorithm followed by morphologic tumor selection to automatically detect both benign and malignant tumors on 3-D DCE-MRI study. In conclusion, the developed DCE-MRI computer-aided detection (CADe) system might potentially assist physicians to detect mass before clinic diagnosis.

Acknowledgments The authors would like to thank the National Science Council of the Republic of China (NSC 101-2221-E-002-068-MY3), Ministry of Economic Affairs (102-EC-17-A-19-S1-164), and Ministry of Education (AE-00-00-06) of the Republic of China for the financial support

References

1. Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Waldron W, et al: SEER Cancer Statistics Review, 1975–2008. National Cancer Institute, Bethesda, 2011
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: Global cancer statistics. CA Cancer J Clin 61:69–90, 2011
3. Roubidoux MA, Helvie MA, Lai NE, Paramagul C: Bilateral breast cancer: early detection with mammography. Radiology 196:427–431, 1995
4. Kolb TM, Lichy J, Newhouse JH: Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology 225:165–175, 2002
5. Kelly KM, Dean J, Comulada WS, Lee SJ: Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. Eur Radiol 20:734–742, 2010
6. Mogatadakala KV, Donohue KD, Piccoli CW, Forsberg F: Detection of breast lesion regions in ultrasound images using wavelets and order statistics. Med Phys 33:840–849, 2006
7. Muradali D, Ghai S, Bukhanov K, Kulkarni S: Nonenhancing breast malignancies on MRI: sonographic and pathologic correlation. Am J Roentgenol 185:481–487, 2005
8. Teifke A, Hlawatsch A, Beier T, Vomweg TW, Schadmand S, Schmidt M, et al: Undetected malignancies of the breast: dynamic contrast-enhanced MR imaging at 1.0 T. Radiology 224:881–888, 2002
9. Bluemke DA, Gatsonis CA, Chen MH, DeAngelis GA, DeBruhl N, Harms S, et al: Magnetic resonance imaging of the breast prior to biopsy. JAMA 292:2735–2742, 2004
10. Warner E, Plewes DB, Shumak RS, Catzavelos GC, Di Prospero LS, Yaffe MJ, et al: Comparison of breast magnetic resonance imaging, mammography, and ultrasound for surveillance of women at high risk for hereditary breast cancer. J Clin Oncol 19:3524–3531, 2001
11. Yeh ED: Breast magnetic resonance imaging: current clinical indications. Magn Reson Imaging Clin N Am 18:155–169, 2010. vii
12. Lin M, Chen J-H, Nalcioglu O, Su M-YL: Automatic template-based breast segmentation on MRI using nonrigid registration algorithms. In: International Society for Magnetic Resonance in Medicine (ISMRM) conference, 2012, pp. 2569, City
13. Galbraith SM, Lodge MA, Taylor NJ, Rustin GJS, Bentzen S, Stirling JJ, et al: Reproducibility of dynamic contrast-enhanced MRI in human muscle and tumours: comparison of quantitative and semi-quantitative analysis. NMR Biomed 15:132–142, 2002
14. Su MY, Nie K, Chen JH, Chan S, Chau MKI, Yu HJ, et al: Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI. Med Phys 35:5253–5262, 2008
15. Kroon D-J, Slump CH: MRI modality transformation in demon registration. In: IEEE International Symposium on Biomedical Imaging, ISBI '09. IEEE Signal Processing Society, City, 2009, pp. 963–966
16. Hayton P, Brady M, Tarassenko L, Moore N: Analysis of dynamic MR breast images using a model of contrast enhancement. Med Image Anal 1:207–224, 1997
17. Chang YC, Huang YH, Huang CS, Chang PK, Chen JH, Chang RF: Classification of breast mass lesions using model-based analysis of the characteristic kinetic curve derived from fuzzy c-means clustering. Magn Reson Imaging 20:312–322, 2012
18. Frangi AF, Niessen WJ, Vincken KL, Viergever MA: Multiscale vessel enhancement filtering, Medical Image Computing and Computer-Assisted Intervention - Miccai'98 1496, 1998, pp. 130–137
19. Jiamin L, White JM, Summers RM: Automated detection of blob structures by Hessian analysis and object scale. In: Image Processing (ICIP), 2010 17th IEEE International Conference on, City, 2010, pp. 841–844
20. Jia T, Zhao D-Z, Wei Y, Zhu X-H, Wang X: Computer-aided lung nodule detection based on CT images. In: Complex Medical Engineering, 2007 CME 2007 IEEE/ICME International Conference on, City, 2007, pp. 816–819
21. Okada K, Akdemir U: Blob segmentation using joint space-intensity likelihood ratio test: application to 3D tumor segmentation. In: Computer Vision and Pattern Recognition, 2005 CVPR 2005 I.E. Computer Society Conference on, vol. 2, City, 2005, pp. 437–444
22. Rosenfel A: Connectivity in digital pictures. J ACM 17:146, 1970
23. Rosenfel A, Pfaltz JL: Sequential operations in digital picture processing. J ACM 13:471, 1966
24. Chen JH, Chang RF, Chang-Chien KC, Takada E, Huang CS, Chou YH, et al: Rapid image stitching and computer-aided detection for multipass automated breast ultrasound. Med Phys 37:2063–2073, 2010
25. Chen W, Giger ML, Li H, Bick U, Newstead GM: Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. Magn Reson Med 58:562–571, 2007
26. Chang RF, Moon WK, Shen YW, Huang CS, Chiang LR: Computer-aided diagnosis for the classification of breast masses in automated whole breast ultrasound images. Ultrasound Med Biol 37:539–548, 2011
27. Meinel LA, Stolpen AH, Berbaum KS, Fajardo LL, Reinhardt JM: Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. J Magn Reson Imaging 25:89–95, 2007
28. Haralick RM: Statistical and structural approaches to texture. Proc IEEE 67:786–804, 1979
29. Field AP: Discovering Statistics Using SPSS. SAGE Publications, London, 2009
30. Bengio Y, Grandvalet Y: No unbiased estimator of the variance of K-fold cross-validation. J Mach Learn Res 5:1089–1105, 2004
31. Yu SY, Guan L: A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films. IEEE Trans Med Imaging 19:115–126, 2000
32. Chakraborty DP: Validation and statistical power comparison of methods for analyzing free-response observer performance studies. Acad Radiol 15:1554–1566, 2008
33. Chakraborty DP, Breatnach ES, Yester MV, Soto B, Barnes GT, Fraser RG: Digital and conventional chest imaging: a modified ROC study of observer performance using simulated nodules. Radiology 158:35–39, 1986
34. Shimauchi A, Jansen SA, Abe H, Jaskowiak N, Schmidt RA, Newstead GM: Breast cancers not detected at MRI: review of false-negative lesions. Am J Roentgenol 194:1674–1679, 2010
35. Anderson BO, Shyyan R, Eniu A, Smith RA, Yip CH, Bese NS, et al: Breast cancer in limited-resource countries: an overview of the Breast Health Global Initiative 2005 guidelines. Breast J 12(Suppl 1):S3–S15, 2006
36. Kuhl CK, Schrading S, Bieling HB, Wardelmann E, Leutner CC, Koenig R, et al: MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. Lancet 370:485–492, 2007
37. Renz DM, Bottcher J, Diekmann F, Poellinger A, Maurer MH, Pfeil A, et al: Detection and classification of contrast-enhancing masses by a fully automatic computer-assisted diagnosis system for breast MRI. J Magn Reson Imaging 35:1077–1088, 2012

38. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al: Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. Radiology 246:376–383, 2008
39. Chen JH, Moon WK, Shen YW, Huang CS, Luo SC, Kuzucan A, et al: Comparative study of density analysis using automated whole breast ultrasound and MRI. Med Phys 38:382–389, 2011
40. Dorr RT, Alberts DS: Quantitation of ellipsoid tumor areas using a circumferential measuring device. Med Oncol Tumor Pharmacother 5:249–251, 1988
41. Liberman L, Mason G, Morris EA, Dershaw DD: Does size matter? Positive predictive value of MRI-detected breast lesions as a function of lesion size. AJR Am J Roentgenol 186: 426–430, 2006