

## Supplementary

**Table S1** base clinical and gray-scale ultrasound imaging features and their associations with benign and malignant (pathology results) of breast tumors

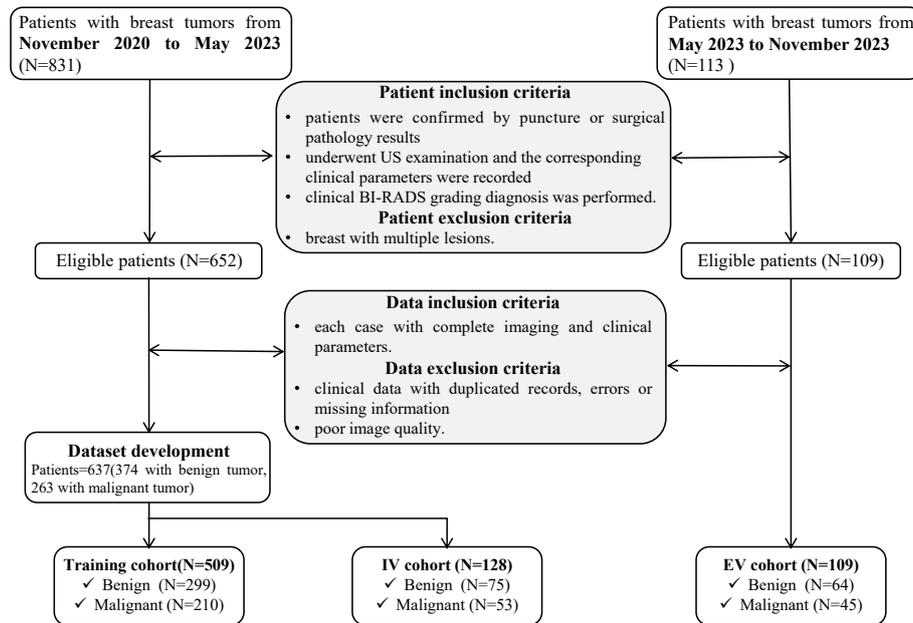
Feature	Total (n=746)	Benign (n=438)	Malignant (n=308)	P value
Age (years)	44.23±12.67	38.77±11.02	52.00±10.66	<0.001**
Dist_LesionToNipple (mm)	11.71±12.10	10.57±12.00	13.33±12.07	0.002*
Dist_LesionToSurface (mm)	6.28±5.67	6.12±5.83	6.51±5.43	0.358
Aspect_Ratio	0.62±0.22	0.58±0.19	0.66±0.25	<0.001**
Tissue composition				<0.001**
Fat	3 (0.40)	3 (0.68)	0 (0)	
Fibroglandular	303 (40.62)	265 (60.50)	38 (12.34)	
Heterogeneous	440 (58.98)	170 (38.81)	270 (87.66)	
Shape				<0.001**
Oval	298 (39.95)	293 (66.89)	5 (1.62)	
Round	6 (0.80)	5 (1.14)	1 (0.32)	
Irregular	442 (59.25)	140 (31.96)	302 (98.05)	
Orientation				<0.001**
Parallel	428 (57.37)	369 (84.25)	59 (19.16)	
Not parallel	318 (42.63)	69 (15.75)	249 (80.84)	
Margin				<0.001**
Circumscribed				<0.001**
No	460 (61.66)	161 (36.76)	299 (97.08)	
Yes	286 (38.34)	277 (63.24)	9 (2.92)	
Indistinct				0.034
No	565 (75.74)	319 (72.83)	246 (79.87)	
Yes	181 (24.26)	119 (27.17)	62 (20.13)	
Angular				<0.001**
No	520 (69.71)	420 (95.89)	100 (32.47)	
Yes	226 (30.29)	18 (4.11)	208 (67.53)	
Microlobulated				<0.001**
No	529 (70.91)	404 (92.24)	125 (40.58)	
Yes	217 (29.09)	34 (7.76)	183 (59.42)	
Spiculated				<0.001**
No	663 (88.87)	434 (99.09)	229 (74.35)	
Yes	83 (11.13)	4 (0.91)	79 (25.65)	
Posterior features				<0.001**
No posterior features	727 (97.45)	438 (100.00)	289 (93.83)	
Enhancement	1 (0.13)	0	1 (0.32)	

Table S1 (continued)

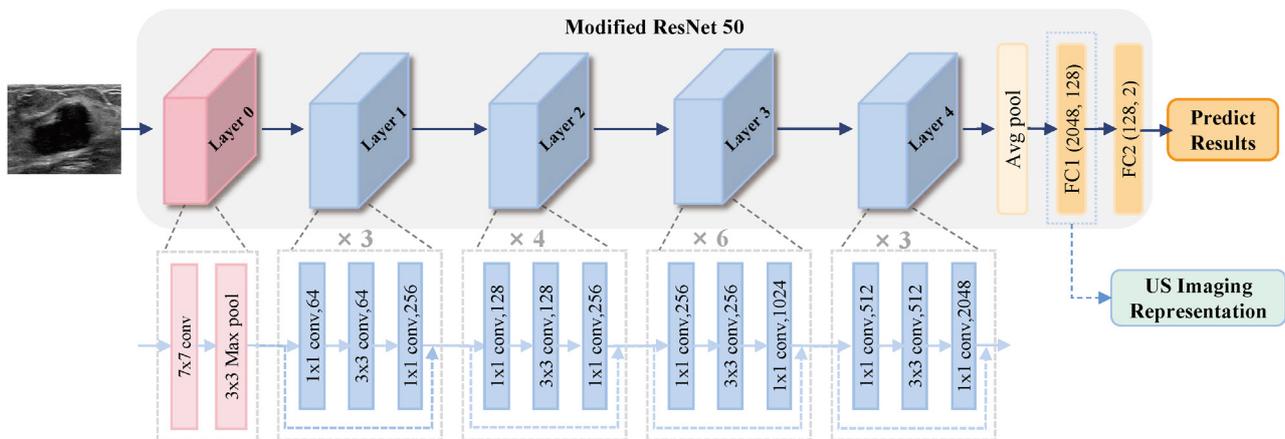
Table S1 (continued)

Feature	Total (n=746)	Benign (n=438)	Malignant (n=308)	P value
Shadowing	15 (2.01)	0	15 (4.87)	
Combined pattern	3 (0.40)	0	3 (0.97)	
Echo pattern				0.188
Anechoic	8 (1.07)	7 (1.60)	1 (0.32)	
Complex cystic/solid	17 (2.28)	12 (2.74)	5 (1.62)	
Hypoechoic	711 (95.31)	411 (93.84)	300 (97.40)	
isoechoic	3 (0.40)	2 (0.46)	1 (0.32)	
Heterogeneous	7 (0.94)	6 (1.37)	1 (0.32)	
Micro-calcifications				<0.001**
No	495 (66.35)	386 (88.13)	109 (35.39)	
Yes	251 (33.65)	52 (11.87)	199 (64.61)	
Associate features				<0.001**
Architectural distorted				<0.001**
No	723 (96.92)	436 (99.54)	287 (93.18)	
Yes	23 (3.08)	2 (0.46)	21 (6.82)	
Duct changes				0.334
No	681 (91.29)	404 (92.24)	277 (89.94)	
Yes	65 (8.71)	34 (7.76)	31 (10.06)	
Skin changes				0.0598
No	742 (99.46)	438 (100.00)	304 (98.70)	
Yes	4 (0.54)	0	4 (1.30)	
Edema				0.332
No	744 (99.73)	438 (100.00)	306 (99.35)	
Yes	2 (0.27)	0	2 (0.65)	
CDFI features				<0.001**
No blood flow	474 (63.54)	391 (89.27)	83 (26.95)	
Internal vascularity, Adler I	91 (12.20)	35 (7.99)	56 (18.18)	
Internal vascularity, Adler II	155 (20.78)	8 (1.83)	147 (47.73)	
Internal vascularity, Adler III	19 (2.55)	0	19 (6.17)	
Perifocal	7 (0.94)	4 (0.91)	3 (0.97)	

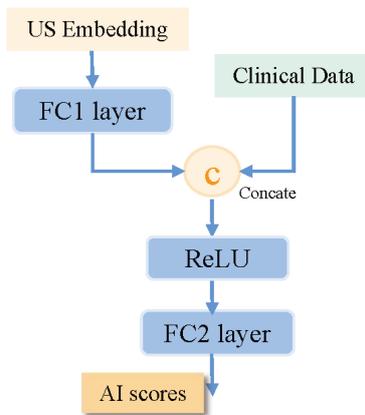
Data are expressed as the mean ± standard deviation or number (percentage). \*, P<0.05; \*\*, P<0.01. Adler 0: no obvious blood flow signals, Adler I: one or two small spot-like blood flow signals, Adler II: strip blood flow signals could be seen, Adler III: reticular blood flow signals could be detected.



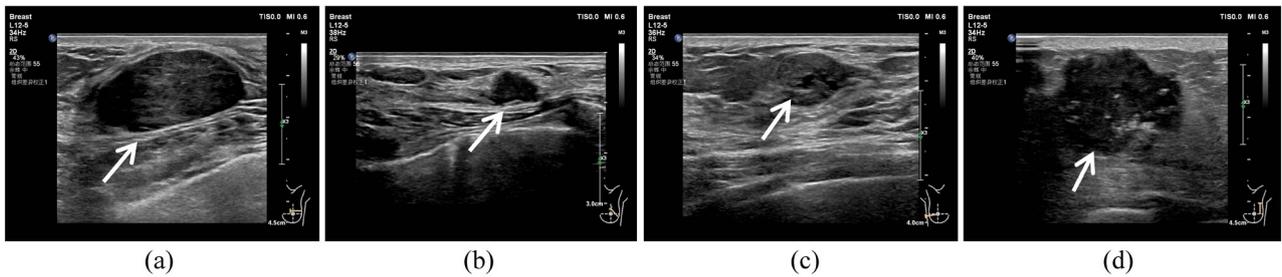
**Figure S1** Flowchart of patients collection and selection of data\_BM for DLR\_BM model construction. US, ultrasound; BI-RADS, Breast Imaging Reporting and Data System; IV, internal validation; EV, external validation; DLR, deep learning radiomics.



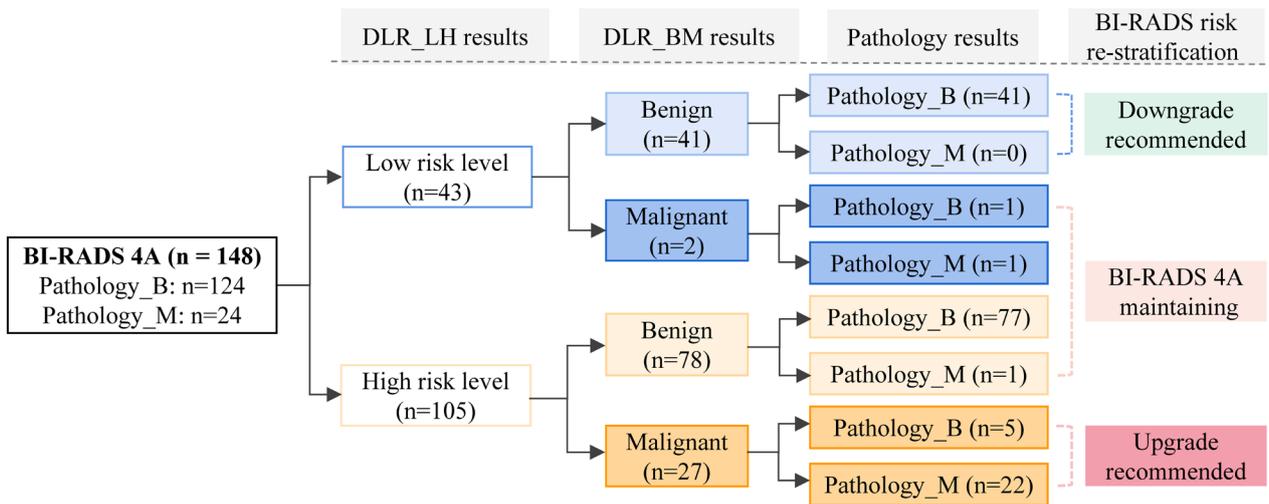
**Figure S2** Architecture of the modified ResNet-50 model. The fully connected layer of the original ResNet-50 was replaced with two new fully connected layers FC1 and FC2, FC1 output the 128-dimensional US embedding, and FC2 output the predictive probability of each category. The ResNet-50 model composed of 5 primary layers and each primary layer consisted of some bottleneck with sub layers: convolutional layer, activation function layer, batch normalization layer, pooling layer and fully connected layer. FC, fully connected layer; US, ultrasound.



**Figure S3** Architecture of the Multi-Layer Perceptron based classifier FC1 layer mapped the 128-dimensional US embedding into a 16-dimensional feature vector, and then concatenated with the normalized 1-dimensional clinical data (patient age) to form the integrated feature maps. ReLU is the rectified linear unit activation function. The FC2 layers output the predictive probability of each category as AI scores. FC, fully connected layer; US, ultrasound; ReLU, rectified linear unit; AI, artificial intelligence.



**Figure S4** Some typical cases with breast tumors. (A) a 32-year-old women with tumor of BI-RADS 3 and pathological benign; (B) a 44-year-old women with tumor of BI-RADS 3 and pathological malignant; (C) a 34-year-old women with tumor of BI-RADS 4A and pathological benign; (D) a 64-year-old women with tumor of BI-RADS 5 and pathological malignant. Consistent with the BI-RADS lexicon descriptors, Tumors with low risk level are more likely to have parallel orientation, circumscribed margin and oval shape, while tumors with high risk level are more likely to have an angular, instinct micro-lobulated or spiculated margin, irregular shape and always along with micro-calcification. The arrows point to the lesion areas. BI-RADS, Breast Imaging Reporting and Data System.



**Figure S5** Risk re-stratification process of the whole BI-RADS 4A patients of this study by the DLR\_LH and DLR\_BM model. Lesions predicted as low risk level by DLR\_LH and as benign tumor by DLR\_BM were recommended to downgrade to BI-RADS 3, and lesions predicted as high risk level by DLR\_LH and as malignant tumor by DLR\_BM were recommended to upgrade to a higher grade. Finally, 27.7% (41/148) of BI-RADS 4A lesions were downgraded to BI-RADS 3, 18.2% (27/148) lesions were upgraded to a higher risk level with malignancy probability of 81.5% (22/27) and 54.1% (80/148) lesions should maintain the risk level of BI-RADS 4A. BI-RADS, Breast Imaging Reporting and Data System.