



OPEN ACCESS

EDITED BY

Andreas Kanavos,
Ionian University, Greece

REVIEWED BY

Zuhal Hamd,
Princess Nourah bint Abdulrahman University,
Saudi Arabia
Andi Muh. Maulana,
Universitas Muhammadiyah Purwokerto,
Indonesia

*CORRESPONDENCE

Shanhu Hao

✉ haoshanhu3257@163.com

Guoxu Zhang

✉ zhangguoxu_502@163.com

RECEIVED 18 July 2025

ACCEPTED 30 September 2025

PUBLISHED 23 October 2025

CITATION

Lu G, Tian R, Yang W, Liu D, Chen W, Liang J, Peng Q, Hao S and Zhang G (2025) Intra-tumor and peritumoral radiomics and deep learning based on ultrasound for differentiating fibroadenoma and phyllodes tumor: a multicenter study. *Front. Oncol.* 15:1668793. doi: 10.3389/fonc.2025.1668793

COPYRIGHT

© 2025 Lu, Tian, Yang, Liu, Chen, Liang, Peng, Hao and Zhang. 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.

Intra-tumor and peritumoral radiomics and deep learning based on ultrasound for differentiating fibroadenoma and phyllodes tumor: a multicenter study

Guoxiu Lu¹, Ronghui Tian², Wei Yang³, Dongmei Liu⁴, Wenjing Chen⁵, Jingjing Liang¹, Qi Peng¹, Shanhu Hao^{1*} and Guoxu Zhang^{1*}

¹Department of Nuclear Medicine, General Hospital of Northern Theater Command, Shenyang, Liaoning, China, ²School of Software, Shenyang University of Technology, Shenyang, Liaoning, China,

³Department of Radiology, Cancer Hospital of China Medical University, Liaoning Cancer Hospital and Institute, Shenyang, Liaoning, China, ⁴Department of Ultrasound, Beijing Shijitan Hospital, Capital Medical University, Beijing, China, ⁵Department of Research and Development, United Imaging Intelligence (Beijing) Co., Ltd., Beijing, China

Purpose: To develop and validate an integrated intra-tumoral (ITR) and peritumoral (PTR) radiomics-deep learning model based on ultrasound (US) imaging for accurately differentiating fibroadenomas (FA) from phyllodes tumors (PT) and further classifying PT into benign, borderline, and malignant subtypes.

Methods: This multicenter retrospective study enrolled 300 patients (141 FA, 159 PT) from three institutions. US images were analyzed using manual segmentation of ITR and PTR (4mm, 8mm, 12mm, 16mm expansions). A total of 114 radiomics features were extracted per region using PyRadiomics. Five deep learning models (CNN, MLP, ViT, GAN, RNN) and six machine learning classifiers were evaluated. Optimal features were selected via LASSO and Boruta algorithms. Integrated models combining radiomics (ITR ± PTR) with clinical factors (diameter, Bi-RADS) were developed. Performance was assessed using AUC, accuracy, sensitivity, specificity, F1-score, and biopsy reduction rate. Internal validation used a 7:3 random split stratified by center and pathology. External validation was performed on a per-center hold-out basis.

Results: The combined model (ITR + 8mm PTR + clinical) achieved the highest performance for FA/PT differentiation (AUC: 0.960; accuracy: 96.0%; sensitivity: 96.0%; specificity: 94.5%). For PT subtyping (benign/borderline/malignant), the model attained an AUC of 0.874 (accuracy: 77.2%). The integrated model significantly reduced unnecessary biopsy rates by 11.7% overall (18.1% for PT cases). Peritumoral analysis (8mm PTR) contributed critically to model performance, likely capturing stromal interactions at the tumor periphery.

Conclusion: Integrating intra-tumoral, peritumoral (8mm), and clinical US radiomics features enables highly accurate non-invasive differentiation of FA

and PT and stratification of PT subtypes. This approach reduces diagnostic ambiguity in Bi-RADS 4 lesions and decreases unnecessary biopsies, demonstrating significant clinical utility for precision diagnosis of breast fibroepithelial tumors.

KEYWORDS

fibroadenoma, phyllodes tumor, intratumoral, peritumoral, radiomics, deep learning, ultrasound

1 Introduction

Accurate differential diagnosis of breast tumors remains a core clinical challenge. Radiomics, through high-throughput extraction of medical imaging features, has established a new paradigm for non-invasive diagnosis (1, 2). Our previous research developed a differentiation model between benign and malignant breast tumors using multimodal deep learning radiomics (including mammography, MRI, and ultrasound), addressing the fundamental classification issue (3). However, for common benign breast tumors such as fibroadenomas (FA) and potentially malignant phyllodes tumors (PT), there exist distinct subtypes with significant differences in biological behavior and prognosis (4, 5). Breast ultrasound plays a crucial role in differentiating between FA and PT (6).

Breast fibroepithelial lesions encompass a heterogeneous group of biphasic tumors, including FA and PT. PTs were first identified in 1838 and described as cystosarcomas phyllodes due to their leaf-like appearance (7, 8). These tumors account for 2.5% of all breast fibroepithelial lesions and represent less than 1% of all breast tumors (9, 10). Notably, PT exhibits higher incidence rates in Asian populations, particularly among young women. Based on histological features (stromal cellularity, atypia, mitotic activity, stromal overgrowth, and margin characteristics), PTs are classified into three subtypes: benign, borderline, and malignant (5, 11). PT typically recurs locally within 2–3 years after diagnosis, with total recurrence rates of 10–17% for benign, 14–25% for borderline, and 23–30% for malignant PTs (12, 13).

Therefore, extensive local resection is required for malignant PT. Metastasis in PT indicates poor prognosis, increased mortality risk, and reduced survival rates. In patients with benign or borderline PT, distant metastasis is extremely rare. The occurrence of distant metastasis in benign or borderline PT is likely due to insufficient sample size and misclassification. Thus, precise tumor grading, particularly accurate diagnosis of malignant PT, becomes crucial. Patients with malignant PT should consider systemic treatments such as chemotherapy or targeted therapy, while those with borderline PT have minimal or no metastatic risk, allowing safe surgical intervention (14). Accurate preoperative diagnosis of FA, PT cell tumors, and adenomas not only aids in formulating precise surgical plans and determining appropriate

tumor margins and axillary dissection but also helps avoid overtreatment or under-treatment (15). Currently, histological morphology remains the sole diagnostic criterion for malignant PT. However, since PT themselves represent a continuous disease spectrum with overlapping histological features, and given the numerous subjective evaluation parameters used in current grading systems, pathologists lack consensus in classifying malignant PT, which may not fully align with their biological behavior. Previous studies have shown that older age, rapid growth of lesions, and tumors larger than 3 cm are more likely to become borderline and malignant PT (16). Another study observed that there was about 60% inconsistency between biopsy pathology and resection pathology, which may be attributed to insufficient sample size and tumor heterogeneity (17). Therefore, the existing grading criteria for FA and PT are challenged, and accurate and reproducible grading is worth further exploration.

The characteristics of tumor heterogeneity are hemorrhagic areas, cystic changes, high cell density, necrosis and mucinous changes (18). In recent years, with the development of radiological examination techniques, including breast ultrasound, mammography and magnetic resonance imaging have been widely used in the diagnosis and treatment evaluation of FA and PT of the breast. However, traditional ultrasound imaging demonstrates limited potential in predicting FA and PT, and its classification. At present, tumor cells and tumor microenvironment (Tumor microenvironment, TME) are considered to be the key factors in cancer occurrence and development, as well as potential targets for treatment (19). Tumor cells coexist with a variety of cellular components in the tumor microenvironment, forming a more complex tumor immune microenvironment composed of immune infiltrating cells than normal healthy tissues, which has become a hot spot in the diagnosis and treatment of breast tumors (20).

Building on previous research, we integrate radiomics and deep learning to comprehensively evaluate tumor sites using ultrasound imaging for distinguishing these challenging entities. This study explores the deep information mining capabilities of ultrasound, integrating intra-tumor (ITR) and peritumoral (PTR) data to achieve precise differentiation between FA and PT and subtyping of PT. The novelty of this work lies in: 1) the multicenter design; 2) the combined analysis of ITR, PTR (across multiple expansions), and clinical features; 3) the direct comparison of radiomics and

deep learning approaches; and 4) the analysis of potential biopsy reduction.

2 Materials and methods

2.1 Patient population and study design

This retrospective, multicenter study was reviewed and approved by the Ethics Committee of the General Hospital of the Northern Theater Command (Approval No.: Y (2404)-030). The need for informed consent was waived. Data were extracted from ultrasound imaging and clinical pathology records of three medical centers (General Hospital of the Northern Theater Command, Liaoning Provincial Cancer Hospital, and The Fourth Affiliated Hospital of China Medical University) between January 2018 and May 2024.

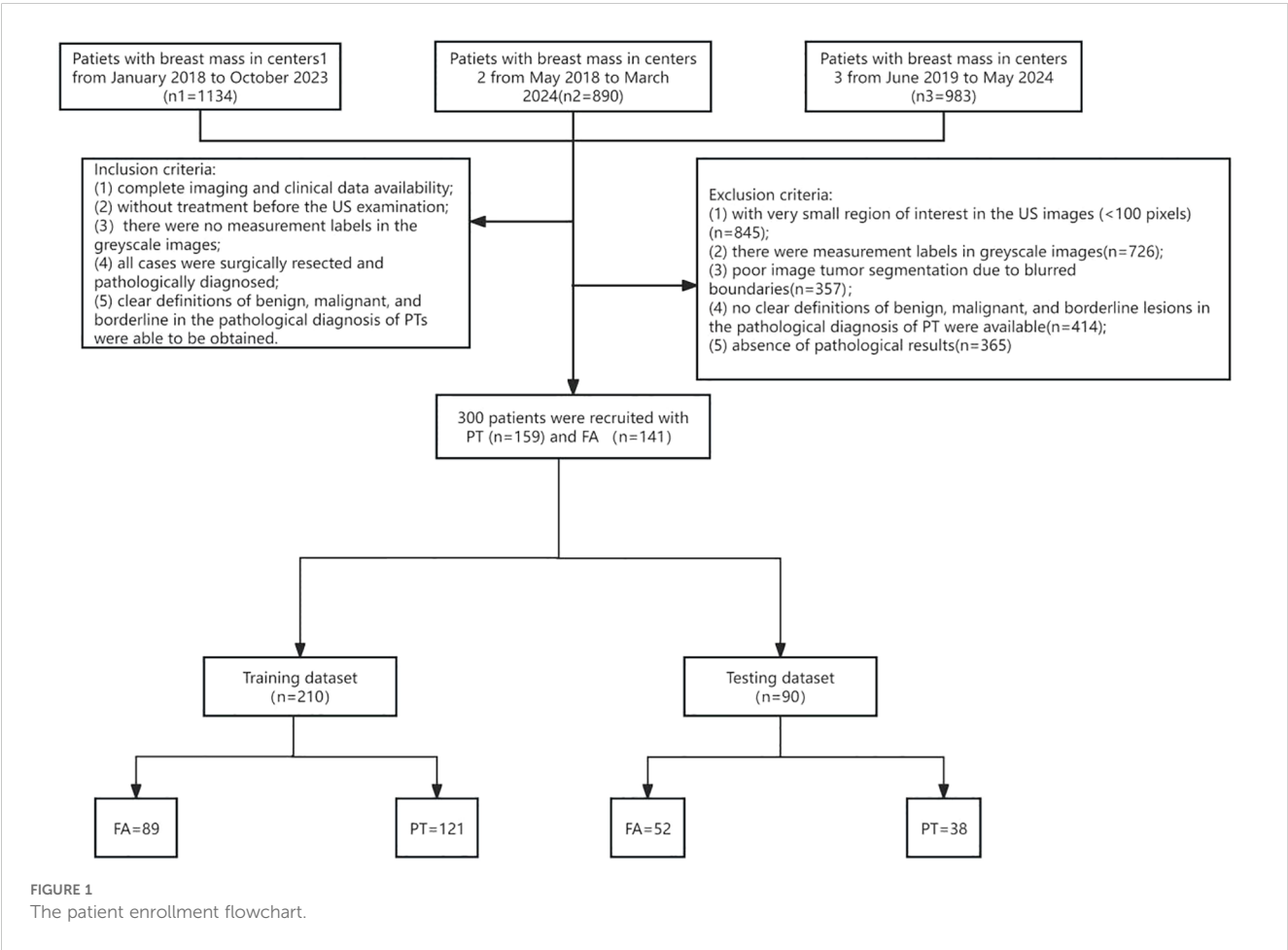
Inclusion criteria were: (1) Complete availability of imaging and clinical data; (2) No prior treatment before ultrasound examination; (3) No measurement markers in ultrasound grayscale images; (4) All cases underwent biopsy or surgical resection with pathological confirmation; (5) Pathological diagnosis of PT could clearly define subtypes: benign, borderline, and malignant. Exclusion criteria included:(1) Small ROI (<100 pixels); (2) Measurement markers in images; (3) Poor tumor segmentation due to blurred borders; (4) Unclear pathological diagnosis;

(5) Lack of definitive pathological results. Based on these criteria, 300 female breast tumor patients were enrolled: 141 with FA and 159 with PT. The patient enrollment flowchart is shown in Figure 1.

2.2 Ultrasound image acquisition and preprocessing

Breast ultrasound examinations were performed by specialists using systems(Mindray DC-7, GE Logiq E9, GE Voluson E8, Philips IU22, GE Logiq E20) equipped with 3–13 MHz linear array transducers. Examination followed ACR BI-RADS 5th edition guidelines (21). Scanning parameters (depth: 4–5 cm; gain: 10–25 dB; dynamic range: 70 dB; frame rate: 26 fps) were adjusted per patient based on habitus and lesion location, following standard clinical protocols at each center. Patients were positioned supine or lateral. The focal area was centered on the lesion. Images were stored in PACS. For each lesion, a specialist selected five representative images (longest axis, perpendicular, three other clear angles).

ROI delineation and image segmentation were performed manually using ITK-SNAP (v3.80) by three ultrasound specialists with >10 years of experience, guided by a senior physician (>20 years experience). Inter- and intra-observer variability was assessed using Dice Similarity Coefficient (DSC) and Intraclass Correlation



Coefficient (ICC); results indicated good to excellent agreement (DSC: 0.78-0.93, ICC: 0.832-0.949), and no statistically significant bias was found in feature extraction between observers ($p > 0.05$ for most features). The ITR was delineated layer-by-layer, avoiding necrosis. PTR was generated by expanding the ITR outward by 4mm, 8mm, 12mm, and 16mm using the Onekey AI platform (22) and MATLAB 2016b (23). The rationale for selecting these PTR thicknesses was based on prior literature suggesting stromal involvement within these distances (24, 25) and exploratory analysis showing varying predictive power across these ranges (see Sensitivity Analysis in Section 2.6). PTR extending beyond breast parenchyma was manually removed. Disagreements were resolved by consensus. Examples are shown in Figures 2–4.

2.3 Feature extraction

Using PyRadiomics (v3.0.1) (26), 114 radiomics features were extracted per ROI: 18 first-order statistics, 14 shape-based, 75 textural features (GLCM, GLRLM, GLSZM, GLDM, NGTDM) (21, 27). Feature extraction settings were default PyRadiomics parameters unless otherwise specified (binWidth: 25; force2D: True).

2.4 Feature selection and model construction

High-dimensional features underwent dimensionality reduction to prevent overfitting. Z-score normalization was applied. The Boruta

algorithm (22) was used for feature selection, comparing original features against shadow features. LASSO regression was also employed for optimal feature subset selection.

A two-stage classification model was developed: 1) Diagnostic network differentiating FA from PT using ITR, PTR, and combined features; 2) Grading network classifying PT into benign, borderline, malignant. The models were constructed using six machine learning classifiers: Random Forest (RF; $n_{\text{estimators}}=100$), Support Vector Machine (SVM; kernel='rbf', $C = 1.0$), XGBoost (XGB; max_depth=6, learning_rate=0.1), LightGBM (LGBM; num_leaves=31), Decision Tree (DT; max_depth=5), and Logistic Regression (LR; solver='liblinear'). Hyperparameters were optimized using 5-fold cross-validation GridSearch within the training set. Five deep learning models [CNN (28), MLP (29), ViT (17), GAN (30), RNN (31)] were also implemented using standard architectures and trained from scratch using Adam optimizer (learning_rate=0.001), batch size=16, for 100 epochs with early stopping.

2.5 Data split and validation strategy

To mitigate the risk of data leakage, patient-level splitting was strictly enforced. The dataset was first stratified by center and pathology label (FA, PT benign, PT borderline, PT malignant). Patients were then randomly split 7:3 into training (210 patients) and internal testing (90 patients) sets, ensuring no patient's images appeared in both sets. For external validation, each center's data was held out iteratively: models trained on data from two centers were tested on the third center's data. This per-center hold-out test provides

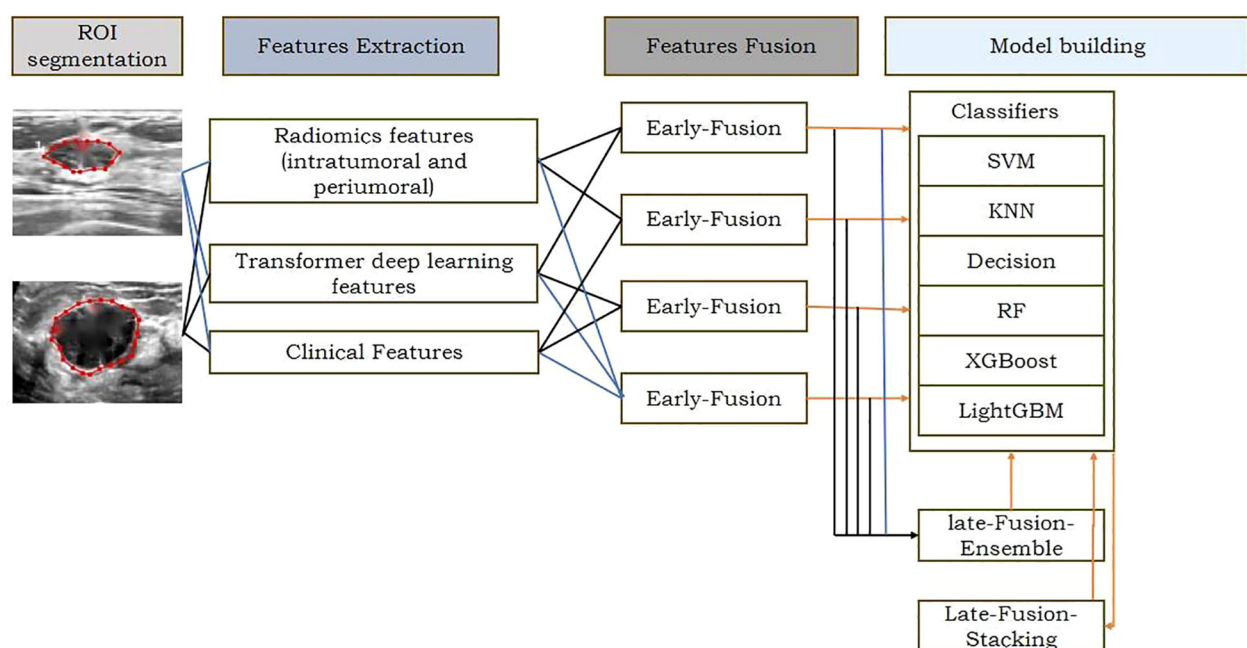


FIGURE 2

The workflow of ROI extraction and preprocessin DLR: conventional radiomic features were extracted from US datas. Feature selection and fusion techniques were applied to reduce dimensionality and integrate complementary information. Classification modeling was done using 6 machine learning algorithms.

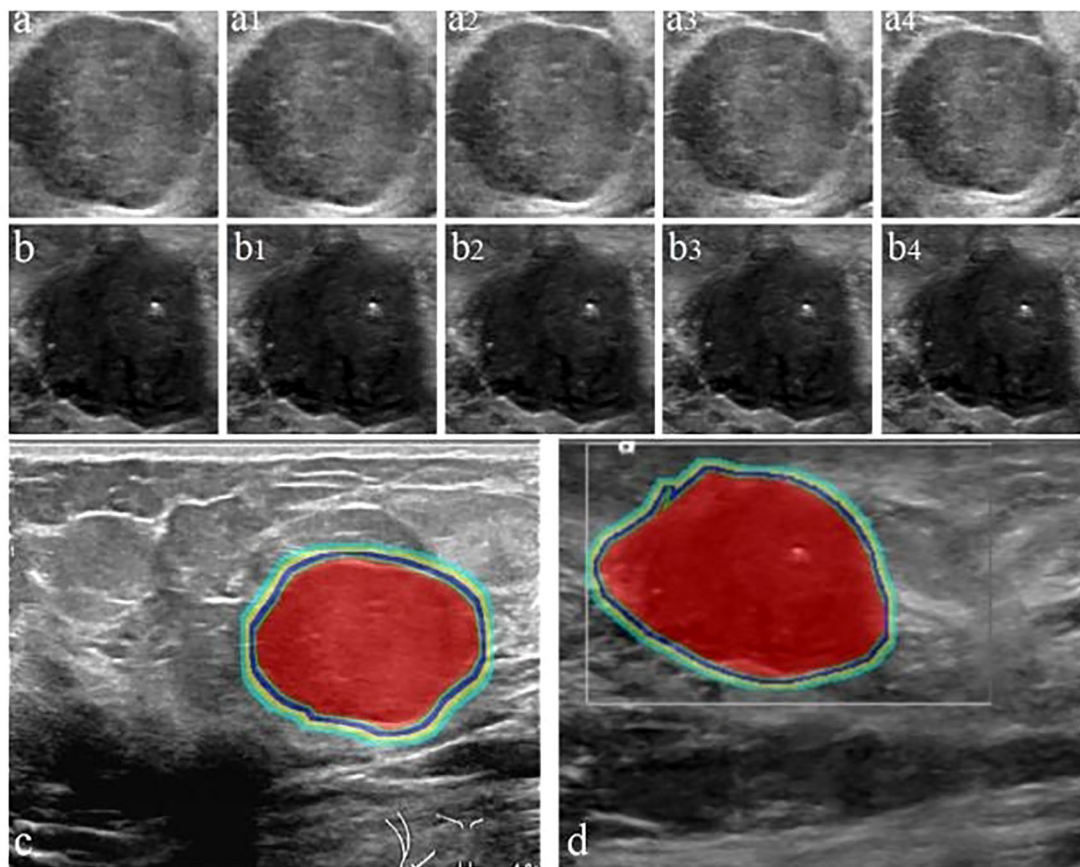


FIGURE 3

Workflow for delineating ITR and PTR in FA and PT patients under different ROI segmentation schemes: FA (a) and PT (b); red areas represent ITR, yellow-green areas represent 4 mm, 8 mm, 12 mm, and 16 mm PTR in FA (c) and PT (d). ITR: intratumoral area; PTR: peritumoral area.

a robust estimate of generalizability. The sample size of 300 patients was based on a pragmatic sample availability from the collaborating centers over the study period. A *post-hoc* power analysis based on the observed AUC (0.96) for the primary outcome (FA vs. PT) indicated sufficient power (>0.95) at $\alpha=0.05$. Performance was assessed via AUC, accuracy, sensitivity, specificity, F1-score. Calibration was evaluated using Hosmer-Lemeshow test and calibration curves. Decision curve analysis (DCA) assessed clinical utility. Five-fold cross-validation was performed on the training set for hyperparameter tuning and internal performance estimation.

2.6 Statistical analysis and sensitivity analysis

Statistical analysis used R (v4.3.1) and SPSS (v26.0). Non-normal continuous data summarized as median (IQR); compared using Mann-Whitney U/Kruskal-Wallis tests. Categorical variables as n (%); compared using Chi-square/Fisher's exact test. Model performance metrics reported with 95% confidence intervals (CI) calculated via bootstrap (2000 repetitions). AUC comparisons used DeLong test. Inter-observer agreement used ICC and Cohen's Kappa. $P<0.05$ significant. A sensitivity analysis was performed for the PTR thickness

parameter. Models were built and evaluated using only ITR, and ITR combined with each PTR thickness (4, 8, 12, 16 mm). The 8mm PTR expansion yielded the highest average AUC across classifiers for the FA vs. PT task (See [Supplementary Table S1](#)), justifying its selection as the optimal PTR width for the combined model.

3 Results

3.1 Clinical data characteristics

Among the 300 breast mass patients enrolled in the study, 141 were ultimately diagnosed with FA and 159 with PT through postoperative pathological examination. A randomized 7:3 split was used to divide patients into a training set (210 cases) and a validation set (90 cases). No significant differences were observed between the training and validation sets in terms of age, lesion diameter, location, menopausal status, clinical symptoms, growth rate, hardness, mobility, margins, echogenicity, blood flow, or BI-RADS classification (all $P>0.05$). Each patient had a single lesion in one breast and underwent surgical treatment: 85 patients (28.3%) received breast biopsy, 110 patients (36.7%) underwent local excision, 73 patients (24.3%) had wide excision, and 32 patients

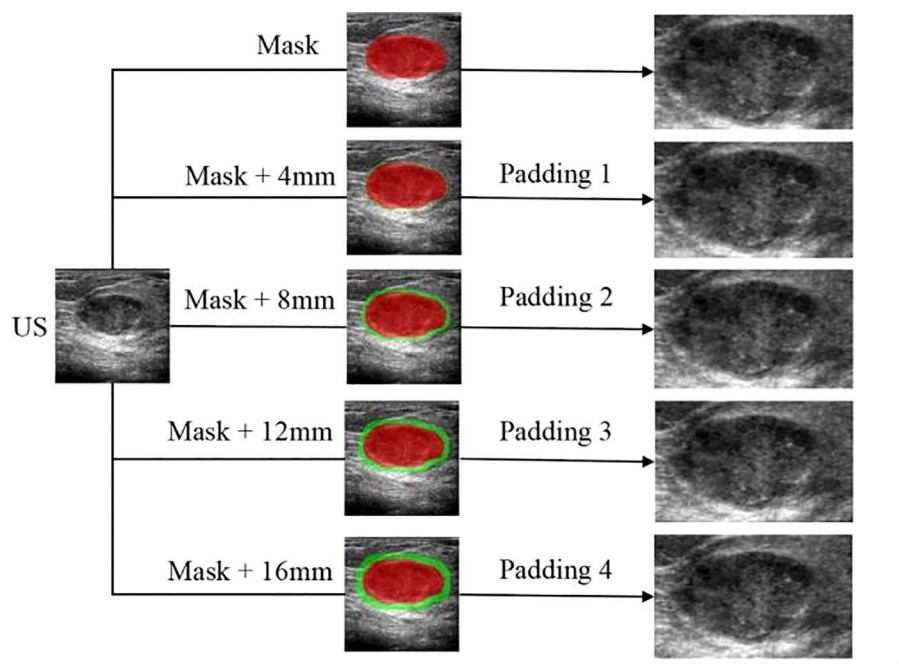


FIGURE 4

The workflow of the radiomics analysis of ITR and PTR: The ROI delineation of the tumor was increased by 4mm layer by layer to 16mm; The peritumoral lesion image gradually increased by 4mm to 16mm.

(10.6%) underwent radical mastectomy. Detailed demographic data and lesion characteristics are summarized in Table 1.

3.2 Feature consistency and selection

Dice coefficients for ROI segmentation were 0.78-0.92, 0.80-0.93, 0.84-0.92 for the three physicians. ICCs were 0.832-0.949 (intra-observer) and 0.742-0.925 (inter-observer), indicating good reproducibility. LASSO selected 5 optimal features from ITR and 10 from 8mm PTR (1 first-order, 9 higher-order).

3.3 Predictive performance of the imaging radiomics model

Performance metrics for Models 1–7 are detailed in Table 2. The combined Model 7 (ITR + 8mm PTR + Clinical) performed best for FA/PT differentiation (AUC: 0.960, Accuracy: 96.0%, Sensitivity: 96.0%, Specificity: 94.5%). For PT subtyping, Model 7 (Light GBM classifier) achieved an AUC of 0.874 (95% CI: 0.798-0.950), Accuracy: 77.2%.

3.4 Predictive performance of deep learning models

Performance of DL models is shown in Table 3. GAN performed best for FA/PT (AUC: 0.976). MLP performed best for

PT subtyping (AUC: 0.950). The relatively lower performance of CNN and ViT compared to GAN might be attributed to the GAN's potential to learn more robust feature representations from limited data through its adversarial training paradigm, whereas standard CNNs and ViTs may require larger datasets to achieve optimal performance in this specific task.

3.5 Subgroup analysis based on BI-RADS

To evaluate the impact of BI-RADS classification, patients were divided into three subgroups: Grade 3 (n=86, 42 PT cases), Grade 4 (including Grade 4a [76 cases], Grade 4b [68 cases], and Grade 4c [47 cases]), and Grade 5 (n=23, 14 PT cases). In all three subgroups (n=191, 117 PT cases and n=23, 14 PT cases), Model 7 (ITR + PTR8 + clinical) showed high diagnostic accuracy across BI-RADS subgroups: Grade 3: 84.0%, Grade 4: 66.0%, Grade 5: 88.0%. Specificity was highest for Grade 5 lesions (96.6%). Notably, in the Grade 4 lesion subgroup, Model 7 showed enhanced specificity when incorporating intratumoral and peritumoral (8mm) clinical features. This highlights the potential of integrating tumor-intrusion, peritumoral, and clinical characteristics for addressing diagnostic challenges in BI-RADS Grade 4 lesions.

3.6 Biopsy reduction analysis

Model 7 achieved an overall biopsy reduction rate of 11.7% (10/85), with an 18.1% (6/33) reduction for PT cases. Confidence

TABLE 1 The baseline information.

Items	Training set (n=210)				Test set (n=90)			
	FA(n=89)	PT(n=121)	Value	P	FA(n=38)	PT(n=52)	Value	P
Case	89(42.4%)	121(57.6%)			38(42.2%)	52(57.8%)	0.001	0.980
Images	280	395			123	165	0.011	0.724
Age	48.03+1.29	50.20+1.77	0.829	0.408	57.16+2.47	51.13+2.44	1.696	0.093
Diameter	3.22+0.14	3.52+0.13	1.592	0.112	3.75+0.24	3.12+0.21	1.964	0.053
Position*							6.027	0.925
Centre	3(3.4%)	3(2.5%)			1(2.6%)	1(1.9%)		
UIQ	21(23.6%)	32(26.4%)			10(26.3%)	16(30.8%)		
LIQ	24(27.0%)	33(27.3%)			6(15.8%)	12(23.1%)		
LOQ	22(24.7%)	26(21.5%)			11(28.9%)	8(15.4%)		
UOQ	19(21.3%)	27(22.3%)			10(26.3%)	15(28.8%)		
Menopause							1.364	0.714
Yes	51(57.3%)	63(52.1%)			18(47.4%)	26(50.0%)		
No	38(42.7%)	58(47.9%)			20(52.6%)	26(50.0%)		
Clinical symptom							1.946	0.583
Positive	43(48.3%)	56(46.3%)			14(36.8%)	21(40.4%)		
Negative	46(51.7%)	65(53.7%)			24(63.2%)	31(59.6%)		
Growth							1.713	0.634
Fast	40(44.9%)	64(52.9%)			17(44.7%)	24(46.2%)		
Slow	49(55.1%)	57(47.1%)			21(55.3%)	28(53.8%)		
Hardness							12.187	0.051
Soft	42(47.2%)	38(31.4%)			13(34.2%)	20(38.5%)		
Middle	25(28.1%)	36(29.8%)			17(44.7%)	20(38.5%)		
Hard	22(24.7%)	47(38.8%)			8(21.1%)	12(23.0%)		
Activity							0.675	0.879
Good	50(56.2%)	63(52.1%)			22(57.9%)	27(51.9%)		
Bad	39(43.8%)	58(47.9%)			16(42.1%)	25(48.1%)		
Border							3.280	0.350
Clear	64(71.9%)	84(69.4%)			26(68.4%)	30(57.7%)		
Irregular	25(28.1%)	37(30.6%)			12(31.6%)	22(42.3%)		
Echo							2.944	0.400
Even	45(50.6%)	75(62.0%)			23(60.5%)	29(55.8%)		
Inhomogeneous	44(49.4%)	46(38.0%)			15(39.5%)	23(44.2%)		
Blood flow							2.459	0.483
Yes	42(47.2%)	54(44.6%)			18(47.4%)	18(34.6%)		
No	47(52.8%)	67(55.4%)			20(52.6%)	34(65.4%)		
BiRADS							11.967	0.448
3	36(40.4%)	32(26.4%)			8(21.1%)	10(19.2%)		

(Continued)

TABLE 1 Continued

Items	Training set (n=210)				Test set (n=90)			
	FA(n=89)	PT(n=121)	Value	P	FA(n=38)	PT(n=52)	Value	P
4a	22(24.7%)	29(24.0%)			9(23.7%)	16(30.8%)		
4b	15(16.9%)	31(25.6%)			10(26.3%)	12(23.1%)		
4c	10(11.2%)	19(15.7%)			8(21.1%)	10(19.2%)		
5	6(6.7%)	10(8.3%)			3(7.9%)	4(7.7%)		

*FA, Fibroadenoma; PT, Phyllodes Tumor; UOQ, upper outer quadrant; UIQ, upper inner quadrant; LIQ, lower inner quadrant; LOQ, lower outer quadrant.

intervals (95%) for the reduction rates were calculated using the Clopper-Pearson exact method: Overall: 11.7% (5.7%-20.6%); PT: 18.1% (7.0%-35.5%). Results are in [Table 4](#).

4 Discussion

Compared with traditional visual-based image evaluation methods, radiological features derived from imaging can uncover additional latent characteristics. This radiomics-based approach demonstrated potential in distinguishing between FA and PT. Some findings revealed that the optimal combination of imaging radiomics features included high width-to-height ratio, edge blurriness, machine learning, energy, gray entropy, and intramural calcification could achieve good performance. In this study, we developed a model using radiomics features extracted from 114 breast lesions within each ROI through grayscale co-occurrence matrix (GLCM), grayscale run-length matrix (GLRLM),

grayscale size zone matrix (GLSZM), grayscale-dependent matrix (GLDM), and neighborhood grayscale difference matrix (NGTDM). The diagnostic efficacy metrics for FA and PT showed accuracy, AUC, sensitivity, specificity, precision, recall, and F1 values of 84.8%, 0.935, 86.7%, 86.2%, 91.3%, and 84.0% respectively. Our model demonstrated slightly higher accuracy than some studies ([17](#), [30](#)), likely due to the distinct radiomics feature extraction. Domestic research indicates that AI-assisted diagnosis through radiomics and deep learning can identify subtle morphological and textural differences between PT and FA, thereby enhancing diagnostic efficacy.

Previous radiomics analyses primarily focused on visual tumor boundaries, but recent studies have also emphasized peritumoral regional information. Researchers propose that the peritumoral area may be the earliest site of tissue lesion development and plays a critical role in determining tumor progression and treatment response ([32](#)). However, inflammatory tissues (characterizing the transition between tumors and normal parenchyma) introduce

TABLE 2 Results of radiomic classification utilizing intratumoral and peritumoral information of US imaging.

Model	Accuracy	AUC	Sensitivity	Specificity	Accuracy	Recall	F1	Classifier
1(ITR)	0.848	0.935	0.867	0.862	0.913	0.840	0.875	XG Boost
2(PTR4)	0.550	0.573	0.467	0.800	0.875	0.467	0.609	Light GBM
3(PTR8)	0.750	0.747	0.733	0.800	0.917	0.733	0.815	RF
4(PTR12)	0.450	0.400	0.333	0.800	0.833	0.333	0.476	SVM
5(PTR16)	0.550	0.360	0.600	0.400	0.750	0.600	0.667	RF
6(ITR+PTR8)	0.894	0.958	0.906	0.931	0.945	0.932	0.927	XGB
7(ITR+PTR8+clinical)	0.960	0.960	0.960	0.945	0.971	0.960	0.965	LGBM

*ITR, Intratumoral Region; PTR, Peritumoral Region; XGB, eXtreme Gradient Boosting; LGBM, Light Gradient Boosting Machine; RF, Random Forest; SVM, Support Vector Machine

TABLE 3 Results of deep learning information of US imaging.

Model	Accuracy	AUC	Sensitivity	Specificity	Accuracy	Recall	F1
CNN	0.531	0.498	0.421	0.692	0.667	0.421	0.516
MLP	0.746	0.944	0.630	0.577	0.577	0.789	0.667
ViT	0.625	0.401	0.947	0.154	0.621	0.947	0.750
GAN	0.896	0.976	0.870	0.952	0.976	0.870	0.920
RNN	0.750	0.733	0.733	0.800	0.917	0.733	0.815

*CNN, Convolutional Neural Network; MLP, Multilayer Perceptron; ViT, Vision Transformer; GAN, Generative Adversarial Network; RNN, Recurrent Neural Network.

TABLE 4 Performance of different models in reducing the rate of lesion biopsy.

Model	Biopsy reduction rate	FA	PT	Values	P value
1(TR)	5.8% (5/85)	5.7% (3/52)	6.1% (2/33)	0.035	0.852
6(ITR+PTR8)	8.2%(7/85)	7.7% (4/52)	9.0% (3/33)	0.004	0.949
7(ITR+PTR8+clinical)	11.7%(5.7%-20.6%)	7.7% (5.5%-14.8%)	18.1%(7.0%-35.5%)	1.191	0.275

certain misleading factors into radiographic results. Therefore, clinical visual analysis alone often fails to detect peritumoral regions, whose extent may influence tumor diversity (33). Given the varying complexities of different diseases and pathological issues, standardized end-to-end solutions are typically unavailable for comprehensive analysis. Instead, parameterized approaches tailored to specific tumor locations and imaging modalities are often required. Zhang (34) developed a linear discriminant analysis (LDA) model integrating three radiomics features (from ITR, 5mm PTR, and ITR + 10mm PTR) with two clinical factors (age and BI-RADS classification), demonstrating strong predictive power in both internal and external test datasets. This model, which combines intratumoral and peritumoral radiomics features with clinical factors, successfully predicted malignant BI-RADS 4 lesions in contrast-enhanced mammography using AUC values of 0.907 and 0.904 respectively. Our study demonstrates that integrating ITR, PTR (8mm), and clinical features achieves excellent performance in differentiating FA from PT and grading PT subtypes using US radiomics. The optimal 8mm PTR likely captures critical stromal alterations and tumor-host interactions at the invasive edge, a known hotspot for biological activity in breast tumors (35, 36).

In our model 7, which integrates intratumoral +8mmPTR with clinical features (diameter and BI-RADS classification), we observed slightly lower sensitivity and accuracy compared to the previous model in FA and PT cases classified as BI-RADS 4. However, it demonstrated higher specificity. These differences may stem from subtle variations in the microscopic structures of FA and PT within the selected multicenter ultrasound imaging data. Huang (37) emphasized not only clinically-derived models but also the importance of comprehensive radiomics analysis for accurate breast nodule characterization. Our findings reveal that the combined model integrating ITR + 8mm PTR+clinical features achieves optimal diagnostic precision. These results demonstrate that ultrasound-based intratumoral/peritumoral features combined with clinical characteristics exhibit superior diagnostic efficacy in both FA/PT classification models and pathological subtypes grading models for PT, thereby enhancing diagnostic accuracy and supporting clinical decision-making. Our model significantly reduced the potential need for biopsies, especially for PT lesions. This aligns with the goal of precision medicine to minimize invasive procedures (32). And potential clinical integration could involve PACS-integrated software for automatic feature extraction and model inference, providing real-time decision support during ultrasound examination and potentially reducing workflow interruptions.

This study also has some limitations. A key limitation is the retrospective design and class imbalance, particularly for borderline and malignant PT subtypes, which may introduce selection bias and affect model generalizability. Future prospective studies with larger, balanced cohorts are needed. Furthermore, this study utilized only ultrasound. While US is crucial, incorporating multimodal imaging (mammography, MRI) in future work could potentially enhance performance further.

5 Conclusion

This study established ultrasound-based intra-tumoral, peritumoral, and clinical radiomics features. Diagnostic efficacy for FA and PT was first evaluated. Building on this foundation, further classified PT into benign, borderline, and malignant subtypes, and analyzed performance across different BI-RADS grades, and identified ITR and PTR characteristics associated with reduced biopsy rates. In conclusion, the proposed US-based radiomics model integrating intra-tumoral, peritumoral (8mm), and clinical features serves as an effective non-invasive tool for differentiating FA from PT and classifying PT subtypes. It shows particular value in managing BI-RADS 4 lesions and reducing unnecessary biopsies. Future work should focus on large-scale, prospective, multicenter validation and exploration of multimodal integration.

Data availability statement

The datas in this article would be apply to the authors. Requests to access the datasets should be directed to guoxiu_16@sina.com.

Ethics statement

The studies involving human participants were reviewed and approved by the General Hospital of Northern Theater Command Ethics (No. Y(2404)-030). The need for informed consent was waived for the use of the imaging data in this study. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

GL: Formal analysis, Methodology, Validation, Data curation, Supervision, Project administration, Conceptualization, Software, Investigation, Writing – original draft, Resources, Funding acquisition, Visualization, Writing – review & editing. RT: Writing – review & editing, Validation, Methodology, Formal analysis, Writing – original draft, Data curation. WY: Methodology, Writing – original draft, Writing – review & editing, Resources, Data curation. DL: Supervision, Writing – review & editing, Writing – original draft, Validation. WC: Writing – review & editing, Data curation. JL: Writing – review & editing, Data curation. QP: Writing – review & editing, Data curation. SH: Validation, Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Supervision, Project administration, Data curation, Software, Investigation, Conceptualization, Resources, Visualization, Writing – original draft. GZ: Resources, Funding acquisition, Writing – review & editing, Formal analysis, Software, Visualization, Writing – original draft, Methodology, Data curation, Conceptualization, Project administration, Validation, Supervision, Investigation.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was partially supported by grants from The Research Program of General hospital of Northern Theater Command, New Grant Number: ZZKY2024001; The Applied Basic Research Program of Liaoning Province (2022JH2/101500011); Livelihood Science and Technology Plan Joint Plan Project of Liaoning Province (2021JH2/10300098). Science and Technology Research Program of China Railway Corporation (J2024Z605). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. Qi YJ, Su GH, You C, Zhang X, Xiao Y, Jiang YZ, et al. Radiomics in breast cancer: Current advances and future directions. *Cell Rep Med*. (2024) 5:101719. doi: 10.1016/j.xcrm.2024.101719
2. Tagliafico AS, Piana M, Schenone D, Lai R, Massone AM, Houssami N, et al. Overview of radiomics in breast cancer diagnosis and prognostication. *Breast (Edinburgh Scotland)*. (2020) 49:74–80. doi: 10.1016/j.breast.2019.10.018
3. Lu G, Tian R, Yang W, Liu R, Liu D, Xiang Z, et al. Deep learning radiomics based on multimodal imaging for distinguishing benign and Malignant breast tumours. *Front Med*. (2024) 11:1402967. doi: 10.3389/fmed.2024.1402967
4. Chen Z, Zhang Y, Li W, Gao C, Huang F, Cheng L, et al. Single cell profiling of female breast fibroadenoma reveals distinct epithelial cell compositions and therapeutic targets. *Nat Commun*. (2023) 14:3469. doi: 10.1038/s41467-023-39059-3
5. Bogach J, Shakeel S, Wright FC, Hong NJL. Phyllodes tumors: A scoping review of the literature. *Ann Surg Oncol*. (2022) 29:446–59. doi: 10.1245/s10434-021-10468-2
6. Park HL, Pyo YC, Kim KY, Park JS, Shin JE, Kim HR, et al. Recurrence rates and characteristics of phyllodes tumors diagnosed by ultrasound-guided vacuum-assisted breast biopsy (VABB). *Anticancer Res*. (2018) 38:5481–7. doi: 10.21873/anticancer.12881
7. Seow DYB, Tay TKY, Tan PH. Fibroepithelial lesions of the breast: A review of recurring diagnostic issues. *Semin Diagn Pathol*. (2022) 39:333–43. doi: 10.1053/j.semdp.2022.04.001
8. Lissidini G, Mulè A, Santoro A, Papa G, Nicosia L, Cassano E, et al. Malignant phyllodes tumor of the breast: a systematic review. *Pathologica*. (2022) 114:111–20. doi: 10.32074/1591-951X-754
9. Saxena P, Lalchandani A, Dausage C. Recurrent phyllodes tumour of breast infiltrating the latissimus dorsi reconstruction flap. *BMJ Case Rep*. (2020) 13(12): e238306. doi: 10.1136/bcr-2020-238306
10. Jang JH, Choi MY, Lee SK, Kim S, Kim J, Lee J, et al. Clinicopathologic risk factors for the local recurrence of phyllodes tumors of the breast. *Ann Surg Oncol*. (2012) 19:2612–7. doi: 10.1245/s10434-012-2307-5
11. Rana C, Kamal N, Mishra P, Singh A, Ramakant P, Mishra A, et al. Cellular fibroadenoma versus phyllodes tumors: A pre-operative diagnostic approach based on radiological and cytological features. *Diagn cytopathology*. (2022) 50:375–85. doi: 10.1002/dc.24965
12. Tan BY, Tan PH. A diagnostic approach to fibroepithelial breast lesions. *Surg Pathol Clinics*. (2018) 11:17–42. doi: 10.1016/j.path.2017.09.003
13. Choi J, Koo JS. Comparative study of histological features between core needle biopsy and surgical excision in phyllodes tumor. *Pathol Int*. (2012) 62:120–6. doi: 10.1111/j.1440-1827.2011.02761.x
14. Choi N, Kim K, Shin KH, Kim Y, Moon HG, Park W, et al. Malignant and borderline phyllodes tumors of the breast: a multicenter study of 362 patients (KROG 16-08). *Breast Cancer Res Treat*. (2018) 171:335–44. doi: 10.1007/s10549-018-4838-3

Conflict of interest

Author WC was employed by United Imaging Intelligence Beijing Co., Ltd.

The remaining 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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2025.1668793/full#supplementary-material>

15. Kim JH, Ko ES, Lim Y, Lee KS, Han BK, Ko EY, et al. Breast cancer heterogeneity: MR imaging texture analysis and survival outcomes. *Radiology*. (2017) 282:665–75. doi: 10.1148/radiol.2016160261
16. Cui WJ, Wang C, Jia L, Ren S, Duan SF, Cui C, et al. Differentiation between G1 and G2/G3 phyllodes tumors of breast using mammography and mammographic texture analysis. *Front Oncol*. (2019) 9:433. doi: 10.3389/fonc.2019.00433
17. Sim Y, Lee SE, Kim EK, Kim S. A radiomics approach for the classification of fibroepithelial lesions on breast ultrasonography. *Ultrasound Med Biol*. (2020) 46:1133–41. doi: 10.1016/j.ultrasmedbio.2020.01.015
18. Abunasser BS, Al-Hiealy MRJ, Zaout IS, Abu-Naser SS. Convolution neural network for breast cancer detection and classification using deep learning. *Asian Pacific J Cancer prevention: APJCP*. (2023) 24:531–44. doi: 10.31557/APJCP.2023.24.2.531
19. Tian R, Lu G, Zhao N, Qian W, Ma H, Yang W, et al. Constructing the optimal classification model for benign and Malignant breast tumors based on multifeature analysis from multimodal images. *J Imaging Inf Med*. (2024) 37:1386–400. doi: 10.1007/s10278-024-01036-7
20. Süküt Y, Yurdakurban E, Duran GS. Accuracy of deep learning-based upper airway segmentation. *J Stomatology Oral Maxillofac Surg*. (2025) 126:102048. doi: 10.1016/j.jormas.2024.102048
21. Yan H, Wang Q, Zhao F, Kang D, Qiao Y. The evidence and concerns about screening ultrasound for breast cancer. *Cancer Biol Med*. (2025) 22:295–300. doi: 10.20892/j.issn.2095-3941.2024.0562
22. Dong TF, Zhou CJ, Huang ZY, Zhao H, Wang XL, Yan SJ, et al. MLP-UNet: an algorithm for segmenting lesions in breast and thyroid ultrasound images. *Comput assisted Surg (Abingdon England)*. (2025) 30:2523266. doi: 10.1080/24699322.2025.2523266
23. Umamaheswari T, Babu YMM, ViT-MAENB7: An innovative breast cancer diagnosis model from 3D mammograms using advanced segmentation and classification process. *Comput Methods programs biomedicine*. (2024) 257:108373. doi: 10.1016/j.cmpb.2024.108373
24. Zhu Y, Zhang S, Wei W, Yang L, Wang L, Wang Y, et al. Intra- and peritumoral radiomics nomogram based on DCE-MRI for the early prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer. *Front Oncol*. (2025) 15:1561599. doi: 10.3389/fonc.2025.1561599
25. Cheng MY, Wu CG, Lin YY, Zou JC, Wang DQ, Haffty BG, et al. Development and validation of a multivariable risk model based on clinicopathological characteristics, mammography, and MRI imaging features for predicting axillary lymph node metastasis in patients with upgraded ductal carcinoma in situ. *Gland Surg*. (2025) 14:738–53. doi: 10.21037/gs-2025-89
26. Veerlapalli P, Dutta SR. A hybrid GAN-based deep learning framework for thermogram-based breast cancer detection. *Sci Rep*. (2025) 15:19665. doi: 10.1038/s41598-025-04676-z
27. Yao H, Zhang X, Zhou X, Liu S. Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification. *Cancers*. (2019) 11(12):1901. doi: 10.3390/cancers11121901
28. Guiot J, Vaidyanathan A, Deprez L, Zerka F, Danthine D, Frix AN, et al. A review in radiomics: Making personalized medicine a reality via routine imaging. *Medicinal Res Rev*. (2022) 42:426–40. doi: 10.1002/med.21846
29. Zeng F, Zeng H, Yang J, Huang D, Liu J, Wen C, et al. Differentiation between phyllodes tumor and fibroadenoma of the breast: A radiomics prediction model based on full-field digital mammography & Digital tomosynthesis. *Technol Cancer Res Treat*. (2024) 23:15330338241289474. doi: 10.1177/15330338241289474
30. Niu S, Huang J, Li J, Liu X, Wang D, Wang Y, et al. Differential diagnosis between small breast phyllodes tumors and fibroadenomas using artificial intelligence and ultrasound data. *Quant Imaging Med Surg*. (2021) 11:2052–61. doi: 10.21037/qims-20-919
31. Wu X, Dong D, Zhang L, Fang M, Zhu Y, He B, et al. Exploring the predictive value of additional peritumoral regions based on deep learning and radiomics: A multicenter study. *Med Phys*. (2021) 48:2374–85. doi: 10.1002/mp.14767
32. Packer MDC, Lester SC. Current understanding of phyllodes tumors of the breast: Tumor classification, molecular landscape, and best pathology practice. *Hum Pathol*. (2025) 162:105863. doi: 10.1016/j.humpath.2025.105863
33. Mottola M, Golfieri R, Bevilacqua A. The effectiveness of an adaptive method to analyse the transition between tumour and peritumour for answering two clinical questions in cancer imaging. *Sensors (Basel)*. (2024) 24(4):1156. doi: 10.3390/s24041156
34. Zhang S, Shao H, Li W, Zhang H, Lin F, Zhang Q, et al. Intra- and peritumoral radiomics for predicting Malignant BiRADS category 4 breast lesions on contrast-enhanced spectral mammography: a multicenter study. *Eur Radiol*. (2023) 33:5411–22. doi: 10.1007/s00330-023-09513-3
35. Tsuchiya M, Masui T, Terauchi K, Yamada T, Katayama M, Ichikawa S, et al. MRI-based radiomics analysis for differentiating phyllodes tumors of the breast from fibroadenomas. *Eur Radiol*. (2022) 32:4090–100. doi: 10.1007/s00330-021-08510-8
36. Dominguez-Cejudo MA, Gil-Torralvo A, Cejuela M, Molina-Pinelo S, Salvador Bofill J. Targeting the tumor microenvironment in breast cancer: prognostic and predictive significance and therapeutic opportunities. *Int J Mol Sci*. (2023) 24(23):16771. doi: 10.3390/ijms242316771
37. Huang Z, Mo S, Wu H, Kong Y, Luo H, Li G, et al. Optimizing breast cancer diagnosis with photoacoustic imaging: An analysis of intratumoral and peritumoral radiomics. *Photoacoustics*. (2024) 38:100606. doi: 10.1016/j.pacs.2024.100606