



# Radiomics Based on Digital Mammography Helps to Identify Mammographic Masses Suspicious for Cancer

Guangsong Wang<sup>1</sup>, Dafa Shi<sup>1</sup>, Qiu Guo<sup>1</sup>, Haoran Zhang<sup>1</sup>, Siyuan Wang<sup>1</sup> and Ke Ren<sup>1,2\*</sup>

<sup>1</sup> Xiang'an Hospital, Xiamen University, Xiamen, China, <sup>2</sup> Xiamen Key Laboratory of Endocrine-Related Cancer Precision Medicine, Xiamen, China

## OPEN ACCESS

### Edited by:

Samata Kakkad,  
Merck, United States

### Reviewed by:

Ning Mao,  
Peking University People's Hospital,  
China  
Stacey Winham,  
Mayo Clinic, United States

### \*Correspondence:

Ke Ren  
renke816@163.com

### Specialty section:

This article was submitted to  
Cancer Imaging and  
Image-directed Interventions,  
a section of the journal  
Frontiers in Oncology

Received: 25 December 2021

Accepted: 04 March 2022

Published: 01 April 2022

### Citation:

Wang G, Shi D, Guo Q, Zhang H,  
Wang S and Ren K (2022) Radiomics  
Based on Digital Mammography Helps  
to Identify Mammographic Masses  
Suspicious for Cancer.  
Front. Oncol. 12:843436.  
doi: 10.3389/fonc.2022.843436

**Objectives:** This study aims to build radiomics model of Breast Imaging Reporting and Data System (BI-RADS) category 4 and 5 mammographic masses extracted from digital mammography (DM) for mammographic masses characterization by using a sensitivity threshold similar to that of biopsy.

**Materials and Methods:** This retrospective study included 288 female patients (age,  $52.41 \pm 10.31$ ) who had BI-RADS category 4 or 5 mammographic masses with an indication for biopsy. The patients were divided into two temporal set (training set, 82 malignancies and 110 benign lesions; independent test set, 48 malignancies and 48 benign lesions). A total of 188 radiomics features were extracted from mammographic masses on the combination of craniocaudal (CC) position images and mediolateral oblique (MLO) position images. For the training set, Pearson's correlation and the least absolute shrinkage and selection operator (LASSO) were used to select non-redundant radiomics features and useful radiomics features, respectively, and support vector machine (SVM) was applied to construct a radiomics model. The receiver operating characteristic curve (ROC) analysis was used to evaluate the classification performance of the radiomics model and to determine a threshold value with a sensitivity higher than 98% to predict the mammographic masses malignancy. For independent test set, identical threshold value was used to validate the classification performance of the radiomics model. The stability of the radiomics model was evaluated by using a fivefold cross-validation method, and two breast radiologists assessed the diagnostic agreement of the radiomics model.

**Results:** In the training set, the radiomics model obtained an area under the receiver operating characteristic curve (AUC) of 0.934 [95% confidence intervals (95% CI), 0.898–0.971], a sensitivity of 98.8% (81/82), a threshold of 0.22, and a specificity of 60% (66/110). In the test set, the radiomics model obtained an AUC of 0.901 (95% CI, 0.835–0.961), a sensitivity of 95.8% (46/48), and a specificity of 66.7% (32/48). The radiomics model had relatively stable sensitivities in fivefold cross-validation (training set,  $97.39\% \pm 3.9\%$ ; test set,  $98.7\% \pm 4\%$ ).

**Conclusion:** The radiomics method based on DM may help reduce the temporarily unnecessary invasive biopsies for benign mammographic masses over-classified in BI-RADS category 4 and 5 while providing similar diagnostic performance for malignant mammographic masses as biopsies.

**Keywords:** breast (diagnostic), breast cancer, Mammography, Radiomic analysis, BI-RADS (Breast imaging reporting and data system)

## INTRODUCTION

In 2020, female breast cancer (BC) became the most common type of cancer with an estimated 2.3 million new cases (11.7%), followed by lung cancer (11.4%) (1). Treatment of BC relies on conducting an accurate diagnosis, including histological, molecular, and clinical phenotypes. Non-invasive imaging techniques such as mammography, ultrasound, and magnetic resonance (MR) are available for qualitative and quantitative analysis of BC in clinical practice. The American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) is a standardized assessment structure that enables radiologists to clearly and concisely communicate results of breast imaging to referring physicians (2). In the fifth edition of the BI-RADS atlas (3), category 4 and 5 breast lesions are defined as suspicious cancerous lesions, and a biopsy is recommended for further diagnosis. Recent studies have shown that a large number of benign lesions are present in category 4 and 5 breast lesions, particularly in the mammography reporting system, exposing these patients to invasive biopsies (4–6). Depending on the technique, the sensitivity values of biopsy results ranged from 87% to >97% (7–9).

Radiomics is a high-throughput image mining technique that aims to enhance the predictive power of medical images by quantifying the morphology, intensity distribution, and texture patterns of lesions. Recent investigators have examined the role of mammography, ultrasound, and MR radiomics in the prediction of molecular subtypes (10–12), lymph node metastasis (13–15), response to neoadjuvant chemotherapy (16–18), recurrence risk (19, 20), and disease-free survival (21, 22) of BC. However, with the initiative of precision medicine (23–25), the reduction in overdiagnosis and overtreatment of breast lesions through non-invasive radiomics method is also a topic worth investigating.

There are four main findings of breast lesions on diagnostic mammography images: masses, calcifications, architectural distortion, and asymmetries. One large sample study (26) showed that BC most often presented as mass at 56%, followed by calcifications at 29%, asymmetry at 12%, and architectural deformities at 4%, and another small sample study (27) suggest that approximately 50% of breast lesions presenting as a mass were ultimately confirmed benign lesions. In addition, mass may be the only finding or one of the combined findings of breast lesions in mammography (we defined these masses as mammographic masses). Although experienced radiologists have a high diagnostic accuracy in identifying benign and malignant mammographic masses, less experienced radiologists

sometimes make excess errors (28) such as benign mammographic masses are over-classified as BIRADS category 4 or even 5.

A previous study has shown that combining both craniocaudal (CC) and mediolateral oblique (MLO) positions radiomics data had good classification performance between HER2-enriched BC and non-HER2-enriched BC (11). Here, we combined both CC and MLO positions radiomics data aimed to explore a model with a sensitivity more than 98% for the characterization of BI-RADS category 4 and 5 mammographic masses, thereby achieving a reduction in biopsies of benign lesions at a very low rate of missed malignant lesions.

## MATERIALS AND METHODS

### Patients

This retrospective study was granted approval by the local institutional board, and written informed consent was waived. A total of 288 patients' clinical and mammographic images data were included in this study from December 2018 to February 2021. The inclusion criteria were as follows: (a) patients who had suspected breast tumor accepted mammography and (b) patients with mass as defined by BI-RADS mammography lexicon (occupancy structures with protrude outward in contour on both CC and MLO position images) and classified in category 4 or 5. The exclusion criteria were as follows: (a) patients without a clear benign or malignant pathological result; (b) patients who had multifocal or bilateral mammographic masses; (c) patients accepted biopsy before mammography examination; (d) patients underwent any treatment before mammography screening, including surgery, chemotherapy or radiotherapy, and anti-HER2 therapy.

### Imaging and Saving Acquisition

All patients were examined with a GE Senographe Essential (GE Medical Systems, Waukesha, WI), and all mammographic images were saved at 12-bit quantization level and 100- $\mu$ m pixel size. The mammographic images were not further processed or normalized (29, 30).

### Radiomics Analysis of Mammographic Masses

Both CC and MLO position images of all patients were used to conduct mammographic mass masking, and the CC radiomics features and MLO radiomics features were extracted as separate features. Two breast radiologists (radiologist 1, 4 years'

experience; radiologist 2, 10 years' experience) who were blinded to the pathological results manually masked the masses in 3Dslicer 4.10.2 ([www.slicer.org](http://www.slicer.org)) (**Figure 1**). A total of 188 radiomics features were extracted from mammographic masses by using the Pyradiomics python package, including 4 shape features, 34 density features, and 150 texture features. Shape features were used to quantify the size and regularity of the mammographic masses, including maximum two-dimensional (2D) diameter and perimeter to surface ratio, with a lower value of perimeter-to-surface ratio indicating a more regular mammographic mass. First-order features refer to radiologically relevant information about the density of mammographic masses such as mean value and kurtosis. Texture features were calculated based on gray-level co-occurrence matrix, gray-level dependence matrix, gray-level run-length matrix, gray-level size zone matrix, and neighborhood gray-tone difference matrix, which were used to quantify the randomness, correlations, variation, homogeneity, and heterogeneity of mammographic masses. The detailed formulae for the calculation of the radiomics features can be found here (31), and the data from radiologist 1 were used to build a radiomics model.

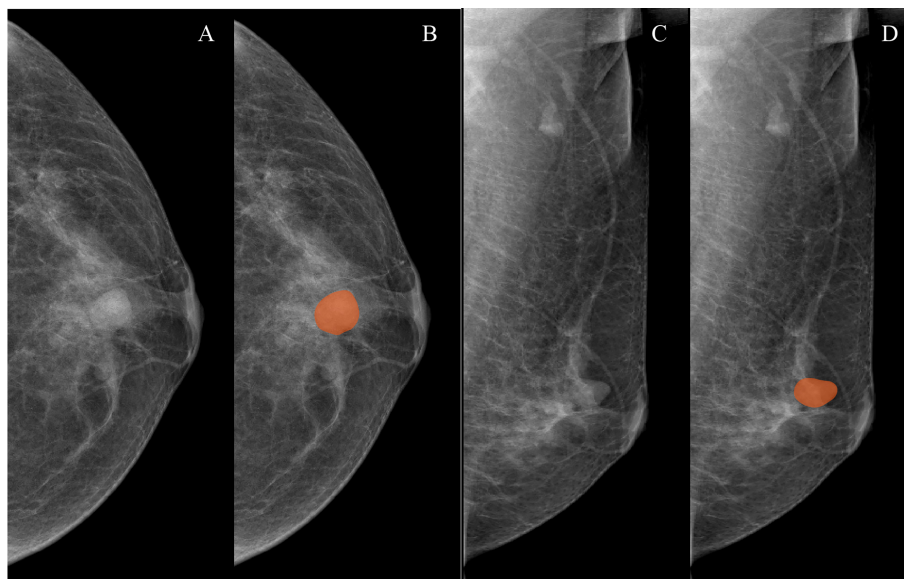
### Patients Grouping and Feature Selection

The patients were divided into two temporal sets based on the order in which they accepted their mammography examinations. The training set consisted of the first two-thirds of patients, and the independent test set comprised the last one-third of patients. In order to avoid some potential bias such as model over-fitting, we applied Pearson's correlation and the least absolute shrinkage and selection operator (LASSO) regression to screen out non-redundant and useful radiomics features in the training set,

respectively. For the Pearson's correlation method, each radiomics feature generated 187 correlation coefficients and 1 corresponding mean absolute correlation coefficient. If two radiomics features had a coefficient exceeding 0.8, the radiomics feature with the larger mean absolute correlation coefficient was deleted. This was implemented in R software version 4.0.1 with package "caret." For the LASSO regression (alpha=1, no elastic net), a 10-fold cross-validation method with 1 standard error of the minimum mean-square error criteria was used to select radiomics features (32), and corresponding  $\lambda$  values were also be calculated. In this study, the radiomics features with non-zero coefficient at the suitable value of parameter  $\lambda$  were determined as useful radiomics features. This was implemented in R software version 4.0.1 with "glmnet" package.

### Radiomics Model Construction and Testing

The malignant mammographic masses were coded as 1, and the benign mammographic masses were coded as 0. The support vector machines (SVMs) with linear kernel (output predicted probability and other parameters are default parameters) were used to construct a radiomics model in this study because of popularity and efficiency in BC (33). This was done in R software version 4.0.1 with "e1071" package. The useful radiomics features were used to construct a radiomics model in the training set for distinguishing between benign and malignant mammographic masses. The receiver operating characteristic curve (ROC) analysis was used to evaluate the classification performance of the radiomics model, including the area under the receiver operating characteristic curve (AUC), the probability threshold value (cut-point) of higher than 98% sensitivity, and corresponding specificity (5). The independent test set was



**FIGURE 1** | Examples of mammographic masses masking on digital mammography images. **(A, C)** Craniocaudal (CC) position images and mediolateral oblique position (MLO) images, respectively. **(B, D)** Manually drawn areas of mammographic masses masking on CC and MLO images, respectively.

used to test the results of the training set by using the “predict” function, and the probability threshold value was also tested. The ROC analysis was implemented in R software version 4.0.1 with package “pROC.” The workflow of this study is reported in **Figure 2**.

## Relative Importance of Useful Radiomics Features

The magnitude of the coefficients in the LASSO algorithm were used to measure the relative weight of useful features as previously described (34, 35). Furthermore, the useful features were grouped by category, such as shape feature group, first-order feature group, and texture feature group, which were added sequentially to the final SVM model, and the AUC of each addition was calculated to assess whether all three groups of useful features contributed to the model.

## Radiomics Model Stability

In this study, a fivefold cross-validation method was used to evaluate the stability of radiomics model. In order not to break the concealment of the test set data, a fivefold cross-validation method was performed separately in the training and test sets. Specifically, the training and test sets were divided into 10 subsets each by fivefold cross-validation method. The “predict” function of R was used to test the diagnostic performance of the radiomics model in these subsets, and the mean and standard deviation of AUCs, sensitivities, and specificities were calculated to measure the stability of the radiomics model. The fivefold cross-validation method was performed in R software version 4.0.1 with package “caret.”

## Reproducibility Assessment

The intra-class correlation coefficients (ICCs) of interobserver (radiologist 1 vs. radiologist 2) were calculated to evaluate the reproducibility of the radiomics features extraction. All data from radiologists 1 and 2 were used as separate test sets so that a kappa value for both radiologists could be calculated to assess the diagnostic reproducibility of the radiomics model.

## Statistical Analysis

All statistical analyses were performed in R software version 4.0.1. All ROC analysis were based on package “pROC,” and the differences of AUC were calculated on Delong’s test. All confidence intervals (CI) were derived from 1,000 bootstrap replicates. All statistical tests were two-sided, and the Bonferroni’s method was used to adjust for multiple comparisons.

## RESULTS

### Clinical Data of Patients

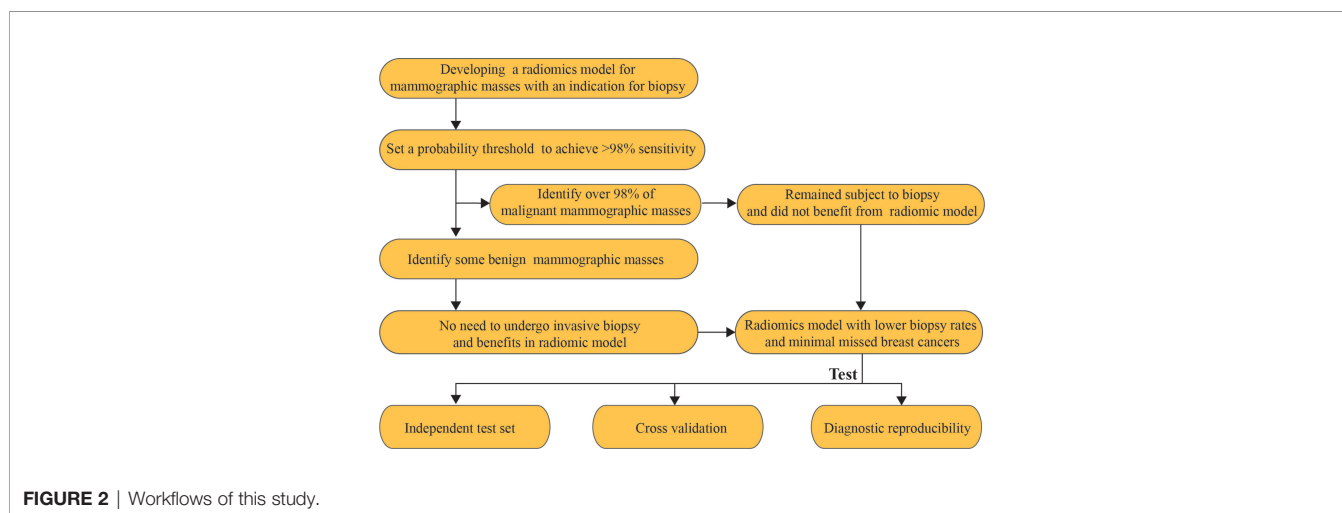
This study included 288 female patients (age,  $52.41 \pm 10.31$ ) with solitary BI-RADS category 4 or 5 mammographic masses, including 130 cases of malignant mammographic masses and 158 cases of benign mammographic masses. The malignant mammographic masses include invasive ductal carcinoma ( $n=51$ ), invasive lobular carcinoma ( $n=37$ ), mucinous carcinoma ( $n=22$ ), and ductal carcinoma *in situ* ( $n=20$ ), while benign mammographic masses include fibroadenomas ( $n=87$ ), adenosis ( $n=38$ ), and hyperplasia ( $n=33$ ).

### Patients Grouping and Feature Selection

The training set included 82 malignant and 110 benign mammographic masses; the independent test set consisted of 48 malignant and 48 benign mammographic masses. Baseline characteristics of study population in training and test sets are reported in **Table 1**, including age, mass size, mass shape, mass margin, breast density, and BI-RADS category. The Pearson’s correlation method screened out 32 non-redundant radiomics features (**Supplementary Figure S1**), and the LASSO method further selected 14 useful radiomics features (**Figures 3A, B**).

### Radiomics Model Construction and Testing

In the training set, the radiomics model obtained an AUC of 0.934 (95% CI, 0.898–0.971), a threshold of 0.22, a sensitivity of 98.8% [81/82], and a specificity of 60% [66/110]. In the test set, the

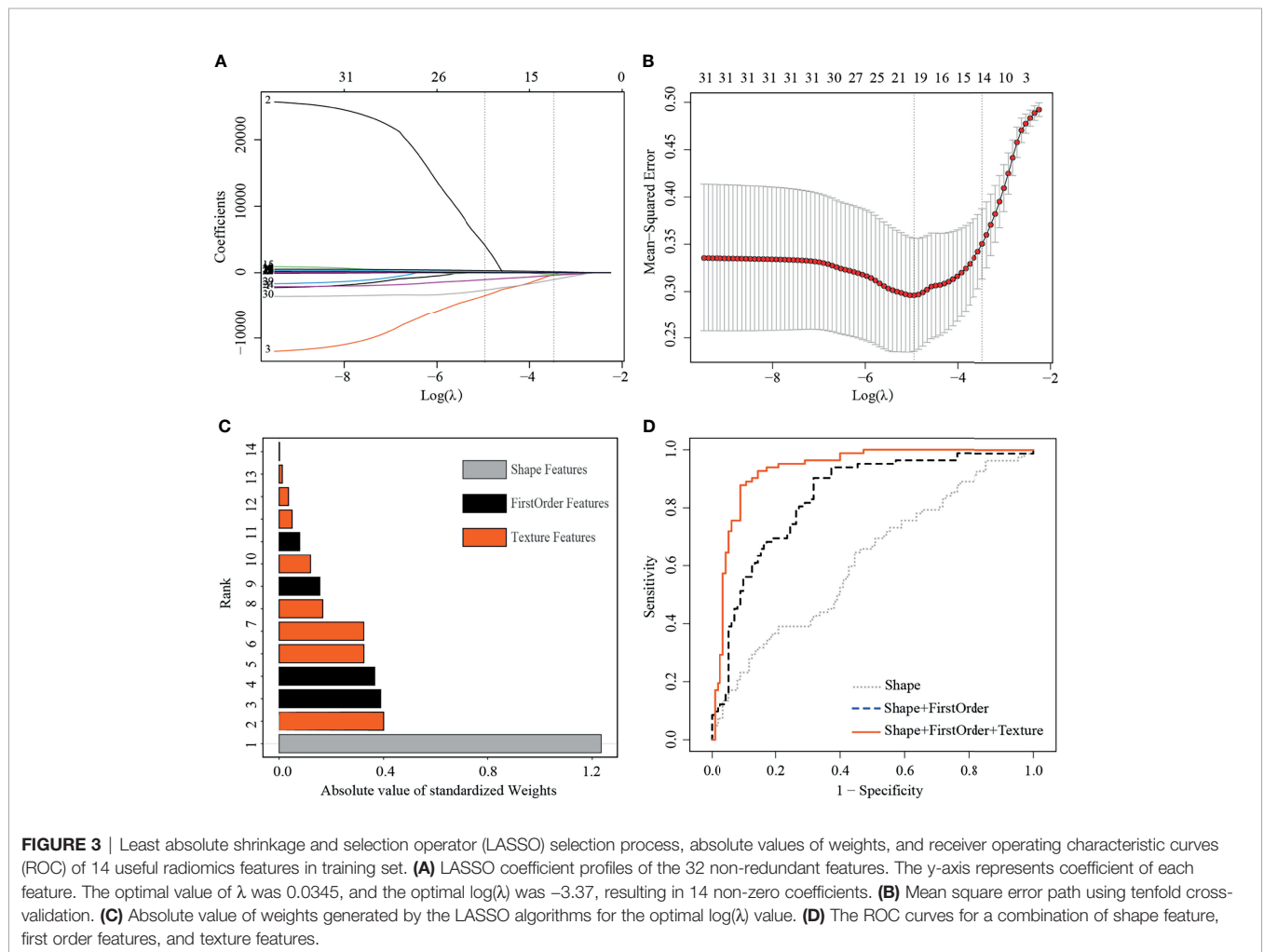


**TABLE 1 |** Baseline characteristics of study population in training and test sets.

Characters	Training set			Test set		
	Malignant(n= 82)	Benign(n= 110)	p-value	Malignant(n= 48)	Benign(n= 48)	p-value
Age (mean ± SD, years)	54.2 ± 10.94	50.38 ± 9.71	0.013	56.66 ± 10.33	49.75 ± 8.76	<0.001
Size (mean ± SD, cm)	1.5 ± 0.51	1.34 ± 0.42	0.021	1.69 ± 0.58	1.3 ± 0.42	<0.001
Shape			0.002			<0.001
Round or oval	10 (12.2%)	34 (30.9%)		5 (10.4%)	20 (41.7%)	
Irregular	72 (87.8%)	76 (69.1%)		43 (89.6%)	28 (58.3%)	
Margin			<0.001			<0.001
Circumscribed	5 (6.1%)	16 (14.5%)		3 (6.3%)	12 (25%)	
Ill-defined	26 (31.7%)	82 (74.5%)		22 (45.8%)	33 (68.8%)	
Spiculated	51 (62.2%)	12 (10.9%)		23 (47.9%)	3 (6.3%)	
Breast density			0.169			0.094
Entirely fatty	7 (8.5%)	15 (13.6%)		4 (8.3%)	7 (14.6%)	
Scattered fibroglandular	37 (45.1%)	58 (52.7%)		22 (45.8%)	29 (60.4%)	
Heterogeneously dense	38 (46.3%)	37 (33.6%)		22 (45.8%)	12 (25%)	
Extremely dense	...	...				
BI-RADS category			<0.001			0.031
4	64 (78%)	105 (95.5%)		38 (79.2%)	46 (95.8%)	
5	18 (22%)	5 (4.5%)		10 (20.8%)	2 (4.2%)	

Student's t-test for normally distributed continuous variable (age and size); Pearson's chi-square test for categorical variables (shape, margin, breast density, BI-RADS categories, and pathological results).

SD, standard deviation; BI-RADS, breast imaging reporting and data system.



radiomics model obtained an AUC of 0.901 (95% CI, 0.835–0.961), a sensitivity of 95.8% [46/48], and a specificity of 66.7% [32/48] (Figure 4A).

## Relative Importance of Useful Radiomics Features

For the useful single radiomics feature, perimeter-to-surface ratio from shape features had the largest absolute weight value of 1.234, followed by coarseness from texture features of 0.4 and mean from first-order features of 0.389 (Figure 3C, details in Table 2). For useful single-category radiomics features, texture feature group obtained the largest absolute weight value of 1.435, followed by shape feature group of 1.234, and lastly by first-order feature group of 0.96 (Figure 3C, details in Table 2). When shape feature group, first-order feature group, and texture feature group were added sequentially to the final SVM model, ROC analysis showed a significant increase in AUC for each addition (shape feature group, AUC=0.613; shape feature group plus first-order feature group, AUC=0.835; first-order feature group plus and first-order feature group plus texture feature group, AUC=0.934;  $p < 0.001$  for each addition) (Figure 3D, details in Table 3).

## Radiomics Model Stability

In the training set, the average AUC was  $0.9 \pm 0.038$ , average sensitivity was  $97.39\% \pm 3.9\%$ , and average specificity was  $50\% \pm 12.5\%$ . In the test set, the average AUC was  $0.915 \pm 0.062$ , average sensitivity was  $98.7\% \pm 4\%$ , and average specificity was  $36.7\% \pm 8.6\%$  (Figure 4B, details in Table 4 and Supplementary Figure S2).

## Reproducibility Assessment

Of the 188 extracted radiomics feature, more than 95% [180/188] radiomics features obtained good reproducibility ( $ICC > 0.75$ ) between radiologist 1 and radiologist 2. The radiomics model

**TABLE 2** | Absolute value of weights of selected useful radiomics features in training set.

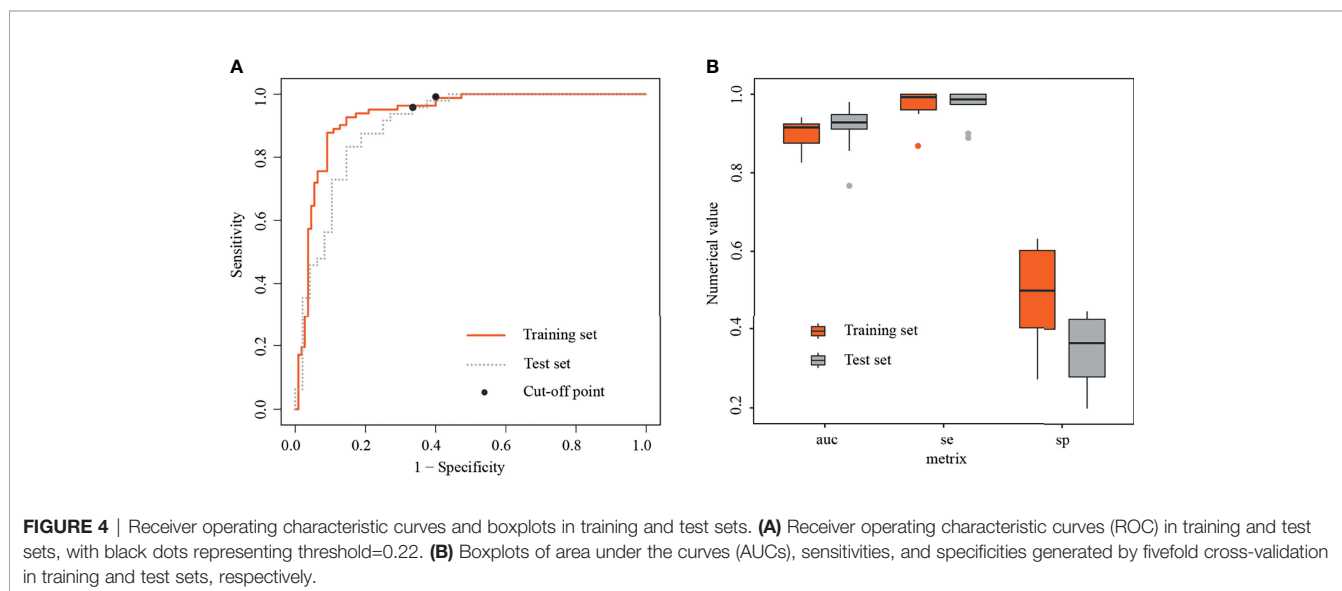
Feature category	Feature name	Absolute weights
Shape	MLO/Sphericity	1.234
	FirstOrder	0.389
	MLO/Uniformity	0.366
Texture	CC/Skewness	0.156
	MLO/Mean	0.079
	MLO/NGTDM Coarseness	0.4
	MLO/GLSZM GrayLevelNonUniformity	0.325
	CC/GLDM	0.324
	SmallDependenceHighGrayLevelEmphasis	
	MLO/GLDM	0.167
	LargeDependenceLowGrayLevelEmphasis	
	CC/NGTDM Coarseness	0.12
	CC/NGTDM Busyness	0.05
MLO/GLCM InverseDifferenceNormalized	0.036	
CC/GLSZM ZoneVariance	0.012	
CC/GLCM ClusterProminence	0.001	

CC, craniocaudal; MLO, mediolateral oblique; NGTDM, neighborhood gray-tone difference matrix; GLSZM, gray-level size zone matrix; GLDM, gray-level dependence matrix; GLCM, gray-level co-occurrence matrix.

obtained good agreement between radiologist 1 and radiologist 2 (Cohen's kappa=0.748; 95% CI, 0.67–0.825;  $p < 0.001$ ).

## DISCUSSION

In this study, we built a radiomics model with similar sensitivity to biopsy for predicting malignancy of BI-RADS category 4 and 5 mammographic masses by using the combination of CC and MLO position images from DM. In both training and test sets, the radiomics model obtained specificity by over 60% while preserving sensitivity more than 95.8%. Both AUC and sensitivity were relatively stable, while the specificity was not so stable. These experimental results suggest that the non-invasive



**TABLE 3** | Classification performance of selected shape feature, first order features, and texture features in classifying malignancies and benign lesions in the training set.

Features	AUC*	p-value†
Shape	0.613 (0.528, 0.695)	...
Shape+FirstOrder	0.835 (0.774, 0.888)	...
Shape+FirstOrder+Texture	0.934 (0.898, 0.968)	...
Shape vs (Shape+FirstOrder)	...	<0.001 (0.017)
Shape vs (Shape+FirstOrder+Texture)	...	<0.001 (0.017)
(Shape+FirstOrder) vs (Shape+FirstOrder+Texture)	...	<0.001 (0.017)

\*Numbers in parentheses are 95% confidence intervals.

†Numbers in parentheses are the significance level.

AUC, area under the receiver operating characteristic curve.

**TABLE 4** | Classification performance of radiomics model in fivefold cross-validation.

Data set	Fivefold CV	Pathology results			AUC	Sensitivity (%)	Specificity (%)
		Malignant	Benign	p			
Training set	-Fold 1	67	87	0.653	0.941 (0.898, 0.977)	98.5 [66/67]	50.6 [44/87]
	Fold 1	15	23		0.910 (0.806, 0.997)	100 [15/15]	39.1 [9/23]
	-Fold 2	65	88	0.901	0.927 (0.881, 0.965)	100 [65/65]	30.7 [27/88]
	Fold 2	17	22		0.912 (0.786, 1)	100 [17/17]	27.2 [6/22]
	-Fold 3	62	91	0.302	0.926 (0.878, 0.967)	95.2 [59/62]	58.2 [53/91]
	Fold 3	20	19		0.856 (0.713, 0.966)	95 [19/20]	63.2 [12/19]
	-Fold 4	67	87	0.653	0.920 (0.871, 0.962)	98.5 [66/67]	63.2 [55/87]
	Fold 4	15	23		0.826 (0.672, 0.954)	100 [15/15]	43.5 [10/23]
	-Fold 5	67	87	0.653	0.919 (0.872, 0.960)	100 [67/67]	49.4 [43/87]
	Fold 5	15	23		0.864 (0.742, 0.965)	86.7 [13/15]	60.9 [14/23]
	Mean± SD	...	...		0.9 ± 0.038	97.39 ± 3.9	50 ± 12.5
Test set	-Fold 1	38	39	0.798	0.919 (0.846, 0.976)	97.4 [37/38]	33.3 [13/39]
	Fold 1	10	9		0.967 (0.867, 1)	100 [10/10]	44.4 [4/9]
	-Fold 2	38	38	1	0.912 (0.837, 0.968)	100 [38/38]	23.7 [9/38]
	Fold 2	10	10		0.980 (0.920, 1)	100 [10/10]	20 [2/10]
	-Fold 3	39	38	0.798	0.947 (0.883, 0.994)	97.4 [38/39]	44.7 [17/38]
	Fold 3	9	10		0.856 (0.667, 1)	100 [9/9]	40 [4/10]
	-Fold 4	39	38	0.798	0.937 (0.879, 0.981)	100 [39/39]	26.3 [10/38]
	Fold 4	9	10		0.911 (0.688, 1)	88.9 [8/9]	40 [4/10]
	-Fold 5	38	39	0.798	0.949 (0.887, 0.993)	97.4 [37/38]	43.6 [17/39]
	Fold 5	10	9		0.767 (0.500, 1)	90 [9/10]	33.3 [3/9]
	Mean± SD	...	...		0.915 ± 0.062	98.7 ± 4	36.7 ± 8.6

CV, cross validation; AUC, area under the receiver operating characteristic curve; SD, standard deviation.

imaging radiomics method could achieve similar sensitivity to biopsy while avoiding some benign mammographic masses to undergo unnecessary invasive biopsy.

We are aware of several papers that use mammographic radiomics data to differentiate between benign and malignant breast lesions (30, 36–38), and we are aware that several of these papers incorporated morphological features (36, 37), and several that did not (30, 38). However, most breast lesions that present alone as calcification, architectural distortion, or asymmetries tend to lack distinct grayscale contrast boundaries, which compromises the masking of tumors. Besides, non-invasive imaging techniques with similar biopsy sensitivity are important for differentiating between benign and malignant BI-RADS category 4 and 5 breast lesions for which invasive biopsy was already indicated. To the best of our knowledge, this is the first study to use mammographic radiomics data to predict the malignancy of BI-RADS category 4 and 5 mammographic masses with a sensitivity similar to that of a biopsy, and we believe that the radiomics model is more limited but more relevant.

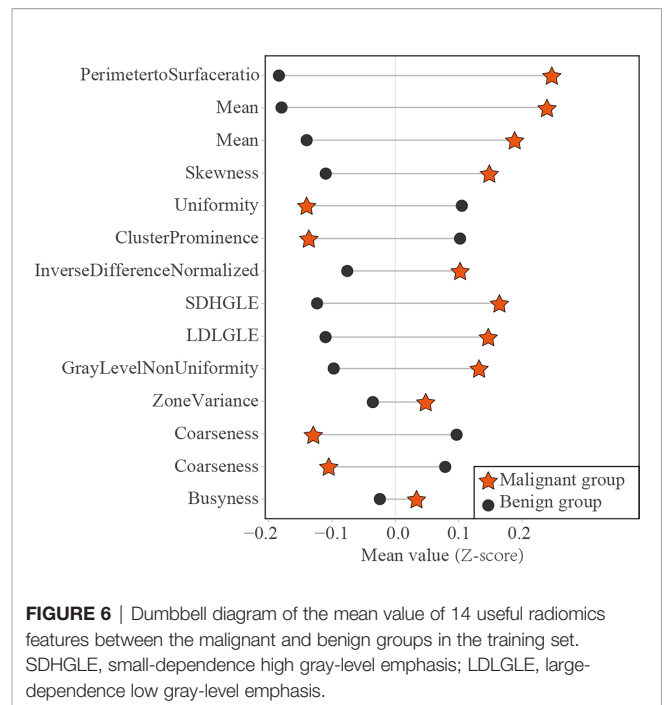
When using radiomics data to predict the malignancy of breast lesions, commonly used mammography images include CC position alone (37, 39), mixed CC and MLO position (30, 36), and combination of CC position and MLO position (40, 41). Gupta et al. demonstrated that the corresponding first-order and texture features of mammographic masses between CC position and MLO position were not strongly correlated (42), suggesting that the inclusion of first-order and texture features from multiple mammographic positions may impact the accuracy of diagnosis of mammographic masses. Ma et al. has shown that by combining both CC and MLO position, radiomics data had higher classification performance between HER2-enriched BC and non-HER2-enriched BC than using CC position alone and MLO position alone (11). Hence, we applied both CC and MLO position radiomics data to predict the malignancy of mammographic masses suspicious for cancer, and the results showed that this method had good classification performance.

Recent studies have shown that the use of a single random training–test set split may lead to unreliable results in small

sample radiomics machine learning studies (43). In our study, we divided the training and test sets based on the chronological order in which patients underwent mammography examinations, and we performed random fivefold cross-validation in training and test sets, respectively, to assess the stability of the radiomics model. Although the specificity gives a large error (12.5%, 8.6%), the AUC (0.038, 0.062) and sensitivity (3.9%, 4%) were relatively stable. This is consistent with previous findings that cross-validation may lead to large error bars for small sample sizes (44).

Of the 188 mammographic radiomics features, the feature selection method screened out 14 useful features in the training set (n=192, ratio 14:1). As suggested by Gillies et al., each radiomics feature requires at least 10 samples to support in a classifier (45). Previous publication has shown that maximum 2D diameter is a useful feature for predicting BC (37). In this study, we found a significant positive correlation between maximum 2D diameter and gray-level size zone matrix-based gray level non-uniformity in the CC position images ( $r=0.82$ ,  $p<0.001$ ) and MLO position images ( $r=0.924$ ,  $p<0.001$ ) of the training set (Figure 5). Thus, the maximum 2D diameter feature was removed when filtering features in the Pearson’s correlation analysis. It is worth noting that although the size feature was not included in our radiomics model, we do not consider it unimportant in predicting malignancy of mammographic masses suspicious for cancer.

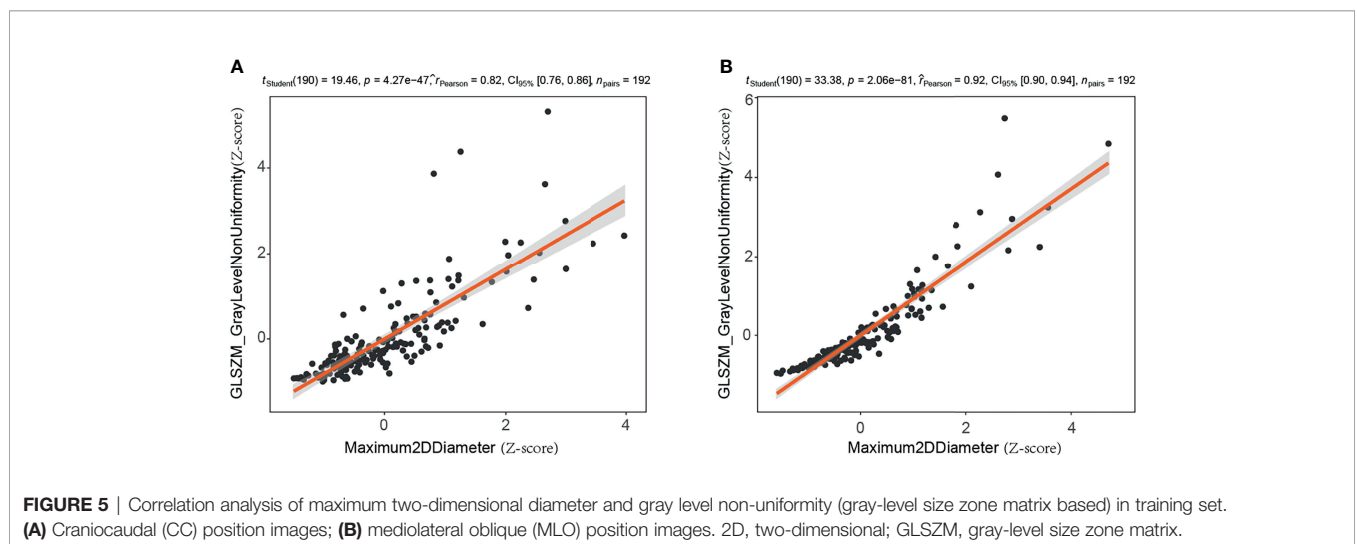
Of the 14 useful mammographic radiomics features, perimeter-to-surface ratio had the greatest weight value of 1.234, suggesting that morphological features remain important in identifying benign lesions or malignancies for mammographic masses suspicious for cancer. The mean density obtained moderate weight values of 0.389 and 0.079, which was higher in the malignant group than in the benign group, indicating that the malignant mammographic masses tend to be higher density than benign mammographic masses. In addition, the malignant group had higher value for inverse difference normalized, zone variance, gray level non-uniformity, and busyness, and lower value for uniformity, skewness,



**FIGURE 6** | Dumbbell diagram of the mean value of 14 useful radiomics features between the malignant and benign groups in the training set. SDHGLE, small-dependence high gray-level emphasis; LDLGLE, large-dependence low gray-level emphasis.

coarseness, small-dependence high gray-level emphasis, and large-dependence low-gray level emphasis, indicating more heterogeneity of malignant mammographic masses in the density and texture patterns (31) (Figure 6). This experimental result is consistent with the pathological fact that malignant tumors tend to be heterogeneous, while benign lesions tend to be homogeneous (46).

Admittedly, our study has several limitations. First, mammographic radiomics data for this study were collected from a single center with a limited number of participants, and further multicenter testing is needed. Second, the mammographic masses were masked by manual method; however, good inter-observer reproducibility was obtained in feature extraction and model diagnosis. Some publications indicate that semi-automatic



**FIGURE 5** | Correlation analysis of maximum two-dimensional diameter and gray level non-uniformity (gray-level size zone matrix based) in training set. (A) Craniocaudal (CC) position images; (B) mediolateral oblique (MLO) position images. 2D, two-dimensional; GLSZM, gray-level size zone matrix.



segmentation method had higher inter-observer reproducibility (47, 48). Further work is needed to construct a semi-automatic segmentation method for mammographic masses. Finally, the radiomics model was constructed by 2D images, which may lose some important information of mammographic masses. However, some publications showed that 2D radiomics features had higher classification performance than 3D radiomics features in lung cancer (47, 48).

In conclusion, a mammographic radiomics model combining both CC and MLO position images had excellent sensitivity and moderate specificity in differentiating malignancies and benign lesions for BI-RADS category 4 and 5 mammographic masses. It may help reduce the temporarily unnecessary invasive biopsies for benign mammographic masses while providing similar diagnostic performance for malignant mammographic masses as biopsies.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Xiang'an Hospital of

Xiamen University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

Study concept and design: GW and DS. Acquisition of data: HZ, QG, and SW. Analysis of data: GW, DS. Drafting of the manuscript: GW, DS, HZ, and SW. Critical revision: QG and KR. Statistical analysis: GW and DS. Study supervision: QG and KR. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Scientific Research Foundation for Advanced Talents, Xiang'an Hospital of Xiamen University (no. PM201809170011).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.843436/full#supplementary-material>

## REFERENCES

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
- Rao AA, Feneis J, Lalonde C, Ojeda-Fournier H. A Pictorial Review of Changes in the BI-RADS Fifth Edition. *Radiographics* (2016) 36:623–39. doi: 10.1148/rg.2016150178
- Balleyguier C, Arfi-Rouche J, Boyer B, Gauthier E, Helin V, Loshkajian A, et al. A New Automated Method to Evaluate 2D Mammographic Breast Density According to BI-RADS® Atlas Fifth Edition Recommendations. *Eur Radiol* (2019) 29:3830–8. doi: 10.1007/s00330-019-06016-y
- Strobel K, Schradang S, Hansen NL, Barabasch A, Kuhl CK. Assessment of BI-RADS Category 4 Lesions Detected With Screening Mammography and Screening US: Utility of MR Imaging. *Radiology* (2015) 274:343–51. doi: 10.1148/radiol.14140645
- Bickelhaupt S, Jaeger PF, Laun FB, Lederer W, Daniel H, Kuder TA, et al. Radiomics Based on Adapted Diffusion Kurtosis Imaging Helps to Clarify Most Mammographic Findings Suspicious for Cancer. *Radiology* (2018) 287:761–70. doi: 10.1148/radiol.2017170273
- Ali M, D'Amico NC, Interlenghi M, Maniglio M, Fazzini D, Schiaffino S, et al. A Decision Support System Based on BI-RADS and Radiomic Classifiers to Reduce False Positive Breast Calcifications at Digital Breast Tomosynthesis: A Preliminary Study. *Appl Sci* (2021) 11:2503. doi: 10.3390/app11062503
- Ciatto S, Houssami N, Ambrogetti D, Bianchi S, Bonardi R, Brancato B, et al. Accuracy and Underestimation of Malignancy of Breast Core Needle Biopsy: The Florence Experience of Over 4000 Consecutive Biopsies. *Breast Cancer Res Treat* (2007) 101:291–7. doi: 10.1007/s10549-006-9289-6
- Wang M, He X, Chang Y, Sun G, Thabane L. A Sensitivity and Specificity Comparison of Fine Needle Aspiration Cytology and Core Needle Biopsy in Evaluation of Suspicious Breast Lesions: A Systematic Review and Meta-Analysis. *Breast* (2017) 31:157–66. doi: 10.1016/j.breast.2016.11.009
- Youk JH, Kim EK, Kim MJ, Lee JY, Oh KK. Missed Breast Cancers at US-Guided Core Needle Biopsy: How to Reduce Them. *Radiographics* (2007) 27:79–94. doi: 10.1148/rg.271065029
- Leithner D, Horvat JV, Marino MA, Bernard-Davila B, Jochelson MS, Ochoa-Albiztegui RE, et al. Radiomic Signatures With Contrast-Enhanced Magnetic Resonance Imaging for the Assessment of Breast Cancer Receptor Status and Molecular Subtypes: Initial Results. *Breast Cancer Res* (2019) 21:106. doi: 10.1186/s13058-019-1187-z
- Ma W, Zhao Y, Ji Y, Guo X, Jian X, Liu P, et al. Breast Cancer Molecular Subtype Prediction by Mammographic Radiomic Features. *Acad Radiol* (2019) 26:196–201. doi: 10.1016/j.acra.2018.01.023
- Guo Y, Hu Y, Qiao M, Wang Y, Yu J, Li J, et al. Radiomics Analysis on Ultrasound for Prediction of Biologic Behavior in Breast Invasive Ductal Carcinoma. *Clin Breast Cancer* (2018) 18:e335–e44. doi: 10.1016/j.clbc.2017.08.002
- Han L, Zhu Y, Liu Z, Yu T, He C, Jiang W, et al. Radiomic Nomogram for Prediction of Axillary Lymph Node Metastasis in Breast Cancer. *Eur Radiol* (2019) 29:3820–9. doi: 10.1007/s00330-018-5981-2
- Sun Q, Lin X, Zhao Y, Li L, Yan K, Liang D, et al. Deep Learning vs. Radiomics for Predicting Axillary Lymph Node Metastasis of Breast Cancer Using Ultrasound Images: Don't Forget the Peritumoral Region. *Front Oncol* (2020) 10:53. doi: 10.3389/fonc.2020.00053
- Yang J, Wang T, Yang L, Wang Y, Li H, Zhou X, et al. Preoperative Prediction of Axillary Lymph Node Metastasis in Breast Cancer Using Mammography-Based Radiomics Method. *Sci Rep* (2019) 9:4429. doi: 10.1038/s41598-019-40831-z
- Wang Z, Lin F, Ma H, Shi Y, Dong J, Yang P, et al. Contrast-Enhanced Spectral Mammography-Based Radiomics Nomogram for the Prediction of Neoadjuvant Chemotherapy-Insensitive Breast Cancers. *Front Oncol* (2021) 11:605230. doi: 10.3389/fonc.2021.605230

17. Liu Z, Li Z, Qu J, Zhang R, Zhou X, Li L, et al. Radiomics of Multiparametric MRI for Pretreatment Prediction of Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer: A Multicenter Study. *Clin Cancer Res* (2019) 25:3538–47. doi: 10.1158/1078-0432.Ccr-18-3190
18. DiCenzo D, Quiaoit K, Fatima K, Bhardwaj D, Sannachi L, Gangeh M, et al. Quantitative Ultrasound Radiomics in Predicting Response to Neoadjuvant Chemotherapy in Patients With Locally Advanced Breast Cancer: Results From Multi-Institutional Study. *Cancer Med* (2020) 9:5798–806. doi: 10.1002/cam4.3255
19. Yu F, Hang J, Deng J, Yang B, Wang J, Ye X, et al. Radiomics Features on Ultrasound Imaging for the Prediction of Disease-Free Survival in Triple Negative Breast Cancer: A Multi-Institutional Study. *Br J Radiol* (2021) 94:20210188. doi: 10.1259/bjr.20210188
20. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. *Radiology* (2016) 281:382–91. doi: 10.1148/radiol.2016152110
21. Park H, Lim Y, Ko ES, Cho HH, Lee JE, Han BK, et al. Radiomics Signature on Magnetic Resonance Imaging: Association With Disease-Free Survival in Patients With Invasive Breast Cancer. *Clin Cancer Res* (2018) 24:4705–14. doi: 10.1158/1078-0432.Ccr-17-3783
22. Xiong L, Chen H, Tang X, Chen B, Jiang X, Liu L, et al. Ultrasound-Based Radiomics Analysis for Predicting Disease-Free Survival of Invasive Breast Cancer. *Front Oncol* (2021) 11:621993. doi: 10.3389/fonc.2021.621993
23. Collins FS, Varmus H. A New Initiative on Precision Medicine. *N Engl J Med* (2015) 372:793–5. doi: 10.1056/NEJMp1500523
24. Hodson R. Precision Medicine. *Nature* (2016) 537:549. doi: 10.1038/537549a
25. Jameson JL, Longo DL. Precision Medicine—Personalized, Problematic, and Promising. *N Engl J Med* (2015) 372:2229–34. doi: 10.1056/NEJMs1503104
26. Venkatesan A, Chu P, Kerlikowske K, Sickles EA, Smith-Bindman R. Positive Predictive Value of Specific Mammographic Findings According to Reader and Patient Variables. *Radiology* (2009) 250:648–57. doi: 10.1148/radiol.2503080541
27. Varela C, Karssemeijer N, Hendriks JH, Holland R. Use of Prior Mammograms in the Classification of Benign and Malignant Masses. *Eur J Radiol* (2005) 56:248–55. doi: 10.1016/j.ejrad.2005.04.007
28. Boyer B, Canale S, Arfi-Rouche J, Monzani Q, Khaled W, Balleyguier C. Variability and Errors When Applying the BIRADS Mammography Classification. *Eur J Radiol* (2013) 82:388–97. doi: 10.1016/j.ejrad.2012.02.005
29. Wang L, Yang W, Xie X, Liu W, Wang H, Shen J, et al. Application of Digital Mammography-Based Radiomics in the Differentiation of Benign and Malignant Round-Like Breast Tumors and the Prediction of Molecular Subtypes. *Gland Surg* (2020) 9:2005–16. doi: 10.21037/gs-20-473
30. Li Z, Yu L, Wang X, Yu H, Gao Y, Ren Y, et al. Diagnostic Performance of Mammographic Texture Analysis in the Differential Diagnosis of Benign and Malignant Breast Tumors. *Clin Breast Cancer* (2018) 18:e621–e7. doi: 10.1016/j.clbc.2017.11.004
31. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77:e104–e7. doi: 10.1158/0008-5472.Can-17-0339
32. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Ser B-Methodol* (1996) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x
33. Ozer ME, Sarica PO, Arga KY. New Machine Learning Applications to Accelerate Personalized Medicine in Breast Cancer: Rise of the Support Vector Machines. *Omic* (2020) 24:241–6. doi: 10.1089/omi.2020.0001
34. Li JW, Cao YC, Zhao ZJ, Shi ZT, Duan XQ, Chang C, et al. Prediction for Pathological and Immunohistochemical Characteristics of Triple-Negative Invasive Breast Carcinomas: The Performance Comparison Between Quantitative and Qualitative Sonographic Feature Analysis. *Eur Radiol* (2022) 32(3):1590–600. doi: 10.1007/s00330-021-08224-x
35. Li Z, Chen F, Zhang S, Ma X, Xia Y, Shen F, et al. The Feasibility of MRI-Based Radiomics Model in Presurgical Evaluation of Tumor Budding in Locally Advanced Rectal Cancer. *Abdom Radiol (NY)* (2022) 47:56–65. doi: 10.1007/s00261-021-03311-5
36. Mao N, Yin P, Wang Q, Liu M, Dong J, Zhang X, et al. Added Value of Radiomics on Mammography for Breast Cancer Diagnosis: A Feasibility Study. *J Am Coll Radiol* (2019) 16:485–91. doi: 10.1016/j.jacr.2018.09.041
37. Li H, Mendel KR, Lan L, Sheth D, Giger ML. Digital Mammography in Breast Cancer: Additive Value of Radiomics of Breast Parenchyma. *Radiology* (2019) 291:15–20. doi: 10.1148/radiol.2019181113
38. Karahaliou A, Skiadopoulos S, Boniatis I, Sakellaropoulos P, Likaki E, Panayiotakis G, et al. Texture Analysis of Tissue Surrounding Microcalcifications on Mammograms for Breast Cancer Diagnosis. *Br J Radiol* (2007) 80:648–56. doi: 10.1259/bjr/30415751
39. Acciavatti RJ, Cohen EA, Maghsoudi OH, Gastounioti A, Pantalone L, Hsieh MK, et al. Incorporating Robustness to Imaging Physics Into Radiomic Feature Selection for Breast Cancer Risk Estimation. *Cancers (Basel)* (2021) 13(21):5497. doi: 10.3390/cancers13215497
40. Lei C, Wei W, Liu Z, Xiong Q, Yang C, Yang M, et al. Mammography-Based Radiomic Analysis for Predicting Benign BI-RADS Category 4 Calcifications. *Eur J Radiol* (2019) 121:108711. doi: 10.1016/j.ejrad.2019.108711
41. Wang S, Sun Y, Li R, Mao N, Li Q, Jiang T, et al. Diagnostic Performance of Perilesional Radiomics Analysis of Contrast-Enhanced Mammography for the Differentiation of Benign and Malignant Breast Lesions. *Eur Radiol* (2022) 32:639–49. doi: 10.1007/s00330-021-08134-y
42. Gupta S, Markey MK. Correspondence in Texture Features Between Two Mammographic Views. *Med Phys* (2005) 32:1598–606. doi: 10.1118/1.1915013
43. An C, Park YW, Ahn SS, Han K, Kim H, Lee SK. Radiomics Machine Learning Study With a Small Sample Size: Single Random Training-Test Set Split may Lead to Unreliable Results. *PLoS One* (2021) 16:e0256152. doi: 10.1371/journal.pone.0256152
44. Varoquaux G. Cross-Validation Failure: Small Sample Sizes Lead to Large Error Bars. *Neuroimage* (2018) 180:68–77. doi: 10.1016/j.neuroimage.2017.06.061
45. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
46. Wang X, Wan Q, Chen H, Li Y, Li X. Classification of Pulmonary Lesion Based on Multiparametric MRI: Utility of Radiomics and Comparison of Machine Learning Methods. *Eur Radiol* (2020) 30:4595–605. doi: 10.1007/s00330-020-06768-y
47. Shen C, Liu Z, Guan M, Song J, Lian Y, Wang S, et al. 2D and 3D CT Radiomics Features Prognostic Performance Comparison in Non-Small Cell Lung Cancer. *Transl Oncol* (2017) 10:886–94. doi: 10.1016/j.tranon.2017.08.007
48. Lacroix M, Frouin F, Dirand AS, Nioche C, Orlhac F, Bernaudin JF, et al. Correction for Magnetic Field Inhomogeneities and Normalization of Voxel Values Are Needed to Better Reveal the Potential of MR Radiomic Features in Lung Cancer. *Front Oncol* (2020) 10:43. doi: 10.3389/fonc.2020.00043

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Shi, Guo, Zhang, Wang and Ren. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.