



HAL
open science

A Channel-Dimensional Feature-Reconstructed Deep Learning Model for Predicting Breast Cancer Molecular Subtypes on Overall b -Value Diffusion-Weighted MRI

Xin-xiang Zhou, Lan Zhang, Quan-xiang Cui, Hui Li, Xi-qiao Sang, Hong-xia Zhang, Yue-Min Zhu, Zi-xiang Kuai

► **To cite this version:**

Xin-xiang Zhou, Lan Zhang, Quan-xiang Cui, Hui Li, Xi-qiao Sang, et al.. A Channel-Dimensional Feature-Reconstructed Deep Learning Model for Predicting Breast Cancer Molecular Subtypes on Overall b -Value Diffusion-Weighted MRI. *Journal of Magnetic Resonance Imaging*, In press, 10.1002/jmri.28895 . hal-04212764

HAL Id: hal-04212764

<https://hal.science/hal-04212764v1>

Submitted on 20 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Channel-Dimensional Feature-Reconstructed Deep Learning Model for Predicting Breast Cancer Molecular Subtypes on Overall *b*-Value Diffusion-Weighted MRI

X. X. Zhou, L. Zhang, Q. X. Cui, H. Li, X. Q. Sang, H. X. Zhang, Y. M. Zhu, Z. X. Kuai

Abstract

Background: Dynamic contrast-enhanced (DCE) MRI commonly outperforms diffusion-weighted (DW) MRI in breast cancer discrimination. However, the side effects of contrast agents limit the use of DCE-MRI, particularly in patients with chronic kidney disease.

Purpose: To develop a novel deep learning model to fully exploit the potential of overall *b*-value DW-MRI without the need for a contrast agent in predicting breast cancer molecular subtypes and to evaluate its performance in comparison with DCE-MRI.

Study Type: Prospective.

Subjects: 486 female breast cancer patients (training/validation/test: 64%/16%/20%).

Field Strength/Sequence: 3.0 T / DW-MRI (13 *b*-values) and DCE-MRI (one pre-contrast and five post-contrast phases).

Assessment: The breast cancers were divided into four subtypes: luminal A, luminal B, HER2+, and triple negative. A channel-dimensional feature-reconstructed (CDFR) deep neural network (DNN) was proposed to predict these subtypes using pathological diagnosis as the reference standard. Additionally, a non-CDFR DNN (NCDFR-DNN) was built for comparative purposes. A mixture ensemble DNN (ME-DNN) integrating two CDFR-DNNs was constructed to identify subtypes on multiparametric MRI (MP-MRI) combining DW-MRI and DCE-MRI.

Statistical Tests: Model performance was evaluated using accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). Model comparisons were performed using the one-way analysis of variance with least significant difference post hoc test and the DeLong test. $P < 0.05$ was considered significant.

Results: The CDFR-DNN (accuracies, 0.79~0.80; AUCs, 0.93~0.94) demonstrated significantly improved predictive performance than the NCDFR-DNN (accuracies, 0.76~0.78; AUCs, 0.92~0.93) on DW-MRI. Utilizing the CDFR-DNN, DW-MRI attained the predictive performance equal ($P = 0.065 \sim 1.000$) to DCE-MRI (accuracies, 0.79~0.80; AUCs, 0.93~0.95). The predictive performance of the ME-DNN on MP-MRI (accuracies, 0.85~0.87; AUCs, 0.96~0.97) was superior to those of both the CDFR-DNN and NCDFR-DNN on either DW-MRI or DCE-MRI.

Data Conclusion: The CDFR-DNN enabled overall *b*-value DW-MRI to achieve the predictive performance comparable to DCE-MRI. MP-MRI outperformed DW-MRI and DCE-MRI in subtype prediction.

Keywords

Breast cancer; Molecular subtype; Deep learning; Diffusion-weighted; Dynamic contrast-enhanced; Magnetic resonance imaging

INTRODUCTION

Breast cancer is heterogeneous at the molecular level, which leads to variations in clinical outcomes.^{1,2} Thus, a pretherapeutic determination for molecular subtypes is needed for patient-tailored managements. Clinically, the identification of molecular subtypes before treatment is mainly based on gene expression profiling or immunohistochemical (IHC) surrogates from biopsy samples.^{3,4} However, the invasive testing frequently causes patients tremendous discomfort and can only capture a snapshot of a heterogenous tumor vulnerable to sampling bias.⁵ Therefore, a better alternative method is required to noninvasively identify molecular subtypes of breast tumor in its entirety.

The rapid developments of radiogenomics that links medical imaging markers to tumor genotypic configurations provides an opportunity for detecting molecular mechanism in a non-invasive way.^{6,7} Meanwhile, advances in computer technology have brought deep learning capable of automatically extracting high-dimensional quantitative features from radiological images to the forefront of undertaking complex clinical challenges.⁸ Indeed, recent application of deep learning to assess radiogenomics of breast dynamic contrast-enhanced (DCE) MRI has yielded encouraging results.⁹⁻¹² As an important MRI technique, DCE-MRI not only has high spatial resolution and interobserver reproducibility, but it can also provide pharmacokinetic data.¹³ Nevertheless, the side effects and contraindications of gadolinium-based contrast agents to some extent limit the application of DCE-MRI.¹⁴

Various MRI approaches without the use of contrast agents have recently gained attention in the research focused on the prediction of subtype of breast cancer.¹⁵⁻¹⁷ Among these imaging approaches, diffusion-weighted (DW) MRI with multiple low and/or high b -values has emerged as the most promising technique.^{16,18} Some important biomarkers (e.g., microstructure, heterogeneity and perfusion) for identifying molecular subtypes can be derived from multi- b -value DW signals through establishing suitable mathematical models, such as the intravoxel incoherent motion, diffusion kurtosis and stretched exponential models.^{16,19,20} Considering extraordinary capacity of deep learning in data mining, it could be expected that the performance of DW-MRI on subtype prediction may be further improved by constructing appropriate deep learning model and be comparable to that of DCE-MRI.

Therefore, this study aimed to develop a novel deep learning model to fully exploit the potential of DW-MRI covering a broad spectrum of b -values (including both low and high b -values) in predicting breast cancer molecular subtypes and to compare its predictive performance to that of DCE-MRI. Additionally, the study aimed to investigate whether a multiparametric MRI (MP-MRI) protocol with the combination of the two imaging techniques has improved performance in the prediction of breast cancer molecular subtypes than any imaging approach used alone.

MATERIALS AND METHODS

Subjects

This prospective study was approved by our institutional review board and written informed consent was obtained. Between July 2018 and December 2021, consecutive 596 female patients were recruited with the following inclusion criteria: 1) aged 18 years or older; 2) diagnosed with breast cancer at mammography or ultrasonography. One hundred and ten patients were excluded from recruitment with the following exclusion criteria: 1) preoperative interventions and therapies ($n=22$); 2) poor image quality caused by obvious artifacts ($n=3$); 3) the maximum tumor diameter lower than 5 mm ($n=18$); 4) incomplete histopathologic results ($n=67$). For patients with multiple lesions in the ipsilateral breast, the largest lesion was selected for analysis. Six patients had simultaneous bilateral lesions and each lesion was evaluated separately. Finally, a total of 486 patients with 492 lesions were enrolled in the

study (Figure 1).

MRI Protocol

For all enrolled patients, MRI was performed on a 3.0T scanner (Ingenia, Philips) with a 7-channel breast coil. DW images were first acquired using a fat suppressed single-shot spin-echo echo planar imaging sequence with monopolar diffusion-encoding gradients [repetition time (TR)/echo time (TE)=6443/77 msec, flip angle=90°, field of view (FOV)=324×324 mm², reconstruction matrix size=352×352, slice thickness/gap=5/1 mm, b -values=0, 10, 25, 50, 75, 100, 200, 400, 600, 800, 1000, 1500 and 2000 s/mm²]. Then, DCE images were acquired using an enhanced T1 high resolution isotropic volume excitation (e-THRIVE) sequence with fat suppression [TR/TE=4.8/2.1 msec, flip angle=12°, FOV=350×350 mm², reconstruction matrix size=784×784, slice thickness/gap=1/0 mm]. The DCE images consisted of one pre-contrast and five post-contrast phases. The contrast agent (Gadovist, Bayer; 0.1 mmol/kg and 2 mL/sec) was injected after the pre-contrast images were acquired. Total scan times were 6 minutes and 46 seconds for DW-MRI and 8 minutes and 57 seconds for DCE-MRI.

Lesion Annotation

All lesions were annotated by two board-certified radiologists (L Z and Q-X C; 6 and 8 years of breast diagnosis experience, respectively) for corresponding molecular subtypes based on expression of IHC surrogates [estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67] from surgical pathological results. An Allred score from the IHC greater than 2 was considered to be positive for ER and PR. For the determination of HER2 status, an IHC HER2 score of 3+, or a score of 2+ with an additional condition of HER2 gene amplification by the fluorescence in situ hybridization was considered to be positive. Positive Ki-67 was defined as expression ≥14%. In this study, molecular subtypes were categorized into luminal A (ER and/or PR+, HER2-, Ki-67-), luminal B (ER and/or PR+, HER2-, Ki-67+), HER2+ and triple negative (TN) (ER and PR-, HER2-) subtypes.^{21,22}

Lesion Segmentation

Image registration was performed between different b -value images of DW-MRI and between different dynamic phase images of DCE-MRI for each lesion using an efficient subpixel image rigid transformation algorithm.²³ This image registration algorithm was carried out in MATLAB software (v. R2014a; MathWorks, Natick, MA, USA) with the use of openly available code at the website: <https://www.mathworks.cn/matlabcentral/fileexchange/18401-efficient-subpixel-image-registration-by-cross-correlation>. The lesion volumes of interest (VOIs) were extracted by stacking up regions of interest (ROIs) manually delineated in slice-by-slice manner from both the DW and DCE images that were best suited to view lesion boundaries. The extracted VOIs were propagated to other b -value images and dynamic phase images on which the lesion had not directly been segmented. The VOI delineation was performed by two radiologists (H L and Z-X K; with 5 and 7 years of breast MRI experience, respectively) using 3D slicer software (v.4.11.20210226; Surgical Planning Laboratory, Brigham and Women's Hospital, Boston, MA, USA). Both radiologists were blinded to the pathologic results. A third radiologist (H-X Z; 24 years of breast MRI experience) supervised the process and made final decisions in cases of disagreement. A representative segmentation was shown in Figure 2.

Image Preprocessing

Given the limited number of enrolled lesions in the multi-class task based on deep learning, the slice of each lesion was used as an independent sample. Considering that more slices per lesion were acquired in DCE-MRI compared to DW-MRI, only some slices were sampled from DCE-MRI to ensure the same sample size as in DW-MRI.

The sampling process involved the following steps. First, the remainder (m) of dividing

the number of DCE-MRI slices by the number of DW-MRI slices (n) for one lesion was calculated. Second, m slices were randomly removed from the top and/or bottom of DCE-MRI slice stack. The remaining slices were then divided equally into n sub-stacks, each corresponding to a DW-MRI slice. Third, the previously removed DCE-MRI slices were inserted back into their adjacent sub-stacks. Finally, from each sub-stack, a slice that was closest to the corresponding DW-MRI slice within the ROI size (mm^2) was extracted. After completing the above steps for all lesions, the total number of sampled DCE-MRI slices was exactly equal to the number of DW-MRI slices.

For each sample slice, the DW and DCE images were cropped around the center of the lesion to create a series of fixed-size square patches (250×250 pixels for DCE-MRI and 124×124 pixels for DW-MRI), respectively. The square patch sizes were determined according to the diameter of the largest lesion among all lesions. The square patches across different b -values and different dynamic phase images were normalized by logarithmic transformation and min-max normalization, respectively.

Several real-time data augmentation methods were implemented on the square patches to prevent overfitting, including: (1) random rotation between 0° and 90° , (2) random width or height shifting in 10%, (3) random shearing of 0.1, and (4) random flipping in horizontal and vertical axes. Ultimately, the square patches were resized to 32×32 pixels and were normalized to have mean=0 and standard deviation=1.

Deep Neural Network (DNN)

A novel channel-dimensional feature-reconstructed (CDFR) DNN was proposed to predict molecular subtypes on DW-MRI or DCE-MRI. An overview of this network was exhibited in Figure 3. The CDFR-DNN consisted of an input layer, two CDFR convolutional block, a conventional convolutional block, and an output layer, with a total of 14 convolutional layers and one fully-connected (FC) layer. The input size was $32 \times 32 \times 13$ for DW-MRI or $32 \times 32 \times 6$ for DCE-MRI. The CDFR convolutional block was the core of network and was composed of two parts: the extracted and the reconstructed part. The size of all convolutional kernels in the reconstructed part was set as $1 \times 1 \times n$ (where n is the number of channels) to focus on the channel-dimensional feature reconstruction. The mean-squared error was computed between the input of extracted part and the output of reconstructed part and used as a CDFR loss in the training of network. After the CDFR module, a 2×2 max pooling operation was implemented to down-sampling the feature maps. Each convolutional layer was followed by a batch normalization layer and an exponential linear unit activation layer to ease optimization and regularize the network. At the end of network, a global max pooling layer, a 4-way FC layer, and a SoftMax layer were utilized to produce the final likelihood values.

For optimizing the CDFR-DNN, a joint training on the sum of a categorical cross-entropy loss and two CDFR losses was implemented by using an Adam optimizer with a learning rate of 0.001 over a fixed number of 1500 epochs. The weights of network were initialized with ‘‘He Uniform Variance Scaling Initializer Method’’.²⁴ The batch size was set as 32. L2 regularization with a limited squared magnitude of the kernel weights was used to prevent overfitting. To assess the CDFR effect on the subtype prediction, the reconstructed parts were removed from the network architecture. This pruned network (referred to as non-CDFR DNN or NCDFR-DNN for short) was then compared with the CDFR-DNN in terms of the performance of predicting breast cancer molecular subtypes using pathological diagnosis as the reference standard. The NCDFR-DNN was trained only on the categorical cross-entropy loss. All of the above trainings were conducted on DW-MRI and DCE-MRI, respectively.

For the subtype prediction on MP-MRI, a mixture ensemble DNN (ME-DNN) integrating two CDFR-DNNs was constructed, as illustrated in Figure 4. The ME-DNN was constructed through integrating the two well-trained CDFR-DNNs on DW-MRI and DCE-MRI respectively at the level of global features. Specifically, the two feature vectors generated by

the global max pooling operations of the two CDFR-DNNs were concatenated together and then fed as a whole into a new output layer that contained three wide FC layers and a SoftMax layer with four output nodes producing the class probabilities.

For optimizing the ME-DNN on MR-MRI, the new introduced output layer was firstly trained by using an Adam optimizer with 0.001 learning rate for 20 epochs. Subsequently, an end-to-end fine-tuning on the entire network with a learning rate of 0.0001 for 20 epochs was carried out to achieve the best performance.

Code for modeling was written in Python with open-source deep learning library PyTorch (version 1.7.1). The model training was run on a windows 10 workstation with NVIDIA GeForce RTX3090 GPU (24 GB), Intel Core-i9 10900k CPU (3.70GHz) and 64GB RAM.

DNN Evaluation

An independent test set was created by extracting 20% of cases from each molecular subtype group. The remaining cases were used to train DNN with an 80%-20% training-validation split and fivefold cross-validation (CV) where each fold roughly had the same ratio of the four subtypes.

Statistical Analyses

All statistical analyses were conducted using SPSS Statistics (version 25.0; IBM, Chicago, IL, USA) and R 4.0.2. The differences in patient demographics between subtypes were evaluated using the Kruskal-Wallis H, chi-square or Fisher's exact test, according to distribution of variables and number of cases. Model performance was quantified by calculating the accuracy in 4-way subtype classification, and the sensitivity, specificity and area under the receiver-operating characteristic (ROC) curve (AUC) in binary subtype classification (Luminal A vs. non-Luminal A; Luminal B vs. non-Luminal B; HER2+ vs. HER2-; and TN vs. non-TN). The 95% confidence intervals (CIs) for AUC were calculated using a 10^6 -sample bootstrapping method, while for sensitivity and specificity using the Wilson Score method. The model performances were compared in 4-way subtype classification and in binary subtype classification by using the one-way analysis of variance (ANOVA) with least significant difference (LSD) post hoc test and the DeLong test, respectively. $P < 0.05$ was considered to be statistically significant.

RESULTS

The clinical-pathologic characteristics of all patients were listed in Table 1. The Kruskal-Wallis H test showed that there were no significant differences between subtypes for the patient ages ($P=0.532$) and lesion diameters ($P=0.178$). Likewise, no significant differences were found in the categorical characteristics between subtypes ($P=0.080\sim 0.621$) by the chi-square or Fisher's exact test, excepting the pathological type.

Table 2 summarizes the 4-way classification accuracy of models for differentiating subtypes on DW-MRI, DCE-MRI, and MP-MRI using the CV. Table 3 lists the comparative results between models in the independent testing set. The CDFR-DNN (accuracies, 0.79~0.80) demonstrated significantly improved classification accuracy relative to the NCDFR-DNN (accuracies, 0.76~0.78) on DW-MRI, although there was no evidence of a difference in the classification accuracy between the two models on DCE-MRI ($P=0.477$). Meanwhile, the CDFR-DNN yielded comparable classification accuracy ($P=1.000$) on DW-MRI (accuracies, 0.79~0.80) than on DCE-MRI (accuracies, 0.79~0.80). When evaluating the predictive performance of the ME-DNN on MP-MRI (accuracies, 0.85~0.87), which showed significantly higher classification accuracy compared with both the CDFR-DNN and the NCDFR-DNN on either DW-MRI or DCE-MRI. In addition, a significant difference of classification accuracy was also identified among all of models by the one-way ANOVA.

Table 4 reports the sensitivity, specificity and AUC of models on the DW-MRI, DCE-MRI and MP-MRI independent tests for binary subtype classification. The ROC curves were

depicted in Figure 5. The comparison of performance between models were presented in Figure 6. It can be observed that the CDFR-DNN (AUC, 0.93~0.94) was significantly superior to the NCDFR-DNN (AUC, 0.92~0.93) on DW-MRI for luminal A vs. non-luminal A, luminal B vs. non-luminal B, and HER2+ vs. non-HER2+ subtype expecting a slight improvement without significant difference ($P=0.083$) for TN vs. non-TN subtype. On DCE-MRI, there were no significant differences ($P=0.266\sim 0.800$) between the two models for the binary classification. When comparing the performance of models between different imaging approaches, the CDFR-DNN did not demonstrate significant difference ($P=0.065\sim 0.849$) between DW-MRI and DCE-MRI for binary subtype classification. As for the ME-DNN on MP-MRI, it significantly outperformed both the CDFR-DNN and the NCDFR-DNN on either DW-MRI or DCE-MRI.

DISCUSSION

In this study, a novel deep learning model (CDFR-DNN) was developed to predict breast cancer molecular subtypes. The CDFR-DNN demonstrated comparable performance to DCE-MRI in molecular subtype prediction, using overall b -value DW-MRI without the need for a contrast agent. Besides, MP-MRI demonstrated significantly improved predictive performance than either DW-MRI or DCE-MRI through ensemble learning which integrated two CDFR-DNNs at the level of global features.

Both DW-MRI and DCE-MRI are capable of producing multi-channel images, with adjustable b -values and phases, respectively, and the image data are closely related to one another in the channel dimension.²⁵⁻²⁸ Moreover, it has been shown in the literature that the variations in image intensities across channels on both DW-MRI and DCE-MRI were useful for revealing the underlying characteristics of tumor.²⁵⁻²⁸ Given this, the present study incorporated the CDFR mechanism into the deep learning model. Since all convolutional kernel sizes were $1\times 1\times n$ (the number of channels) in the reconstructed parts, the channel-dimensional features could be retained as more as possible in the feature maps generated by the extracted parts through minimizing the CDFR losses and were ultimately used by the classifier to differentiate molecular subtypes.

Indeed, the CDFR-DNN was significantly superior to the NCDFR-DNN for subtype prediction on DW-MRI. However, it can be also noted that the CDFR effect was not significant on DCE-MRI. This might be attributed to the smaller number of input channels on DCE-MRI (6 channels) compared to DW-MRI (13 channels). Future studies might compare the predictive performances between the CDFR-DNN and the NCDFR-DNN on DCE-MRI with more dynamic phases. However, such comparison was beyond the scope of the present work. In addition, for correlating signal intensities of each pixel across the DW-MRI or DCE-MRI channels, 1×1 convolutions were adopted in the extracted part behind the input layer.

The improved predictive performance on MP-MRI suggested that DW-MRI and DCE-MRI may be complementary in their ability to identify molecular subtypes. This agrees with the literature, where it has been found that DCE-MRI can provide insight into angiogenesis, perfusion and vessel permeability in cancers, while DW-MRI can probe tumor microstructure and heterogeneity.²⁸⁻³¹ These biomarkers have been shown to be valuable for classifying breast cancer molecular subtypes.^{16,18,32} Thus, in clinical practice, combining DW-MRI and DCE-MRI could be an option to maximize classification performance for individuals without a history of gadolinium allergy or chronic kidney disease.

DCE-MRI has been utilized to distinguish molecular subtypes in recent deep learning studies.^{9,10,12} According to Sun et al., DCE-MRI reached an accuracy of 82.6% and an AUC of 0.836 for the differentiation between luminal and non-luminal subtypes through transfer learning.⁹ In the research by Ha et al., DCE-MRI attained an accuracy of 70% for classifying luminal A, luminal B and HER2+ subtypes with a DNN made up of a series of residual and

inception style layers.¹² Additionally, Zhang et al. reported an accuracy of 0.82 for the identification of hormonal receptor positive/HER2-, HER2+, and TN subtypes on an external DCE-MRI validation set.¹⁰ The results obtained from these studies were comparable to those obtained in this work on DCE-MRI.

Limitations

First, this was a single-center study. Further investigations on a multi-institutional cohort are needed to ensure the generalizability of the proposed models. Second, all images were acquired using the same MRI scanner. The performance of the models should be further evaluated on the MRI data from different vendors. The third limitation is a relatively small sample size for the TN subtype. However, this is a common problem for all breast cancer molecular subtype classification studies due to the fact that the distribution of subtypes in the breast cancer population is inherently imbalanced.¹⁰ Finally, this work used IHC surrogate to define the molecular subtypes instead of utilizing genetic analysis.

Conclusion

The proposed novel channel-dimensional feature-reconstructed deep learning model enabled overall b -value DW-MRI to achieve the performance comparable to DCE-MRI in the prediction of breast cancer molecular subtypes. Therefore, overall b -value DW-MRI can be considered as an alternative approach to identify molecular subtypes without the need for contrast agents. In addition, MP-MRI resulting from the combination of DCE-MRI and NME-DWI has demonstrated much greater discriminatory ability than either imaging technique used individually and it might have potential implementation into the standard MRI protocol to maximize the diagnostic accuracy of breast cancer subtype classification in patients where contrast agent is not contraindicated.

DATA AVAILABILITY

Data and code that support the findings of this prospective study are **openly available** at <https://www.kaggle.com/datasets/starzh10/breastcancer-subtypes> and https://github.com/ZH10/DWI_DCE_CDFR-DNN, respectively.

REFERENCES

- [1] Haynes B, Sarma A, Nangia-Makker P, Shekhar MP. Breast cancer complexity: implications of intratumoral heterogeneity in clinical management. *Cancer Metastasis Rev.* 2017;36:547-555.
- [2] Martelotto LG, Ng CK, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. *Breast Cancer Res.* 2014;16:210.
- [3] Guiu S, Michiels S, André F, et al. Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. *Ann Oncol.* 2012;23:2997-3006.
- [4] Sotiriou C, Neo S-Y, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proceedings of the National Academy of Sciences.* 2003;100:10393-10398.
- [5] Lee JY, Lee K-s, Seo BK, et al. Radiomic machine learning for predicting prognostic biomarkers and molecular subtypes of breast cancer using tumor heterogeneity and angiogenesis properties on MRI. *Eur Radiol.* 2022;32:650-660.
- [6] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006.
- [7] Ashraf AB, Daye D, Gavenonis S, et al. Identification of intrinsic imaging phenotypes for breast cancer tumors: preliminary associations with gene expression profiles. *Radiology.* 2014;272:374-384.
- [8] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin.* 2019;69:127-157.
- [9] Sun R, Hou X, Li X, Xie Y, Nie S. Transfer learning strategy based on unsupervised learning and ensemble learning for breast cancer molecular subtype prediction using dynamic contrast-enhanced MRI. *J Magn Reson Imaging.* 2022;55:1518-1534.
- [10] Zhang Y, Chen J-H, Lin Y, et al. Prediction of breast cancer molecular subtypes on DCE-MRI using convolutional neural network with transfer learning between two centers. *Eur Radiol.* 2021;31:2559-2567.
- [11] Sun R, Meng Z, Hou X, et al. Prediction of breast cancer molecular subtypes using DCE-MRI based on CNNs combined with ensemble learning. *Phys Med Biol.* 2021;66:175009.
- [12] Ha R, Mutasa S, Karcich J, et al. Predicting breast cancer molecular subtype with MRI dataset utilizing convolutional neural network algorithm. *J Digit Imaging.* 2019;32:276-282.
- [13] Hu Q, Whitney HM, Li H, Ji Y, Liu P, Giger ML. Improved classification of benign and malignant breast lesions using deep feature maximum intensity projection MRI in breast cancer diagnosis using dynamic contrast-enhanced MRI. *Radiology: Artificial Intelligence.* 2021;3:e200159.
- [14] Runge VM. Safety of the Gadolinium-Based Contrast Agents for Magnetic Resonance Imaging, Focusing in Part on Their Accumulation in the Brain and Especially the Dentate Nucleus. *Invest Radiol.* 2016;51:273-279.
- [15] Horvat JV, Bernard- Davila B, Helbich TH, et al. Diffusion-weighted imaging (DWI) with apparent diffusion coefficient (ADC) mapping as a quantitative imaging biomarker for prediction of immunohistochemical receptor status, proliferation rate, and molecular subtypes of breast cancer. *J Magn Reson Imaging.* 2019;50:836-846.
- [16] Suo S, Cheng F, Cao M, et al. Multiparametric diffusion-weighted imaging in breast lesions: Association with pathologic diagnosis and prognostic factors. *J Magn Reson Imaging.* 2017;46:740-750.
- [17] Gao W, Yang Q, Li X, et al. Synthetic MRI with quantitative mappings for identifying receptor status, proliferation rate, and molecular subtypes of breast cancer. *Eur J Radiol.*

- 2022;148:110168.
- [18] Song SE, Cho KR, Seo BK, et al. Intravoxel incoherent motion diffusion- weighted MRI of invasive breast cancer: Correlation with prognostic factors and kinetic features acquired with computer- aided diagnosis. *J Magn Reson Imaging*. 2019;49:118-130.
- [19] Cho GY, Moy L, Kim SG, et al. Evaluation of breast cancer using intravoxel incoherent motion (IVIM) histogram analysis: comparison with malignant status, histological subtype, and molecular prognostic factors. *Eur Radiol*. 2016;26:2547-2558.
- [20] Yang ZL, Li Y, Zhan CA, et al. Evaluation of suspicious breast lesions with diffusion kurtosis MR imaging and connection with prognostic factors. *Eur J Radiol*. 2021;145:110014.
- [21] Hammond MEH, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med*. 2010;134:e48-e72.
- [22] Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline focused update. *Arch Pathol Lab Med*. 2018;142:1364-1382.
- [23] Guizar-Sicairos M, Thurman ST, Fienup JR. Efficient subpixel image registration algorithms. *Opt Lett*. 2008;33:156-158.
- [24] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*. Santiago; December 7-13, 2015;pp 1026-1034.
- [25] Lee YJ, Kim SH, Kang BJ, et al. Intravoxel incoherent motion (IVIM)- derived parameters in diffusion- weighted MRI: Associations with prognostic factors in invasive ductal carcinoma. *J Magn Reson Imaging*. 2017;45:1394-1406.
- [26] Kang HS, Kim JY, Kim JJ, et al. Diffusion kurtosis MR imaging of invasive breast cancer: Correlations with prognostic factors and molecular subtypes. *J Magn Reson Imaging*. 2022;56:110-120.
- [27] Agner SC, Rosen MA, Englander S, et al. Computerized image analysis for identifying triple-negative breast cancers and differentiating them from other molecular subtypes of breast cancer on dynamic contrast-enhanced MR images: a feasibility study. *Radiology*. 2014;272:91-99.
- [28] Saha A, Harowicz MR, Grimm LJ, et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br J Cancer*. 2018;119:508-516.
- [29] Le BD, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*. 1988;168:497-505.
- [30] Jensen JH, Helpert JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn Reson Med*. 2005;53:1432-1440.
- [31] Bennett KM, Schmainda KM, Bennett RT, Rowe DB, Lu H, Hyde JS. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn Reson Med*. 2003;50:727-734.
- [32] Leithner D, Horvat JV, Marino MA, et al. Radiomic signatures with contrast-enhanced magnetic resonance imaging for the assessment of breast cancer receptor status and molecular subtypes: initial results. *Breast Cancer Res*. 2019;21:106.

Table 1 Clinical and pathologic characteristics of 492 breast cancers from 486 patients.

Characteristics	Luminal A	Luminal B	HER2+	TN
Totals	85	148	184	75
Age (years)	50±5	52±7	48±11	54±9
Lesion diameter (mm)	34±21	33±22	33±23	34±23
BI-RADS				
IV	17 (20.0%)	36 (24.3%)	29 (15.7%)	12 (16.0%)
V	23 (27.1%)	42 (28.4%)	45 (24.5%)	19 (25.3%)
VI	45 (52.9%)	70 (47.3%)	110 (59.8%)	44 (58.7%)
Pathological type				
Invasive ductal cancer	59 (69.4%)	122 (82.5%)	85 (46.2%)	50 (66.7%)
Invasive lobular cancer	11 (12.9%)	11 (7.4%)	63 (34.2%)	19 (25.3%)
Mixed invasive ductal and lobular cancer	14 (16.5%)	12 (8.1%)	25 (13.6%)	0 (0.0%)
Mucinous cancer	1 (1.2%)	3 (2.0%)	0 (0.0%)	0 (0.0%)
Others	0 (0.0%)	0 (0.0%)	11 (6.0%)	6 (8.0%)
Grade				
I	13 (15.3%)	23 (15.5%)	49 (26.6%)	15 (20.0%)
II	56 (65.9%)	84 (56.8%)	98 (53.3%)	40 (53.3%)
III	16 (18.8%)	41 (27.7%)	37 (20.1%)	20 (26.7%)
Affected side				
Right	43 (50.6%)	71 (48.0%)	95 (51.6%)	43 (57.3%)
Left	42 (49.4%)	77 (52.0%)	89 (48.4%)	32 (42.7%)

Note. — Unless otherwise indicated, data in parentheses are percentages. Mean data are expressed with standard deviations. Others include invasive papillary carcinoma, invasive micropapillary carcinoma and metaplastic carcinoma. BI-RADS = breast imaging-reporting and data system, HER2+ = human epidermal growth factor receptor 2 over-expression, TN = triple negative.

Table 2 The 4-way classification accuracy of models on the DW-MRI, DCE-MRI, and MP-DWI training/validation/independent testing sets at each fold of cross-validation.

Fold	DW-MRI						DCE-MRI						MP-MRI		
	CDFR-DNN			NCDFR-DNN			CDFR-DNN			NCDFR-DNN			ME-DNN		
	Train	Val	Test	Train	Val	Test	Train	Val	Test	Train	Val	Test	Train	Val	Test
1	0.99	0.80	0.80	0.99	0.78	0.76	0.98	0.80	0.80	0.98	0.83	0.81	1.00	0.87	0.86
2	0.99	0.82	0.80	0.99	0.81	0.78	0.98	0.80	0.79	0.98	0.80	0.80	1.00	0.87	0.86
3	0.99	0.81	0.79	0.99	0.80	0.77	0.98	0.80	0.80	0.98	0.79	0.78	1.00	0.88	0.85
4	0.98	0.80	0.79	0.98	0.83	0.76	0.98	0.79	0.80	0.97	0.81	0.78	1.00	0.86	0.87
5	0.99	0.80	0.80	0.99	0.81	0.78	0.98	0.82	0.79	0.98	0.81	0.79	1.00	0.87	0.86
Mean	0.99	0.81	0.80	0.99	0.81	0.77	0.98	0.80	0.80	0.98	0.81	0.79	1.00	0.87	0.86

CDFR = channel-dimensional feature-reconstructed, DCE = dynamic contrast-enhanced, DNN = deep neural network, DW = diffusion-weighted, ME = mixture ensemble, MP = multiparametric, NCDFR = non-CDFR, Val = validation.

Table 3 The comparison (P value) of 4-way classification accuracy between models on DW-MRI, DCE-MRI and MP-MRI.

Imaging	Network	DW-MRI		DCE-MRI		MP-MRI
		CDFR-DNN	NCDFR-DNN	CDFR-DNN	NCDFR-DNN	ME-DNN
DW-MRI	CDFR-DNN	1.000	<0.001	1.000	0.477	<0.001
	NCDFR-DNN	-	1.000	<0.001	0.001	<0.001
DCE-MRI	CDFR-DNN	-	-	1.000	0.477	<0.001
	NCDFR-DNN	-	-	-	1.000	<0.001
MP-MRI	ME-DNN	-	-	-	-	1.000

A P value less than .05 was considered statistically significant. CDFR = channel-dimensional feature-reconstructed, DCE = dynamic contrast-enhanced, DNN = deep neural network, DW = diffusion-weighted, ME = mixture ensemble, MP = multiparametric, NCDFR = non-CDFR.

Table 4 The binary classification performances of models in the DW-MRI, DCE-MRI and MP-MRI independent test sets.

	Imaging	Network	Luminal A vs. non-Luminal A	Luminal B vs. non-Luminal B	HER2+ vs. HER2-	TN vs. non-TN
Sensitivity	DW-MRI	CDFR-DNN	0.73 [0.70, 0.75]	0.79 [0.77, 0.81]	0.84 [0.83, 0.86]	0.76 [0.73, 0.78]
		NCDFR-DNN	0.68 [0.65, 0.71]	0.76 [0.74, 0.78]	0.83 [0.82, 0.85]	0.70 [0.67, 0.73]
	DCE-MRI	CDFR-DNN	0.73 [0.70, 0.75]	0.80 [0.78, 0.81]	0.86 [0.84, 0.87]	0.72 [0.70, 0.75]
		NCDFR-DNN	0.74 [0.71, 0.76]	0.78 [0.76, 0.80]	0.86 [0.84, 0.87]	0.71 [0.68, 0.73]
	MP-MRI	ME-DNN	0.84 [0.81, 0.86]	0.85 [0.83, 0.86]	0.89 [0.88, 0.90]	0.83 [0.80, 0.85]
Specificity	DW-MRI	CDFR-DNN	0.96 [0.95, 0.97]	0.91 [0.90, 0.92]	0.88 [0.87, 0.89]	0.96 [0.95, 0.97]
		NCDFR-DNN	0.95 [0.94, 0.96]	0.91 [0.89, 0.92]	0.86 [0.85, 0.87]	0.96 [0.95, 0.97]
	DCE-MRI	CDFR-DNN	0.97 [0.96, 0.98]	0.90 [0.89, 0.91]	0.87 [0.86, 0.88]	0.97 [0.96, 0.98]
		NCDFR-DNN	0.96 [0.94, 0.97]	0.91 [0.90, 0.92]	0.86 [0.85, 0.87]	0.97 [0.96, 0.98]
		MP-MRI	ME-DNN	0.97 [0.96, 0.98]	0.94 [0.93, 0.95]	0.92 [0.91, 0.93]
AUC	DW-MRI	CDFR-DNN	0.94 [0.93, 0.95]	0.93 [0.92, 0.94]	0.93 [0.92, 0.94]	0.94 [0.93, 0.95]
		NCDFR-DNN	0.93 [0.92, 0.94]	0.92 [0.91, 0.93]	0.92 [0.91, 0.93]	0.93 [0.92, 0.94]
	DCE-MRI	CDFR-DNN	0.95 [0.94, 0.96]	0.93 [0.92, 0.93]	0.93 [0.92, 0.94]	0.93 [0.92, 0.94]
		NCDFR-DNN	0.94 [0.92, 0.95]	0.92 [0.91, 0.93]	0.93 [0.92, 0.94]	0.93 [0.92, 0.94]
		MP-MRI	ME-DNN	0.97 [0.96, 0.98]	0.96 [0.94, 0.97]	0.97 [0.96, 0.98]

Note. — Unless otherwise indicated, data in square brackets are 95% confidence intervals. CDFR = channel-dimensional feature-reconstructed, DCE = dynamic contrast-enhanced, DNN = deep neural network, DW = diffusion-weighted, HER2+ = human epidermal growth factor receptor 2 over-expression, ME = mixture ensemble, MP = multiparametric, NCDFR = non-CDFR, TN = triple negative.

Figure 1 Flowchart of study population with exclusion criteria. DCE = dynamic contrast-enhanced, DW = diffusion-weighted.

Figure 2 A segmented case example from a 59-year-old woman with luminal B breast cancer in the left breast. (A) DW image at $b = 0$ s/mm², (B) Square patch and lesion boundary on the $b = 0$ s/mm² image. The square patch is centered at the centroid of the lesion. (C~O) Color-coded square patches of b -value = 0, 10, 25, 50, 75, 100, 200, 400, 600, 800, 1000, 1500, 2000 s/mm². (P) Pre-contrast DCE image. (Q) Square patch and lesion boundary on the pre-contrast DCE image. The square patch is centered at the centroid of the lesion. (R~W) Color-coded square patches of one pre-contrast and five post-contrast phases.

Figure 3 A detailed description of the CDFR-DNN. The CDFR-DNN consisted of an input layer, two CDFR convolutional block, a conventional convolutional block, and an output layer, with a total of 14 convolutional layers and one fully-connected layer. The numbers of the input channels are 13 and 6 for DW-MRI and DCE-MRI, respectively. BN = batch normalization, CDFR = channel-dimensional feature-reconstructed, Conv = convolution, DNN = deep neural network, ELU = exponential linear unit.

Figure 4 Architecture of the ME-DNN, which integrates two well-trained CDFR-DNNs on DW-MRI and DCE-MRI respectively at the level of global features, for predicting breast cancer molecular subtypes on MP-MRI combining DW-MRI and DCE-MRI. CDFR = channel-dimensional feature-reconstructed, DCE = dynamic contrast-enhanced, DNN = deep neural network, DW = diffusion-weighted, HER2+ = human epidermal growth factor receptor 2 over-expression, ME = mixture ensemble, MP = multiparametric, TN = triple negative.

Figure 5 The ROC curves of models in binary molecular subtype classification. CDFR = channel-dimensional feature-reconstructed, DCE = dynamic contrast-enhanced, DNN = deep neural network, DW = diffusion-weighted, HER2+ = human epidermal growth factor receptor 2 over-expression, ME = mixture ensemble, MP = multiparametric, NCDFR = non-CDFR, ROC = receiver-operating characteristic, TN = triple negative.

Figure 6 The comparison (P value) of model performances in binary molecular subtype classification. CDFR = channel-dimensional feature-reconstructed, DCE = dynamic contrast-enhanced, DNN = deep neural network, DW = diffusion-weighted, HER2+ = human epidermal growth factor receptor 2 over-expression, ME = mixture ensemble, MP = multiparametric, NCDFR = non-CDFR, ROC = receiver-operating characteristic, TN = triple negative.