



Application of machine learning with multiparametric dual-energy computed tomography of the breast to differentiate between benign and malignant lesions

Xiaosong Lan^{1#}, Xiaoxia Wang^{1#}, Jun Qi², Huifang Chen¹, Xiangfei Zeng¹, Jinfang Shi¹, Daihong Liu¹, Hesong Shen¹, Jiuquan Zhang¹

¹Department of Radiology, Chongqing University Cancer Hospital & Chongqing Cancer Institute & Chongqing Cancer Hospital, Chongqing, China; ²Department of Thoracic Surgery, Chongqing University Cancer Hospital, School of Medicine, Chongqing University, Chongqing, China

Contributions: (I) Conception and design: J Zhang, X Lan, X Wang, H Chen, D Liu; (II) Administrative support: J Zhang, J Qi; (III) Provision of study materials or patients: J Zhang, X Wang, H Shen; (IV) Collection and assembly of data: X Lan, X Wang; (V) Data analysis and interpretation: X Lan, X Zeng, J Shi, J Qi; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Prof. Jiuquan Zhang, MD. Department of Radiology, Chongqing University Cancer Hospital, No. 181 Hanyu Road, Shapingba District, Chongqing 400030, China. Email: zhangjq_radiol@foxmail.com.

Background: Multiparametric dual-energy computed tomography (mpDECT) is widely used to differentiate various kinds of tumors; however, the data regarding its diagnostic performance with machine learning to diagnose breast tumors is limited. We evaluated univariate analysis and machine learning performance with mpDECT to distinguish between benign and malignant breast lesions.

Methods: In total, 172 patients with 214 breast lesions (55 benign and 159 malignant) who underwent preoperative dual-phase contrast-enhanced DECT were included in this retrospective study. Twelve quantitative features were extracted for each lesion, including CT attenuation (precontrast, arterial, and venous phases), the arterial-venous phase difference in normalized effective atomic number (nZ_{eff}), normalized iodine concentration (NIC), and slope of the spectral Hounsfield unit (HU) curve (λ_{HU}). Predictive models were developed using univariate analysis and eight machine learning methods [logistic regression, extreme gradient boosting (XGBoost), stochastic gradient descent (SGD), linear discriminant analysis (LDA), adaptive boosting (AdaBoost), random forest (RF), decision tree, and linear support vector machine (SVM)]. Classification performances were assessed based on the area under the receiver operating characteristic curve (AUROC). The best performances of the conventional univariate analysis and machine learning methods were compared using the Delong test.

Results: The univariate analysis showed that the venous phase λ_{HU} had the highest AUROC (0.88). Machine learning with mpDECT achieved an excellent and stable diagnostic performance, as shown by the mean classification performances in the training dataset (AUROC, 0.88–0.99) and testing (AUROC, 0.83–0.96) datasets. The performance of the AdaBoost model based on mpDECT was more stable than the other machine learning models and superior to the univariate analysis (AUROC, 0.96 *vs.* 0.88; $P < 0.001$).

Conclusions: The performance of the AdaBoost classifier based on mpDECT data achieved the highest mean accuracy compared to the other machine learning models and univariate analysis in differentiating between benign and malignant breast lesions.

Keywords: Dual-energy computed tomography (DECT); breast neoplasms; machine learning

Submitted Jan 10, 2021. Accepted for publication Jul 30, 2021.

doi: 10.21037/qims-21-39

View this article at: <https://dx.doi.org/10.21037/qims-21-39>

Introduction

Breast cancer is the most commonly diagnosed cancer in women (1). Recognition of the status of breast tumors is essential for reducing unnecessary biopsies of benign tumors and is vital for treatment decision-making. Magnetic resonance imaging (MRI) is a key means for diagnosing breast lesions and subsequently choosing the appropriate therapy. It is sensitive and is increasingly being used for clinical purposes, such as to assess the extent of malignant breast lesions and monitor the response to chemotherapy (2,3). However, MRI is limited by low specificity, high costs, and incompatible implanted devices (4,5). Thus, the accurate differentiation of breast lesions using MRI remains challenging.

Dual-energy computed tomography (DECT) has promising clinical applications in oncological imaging for the characterization of tumors (6), including the differentiation between benign and malignant tumors (7-10). DECT can provide various quantitative parameters for objective quantitative analysis of breast tumors while screening for lung metastases and inflammation lesions compared with mammograms and ultrasounds. Currently, some studies have investigated the role of DECT in the diagnosis of breast cancer via multiple quantitative parameters (11). One study using DECT quantitative parameters showed that the iodine concentration was higher in breast tumors than in the pectoral muscle and normal breast tissue (12). Another study (10) demonstrated that DECT is a reliable imaging technique with good consistency among observers and can be used for the locoregional staging of breast cancers.

Moreover, the iodine concentration and attenuation [at 70 and 40 kiloelectron volts (keV)] of benign tumors were significantly lower than those of malignant breast tumors. Our previous research (13) found that DECT parameters, including normalized iodine concentration (NIC) and normalized effective atomic number (nZ_{eff}), could be used to discriminate the expression status of immunohistochemical biomarkers of breast cancer. However, to the best of our knowledge, there are no related studies on the diagnostic performance of multiparametric DECT (mpDECT) in differentiating between benign and malignant breast lesions.

Machine learning methods can generate predictive

models by extensively searching the model and parameter spaces, which is different from traditional statistical methods that typically consider and evaluate a limited set of assumptions (14-17). One previous study (18) demonstrated that combining low-dose perfusion breast CT parameters and machine learning approaches is a useful noninvasive method for predicting the molecular subtypes and prognostic biomarkers of breast cancer. Another study (19) showed that machine learning with multiparametric MRI of the breast enables early prediction of pathological complete response to neoadjuvant chemotherapy, as well as survival outcomes, with high accuracy. However, to the best of our knowledge, no study has evaluated the diagnostic performance of machine learning using mpDECT in differentiating between benign and malignant breast lesions.

Previous studies have demonstrated the potential for the application of imaging features with machine learning in breast cancer; however, the potential of mpDECT has not yet been fully tapped. Thus, this study aimed to evaluate the diagnostic performance of conventional univariate analysis and machine learning with mpDECT in distinguishing between benign and malignant breast lesions.

Methods

Participant characteristics

The study was approved by the ethics committee of Chongqing University Cancer Hospital (No.: CZLS20200215-A), and individual consent for this retrospective analysis was waived. This study was conducted per the Declaration of Helsinki (as revised in 2013). The inclusion criteria were as follows: (I) dual-phase contrast-enhanced DECT scan of the thorax; (II) pathological biopsy-confirmed breast malignant or benign lesions; (III) women ≥ 18 years who were not pregnant or breastfeeding; and (IV) patients with no history of chemotherapy or radiation therapy in the breast space. The exclusion criteria were as follows: (I) patients with incomplete pathological/medical information; (II) patients who underwent breast mass biopsy within 1 week before the initial CT examination; (III) cases involving invisible target lesions on CT images; (IV) poor image quality (severe motion artifacts or poor signal-to-noise ratio); and (V) patients with

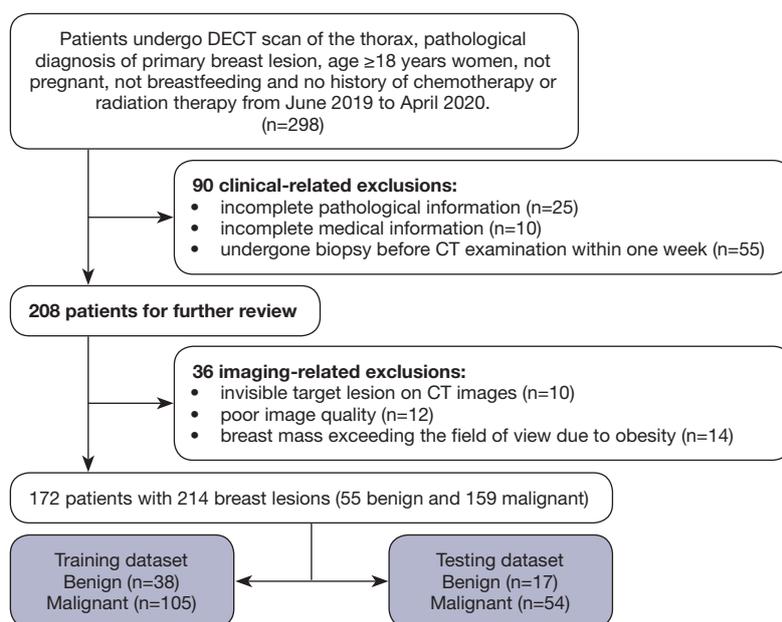


Figure 1 Flowchart of study participant enrolment and selection process. DECT, dual-energy computed tomography; CT, computed tomography.

obesity causing breast mass beyond the field of view. Details regarding patient exclusions can be found in *Figure 1*.

Between June 2019 and April 2020, 172 patients who fulfilled the inclusion criteria were enrolled in our study. All patients underwent DECT scanning once. Among them, 117 patients were diagnosed with malignant breast lesions, 42 patients were diagnosed with unilateral malignant breast lesions (and the other side was diagnosed as benign breast lesions), and 13 patients were diagnosed with benign breast lesions. For all patients, the following information was recorded: age, the largest diameter of the lesion, menstruation state, and histopathological information. The overall workflow chart of this study is shown in *Figure 2*.

DECT image acquisition

DECT data were acquired on a 2.5 generation dual-energy dual-source CT unit (SOMATOM Drive, Siemens Healthineers, Germany). Automatic exposure control (CARE Dose 4D, Siemens Healthineers) was used in our study scans. The scanner settings were as follows: rotation time, 0.28 s; collimation, 64×0.6 mm; pitch, 0.55; reference tube current-time product, 71 milliamperes second for the 100 kilovolt tube and 60 milliamperes second for the Sn140 kilovolt tube; reformatted section increment, 1.5 mm; and reformatted section thickness, 1.5 mm. All participants were

scanned craniocaudally in the supine position. Noncontrast DECT images were obtained first. The iodinated nonionic contrast agent (Ioversol, 320 mg/mL iodine, HENGRUI Medicine, China) was administered through the ulnar vein by a dual-head injector. The dosage was 1.5 mL/kg with a flow rate of 2.5 mL/s, followed by a bolus injection of 30 mL saline administered at the same flow rate. The arterial phase scanning was initiated using a bolus-tracking method with a 100 Hounsfield unit (HU) threshold in the descending aorta and an additional delay of 10 s. The venous phase scan delay time was 25 s after the end of the arterial phase scan.

DECT quantitative features

DECT imaging data were analyzed using viewer software on a syngo.via workstation (syngo.via VB20A, Dual Energy, Siemens Healthineers, Germany). Standard linear-blended images were reconstructed by applying a blending factor of 0.5 ($M_{0.5}$; 50% of the low kV and 50% of the high kV spectrum) for attenuation (HU) measurements. We measured the attenuation in three phases (precontrast, arterial, and venous phases). Dual-energy quantitative parameters were measured in the arterial and venous phases by two radiologists (XXW, with 8 years of experience in breast diagnostic imaging, and XFZ, with 3 years of experience in

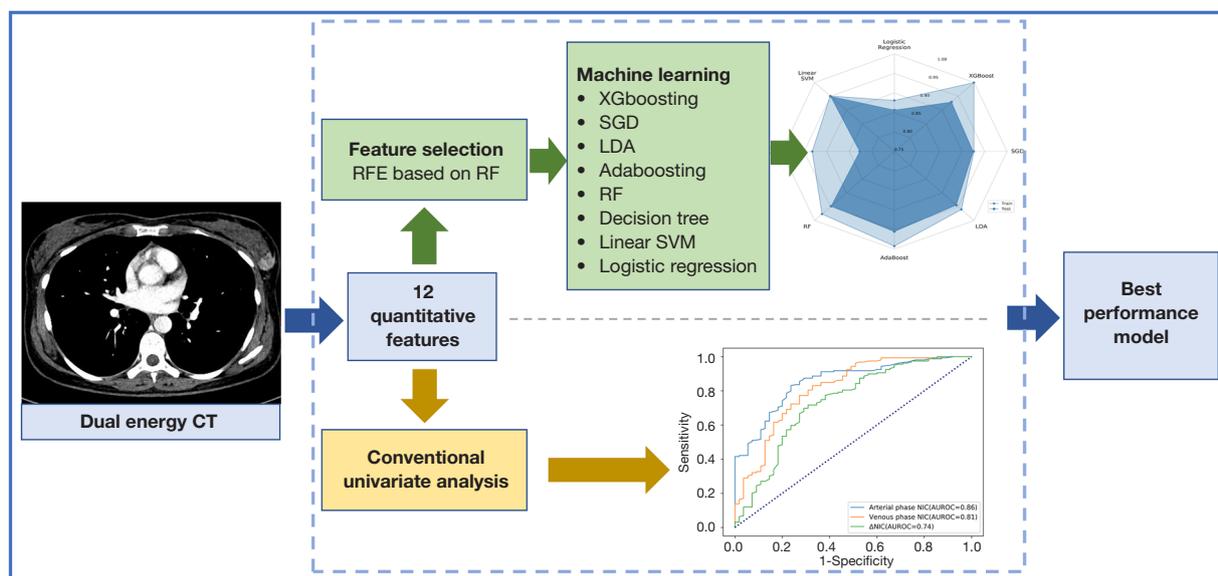


Figure 2 The overall workflow chart of this study. CT, computed tomography; RFE, recursive feature elimination; RF, random forest; XGBoost, extreme gradient boosting; SGD, stochastic gradient descent; LDA, linear discriminant analysis; AdaBoost, adaptive boosting; SVM, linear support vector machine; NIC, normalized iodine concentration; AUROC, area under the receiver operating characteristic curve.

post-processing of DECT) who were blinded to the biopsy results of the breast tumors. A region of interest (ROI) was placed in the breast lesion area as large as possible, excluding the areas of calcification, obvious gross necrosis, or large vessels. The mean area of all ROIs was 421.61 (range, 110.28–1,253.15) mm². The NIC and nZ_{eff} were obtained through the breast lesions iodine concentration (mg/cm³) and effective atomic number dividing by the aortic iodine concentration and the effective atomic number. The slope of the spectral HU curve (λ_{Hu} , HU/keV) was computed according to the following equation (20):

$$\lambda_{Hu} = (HU_{40keV} - HU_{70keV}) / 30keV \quad [1]$$

Eqs. [2-4] for calculating the differences in quantitative DECT features between the arterial phase and venous phase were as follows:

$$\Delta NIC = NIC_{arterial\ phase} - NIC_{venous\ phase} \quad [2]$$

$$\Delta nZ_{eff} = nZ_{eff\ arterial\ phase} - nZ_{eff\ venous\ phase} \quad [3]$$

$$\Delta \lambda_{Hu} = \lambda_{Hu\ arterial\ phase} - \lambda_{Hu\ venous\ phase} \quad [4]$$

Conventional univariate analysis

Univariate analysis of all mpDECT features was performed

to differentiate between benign and malignant breast lesions. The optimal cut-off points of all mpDECT features for predicting benign and malignant breast lesions were determined by the Youden index (21). The area under the receiver operating characteristic curve (AUROC) was used as the classification metric and was used to evaluate the model’s predictive ability.

Machine learning

A recursive feature elimination (RFE) method combined with random forest (RF) was used to select the optimal sequence of features in our study. The concrete implementation was as follows: (I) train the RF model with 10-fold cross-validation (CV); (II) calculate the importance of permutation features; (III) keep the most important features; (IV) repeat steps 1 through 3 until optimal performance is achieved; and (V) select the subset of a feature that predicts the best performance (22).

Eight machine learning models, including logistic regression, extreme gradient boosting (XGBoost), stochastic gradient descent (SGD), linear discriminant analysis (LDA), adaptive boosting (AdaBoost), RF, decision tree, and linear support vector machine (SVM), were applied with the mpDECT to distinguish between benign and malignant

Table 1 Clinicopathological characteristics

Characteristics	Training dataset			Testing dataset			P value
	Malignant (n=105)	Benign (n=38)	P value	Malignant (n=54)	Benign (n=17)	P value	
Age, mean \pm SD, years	53.7 \pm 11.12	50.8 \pm 7.01	0.834	51.8 \pm 5.23	50.5 \pm 8.21	0.104	0.085
Largest diameter, cm	3.01 \pm 1.26	2.71 \pm 0.62	0.773	2.91 \pm 1.21	2.45 \pm 0.73	0.427	0.341
Menstruation state			0.053			0.164	0.162
Premenopausal women	42	20		30	8		
Postmenopausal women	63	18		24	9		
Benign							
Adenosis	–	16		–	7		
Fibroadenoma	–	14		–	6		
Intraductal papilloma	–	7		–	4		
Cyst	–	1		–	0		
Malignant							
Invasive ductal/lobular carcinoma	88	–		46	–		
Medullary carcinoma	1	–		0	–		
Mucinous carcinoma	2	–		1	–		
Phyllodes tumor	2	–		1	–		
Ductal carcinoma <i>in situ</i>	12	–		6	–		

breast lesions. A python (version 3.7.6) library named scikit-learn (version 0.22) was used in our study. The specific parameters of machine learning can be seen in [Figure S1](#).

For more information about each algorithm, see [Appendix 1](#). Each particular learning algorithm was used to provide the best performance model for fitting the input DECT data and correctly predicting benign or malignant breast lesions. Our data were divided into a training group (used to train the model) and a testing group (used to evaluate the model's generalization ability) at a ratio of 67%:33%. Five-fold CV was used to distinguish between the benign and malignant breast lesions groups to improve the performance evaluation and manage the stochasticity in machine learning models. The AUROC was used as the performance metric.

Statistical analysis

Statistical analyses were performed by using commercially statistical software (SPSS software, version 25.0; USA). We randomly selected 30 patients to assess inter-observer agreement in the analysis of the mpDECT features. The

ROIs of mpDECT measurements were repeated twice, with an interval of at least 1 month, following the same procedure. Our study used the intraclass correlation coefficient (ICC) with a two-way random effects model of consistency. In the univariate analyses, 12 mpDECT features were compared using the independent sample *t*-test (normal distribution) or Mann-Whitney U-test (non-normal distribution). Menstruation state was assessed using the χ^2 test. The Delong test was used within the conventional univariate analysis and machine learning models to compare the difference between the AUROCs. The level of significance was defined as $P < 0.05$.

Results

Participant characteristics

In total, 172 participants with 214 histopathologically confirmed breast lesions were included in our study. The patients' clinicopathological characteristics are shown in [Table 1](#). Histopathologically, 159 malignant lesions and 55 benign lesions were diagnosed. No statistically significant

Table 2 Comparison of mpDECT between benign and malignant lesions of the breast and the performance of conventional univariate analysis

Feature	Benign (n=55)	Malignant (n=159)	t/Z value	P value	Cut off	AUROC	Sensitivity (%)	Specificity (%)
Arterial phase NIC	0.033±0.036	0.110±0.068	-8.043	<0.001	0.052	0.86	83.0	76.4
Venous phase NIC	0.176±0.132	0.350±0.144	-7.900	<0.001	0.211	0.81	83.0	67.3
Δ NIC	0.143±0.120	0.240±0.109	-5.547	<0.001	0.178	0.74	71.7	69.0
Arterial phase λ_{HU} (HU/keV)	0.487±1.079	1.714±0.827	-8.701	<0.001	0.930	0.85	86.2	81.8
Venous phase λ_{HU} (HU/keV)	0.927±0.912	2.518±0.962	-10.717	<0.001	1.880	0.88	80.5	83.6
$\Delta\lambda_{\text{HU}}$ (HU/keV)	0.440±0.941	0.814±0.693	-3.129	0.002	-0.030	0.62	91.8	30.9
Arterial phase nZ_{eff}	0.684±0.032	0.725±0.042	-6.537	<0.001	0.720	0.79	61.0	92.7
Venous phase nZ_{eff}	0.824±0.035	0.868±0.052	-5.789	<0.001	0.840	0.79	81.1	67.9
ΔnZ_{eff}	0.140±0.032	0.143±0.054	-0.348	0.728	0.160	0.57	43.4	72.7
Precontrast phase attenuation (HU)	29.55±10.85	39.52±11.92	-5.465	<0.001	37.10	0.74	60.4	76.4
Arterial phase attenuation (HU)	31.70±13.24	54.56±15.55	-9.745	<0.001	42.40	0.87	83.6	81.8
Venous phase attenuation (HU)	38.28±16.00	65.64±18.28	-9.862	<0.001	54.00	0.87	77.4	83.6

The data is represented as means \pm standard deviation. mpDECT, multiparametric dual-energy computed tomography; AUROC, area under the receiver operating characteristic curve; NIC, normalized iodine concentration; λ_{HU} , slope of the spectral Hounsfield unit curve; nZ_{eff} , normalized effective atomic number; HU, Hounsfield unit.

differences were observed between the training and testing datasets in terms of age, largest diameter of the lesion, and menstruation state.

Our study's mean cumulative CT dose index was 5.42 ± 1.94 mGy, while the mean dose length product was 166.12 ± 56.04 mGy cm, and the average effective dose was 2.31 ± 0.78 mSv for each phase.

Conventional univariate analysis

The ICCs of the study for inter-observer variability in terms of mpDECT were 0.930 (0.759–0.990). The mpDECT between benign and malignant breast lesions is shown in *Table 2*. The results showed that, except for ΔnZ_{eff} ($P=0.728$), the 11 other quantitative parameters of malignant lesions were higher than those of benign lesions ($P<0.001$ – 0.002). In the univariate analysis, the venous phase λ_{HU} had the highest AUROC (0.88), followed by arterial phase attenuation (0.87) and venous phase attenuation (0.87) (*Figure S2*). The representative images are shown in *Figure 3*.

Optimum ranking of the features

The mpDECT model was based on RFE incorporated with RF, and the importance of the features in the model

for the prediction of benign and malignant breast lesions is summarized in *Figure 4*. The most relevant features for predicting benign and malignant breast lesions were those quantitative features, including arterial phase λ_{HU} , venous phase λ_{HU} , arterial phase attenuation, arterial phase NIC, venous phase attenuation, and venous phase attenuation phase NIC. According to the ranking of importance scores, we selected six of the most important features to construct models with eight machine learning algorithms.

Diagnostic performance of the eight machine learning models

All models performed excellently, with high AUROC values in training (AUROCs, 0.88–0.99) and testing (AUROCs, 0.83–0.96) datasets (*Figure S3*). *Table 3* and *Figure 5* summarize the AUROCs for all models in the training and testing datasets. Among them, AdaBoost showed the highest performance and outperformed the decision tree models (AUROC 0.96 vs. 0.83; $P=0.034$). There was no statistically significant difference in our study between the performance of the AdaBoost and that of the other six groups of models, including XGBoost, SGD, LDA, AdaBoost, RF, and linear SVM, in the prediction of benign and malignant breast lesions ($P=0.060$ – 0.838).

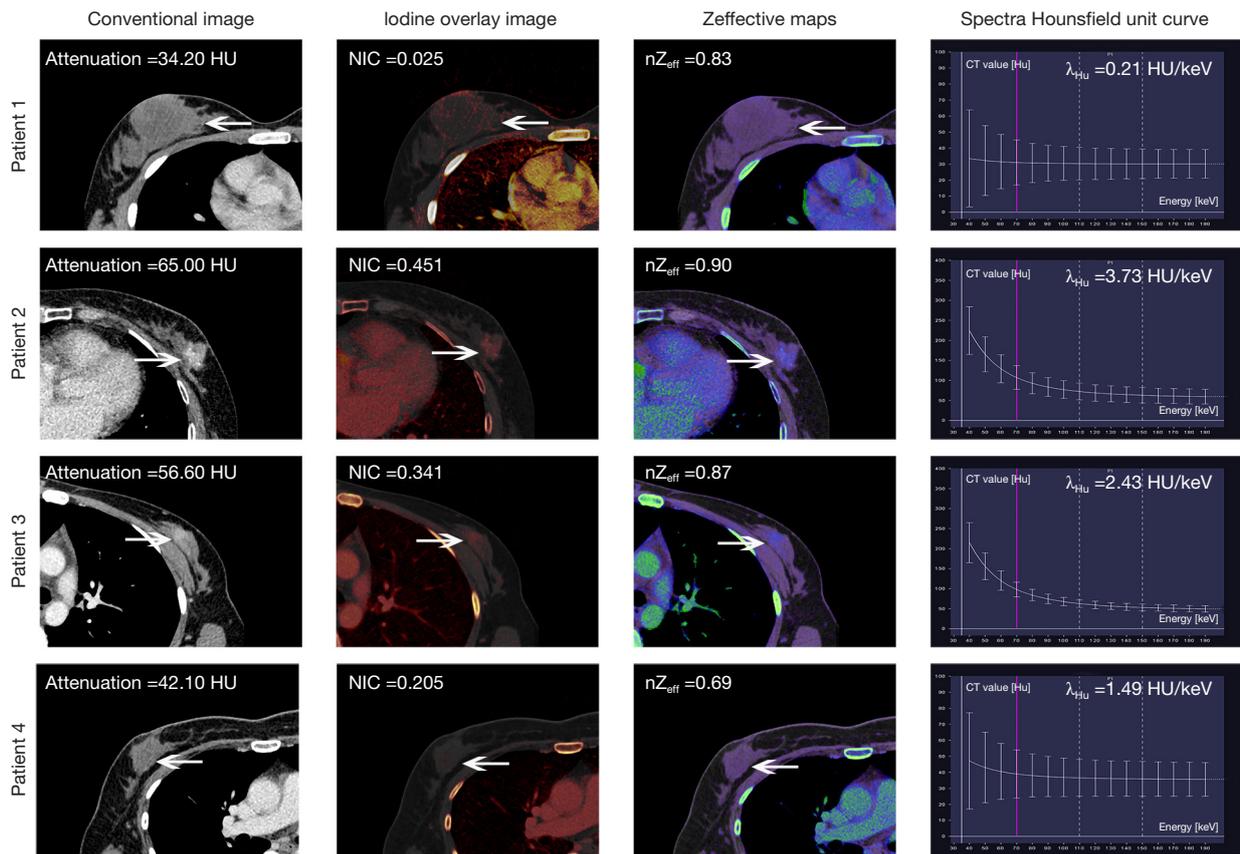


Figure 3 Representative DECT images of contrast-enhanced venous phase in four breast lesion patients. Patient 1: pathological diagnosis of a 34-year-old woman with right breast mucinous carcinoma (arrows). According to the multiple parameters of DECT, benign breast lesions were diagnosed (false negatives). Patient 2: a 52-year-old woman was diagnosed with left breast invasive ductal carcinoma (arrows) on both histopathology and multiple parameters of DECT (true positive). Patient 3: pathological diagnosis of a 45-year-old woman with left breast adenosis (arrows). According to the multiple parameters of DECT, malignant breast lesions were diagnosed (false positive). Patient 4: a 45-year-old woman was diagnosed with right breast fibroadenoma (arrows) on both histopathology and multiple parameters of DECT (true negative). DECT, dual-energy computed tomography; NIC, normalized iodine concentration; λ_{HU} , slope of the spectral Hounsfield unit curve, nZ_{eff} , normalized effective atomic number, HU, Hounsfield unit.

Figure 6 presents the boxplot illustration of the mean performance of five-fold CV in the training dataset along with eight machine learning models in looking for the most stable, high accuracy, and low variance model for predicting benign and malignant breast lesions. AdaBoost was the most stable model (least variance), with AUROC values of 0.99 and 0.96 in the training and testing datasets, respectively (Figure 7). Therefore, we chose the AdaBoost model to compare with the univariate analysis. The AUROC of AdaBoost was significantly higher than that determined by univariate analysis [AUROC of AdaBoost, 0.96 *vs.* AUROC of venous phase λ_{HU} , 0.88 ($P < 0.001$); AUROC of AdaBoost, 0.96 *vs.* AUROC of arterial phase attenuation, 0.87

($P < 0.001$); and AUROC of AdaBoost, 0.96 *vs.* AUROC of venous phase attenuation, 0.87 ($P < 0.001$), respectively].

Discussion

This study applied conventional univariate analysis and machine learning with mpDECT of the breast to distinguish between benign and malignant breast lesions. With the univariate analysis, the venous phase λ_{HU} offered the highest AUROC. Of all machine learning models, the AdaBoost classifier model using mpDECT was more stable than the other machine learning models and outperformed conventional univariate analysis in differentiating between

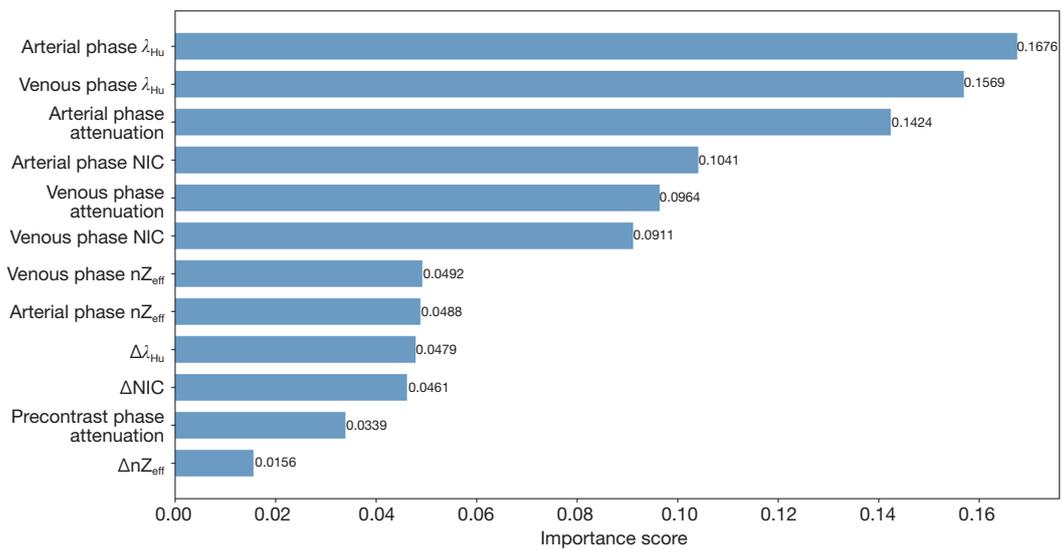


Figure 4 Important features (horizontal) in the mpDECT (vertical) model for predicting malignant and benign breast lesions. mpDECT, multiparametric dual-energy computed tomography; λ_{HU} , slope of the spectral Hounsfield unit curve; NIC, normalized iodine concentration; nZ_{eff} , normalized effective atomic number.

Table 3 Classification performance of mpDECT models using various models in the training and testing datasets

Model	AUROC		Sensitivity		Specificity		F1 score	
	Training dataset	Testing dataset						
Logistic regression	0.88	0.86	0.87	0.92	0.92	0.82	0.88	0.91
XGBoost	0.99	0.93	0.95	0.72	0.97	1.00	0.96	0.75
SGD	0.93	0.92	0.85	0.92	0.92	0.82	0.87	0.91
LDA	0.96	0.94	0.91	0.89	0.92	0.94	0.92	0.90
AdaBoost	0.99	0.96	0.98	0.78	1.00	1.00	0.98	0.80
RF	0.98	0.95	0.97	0.79	0.97	1.00	0.97	0.82
Decision tree	0.93	0.83	0.94	0.89	0.92	0.76	0.94	0.87
Linear SVM	0.95	0.95	0.92	0.85	0.94	1.00	0.93	0.87

mpDECT, multiparametric dual-energy computed tomography; AUROC, area under the receiver operating characteristic curve; XGBoost, extreme gradient boosting; SGD, stochastic gradient descent; LDA, linear discriminant analysis; AdaBoost, adaptive boosting; RF, random forests; SVM, support vector machine.

benign and malignant breast lesions.

DECT can reconstruct multiparametric images and corresponding quantitative parameters as analysis tools and quantitative indicators for clinical diagnosis (23,24). The effective atomic number can quantitatively depict the changes in the X-ray absorption rate for various materials and reflect the atomic number of composite materials; the higher the compound density, the higher

the effective atomic number (25). We concluded that the nZ_{eff} of malignant breast lesions was higher than that of benign lesions. These differences may be due to the varying anatomical structures of malignant and benign lesions caused by the uneven density of breast cell components (1). The iodine concentration can be used to quantitatively reflect the vascularization of various tissues and local corresponding blood volume (6,26,27). New tumor vessels

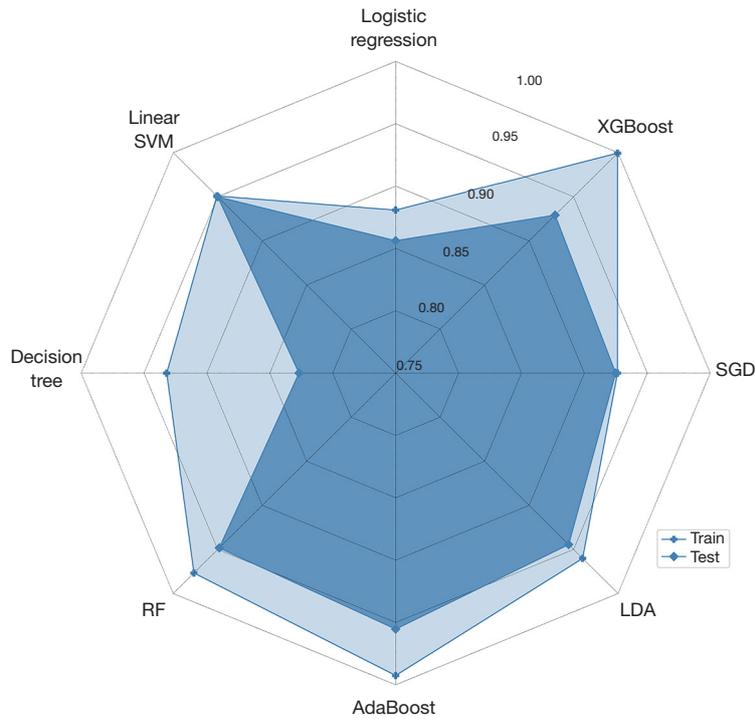


Figure 5 Radar plot illustrations of mean performance in the training and testing datasets to recognize the most stable learning model machine (with high accuracy and low variance) for predicting malignant and benign breast lesions. XGBoost, SGD, LDA, AdaBoost, RF, and linear SVM. XGBoost, extreme gradient boosting; SGD, stochastic gradient descent; LDA, linear discriminant analysis; AdaBoost, adaptive boosting; RF, random forest; SVM, support vector machine.

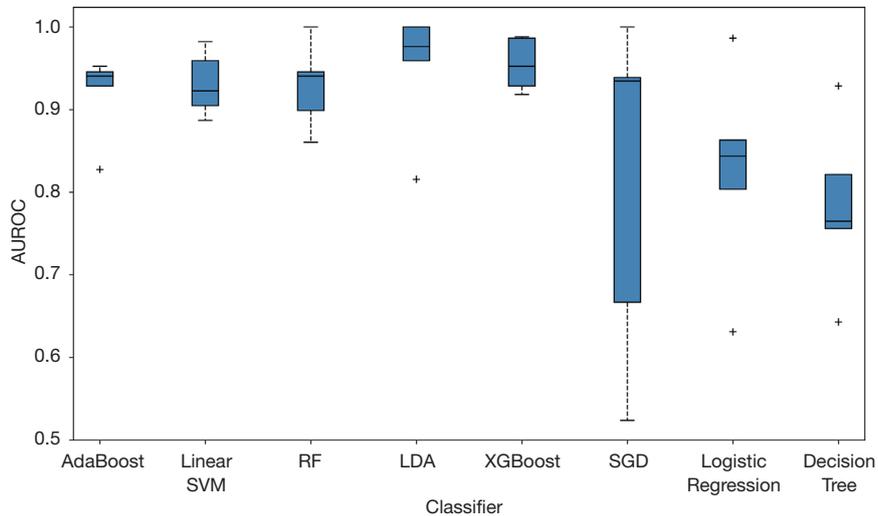


Figure 6 A boxplot illustration of the mean performance of various machine learning algorithms using five-fold CV in the training dataset to predict malignant and benign breast lesions. The AUROC was used as the classification metric. XGBoost, SGD, LDA, AdaBoost, RF, and linear SVM. CV, cross-validation; AUROC, area under the receiver operating characteristic curve; XGBoost, extreme gradient boosting; SGD, stochastic gradient descent; LDA, linear discriminant analysis; AdaBoost, adaptive boosting; RF, random forest; SVM, support vector machine.

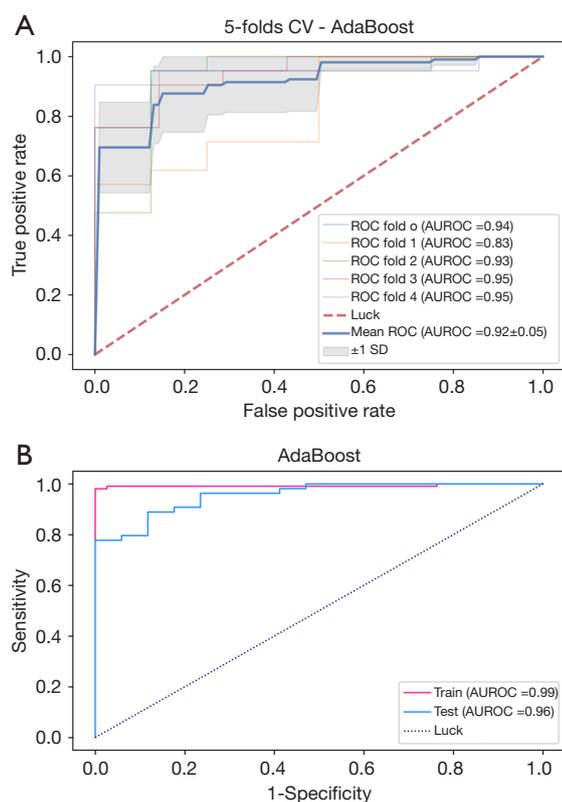


Figure 7 A boxplot illustration of the mean performance of various machine learning algorithms using five-fold CV in the training dataset to predict malignant and benign breast lesions. The AUROC was used as the classification metric. XGBoost, SGD, LDA, AdaBoost, RF, and linear SVM. CV, cross-validation; AUROC, area under the receiver operating characteristic curve; XGBoost, extreme gradient boosting; SGD, stochastic gradient descent; LDA, linear discriminant analysis; AdaBoost, adaptive boosting; RF, random forest; SVM, support vector machine.

in malignant lesions usually contain immature microvessels, increasing blood flow within the tumor (28). We found that the NIC was higher in malignant breast lesions than in benign lesions, which is consistent with the findings of a recent study (29). This result might mean that malignant breast lesions have more underlying microvascular and tumor angiogenesis than benign lesions. λ_{Hu} demonstrates the attenuation changes of the lesion when enhanced by a contrast agent; the faster the spectral curve changes, the higher the proportion of contrast agent in the lesions (30). From our results, we found that the venous phase λ_{Hu} had the highest diagnostic value for differentiating between benign and malignant breast lesions. We obtained similar

results to the most recent study (20), indicating that λ_{Hu} was the best parameter for identifying metastatic sentinel lymph nodes in breast cancer.

RFE is a common integrated tool with strong information feature search capabilities. This algorithm calculates and updates the importance level and eliminates the least important features (31). RF is a good candidate for integrating large amounts of omics data. It can often deal with high-dimensional-related problems, confirm the strong predictors of a specified outcome, and not require assumptions about the underlying model (22). However, high-dimensional datasets have a common problem: the existence of correlated predictors will affect RF's ability to recognize the strongest predictors by reducing the estimated importance scores of relevant variables. The RFE algorithm incorporated with RF was the suggested solution, which presents a promising machine learning algorithm for medical imaging (22). Our study used this method and obtained a ranking of the importance of mpDECT features. Among them, six quantitative features were found to be important for accurately distinguishing between benign and malignant breast lesions.

In contrast to prior research (11), we used multiple quantitative parameters extracted from DECT. We extracted and calculated 12 features per lesion and used eight machine learning algorithms. We demonstrated that the AdaBoost classifier model was more stable than the other machine learning models and outperformed the conventional univariate analysis in differentiating between benign and malignant breast lesions. The AdaBoost method is different from the other seven data mining methods of the machine learning model. It first creates a group of weak classifiers by assigning appropriate extra weight to them and then combines them into a strong model. AdaBoost has unique advantages in accuracy rate and training time compared with other data mining methods (32). To the best of our knowledge, no research has evaluated the value of machine learning approaches in differentiating between benign and malignant breast lesions using mpDECT. The results of our study confirmed the robust performance of the AdaBoost model.

According to the National Comprehensive Cancer Network (NCCN) guidelines for breast cancer (33), chest diagnostic CT with contrast is routinely recommended in clinical stage I–IIb (if directed by signs or symptoms) and IIIa above breast cancer patients for screening lung metastases and other lesions. Primary breast lesions are also included in the screening field. This ‘one scan,

many searches' approach can provide information about lung lesions and provide additional DECT quantitative parameters of primary breast lesions to assist clinical diagnosis. Unfortunately, at present, the clinical value of DECT in breast cancer screening and diagnosis is relatively limited, and we are also looking forward to furthering research that explores the clinical value of DECT in breast cancer screening and diagnosis.

Our study has several limitations that should be noted. Firstly, this study was performed using one DECT device from a single institution. Secondly, the participants underwent thoracic DECT scans to evaluate the breast lesion stage and the presence of potential lung metastasis lesions or other potential lung lesions, which may be considered a selection bias. Finally, fewer participants with benign lesions were included, resulting in an imbalance in the sample size. It should be noted that most patients underwent DECT for breast cancer staging, while patients with benign breast tumors rarely underwent DECT. Therefore, few benign tumors were included in this study.

In conclusion, the performance of AdaBoost based on mpDECT was superior to that of the other machine learning models and conventional univariate analysis in differentiating between benign and malignant breast lesions. The combination of machine learning with mpDECT may provide valuable information for differential diagnoses to guide clinical treatment decisions. This is a key step in achieving precision medicine in breast cancer.

Acknowledgments

The authors thank all volunteers who participated in the study and the staff of the Department of Radiology, Chongqing University Cancer Hospital, Chongqing Cancer Institute, and Chongqing Cancer Hospital in Chongqing, China, for their selfless and valuable assistance. We acknowledge the support of Xiaoyue Zhang from Siemens scientific research.

Funding: This study has received funding from the National Natural Science Foundation of China (Grant No. 82071883), the combination projects of medicine and engineering of the Fundamental Research Funds for the Central Universities in 2019 (Project No. 2019CDYGYB008), the Chongqing key medical research project of a combination of science and medicine (Grant No. 2019ZDXM007), and the 2019 SKY Imaging Research Fund of the Chinese International Medical Foundation (Project No. Z-2014-07-1912-10).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/qims-21-39>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by the ethics committee of Chongqing University Cancer Hospital (No.: CZLS20200215-A) and individual consent for this retrospective analysis was waived. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Haynes B, Sarma A, Nangia-Makker P, Shekhar MP. Breast cancer complexity: implications of intratumoral heterogeneity in clinical management. *Cancer Metastasis Rev* 2017;36:547-55.
- Saslow D, Boetes C, Burke W, Harms S, Leach MO, Lehman CD, Morris E, Pisano E, Schnall M, Sener S, Smith RA, Warner E, Yaffe M, Andrews KS, Russell CA; American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin* 2007;57:75-89.
- Zhang X, Wang D, Liu Z, Wang Z, Li Q, Xu H, Zhang B, Liu T, Jin F. The diagnostic accuracy of magnetic resonance imaging in predicting pathologic complete response after neoadjuvant chemotherapy in patients with different molecular subtypes of breast cancer. *Quant Imaging Med Surg* 2020;10:197-210.
- Heywang-Köbrunner SH, Hacker A, Sedlacek S. Magnetic resonance imaging: the evolution of breast imaging. *Breast*

- 2013;22 Suppl 2:S77-82.
5. Medeiros LR, Duarte CS, Rosa DD, Edelweiss MI, Edelweiss M, Silva FR, Winnikow EP, Simões Pires PD, Rosa MI. Accuracy of magnetic resonance in suspicious breast lesions: a systematic quantitative review and meta-analysis. *Breast Cancer Res Treat* 2011;126:273-85.
 6. Simons D, Kachelriess M, Schlemmer HP. Recent developments of dual-energy CT in oncology. *Eur Radiol* 2014;24:930-9.
 7. Srinivasan A, Parker RA, Manjunathan A, Ibrahim M, Shah GV, Mukherji SK. Differentiation of benign and malignant neck pathologies: preliminary experience using spectral computed tomography. *J Comput Assist Tomogr* 2013;37:666-72.
 8. Tawfik AM, Razek AA, Kerl JM, Nour-Eldin NE, Bauer R, Vogl TJ. Comparison of dual-energy CT-derived iodine content and iodine overlay of normal, inflammatory and metastatic squamous cell carcinoma cervical lymph nodes. *Eur Radiol* 2014;24:574-80.
 9. Wang X, He Y, Fan Z, Wang T, Xie Y, Li J, Ouyang T. Effect of trastuzumab among HER2-positive breast cancer patients that achieved pathologic complete response after neoadjuvant chemotherapy. *Breast Care (Basel)* 2019;14:388-93.
 10. Volterrani L, Gentili F, Fausto A, Pelini V, Megha T, Sardanelli F, Mazzei MA. Dual-energy CT for locoregional staging of breast cancer: preliminary results. *AJR Am J Roentgenol* 2020;214:707-14.
 11. Wang X, Liu D, Zeng X, Jiang S, Li L, Yu T, Zhang J. Dual-energy CT quantitative parameters for the differentiation of benign from malignant lesions and the prediction of histopathological and molecular subtypes in breast cancer. *Quant Imaging Med Surg* 2021;11:1946-57.
 12. Metin Y, Metin NO, Özdemir O, Taşçı F, Kul S. The role of low keV virtual monochromatic imaging in increasing the conspicuity of primary breast cancer in dual-energy spectral thoracic CT examination for staging purposes. *Acta Radiol* 2020;61:168-74.
 13. Wang X, Liu D, Zeng X, Jiang S, Li L, Yu T, Zhang J. Dual-energy CT quantitative parameters for evaluating Immunohistochemical biomarkers of invasive breast cancer. *Cancer Imaging* 2021;21:4.
 14. Wei L, Yang Y, Nishikawa RM, Jiang Y. A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. *IEEE Trans Med Imaging* 2005;24:371-80.
 15. Antropova N, Abe H, Giger ML. Use of clinical MRI maximum intensity projections for improved breast lesion classification with deep convolutional neural networks. *J Med Imaging (Bellingham)* 2018;5:014503.
 16. Nakagawa M, Nakaura T, Namimoto T, Kitajima M, Uetani H, Tateishi M, Oda S, Utsunomiya D, Makino K, Nakamura H, Mukasa A, Hirai T, Yamashita Y. Machine learning based on multi-parametric magnetic resonance imaging to differentiate glioblastoma multiforme from primary cerebral nervous system lymphoma. *Eur J Radiol* 2018;108:147-54.
 17. Wan KW, Wong CH, Ip HF, Fan D, Yuen PL, Fong HY, Ying M. Evaluation of the performance of traditional machine learning algorithms, convolutional neural network and AutoML Vision in ultrasound breast lesions classification: a comparative study. *Quant Imaging Med Surg* 2021;11:1381-93.
 18. Park EK, Lee KS, Seo BK, Cho KR, Woo OH, Son GS, Lee HY, Chang YW. Machine learning approaches to radiogenomics of breast cancer using low-dose perfusion computed tomography: predicting prognostic biomarkers and molecular subtypes. *Sci Rep* 2019;9:17847.
 19. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, Dubsy P, Baltzer P, Clauser P, Kapetas P, Morris EA, Meyer-Baese A, Pinker K. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Invest Radiol* 2019;54:110-7.
 20. Zhang X, Zheng C, Yang Z, Cheng Z, Deng H, Chen M, Duan X, Mao J, Shen J. Axillary sentinel lymph nodes in breast cancer: quantitative evaluation at dual-energy CT. *Radiology* 2018;289:337-46.
 21. Li C, Chen J, Qin G. Partial Youden index and its inferences. *J Biopharm Stat* 2019;29:385-99.
 22. Darst BF, Malecki KC, Engelman CD. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. *BMC Genet* 2018;19:65.
 23. Cui Y, Gao SY, Wang ZL, Li XT, Sun YS, Tang L, Zhang XP. Which should be the routine cross-sectional reconstruction mode in spectral CT imaging: monochromatic or polychromatic? *Br J Radiol* 2012;85:e887-90.
 24. Deniffel D, Sauter A, Dangelmaier J, Fingerle A, Rummeny EJ, Pfeiffer D. Differentiating intrapulmonary metastases from different primary tumors via quantitative dual-energy CT based iodine concentration and conventional CT attenuation. *Eur J Radiol* 2019;111:6-13.
 25. Mileto A, Allen BC, Pietryga JA, Farjat AE, Zarzour

- JG, Bellini D, Ebner L, Morgan DE. Characterization of incidental renal mass with dual-energy CT: diagnostic accuracy of effective atomic number maps for discriminating nonenhancing cysts from enhancing masses. *AJR Am J Roentgenol* 2017;209:W221-30.
26. Apfaltrer P, Meyer M, Meier C, Henzler T, Barraza JM Jr, Dinter DJ, Hohenberger P, Schoepf UJ, Schoenberg SO, Fink C. Contrast-enhanced dual-energy CT of gastrointestinal stromal tumors: is iodine-related attenuation a potential indicator of tumor response? *Invest Radiol* 2012;47:65-70.
 27. Sinn BV, Weber KE, Schmitt WD, Fasching PA, Symmans WF, Blohmer JU, Karn T, Taube ET, Klauschen F, Marmé F, Schem C, Stickeler E, Ataseven B, Huober J, von Minckwitz G, Seliger B, Denkert C, Loibl S. Human leucocyte antigen class I in hormone receptor-positive, HER2-negative breast cancer: association with response and survival after neoadjuvant chemotherapy. *Breast Cancer Res* 2019;21:142.
 28. Li Y, Yang ZG, Chen TW, Chen HJ, Sun JY, Lu YR. Peripheral lung carcinoma: correlation of angiogenesis and first-pass perfusion parameters of 64-detector row CT. *Lung Cancer* 2008;61:44-53.
 29. Comstock CE, Gatsonis C, Newstead GM, Snyder BS, Gareen IF, Bergin JT, Rahbar H, Sung JS, Jacobs C, Harvey JA, Nicholson MH, Ward RC, Holt J, Prather A, Miller KD, Schnall MD, Kuhl CK. Comparison of abbreviated breast MRI vs digital breast tomosynthesis for breast cancer detection among women with dense breasts undergoing screening. *JAMA* 2020;323:746-56.
 30. Wildman-Tobriner B, Middleton MM, Moylan CA, Rossi S, Flores O, Chang ZA, Abdelmalek MF, Sirlin CB, Bashir MR. Association between magnetic resonance imaging-proton density fat fraction and liver histology features in patients with nonalcoholic fatty liver disease or nonalcoholic steatohepatitis. *Gastroenterology* 2018;155:1428-35.e2.
 31. Lei T, Sun H, Kang Y, Zhu F, Liu H, Zhou W, Wang Z, Li D, Li Y, Hou T. ADMET Evaluation in Drug Discovery. 18. Reliable prediction of chemical-induced urinary tract toxicity by boosting machine learning approaches. *Mol Pharm* 2017;14:3935-53.
 32. Qi Z, Meng F, Tian Y, Niu L, Shi Y, Zhang P. Adaboost-LLP: A boosting method for learning with label proportions. *IEEE Trans Neural Netw Learn Syst* 2018;29:3548-59.
 33. Bevers TB, Helvie M, Bonaccio E, Calhoun KE, Daly MB, Farrar WB, et al. Breast cancer screening and diagnosis, version 3.2018, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2018;16:1362-89.

Cite this article as: Lan X, Wang X, Qi J, Chen H, Zeng X, Shi J, Liu D, Shen H, Zhang J. Application of machine learning with multiparametric dual-energy computed tomography of the breast to differentiate between benign and malignant lesions. *Quant Imaging Med Surg* 2022;12(1):810-822. doi: 10.21037/qims-21-39

Machine learning algorithm

Logistic regression models the conditional probability using a logistic function, which depends on the odds, based on the values of the independent variables. Logistic regression tries to minimize the negative log-likelihood of conditional probability via optimization algorithms. It is designed to find cumulative logistic distribution by measuring the relationship between one dependent and one or more independent variables (34).

Boosting algorithms change the training data distribution iteratively using a base classifier to predict the hard-to-classify exemplars. Boosting assigns a weight to each training exemplar, and they can be changed at the end of each round of boosting adaptively (35).

SGD is a method for unconstrained optimization problems that try to learn linear scoring by assigning weights and intercept parameters to minimize the loss function (36).

LDA finds a linear combination of attributes, which separates two or more classes by determining a subspace of lower dimension of the original data. Statistical measures, such as variance and mean, are used to determine separability. LDA maximizes the projected class means by minimizing the class variance in that direction by fitting a Gaussian density to each class with the assumption that all classes share the same covariance matrix (37).

RF are based on the concept of decision trees. A decision tree represents patterns and structures in the input data with hierarchical and sequential nodes that form tree-like structures. A decision tree comprises (I) internal nodes, (II) branches, and (III) terminal nodes. RF constructs many decision trees from a dataset, combines the results from all the trees, and makes a prediction with a majority vote (bootstrap aggregation) to make predictions on classification or regression (19).

Linear SVM is a classifier that finds a hyperplane based on the maximal margin rule to separate the data into two classes. It can be applied to linearly separable data sets and nonlinearly separable data sets using nonlinear kernels. The nonlinear data have to be transformed into a new linear space from its original coordinate space. Thus, in a new coordinate space, the linear decision boundaries could separate the sample data (38). The specific parameter setting of machine learning is shown in [Figure S1](#).

In this study, the optimum ranking of the features based on their importance in the models was reported employing RFE. This method can select features by recursively considering smaller and smaller sets of features by training a classifier on the initial set of features and weights. Then, features with the absolute minimum weights are pruned from the current set features. By repeating this procedure, the desired optimum number of features with the maximum accuracy is identified.

References

34. Falconieri N, Van Calster B, Timmerman D, Wynants L. Developing risk models for multicenter data using standard logistic regression produced suboptimal predictions: a simulation study. *Biom J*. 2020 Jul;62:932-44.
35. Mayr A, Binder H, Gefeller O, Schmid M. The evolution of boosting algorithms. From machine learning to statistical modelling. *Methods Inf Med* 2014;53:419-27.
36. Ilboudo WEL, Kobayashi T, Sugimoto K. Robust Stochastic Gradient Descent With Student-t Distribution Based First-Order Momentum. *IEEE Trans Neural Netw Learn Syst* 2020. [Epub ahead of print]. doi: 10.1109/TNNLS.2020.3041755.
37. Dornaika F, Khoder A. Linear embedding by joint Robust Discriminant Analysis and Inter-class Sparsity. *Neural Netw* 2020;127:141-59.
38. Citak-Er F, Firat Z, Kovanlikaya I, Ture U, Ozturk-Isik E. Machine-learning in grading of gliomas based on multi-parametric magnetic resonance imaging at 3T. *Comput Biol Med* 2018;99:154-60.

```

1 LogisticRegression(penalty='l2', C=0.04, class_weight="balanced", solver='saga', dual=False,
2                   tol=0.1, fit_intercept=True, intercept_scaling=1, random_state=None,
3                   max_iter=1000, multi_class='auto', verbose=0, warm_start=False),
4 XGBClassifier(base_score=0.5, booster=None, colsample_bylevel=1,
5              colsample_bynode=1, colsample_bytree=1, gamma=0, gpu_id=-1,
6              importance_type='gain', interaction_constraints=None,
7              learning_rate=0.300000012, max_delta_step=0, max_depth=6,
8              min_child_weight=1, monotone_constraints=None,
9              objective='binary:logistic', missing=None,
10             n_estimators=100, n_jobs=0, num_parallel_tree=1,
11             random_state=0, reg_alpha=0,
12             reg_lambda=1, scale_pos_weight=1, subsample=1, tree_method=None,
13             validate_parameters=False, verbosity=None),
14 SGDClassifier(alpha=0.0001, average=False, class_weight=None, early_stopping=False,
15              epsilon=0.1, eta0=0.0, fit_intercept=True, l1_ratio=0.15, learning_rate='optimal',
16              loss='hinge', max_iter=1000, n_iter_no_change=5, n_jobs=None, penalty='l2',
17              power_t=0.5, random_state=None, shuffle=True,
18              tol=0.001, validation_fraction=0.1, verbose=0, warm_start=False),
19 LinearDiscriminantAnalysis(covariance_estimator=None, n_components=None, priors=None,
20                            shrinkage=None, solver='svd', store_covariance=False, tol=0.0001),
21 AdaBoostClassifier(base_estimator=None, n_estimators=50, learning_rate=1.0, algorithm='SAMME.R',
22                  random_state=None),
23 RandomForestClassifier(n_estimators=100, criterion='gini', max_depth=10, min_samples_split=2,
24                       min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto',
25                       max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None,
26                       bootstrap=True, oob_score=False, n_jobs=None, random_state=None, verbose=0,
27                       warm_start=False, class_weight=None, ccp_alpha=0.0, max_samples=None),
28 DecisionTreeClassifier(criterion='gini', splitter='best', max_depth=5, min_samples_split=2, min_samples_leaf=1,
29                       min_weight_fraction_leaf=0.0, max_features=None, random_state=None, max_leaf_nodes=None,
30                       min_impurity_decrease=0.0, min_impurity_split=None, class_weight=None,
31                       presort='deprecated', ccp_alpha=0.0),
32 SVC(C=1.0, kernel='linear', degree=3, gamma='scale', coef0=0.0, shrinking=True, probability=False,
33     tol=0.001, cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr',
34     break_ties=False, random_state=None)

```

Figure S1 Specific parameters of machine learning

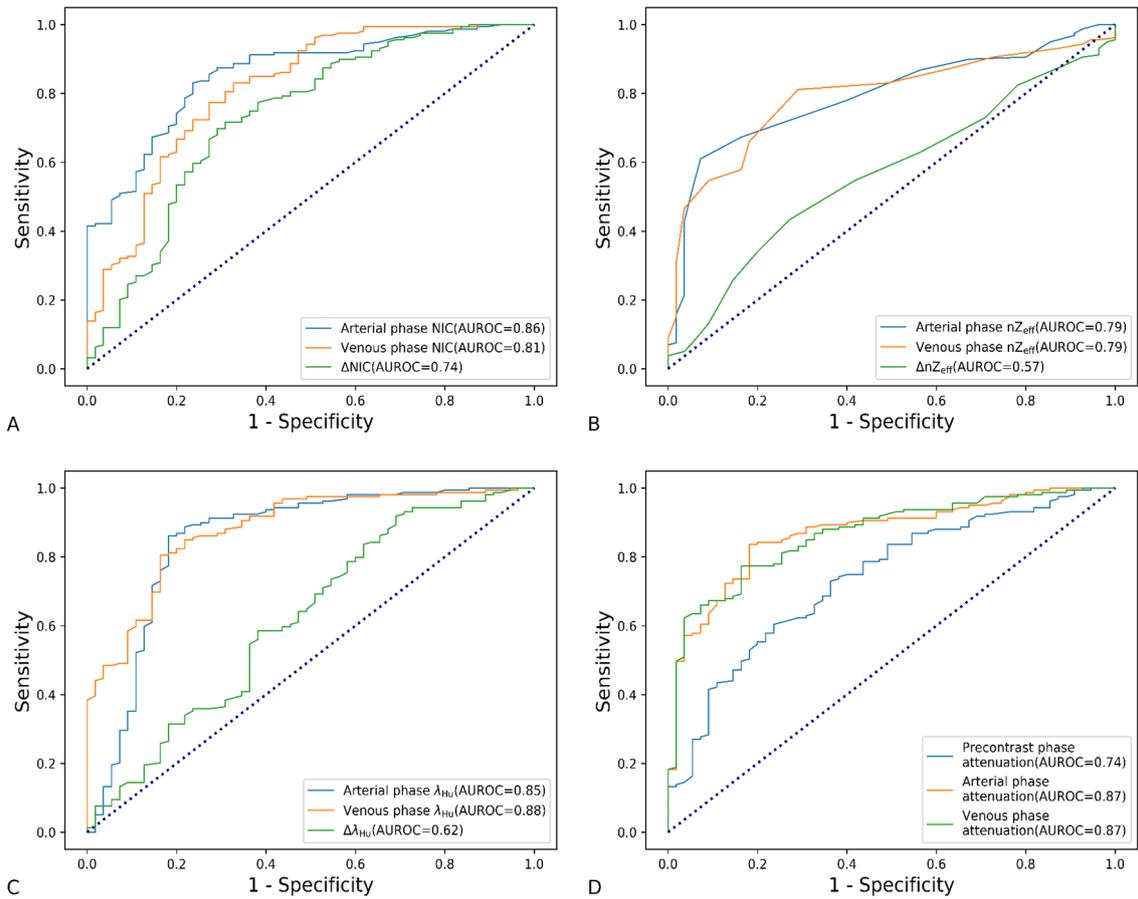
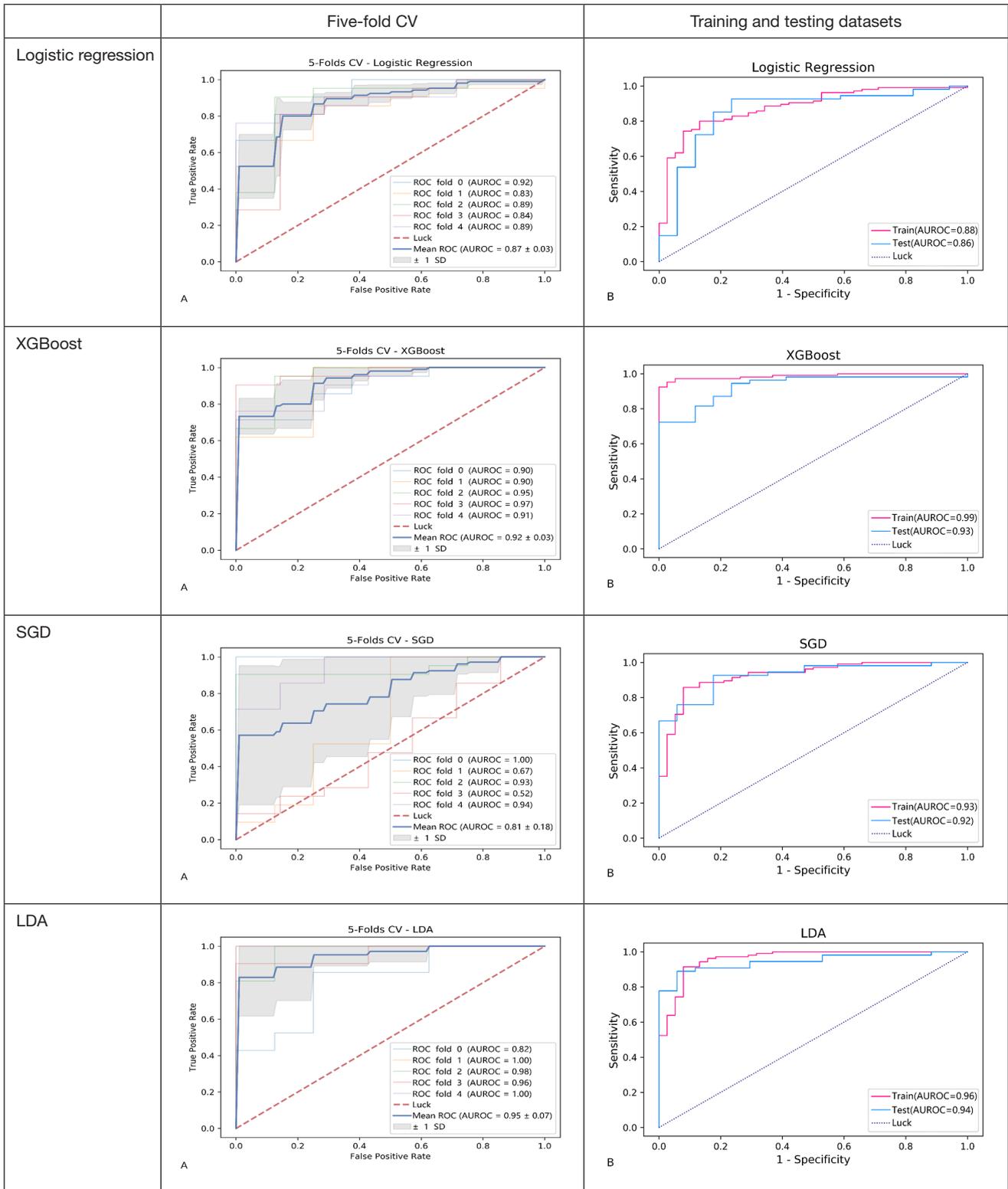


Figure S2 ROC curves of univariate logistic regression model using mpDECT to predict malignant and benign breast lesions. (A) The NIC; (B) the nZ_{eff} ; (C) the λ_{HU} ; and (D) attenuation. ROC, receiver operating characteristic; mpDECT, multiparametric dual-energy computed tomography; NIC, normalized iodine concentration; nZ_{eff} , normalized effective atomic number; λ_{HU} , slope of the spectral Hounsfield unit curve; AUROC, area under the receiver operating characteristic curve.



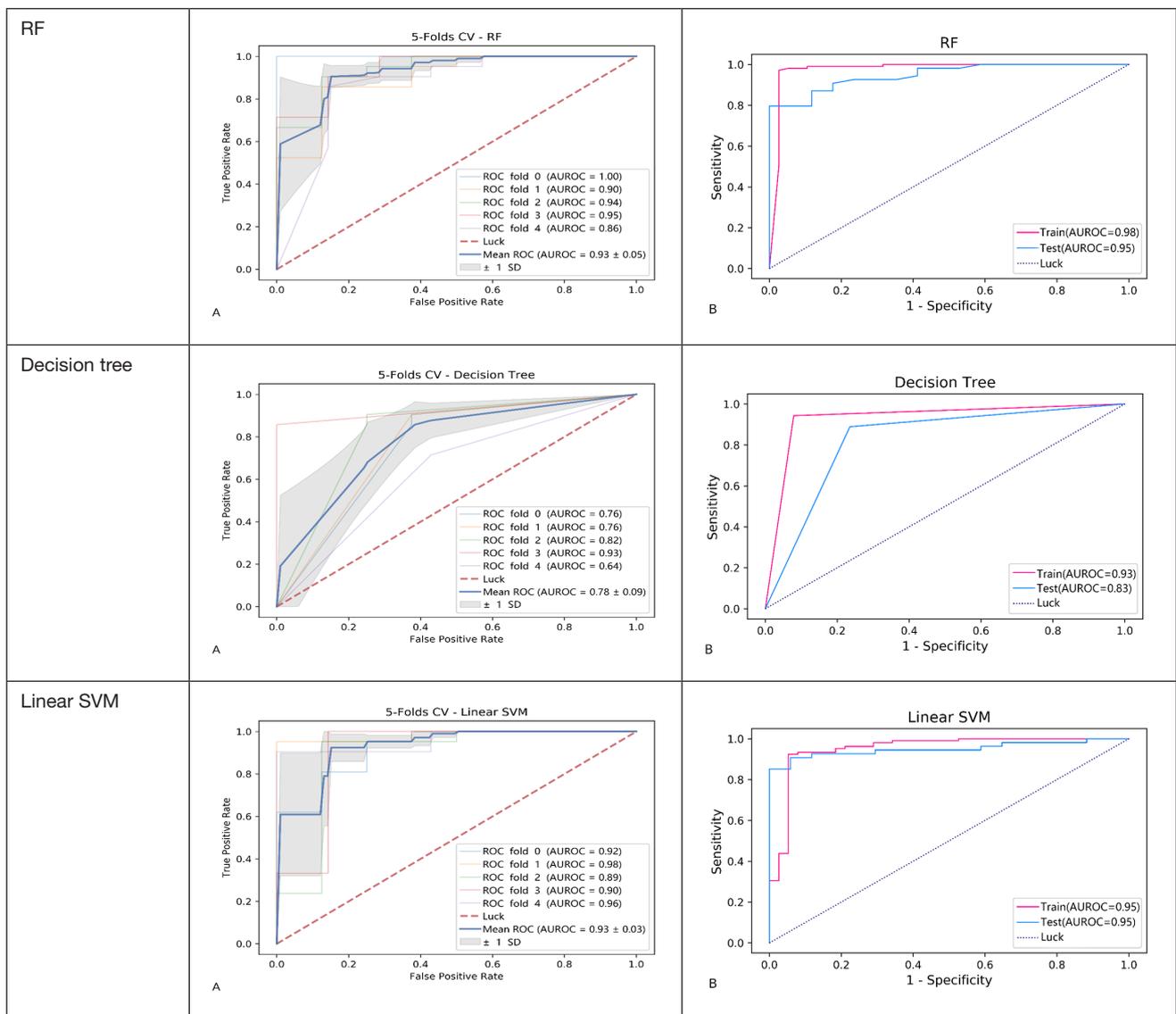


Figure S3 ROC curves of mpDECT model using logistic regression, XGBoost, SGD, LDA, RF, decision tree, and linear SVM classifier using five-fold CV in the prediction of malignant and benign breast lesions. (A) Five-fold CV; (B) training and testing dataset. ROC, receiver operating characteristic; mpDECT, multiparametric dual-energy computed tomography; XGBoost, extreme gradient boosting; SGD, stochastic gradient descent; LDA, linear discriminant analysis; AdaBoost, adaptive boosting; RF, random forest; SVM, support vector machine; CV, cross-validation; AUROC, area under the receiver operating characteristic curve.