Received: 13 July 2022 | Accepted: 13 September 2022

DOI: 10.1111/1759-7714.14673

ORIGINAL ARTICLE

WILEY

Noninvasive assessment of breast cancer molecular subtypes on multiparametric MRI using convolutional neural network with transfer learning

Haolin Yin | **Lutian Bai** | **Huihui Jia** | **Guangwu Lin**

Department of Radiology, Huadong Hospital Affiliated to Fudan University, Shanghai, China

Correspondence
Guangwu Lin, Department of Radiology, Huadong Hospital Affiliated to Fudan University, Shanghai, Shanghai 200040, China.
Email: lingw01000@163.com

Funding information
National Natural Science Foundation of China, Grant/Award Number: 81771816

Abstract

Background: To evaluate the performances of multiparametric MRI-based convolutional neural networks (CNNs) for the preoperative assessment of breast cancer molecular subtypes.

Methods: A total of 136 patients with 136 pathologically confirmed invasive breast cancers were randomly divided into training, validation, and testing sets in this retrospective study. The CNN models were established based on contrast-enhanced T_1-weighted imaging (T_1C), Apparent diffusion coefficient (ADC), and T_2-weighted imaging (T_2W) using the training and validation sets. The performances of CNN models were evaluated on the testing set. The area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and accuracy were calculated to assess the performance.

Results: For the separation of each subtype from other subtypes on the testing set, the T_1C-based models yielded AUCs from 0.762 to 0.920; the ADC-based models yielded AUCs from 0.686 to 0.851; and the T_2W-based models achieved AUCs from 0.639 to 0.697.

Conclusion: T_1C-based models performed better than ADC-based models and T_2W-based models in assessing the breast cancer molecular subtypes. The discriminating performances of our CNN models for triple negative and human epidermal growth factor receptor 2-enriched subtypes were better than that of luminal A and luminal B subtypes.

KEYWORDS
breast cancer, breast MRI, deep learning, molecular subtypes, neural network

INTRODUCTION

Breast cancer is considered the most common and highest-mortality malignant tumor worldwide.[1] Breast cancer, as a highly heterogeneous disease in clinical practice, has four main intrinsic molecular subtypes based on the expression status of several molecular receptors: luminal A, luminal B, human epidermal growth factor receptor 2 (HER2)-enriched, and triple negative (TN).[2] Luminal and HER2-enriched tumors are often considered good candidates for endocrine therapy and targeted antibody therapy, respectively.[3,4] TN tumors have the worst prognosis and responses to chemotherapy.[5] The characteristics of breast cancers that have substantial differences in phenotype, treatment, prognosis, response, and outcome cannot be fully reflected by traditional evaluations based on tumor grade, size, and histology.[6,7] The molecular subtypes of breast cancer are routinely evaluated by immunohistochemical surrogates or gene expression profiling from tissue biopsy.[8,9] This approach has some limitations because the limited number, size, and location of the samples cannot capture the heterogeneity within the whole tumor. Additionally, breast tumor biology may not remain stable over time or during treatment.[10] The heterogeneity within a single tumor may cause therapy resistance and treatment failure.[11–13]

Medical imaging describes a set of noninvasive and alternative approaches for the preoperative evaluation of breast cancer and the monitoring of biological changes during treatment.[14] In this context, magnetic resonance imaging (MRI) is considered a powerful tool for the examination of breast cancer, because it can be used to scan the whole breast to obtain high-spatial-resolution images. Multiparametric MRI consists of T_2-weighted, dynamic contrast enhanced (DCE), diffusion-weighted, and other imaging sequences. Although the DCE-MRI sequence is primarily focused on, complementary sequences such as diffusion-weighted imaging (DWI) can also provide additional functional information via the detection of surrogate markers for tissue microstructure and cell density by measuring the random movement of water molecules.[15] MRI coupled with radiomics has yielded some initial encouraging results regarding the prediction of molecular subtype,[16] tumor histology,[17] recurrence risk,[18] chemotherapy response,[19] and metastatic potential.[20] Convolutional neural networks (CNNs) have displayed excellent performance in the image recognition field. Worldwide efforts have provided many effective and mature CNNs, such as ResNet, VGG, AlexNet, and InceptionV3. Different from radiomics, which extracts handcrafted features to carry out a specific task; CNNs can automatically and adaptively learn multilevel features from medical images to perform end-to-end tasks without handcrafted feature engineering. Additionally, CNNs can scan all image pixels with convolution kernels, which helps perceive the global information within the image. Some studies have indicated that MRI-based deep learning has the potential in predicting molecular subtype in breast cancer.[21,22] However, the performances of multiparametric MRI-based CNNs in the prediction of molecular subtypes of breast cancer still need to be further explored.

We hypothesized that the microstructural heterogeneity of breast cancer among different molecular subtypes would lead to different phenotypes on imaging that can be captured with multiparametric MRI-based CNNs. Therefore, the purpose of our study was to evaluate the performances of multiparametric MRI-based CNNs for the preoperative assessment of breast cancer molecular subtypes.

METHODS

Patients

Our retrospective study was approved by the ethics committee of our hospital with a waiver of written informed consent. A database search was performed for patients who underwent multiparametric breast MRI examinations from January 2015 to June 2019. The inclusion criteria for the study were as follows: (1) histologically confirmed invasive breast cancer and a histologic result of estrogen receptor (ER), progesterone receptor (PR), HER2, and antigen identified by monoclonal antibody (Ki-67) expression; and (2) preoperative multiparametric MRI examination. The exclusion criteria were as follows: (1) preoperative chemotherapy, radiotherapy, or endocrine therapy; (2) lack of complete clinical data; (3) preoperative invasive operation of the breast; and (4) obvious imaging artifacts. After screening, a total of 136 lesions from 136 patients with invasive breast cancers were included (only the largest breast lesion with a corresponding pathological result was included in the analysis for patients with multicentric lesions). According to the classification goals, all patients in our study were randomly divided into the training, validation, and testing sets at a 6:2:2 ratio.

Magnetic resonance image acquisition

All examinations were performed using a 3.0 Tesla magnetic resonance scanner (Verio, Siemens Medical Systems) in the study, with a dedicated 4-channel breast surface coil. All patients were examined in a prone position during the MRI examination. The transverse turbo spin-echo T_2-weighted sequence was performed by adopting the following parameters: repetition time (TR)/echo time (TE), 5300/83 ms; matrix size, 364 × 290; field of view, 34 × 34 cm; slice thickness, 4.5 mm; slice gap, 1 mm; number of excitations (NEX), 1. The transverse DWI sequences were performed at b values of 50 and 850 s/mm^2 by adopting the following parameters: TR/TE, 7300/81 ms; matrix size, 192 × 126; field of view, 34 × 14 cm; slice thickness, 3.5 mm; slice gap, 0.5 mm; NEX, 5. Apparent diffusion coefficient (ADC) maps were calculated on the corresponding workstation by adopting the least squares method with images of b values of 50 and 850 s/mm^2. After T_2W and DWI sequence, a DCE-MRI transverse sequence was performed with six dynamic acquisitions, one before and five after an elbow vein bolus injection of gadolinium dimeglumine (GE Healthcare) equal to 0.1 mmol/kg body weight, followed by a 20-ml saline flush. The first postcontrast dynamic image acquisition started at the 30th second after contrast agent injection. The temporal resolution was 60 s for each phase of the post-contrast dynamic acquisitions. DCE-MRI sequence scan parameters were as follows: TR/TE, 4.26/1.53 ms; flip angle, 6°; matrix, 448 × 372; field of view, 35 × 35 cm; slice thickness, 0.9 mm; slice gap, 0.18 mm; NEX, 1.

Histopathological analysis

Immunohistochemical (IHC) analysis included ER, PR, Ki-67, and HER2. Molecular subtypes were classified as luminal A, luminal B, HER2-enriched, and TN according to a previous study.[23] ER$^+$ and/or PR$^+$, HER2$^-$ and low-Ki-67 lesions were determined as luminal A. ER- and/or PR$^+$, HER2$^-$, and high-Ki-67 lesions were determined as luminal B$^-$, whereas ER$^-$ and/or PR$^+$, and HER2$^+$ lesions were determined as luminal B$^+$; here, the luminal B$^-$ and luminal

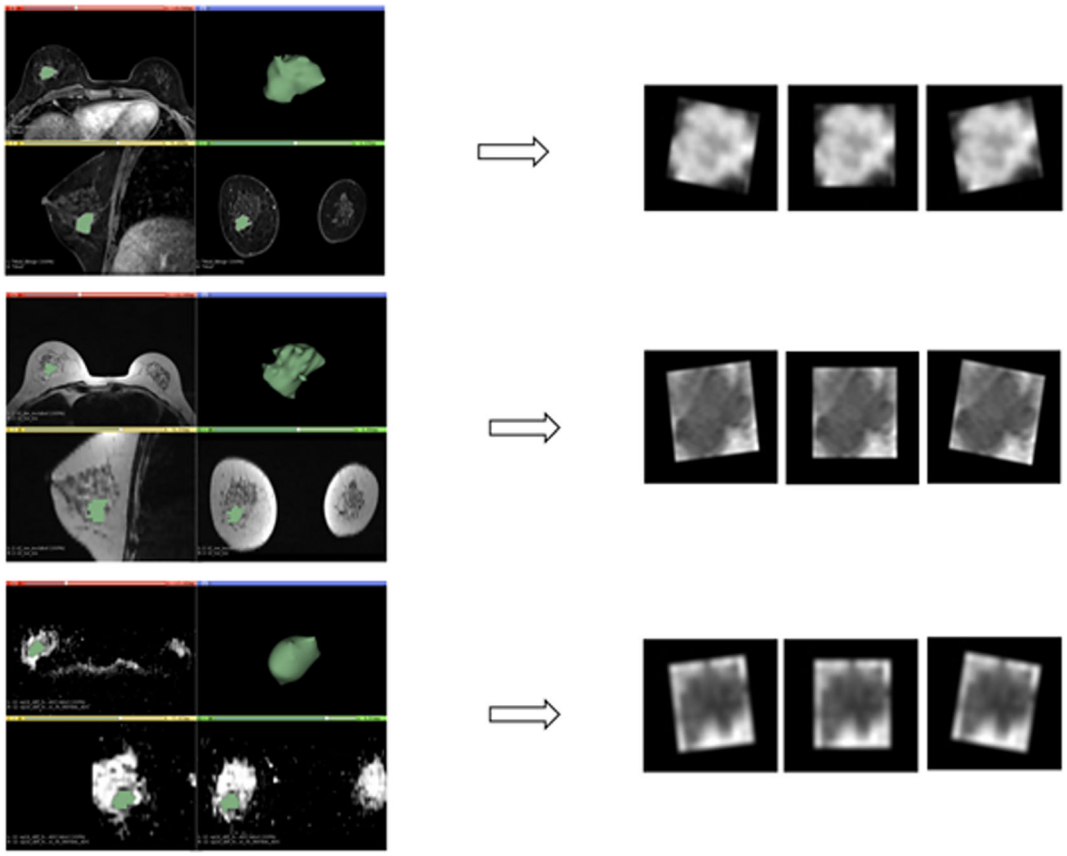

FIGURE 1 Delineation of regions of interest and image preprocessing included block extraction and data augmentation

B^+ were combined into a single type (luminal B) in the deep learning analysis to ensure a certain dataset size. ER^-, PR^-, and $HER2^+$ lesions were determined as HER2-enriched. ER^-, PR^-, and $HER2^-$ lesions were determined TN. Lesions with equivocal HER2 status, were further evaluated using fluorescence in situ hybridization. Lesions with measured gene amplification were classified as $HER2^+$.

Regions of interest delineation

The transverse T_2W, ADC, and contrast-enhanced T_1-weighted ($T_1 + C$) images were used to develop the CNN models. The first transverse postcontrast dynamic images were used as $T_1 + C$, as this phase of DCE-MRI (obtained at the 90th second after contrast agent injection) best shows the breast lesion boundary relative to adjacent tissues and contains rich predicting information. The regions of interest (ROIs) of breast lesions on the transverse $T_1 + C$, T_2W, and ADC images were manually delineated by a radiologist with 6 years of experience interpreting breast MRI. Delineation of the lesion ROIs was performed slice-by-slice for the whole tumor volume (Figure 1) using open-source software (3Dslicer; https://www.slicer.org/). The ROIs of all breast lesions were confirmed by a radiologist with 12 years of experience.

Data preprocessing

Data augmentation was applied to the image sets (Figure 1), using affine transformations, including random rotation from −10 to 10°, stretching from 0.9 to 1.1, and shifting from −10 to 10 pixels. After the image transformations, one subtype set was increased tenfold, and all other subtype sets were augmented to balance the cases (each molecular subtype with the one subtype versus all other subtypes strategy). According to the ROIs of each breast lesion, a block centered at the center of the tumor containing the whole tumor was cropped from the images. All slices of the blocks were reshaped to 224 × 224 by zero-padding.

Deep-Learning analysis

All experiments were conducted in Python (version 3.7.0; Python Software Foundation) by using PyTorch (version: 1.4.0) on a workstation equipped with a GeForce GTX 3080 GPU. ResNet has good stability and performance in many computer vision tasks.[24] Therefore, the ResNet18 architecture pre-trained with ImageNet (http://www.imagenet.org/) was used to establish preoperative prediction models based on the images of the three sequences (Figure 2). Several modifications were applied to the pre-trained CNNs in this

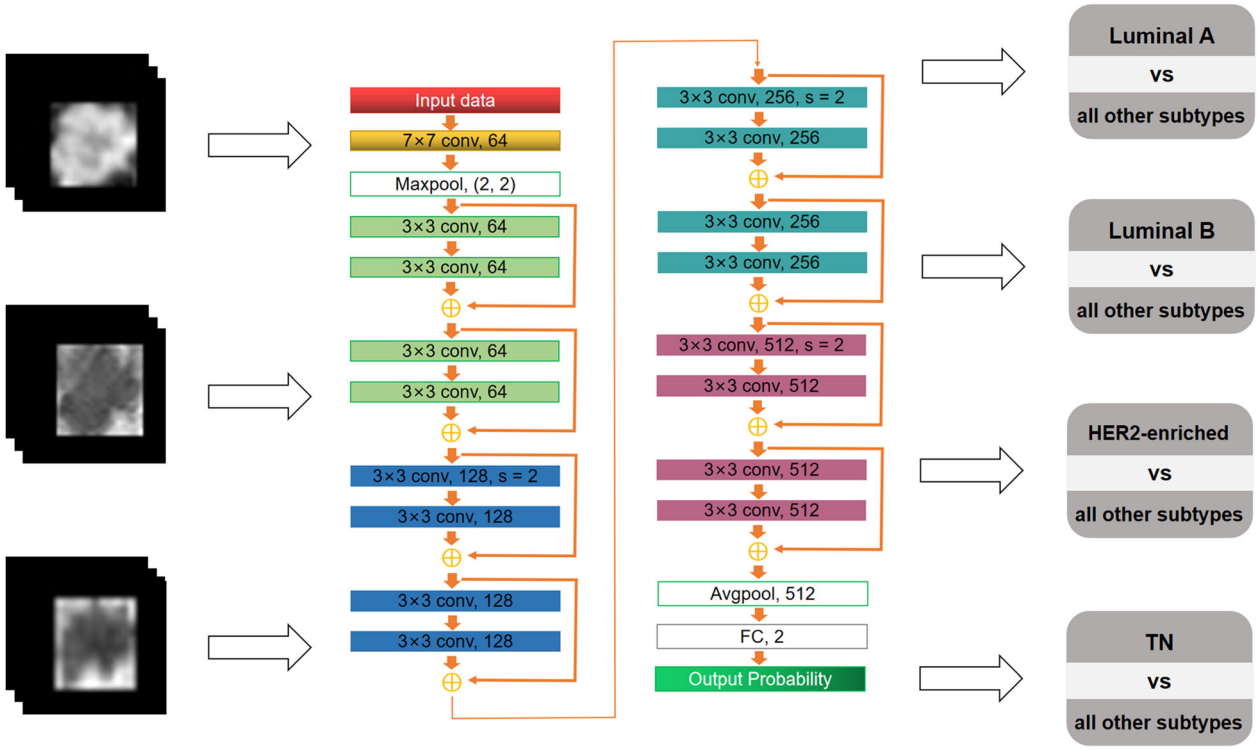

FIGURE 2 The conceptual architecture of the multiparametric MRI-based convolutional neural networks for the preoperative assessment of breast cancer molecular subtypes

study. The number of neurons in the fully connected layer was adjusted from 1000 to 2 to fit the classification task of this study. All images of the three sequences were normalized before they were fed into the networks by subtracting the mean and dividing by the standard deviation. The weights of the fully connected layers and the convolutional layers were adjusted according to the images of the training set, which were fed into the networks. The Adam algorithm was adopted during the training process to minimize the loss (cross-entropy) function, with a mini-batch size of 8. The initial learning rate for the Adam optimizer was set to 0.001 and decayed by a factor of 10 each time when there was no improvement in accuracy for 10 continuous epochs in the validation set. Finally, the CNN model with the highest accuracy of the validation set was selected. The images of the testing set were fed into the trained CNN model to output the probability of every class, and the class with the highest probability was chosen as the classification result.

Statistical analysis

All statistical analyses were performed using SPSS (IBM SPSS, v.25.0) and Python. We assessed the predictive performance of the CNN models for each molecular subtype with a one subtype versus all other subtypes strategy on the test set. The gold standard for the prediction of molecular subtypes was the postoperative histopathology result. The area under the receiver operating characteristic curve (AUC) and its 95% confidence interval (CI), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated. The optimal cutoff value was determined by maximizing Youden's index. Significant differences between AUCs were compared by DeLong's test.[25] Welch's *t*-test or Student's *t*-test was used for continuous variables, and Pearson's χ^2 test was used for categorical variables. A p-value <0.05 was considered statistically significant.

RESULTS

Clinicopathological characteristics

The clinicopathological information of each patient is provided in Table 1. Of the 136 biopsy-proven breast cancers, 35 were histologically confirmed as luminal A, 52 as luminal B, 26 as HER2-enriched, and 23 as TN. The median ages of the patients were 52, 48, 51, and 52 years for the luminal A, luminal B, HER2-enriched, and TN groups, respectively. There was no significant difference in age, tumor size, family history, histologic type or menopausal status among the patients, except for ER status, PR status, HER2 status, and Ki-67.

Discriminating luminal A from other subtypes

The accuracy, sensitivity, specificity, and AUC in discriminating luminal A from the other subtypes are

TABLE 1 Clinicopathological characteristics of the patients

	All (N = 136)	Luminal A (N = 35)	Luminal B (N = 52)	HER2-enriched (N = 26)	TN (N = 23)	*p* value
Age-median (IQR)	51 (45–57)	52 (47–56)	48 (43–55)	51 (46–58)	52 (47–62)	0.116
Tumor size						0.997
≤2.0 cm	99 (72.8)	25 (71.4)	38 (73.1)	19 (73.1)	17 (73.9)	
2.1–4.0 cm	37 (27.2)	10 (28.6)	14 (26.9)	7 (26.9)	6 (26.1)	
Histologic type						0.697
IDC	131 (96.3)	33 (94.3)	50 (96.2)	26 (100)	22 (95.7)	
Other	5 (3.7)	2 (5.7)	2 (3.8)	0 (0)	1 (4.3)	
Menopausal						0.860
Premenopausal	74 (54.4)	18 (51.4)	27 (51.9)	15 (57.7)	14 (60.9)	
Post-menopausal	62 (45.6)	17 (48.6)	25 (48.1)	11 (42.3)	9 (39.1)	
Family history						0.935
No	132 (97.1)	34 (97.1)	51 (98.1)	25 (96.2)	22 (95.7)	
Yes	4 (2.9)	1 (2.9)	1 (1.9)	1 (3.8)	1 (4.3)	
ER status						<0.001
Positive	82 (60.3)	35 (100.0)	47 (90.4)	0 (0.0)	0 (0.0)	
Negative	54 (39.7)	0 (0.0)	5 (9.6)	26 (100.0)	23 (100.0)	
PR status						<0.001
Positive	75 (55.1)	32 (91.4)	43 (82.7)	0 (0.0)	0 (0.0)	
Negative	61 (44.9)	3 (8.6)	9 (17.3)	26 (100.0)	23 (100.0)	
HER2 status						<0.001
Positive	54 (39.7)	0 (0.0)	28 (53.8)	26 (100.0)	0 (0.0)	
Negative	82 (60.2)	35 (100.0)	24 (46.2)	0 (0.0)	23 (100.0)	
Ki-67						<0.001
High	92 (67.6)	0 (0.0)	46 (88.5)	24 (92.3)	22 (95.7)	
Low	44 (32.4)	35 (100.0)	6 (11.5)	2 (7.7)	1 (4.3)	

Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; IDC, infiltrating ductal cancer; IQR, interquartile range; Ki-67, antigen identified by monoclonal antibody; PR, progesterone receptor; TN, triple negative.

summarized in Table 2 and Figure 3. The T_1 + C model yielded an AUC of 0.817 (95% CI, 0.734, 0.900), a sensitivity of 0.673, a specificity of 0.886, a PPV of 0.875, an NPV of 0.696, and an accuracy of 0.771. The ADC model achieved an AUC of 0.759 (95% CI, 0.658, 0.859), a sensitivity of 0.596, a specificity of 0.795, a PPV of 0.775, an NPV of 0.625, and an accuracy of 0.698. The T_2W model demonstrated an AUC of 0.639 (95% CI, 0.526, 0.748), a sensitivity of 0.634, a specificity of 0.591, a PPV of 0.647, an NPV of 0.578, and an accuracy of 0.615. The T_1 + C model and the ADC model performed better than the T_2W model ($p = 0.011$ and $p = 0.126$), and the T_1 + C model performed slightly better than the ADC model ($p = 0.402$).

Discriminating luminal B from other subtypes

The accuracy, sensitivity, specificity, and AUC in discriminating luminal B from the other subtypes are summarized in Table 2 and Figure 4. The T_1 + C model yielded an AUC of 0.762 (95% CI, 0.648, 0.876), a sensitivity of 0.690, a specificity of 0.795, a PPV of 0.714, an NPV of 0.775, and an accuracy of 0.750. The ADC model achieved an AUC of 0.686 (95% CI, 0.557, 0.816), a sensitivity of 0.655, a specificity of 0.641, a PPV of 0.576, an NPV of 0.714, and an accuracy of 0.647. The T_2W model demonstrated an AUC of 0.683 (95% CI, 0.554, 0.811), a sensitivity of 0.655, a specificity of 0.667, a PPV of 0.594, an NPV of 0.772, and an accuracy of 0.661. The T_1 + C model performed slightly better than the ADC model ($p = 0.348$) and the T_2W model ($p = 0.373$), and the ADC model had almost equivalent performance to the T_2W model ($p = 0.969$).

Discriminating HER2-enriched from other subtypes

The accuracy, sensitivity, specificity, and AUC in discriminating HER2-enriched from the other subtypes were summarized in Table 2 and Figure 5. The T_1 + C model yielded an AUC of 0.885 (95% CI, 0.817, 0.953), a sensitivity of 0.800, a specificity of 0.895, a PPV of 0.889, an NPV of 0.811, and an accuracy of 0.847. The ADC model achieved an AUC of 0.757 (95% CI, 0.652, 0.862), a

TABLE 2 The performance characteristics of multiparametric MRI-based convolutional neural networks for the assessment of breast cancer molecular subtypes in the testing set

	ACC	SEN	SPE	PPV	NPV	AUC (95% CI)
Luminal A						
$T_1 + C$	0.771	0.673	0.886	0.875	0.696	0.817 (0.734–0.900)
ADC	0.698	0.596	0.795	0.775	0.625	0.759 (0.658–0.859)
T_2WI	0.615	0.634	0.591	0.647	0.578	0.639 (0.526–0.748)
Luminal B						
$T_1 + C$	0.750	0.690	0.795	0.714	0.775	0.762 (0.648–0.876)
ADC	0.647	0.655	0.641	0.576	0.714	0.686 (0.557–0.816)
T_2WI	0.661	0.655	0.667	0.594	0.722	0.683 (0.554–0.811)
HER-2						
$T_1 + C$	0.847	0.800	0.895	0.889	0.811	0.885 (0.817–0.953)
ADC	0.776	0.760	0.792	0.792	0.760	0.757 (0.652–0.862)
T_2WI	0.684	0.660	0.708	0.702	0.667	0.663 (0.547–0.778)
TN						
$T_1 + C$	0.857	0.800	0.930	0.936	0.784	0.920 (0.863–0.977)
ADC	0.838	0.800	0.884	0.898	0.775	0.851 (0.770–0.932)
T_2WI	0.735	0.709	0.767	0.796	0.673	0.697 (0.588–0.806)

Abbreviations: $T_1 + C$, first axial post contrast dynamic images; T_2WI, T2-weighted images; ADC, apparent diffusion coefficient; ACC, accuracy; AUC, area under the receiver operating characteristic curve; CI, confidence interval; HER2, human epidermal growth factor receptor 2; NPV, negative predictive value; PPV, positive predictive value; SEN, sensitivity; SPE, specificity; TN, triple negative.

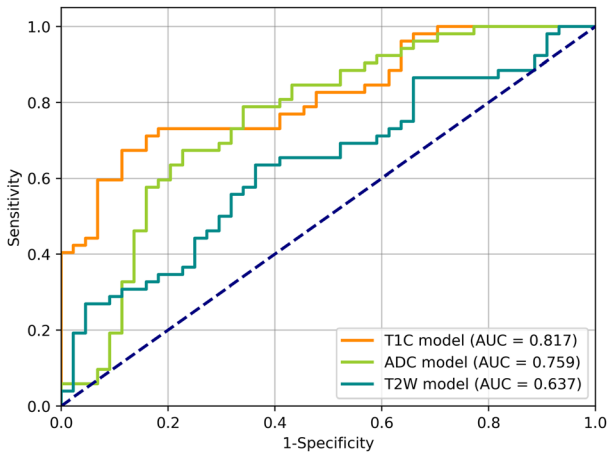

FIGURE 3 Receiver operating characteristic curve analysis of the contrast-enhanced T_1-weighted imaging model (orange line), apparent diffusion coefficient model (green line), and T_2-weighted imaging model (blue line) in predicting luminal A breast cancer in the testing set

sensitivity of 0.760, a specificity of 0.792, a PPV of 0.792, an NPV of 0.760, and an accuracy of 0.776. The T_2W model demonstrated an AUC of 0.663 (95% CI, 0.547, 0.778), a sensitivity of 0.660, a specificity of 0.708, a PPV of 0.702, an NPV of 0.667, and an accuracy of 0.684. The $T_1 + C$ model performed better than the ADC model ($p = 0.035$) and the T_2W model ($p = 0.001$), and the ADC model had a slightly better performance to the T_2W model ($p = 0.260$).

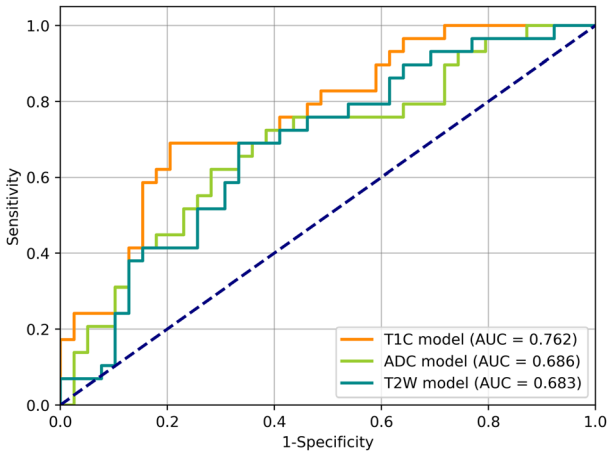

FIGURE 4 Receiver operating characteristic curve analysis of the contrast-enhanced T_1-weighted imaging model (orange line), apparent diffusion coefficient model (green line), and T_2-weighted imaging model (blue line) in predicting luminal B breast cancer in the testing set

Discriminating TN from other subtypes

The accuracy, sensitivity, specificity, and AUC in discriminating TN from the other subtypes are summarized in Table 2 and Figure 6. The $T_1 + C$ model yielded an AUC of 0.920 (95% CI, 0.863, 0.977), a sensitivity of 0.800, a specificity of 0.930, a PPV of 0.936, an NPV of 0.784, and an accuracy of 0.857. The ADC model achieved an AUC of 0.851 (95% CI, 0.770, 0.932), a sensitivity of 0.800, a specificity of 0.884, a PPV of 0.898, an NPV of 0.775, and an accuracy of 0.838. The T_2W model demonstrated an AUC of 0.697 (95% CI, 0.588, 0.806), a sensitivity of 0.709, a specificity of 0.767, a PPV of 0.796, an NPV of 0.673, and an accuracy of 0.735. The $T_1 + C$ model and the ADC model performed better than the T_2W model ($p = 0.001$ and $p = 0.027$), and the $T_1 + C$ model performed slightly better than the ADC model ($p = 0.139$).

DISCUSSION

In this study, we performed comprehensive deep learning analyses to assess the efficiencies of three commonly used breast cancer-examination series on molecular subtypes. To alleviate the problems of limited molecular subtype data for CNN training, we trained CNN models to predict molecular subtypes using transfer learning. Because many image features are composed of universal elements, a CNN initialized with transfer weights can easily outperform a CNN trained from randomly initialized weights.[26] We performed multiple comparisons of the prediction effects among the T_1C, ADC, and T_2W sequences, and found that in the prediction of the four subtypes, the T_1C-based models always outperformed the ADC and T_2W-based models in terms of AUC and accuracy. In addition, our results indicate that the ability of MRI-based CNNs to predict molecular subtypes may also

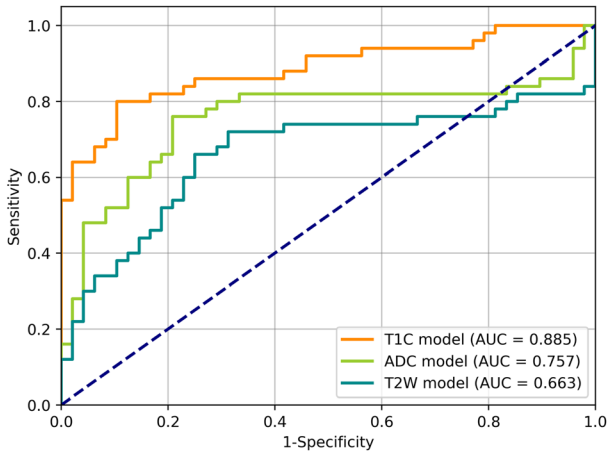

FIGURE 5 Receiver operating characteristic curve analysis of the contrast-enhanced T_1-weighted imaging model (orange line), apparent diffusion coefficient model (green line), and T_2-weighted imaging model (blue line) in predicting HER2-enriched breast cancer in the testing set

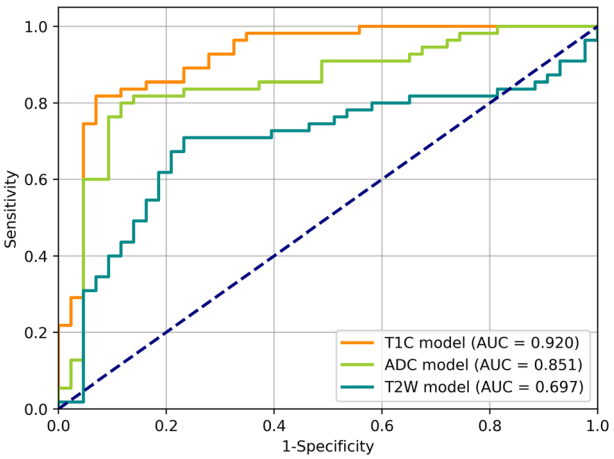

FIGURE 6 Receiver operating characteristic curve analysis of the contrast-enhanced T_1-weighted imaging model (orange line), apparent diffusion coefficient model (green line), and T_2-weighted imaging model (blue line) in predicting triple-negative breast cancer in the testing set

depend on the biological characteristics and imaging features of different molecular subtypes.

With the T_1C models, the best results in terms of AUCs were achieved in the separation of TN and HER2-enriched type from all other types (AUCs, 0.920 and 0.885). Tumors angiogenesis is one of the important factors affecting contrast medium uptake and internal enhancement patterns. Highly vascularized breast tumors tend to show strong contrast enhancement in the first post contrast dynamic phase. HER2 overexpression can increase cell survival, cell proliferation, and invasiveness, as well as neoangiogenesis by improving vascular endothelial growth factor production.[27] The rich angiogenesis of the HER2-enriched subtype leads to rapid early contrast uptake, which can be detected by T_1C.[28] The TN type tends to show a round/oval mass shape and rim enhancement, caused by high angiogenesis in the periphery of the breast tumor.[29,30] Therefore, compared with ADC and T_2W imaging, T_1C imaging can better predict TN- and HER2-enriched breast cancers. The range of AUCs (0.762–0.920) using the T_1C model obtained in our study was generally consistent with that of previous studies reporting AUC values of 0.650–0.910 in the prediction of breast cancer subtypes using contrast-enhanced MRI.[21,22,31] Differences might be attributed to heterogeneity in scanners, patients, classification tasks, and deep learning architectures among the different studies.

Although the ADC-based CNNs achieved lower AUCs than the T_1C-based CNNs, we obtained good results in the separation between TN and all other types (AUC 0.851). TN breast cancer is partially characterized by tumor necrosis, squamous metaplasia or spindle cells, a high nuclear-cytoplasmic ratio, and a high total mitotic count[30,32]; it can also demonstrate an aggressive characteristic. DWI can provide quantitative and qualitative information reflecting cellular changes of TN breast cancers. Areas of necrotic tissue and destruction of cell membrane integrity in tumor lesions may be associated with increased intratumoral water diffusion. This may explain the higher ADC in TN breast cancers when the ADC of the entire breast lesion is measured.[33,34] A previous study showed that radiomic signatures from DWI with ADC mapping allow the separation between TN and all other types with a diagnostic accuracy of 0.736.[35] However, the sample size of the study was very small, and the results may be unstable. Another previous study reported an AUC of 0.804 in the separation of TN from other cancers using DWI-based radiomics.[36] Our study achieved slightly better results than these previous studies.

The results of our study showed that there are still two difficult points in predicting molecular subtypes of breast cancer using MRI-based CNNs. The first is the T_2W-based CNN performed only moderately in predicting the four molecular subtypes (AUC, 0.639–0.697). Iso-/hypointense T_2-signals are considered a feature of breast cancer, which may reflect fibrosis.[37] In contrast, TN breast cancers may show high signal intensity on T_2W images, which may be associated with tumor necrosis.[30,38] However, the differences on the T_2W images of the four molecular subtypes does not seem to be easily captured by CNN. This may be explainable because the signal differences on the T_2W images are small and the image resolution is relatively limited. The second difficult point is that MRI-based CNNs have limited predictive performance for luminal A and luminal B types. Luminal A is considered the most frequent subtype of breast cancers[39] and it can be treated with endocrine therapy; therefore, determining luminal A is clinically important and relevant. Luminal B tumors are resistant to hormone therapy and have a molecular phenotype distinct from that of luminal A tumors. Luminal tumors demonstrate fibrosis and perilesional spiculations.[40] Intratumoral T_2-signal iso-/hypointensity may also be associated with fibrosis and is often observed in luminal lesions.[37] Luminal breast cancers usually show less strong enhancement than other subtypes.[41] Although their performance in predicting luminal A and

luminal B tumors was not particularly outstanding, our T_1C-based CNNs (AUC, 0.762–0.817) still reached a level similar to deep learning and radiomics models in some previous studies (AUC, 0.759–0.860).[22,42,43]

There are some limitations in this study. First, selection bias may be present in this retrospective study, because the patients were not consecutive cases and the number of patients was small. Enlarging the sample size will help to achieve more stable results. Second, molecular subtypes were not confirmed by formal genetic testing, but by IHC surrogates; however, molecular subtypes in clinical practice are often based on IHC surrogates. Third, the proportions of the four molecular subtypes were unbalanced. Although data augmentation was applied to balance data, such imbalance in the original data might have influenced the development of the CNN model.

In conclusion, we established multiple MRI-based CNNs to assess four breast cancer molecular subtypes. Our initial results showed that T_1C performed better than ADC and T_2W in assessing the breast cancer molecular subtypes. The discriminating performances of our CNN models for TN and HER2-enriched breast cancer were better than that of luminal A and luminal B breast cancer. Our study indicates that different MRI sequences may contain different kinds of distinguishing information for the assessment of breast cancer molecular subtypes and that selecting a suitable MRI sequence is helpful to achieve better results for a specific assessment task.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (81771816).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Guangwu Lin https://orcid.org/0000-0001-5085-2111

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
3. Lam SW, Jimenez CR, Boven E. Breast cancer classification by proteomic technologies: current state of knowledge. Cancer Treat Rev. 2014;40(1):129–38.
4. Wang Y, Yin Q, Yu Q, Zhang J, Liu Z, Wang S, et al. A retrospective study of breast cancer subtypes: the risk of relapse and the relations with treatments. Breast Cancer Res Treat. 2011;130(2):489–98.
5. Huber KE, Carey LA, Wazer DE. Breast cancer molecular subtypes in patients with locally advanced disease: impact on prognosis, patterns of recurrence, and response to therapy. Semin Radiat Oncol. 2009;19(4):204–10.
6. Tsoutsou PG, Vozenin MC, Durham AD, Bourhis J. How could breast cancer molecular features contribute to locoregional treatment decision making? Crit Rev Oncol Hematol. 2017;110:43–8.
7. Haynes B, Sarma A, Nangia-Makker P, Shekhar MP. Breast cancer complexity: implications of intratumoral heterogeneity in clinical management. Cancer Metastasis Rev. 2017;36(3):547–55.
8. Ahn HJ, Jung SJ, Kim TH, Oh MK, Yoon HK. Differences in clinical outcomes between luminal A and B type breast cancers according to the St. Gallen Consensus 2013. J Breast Cancer. 2015;18(2):149–59.
9. Orlando L, Viale G, Bria E, Lutrino ES, Sperduti I, Carbognin L, et al. Discordance in pathology report after central pathology review: implications for breast cancer adjuvant treatment. Breast. 2016;30:151–5.
10. Pisco AO, Huang S. Non-genetic cancer cell plasticity and therapy-induced stemness in tumour relapse: 'What does not kill me strengthens me'. Br J Cancer. 2015;112(11):1725–32.
11. Zardavas D, Irrthum A, Swanton C, Piccart M. Clinical management of breast cancer heterogeneity. Nat Rev Clin Oncol. 2015;12(7):381–94.
12. Mann GB, Fahey VD, Feleppa F, Buchanan MR. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. J Clin Oncol. 2005;23(22):5148–54.
13. Burge CN, Chang HR, Apple SK. Do the histologic features and results of breast cancer biomarker studies differ between core biopsy and surgical excision specimens? Breast. 2006;15(2):167–72.
14. Waks AG, Winer EP. Breast cancer treatment: a review. JAMA. 2019;321(3):288–300.
15. Spick C, Pinker-Domenig K, Rudas M, Helbich TH, Baltzer PA. MRI-only lesions: application of diffusion-weighted imaging obviates unnecessary MR-guided breast biopsies. Eur Radiol. 2014;24(6):1204–10.
16. Grimm LJ, Zhang J, Mazurowski MA. Computational approach to radiogenomics of breast cancer: luminal A and luminal B molecular subtypes are associated with imaging features on routine breast MRI extracted using computer vision algorithms. J Magn Reson Imaging. 2015;42(4):902–7.
17. Waugh SA, Purdie CA, Jordan LB, Vinnicombe S, Lerski RA, Martin P, et al. Magnetic resonance imaging texture analysis classification of primary breast cancer. Eur Radiol. 2016;26(2):322–30.
18. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR imaging Radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. Radiology. 2016;281(2):382–91.
19. Aghaei F, Tan M, Hollingsworth AB, Zheng B. Applying a new quantitative global breast MRI feature analysis scheme to assess tumor response to chemotherapy. J Magn Reson Imaging. 2016;44(5):1099–106.
20. Dong Y, Feng Q, Yang W, Lu Z, Deng C, Zhang L, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. Eur Radiol. 2018;28(2):582–91.
21. Ha R, Mutasa S, Karcich J, Gupta N, Pascual Van Sant E, Nemer J, et al. Predicting breast cancer molecular subtype with MRI dataset utilizing convolutional neural network algorithm. J Digit Imaging. 2019;32(2):276–82.
22. Zhang Y, Chen JH, Lin Y, Chan S, Zhou J, Chow D, et al. Prediction of breast cancer molecular subtypes on DCE-MRI using convolutional neural network with transfer learning between two centers. Eur Radiol. 2021;31(4):2559–67.
23. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ, et al. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen international expert consensus on the primary therapy of early breast cancer 2011. Ann Oncol. 2011;22(8):1736–47.
24. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. IEEE Trans Pattern Anal Mach Intell. 2020;42(2):386–97.
25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837.
26. Zhang L, Le L, Nogues I, Summers RM, Liu S, Yao J. DeepPap: deep convolutional networks for cervical cell classification. IEEE J Biomed Health Inform. 2017;21(6):1633–43.

27. Elias SG, Adams A, Wisner DJ, Esserman LJ, van't Veer LJ, Mali WP, et al. Imaging features of HER2 overexpression in breast cancer: a systematic review and meta-analysis. Cancer Epidemiol Biomarkers Prev. 2014;23(8):1464–83.
28. Blaschke E, Abe H. MRI phenotype of breast cancer: kinetic assessment for molecular subtypes. J Magn Reson Imaging. 2015;42(4):920–4.
29. Youk JH, Son EJ, Chung J, Kim JA, Kim EK. Triple-negative invasive breast cancer on dynamic contrast-enhanced and diffusion-weighted MR imaging: comparison with other breast cancer subtypes. Eur Radiol. 2012;22(8):1724–34.
30. Uematsu T, Kasami M, Yuen S. Triple-negative breast cancer: correlation between MR imaging and pathologic findings. Radiology. 2009;250(3):638–47.
31. Zhu Z, Albadawy E, Saha A, Zhang J, Harowicz MR, Mazurowski MA. Deep learning for identifying radiogenomic associations in breast cancer. Comput Biol Med. 2019;109:85–90.
32. Fulford LG, Easton DF, Reis-Filho JS, Sofronis A, Gillett CE, Lakhani SR, et al. Specific morphological features predictive for the basal phenotype in grade 3 invasive ductal carcinoma of breast. Histopathology. 2006;49(1):22–34.
33. Liu S, Ren R, Chen Z, Wang Y, Fan T, Li C, et al. Diffusion-weighted imaging in assessing pathological response of tumor in breast cancer subtype to neoadjuvant chemotherapy. J Magn Reson Imaging. 2015;42(3):779–87.
34. Kim SY, Shin J, Kim DH, Kim EK, Moon HJ, Yoon JH, et al. Correlation between electrical conductivity and apparent diffusion coefficient in breast cancer: effect of necrosis on magnetic resonance imaging. Eur Radiol. 2018;28(8):3204–14.
35. Leithner D, Bernard-Davila B, Martinez DF, Horvat JV, Jochelson MS, Marino MA, et al. Radiomic signatures derived from diffusion-weighted imaging for the assessment of breast cancer receptor status and molecular subtypes. Mol Imaging Biol. 2020;22(2):453–61.
36. Wang Q, Mao N, Liu M, Shi Y, Ma H, Dong J, et al. Radiomic analysis on magnetic resonance diffusion weighted image in distinguishing triple-negative breast cancer from other subtypes: a feasibility study. Clin Imaging. 2021;72:136–41.
37. Yuen S, Monzawa S, Yanai S, Matsumoto H, Yata Y, Ichinose Y, et al. The association between MRI findings and breast cancer subtypes: focused on the combination patterns on diffusion-weighted and T2-weighted images. Breast Cancer. 2020;27(5):1029–37.
38. Li Q, Xiao Q, Yang M, Chai Q, Huang Y, Wu PY, et al. Histogram analysis of quantitative parameters from synthetic MRI: correlations with prognostic factors and molecular subtypes in invasive ductal breast cancer. Eur J Radiol. 2021;139:109697.
39. Ciriello G, Sinha R, Hoadley KA, Jacobsen AS, Reva B, Perou CM, et al. The molecular diversity of luminal A breast tumors. Breast Cancer Res Treat. 2013;141(3):409–20.
40. Taneja S, Evans AJ, Rakha EA, Green AR, Ball G, Ellis IO. The mammographic correlations of a new immunohistochemical classification of invasive breast cancer. Clin Radiol. 2008;63(11):1228–35.
41. Leong LC, Gombos EC, Jagadeesan J, Fook-Chong SM. MRI kinetics with volumetric analysis in correlation with hormonal receptor subtypes and histologic grade of invasive breast cancers. AJR Am J Roentgenol. 2015;204(3):W348–56.
42. Niu S, Jiang W, Zhao N, Jiang T, Dong Y, Luo Y, et al. Intra- and peritumoral radiomics on assessment of breast cancer molecular subtypes based on mammography and MRI. J Cancer Res Clin Oncol. 2022;148(1):97–106.
43. Leithner D, Horvat JV, Marino MA, Bernard-Davila B, Jochelson MS, Ochoa-Albiztegui RE, et al. Radiomic signatures with contrast-enhanced magnetic resonance imaging for the assessment of breast cancer receptor status and molecular subtypes: initial results. Breast Cancer Res. 2019;21(1):106.

How to cite this article: Yin H, Bai L, Jia H, Lin G. Noninvasive assessment of breast cancer molecular subtypes on multiparametric MRI using convolutional neural network with transfer learning. Thorac Cancer. 2022;13(22):3183–91. https://doi.org/10.1111/1759-7714.14673