Transfer Learning-Based DCE-MRI Method for Identifying Differentiation Between Benign and Malignant Breast Tumors

LEILEI ZHOU [1], **ZUOHENG ZHANG** [2], **XINDAO YIN** [1], **HONG-BING JIANG** [3,4], **JIE WANG** [5], **GUAN GUI** [6], **(Senior Member, IEEE), YU-CHEN CHEN** [1], **AND JIN-XIA ZHENG** [5]

[1] Department of Radiology, Nanjing First Hospital, Nanjing Medical University, Nanjing 210006, China
[2] School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China
[3] Department of Medical Equipment, Nanjing First Hospital, Nanjing Medical University, Nanjing 210006, China
[4] Nanjing Emergency Medical Center, Nanjing 210003, China
[5] Department of Radiology, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Care Hospital, Nanjing 210004, China
[6] College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding authors: Jin-Xia Zheng (zhengjinxia878@163.com) and Yu-Chen Chen (chenyuchen1989@126.com)

ABSTRACT In this paper, we propose a transfer learning-based dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) method for classifying fibroadenoma and invasive ductal carcinoma (IDC) in breast tumors. A total of 207 breast tumors from patients were collected and identified by pathologic diagnosis within 15 days after enhanced DCE-MRI examination; 119 patients (average age 50.52 ± 10.33 years) had fibroadenomas, and 88 patients (average age 42.20 ± 10.10 years) had IDCs. Two lesion-level models were built based on the InceptionV3 and VGG19 models, which were pretrained with the ImageNet dataset. The effects of different depths of transfer learning were examined. The network's performance was evaluated through five-fold cross validation. In the lesion-level models, the model based on Inception V3 obtained better results (area under the receiver operating characteristic curve (AUC) = 0.89) when the weights were frozen before layer-276. The other model based on VGG19 obtained better results (AUC = 0.87) when the weights were frozen before layer-13. Compared with the image-level models, both lesion-level models displayed better discrimination (accuracy increased by 13% and 14%) based on the VGG19 and Inception V3 models, respectively. Our research confirms that transfer learning can be utilized to classify fibroadenomas and IDCs in DCE-MRI images. Different depths of transfer learning result in different performances, and our proposed lesion-level model notably improves the classification accuracy.

INDEX TERMS Magnetic resonance imaging, invasive ductal carcinoma, fibroadenoma, transfer learning, lesion classification.

I. INTRODUCTION

Since female breast cancer can be screened early and treated, the death rates decreased by 40% from 1989 to 2016 [1]. Breast cancer is one of the most common diseases in women worldwide and is the second leading cause of cancer-related death [2]. According to the American Cancer Society (ACS), a breast cancer patient who receives treatment for early-stage disease (i.e., stage 0 and stage 1) has a 99% chance of

surviving for at least 5 years after being diagnosed [3], [4]. The key factor for improving breast neoplasm prognosis is early detection. Accurate detection and treatment usually lead to a positive outcome. Currently, various imaging examinations, such as mammograms, ultrasounds, and magnetic resonance imaging (MRI), have increased the incidental detection of breast tumors [5], [6]. In particular, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), which can compactly capture both anatomical and metabolic features, has been demonstrated as a great screening examination for those with a high risk of breast cancer [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Shuhan Shen.

Although the characteristics of benign and malignant breast lesions have been established by conventional imaging techniques [8]–[11], differentiating between benign and malignant breast masses is sometimes difficult for radiologists. Imaging-guided biopsies have been considered the first choice for most breast cancers that cannot be diagnosed accurately with conventional imaging; however, the overlapping appearances of benign and malignant breast tumors have led to a large number of unnecessary breast biopsies, and the majority of lesions yield benign results [12]. Therefore, some imaging approaches are required to assist in accurately differentiating benign from malignant breast tumors and to reduce the number of unnecessary biopsies. Fibroadenoma is the most common benign breast mass, and invasive ductal carcinoma (IDC) is the most common malignancy [13]. It remains difficult to accurately identify fibroadenoma and IDC. Liu *et al.* [14] implemented the addition of histogram analysis to apparent diffusion coefficient maps in DCE-MRI for predicting breast malignancy, in which most lesions were found to be fibroadenoma or IDC. Moreover, a typical breast MRI study consists of several acquired and postprocessed image series, each composed of several dozen slices. Due to this large number of images, reading breast MRI studies and reporting the diagnostic findings are tedious and prone to human errors. Computer-aided interpretation of breast MRI studies can assist radiologists by automatically characterizing the detected lesions and may potentially reduce the radiologist's workload.

A variety of hand-crafted features have been continuously proposed by domain experts: textural [15] (as variance and skewness), morphological [16] (as compactness and perimeter), and dynamic [17] (extracted from the time intensity curve). However, these low-throughput features were selected based on radiologists' expert knowledge, which might limit the potential of the radiomics model. Recently, some studies have used deep learning approaches to perform breast MRI lesion classification. Antropova *et al.* [18] adopted AlexNet [19] pretrained on the ImageNet [20] dataset as a feature extractor by using a support vector machine (SVM) for the malignant/benignant classification task and yielded an area under the ROC curve (AUC) value of 0.85, demonstrating the predictive value of the convolutional neural network (CNN) and transfer learning in the classification of DCE-MRI breast images. However, the ROIs in their training dataset were only extracted from the DCE-MRI slices at the second post-contrast time point, which ignored kinetic and 3D context features. Amit *et al.* [21] proposed a multi-channel representation for DCE-MRI images that could capture both the anatomical and metabolic characteristics of lesions in a single multi-channel image and enabled a high accuracy. However, their dataset was relatively small, and they could not assess the classification in a single patient.

In this paper, we propose a transfer learning-based automatic detection and differentiation method for distinguishing between fibroadenoma and IDC in DCE-MRI. The CNNs used in our model were Inception V3 [22] and VGG19 [23],

which were pretrained on the ImageNet dataset and were adopted to obtain slice image features; then, the multi-slice image features of lesions were merged using two fully connected (FC) layers. Our lesion detection framework provides high detection accuracy and a low false-positive rate by making use of 4-dimensional DCE-MRI data. We hypothesize that models trained with multi-slice images of lesions can obtain better accuracy by combining intra-slice with inter-slice features for differentiating IDC from fibroadenoma.

II. MATERIALS AND METHODS
A. GENERAL INFORMATION
From December 2015 to July 2018, data from a total of 184 female patients with mastopathy were collected from Nanjing Maternity and Child Care Hospital. The collection of this dataset was approved by the Institutional Review Board, and we obtained waived written informed consent. All the patients had undergone surgery within 15 days after MRI examination. The final pathological diagnosis distinguished that 108 patients (average age 50.52 ± 10.33 years) had fibroadenoma, and the other 76 patients (average age 42.20 ± 10.10 years) had IDC. Each case of patient data included the DCE T1 sequence with one precontrast series and five postcontrast series. Three experienced radiologists interpreted the studies and manually delineated the boundaries of the region of interest (ROI) on each relevant slice. The lesion images were then cropped using a rectangular bounding box around the annotated boundaries.

B. ACQUISITION OF MRI IMAGES
All patients underwent imaging with a 1.5 T scanner (Achieva, Philips Medical Systems, Netherlands) equipped with a dedicated four-channel phased-array surface breast coil. Axial DCE images using 3D-T1 gradient echo with fat saturation sequences were acquired (TR/TE: 6.9/3.4 ms; flip angle: 10°; field of view 340 mm×340 mm; matrix: 340×340; thickness: 2 mm; gap: 0; acquisition time: 6 minutes and 25 s; 150 slices spanning entire breast volume). One series (t0) was acquired before and 5 series (t1-t5) were acquired after intravenous injection of 0.1 mmol/kg of a positive paramagnetic contrast agent. An automatic injection system was used (Spectris Solaris EP MR, MEDRAD, Inc., Indianola, PA), and the injection flow rate was set to 2 ml/s, followed by a flush of 10 ml of saline solution at the same rate.

C. IMAGE REPRESENTATION
1) KINETIC GRAPH MEASURE
The kinetic graph shows the pattern of contrast uptake and the temporal location of the DCE sequence, which usually requires manual measurement by radiologists. We used the raw image sequence and encoded program to automatically obtain a kinetic graph (Fig. 1). Then, three-channel image representation was based on this graph.
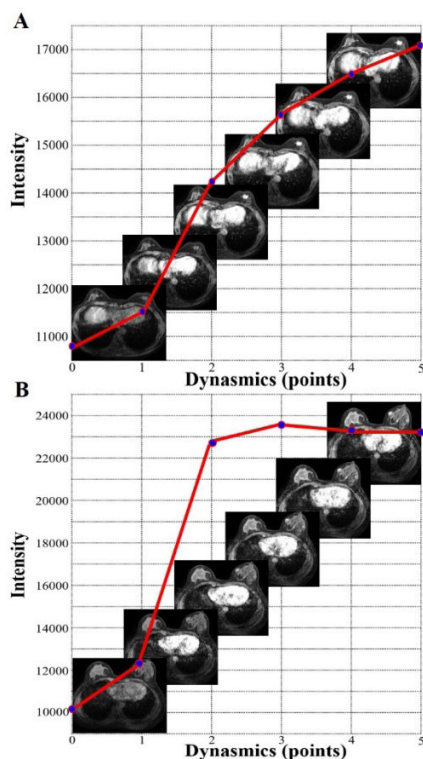
FIGURE 1. Auto-getting kinetic graph. (A) A fibroadenoma lesion with washout kinetics. (B) An IDC lesion with persistent kinetics.

2) THREE-CHANNEL REPRESENTATION

We used our multichannel image representation described in a previous study [24] that could capture both the anatomical and metabolic characteristics of the lesion in a single three-channel image. The three-channel image representation naturally fit the input architecture of the CNNs. The following notations were used for the DCE temporal series: I^{pre} is the precontrast image, I^{peak} is the peak-enhancement image, I^{early} is the initial uptake image, and I^{delay} is the delayed response image. Each slice image is represented by the following three channels:

(1) peak enhancement intensity channel: I^{peak};
(2) contrast uptake channel: $I^{peak} - I^{pre}$;
(3) contrast washout image: $I^{early} - I^{delay}$.

The images were normalized to have a mean of zero and one unit of standard deviation. The training dataset was augmented by adding rotated (rotation range = [−90°, 90°]), two shifts (width shift range = [0, 0.2]; height shift range = [0, 0.2]) and zoom (zoom ratio = 0.2) variants for each image. With sufficient data augmentation, the network could be trained very quickly.

D. LESION DETECTION FRAMEWORK

A flowchart of the study is shown in Fig. 2. We removed the Softmax layers of Inception V3 and VGG19. The other structures of the model remained the same and were initialized by the weights trained on ImageNet. Then, the feature vectors ($1024 \times 1 \times N$) extracted from all the tumor images (N means the number of images of one lesion) of a given lesion were merged into a one-dimensional vector (1024×1) to form the input tensor of the lesion-level model with a Max pooling layer. This layer was used to make the different image sequence lengths uniform. At the end of the model, we added two FC layers (the first layer had 1024 nodes, and the second layer had two nodes for fibroadenoma and IDC) and the Softmax activation to achieve diagnosis at the lesion level.

The two deep learning models were compared in terms of their ability to distinguish fibroadenoma and IDC lesions. Each model was evaluated in a baseline configuration, which included three-channel image representation and data augmentation. In the Keras deep learning framework, VGG19 consists of 25 layers, including 16 convolution layers and 3 fully connected layers. Inception V3 consists of 313 layers, including 13 mixed layers. The shallow layers correspond to general and low-level image features. To explore the effect of freezing different deep layers during transfer learning, we evenly selected the nodal layers in VGG19 and Inception V3. We chose the nodal layers evenly: layer-100, layer-196, layer-276 and layer-310 as the dividing points in Inception V3, as well as layer-9, layer-13, layer-17 and layer-21 as the dividing points in VGG19. The layers prior to the dividing point were frozen, which meant that the weights of these layers were not updated but that others could be trained during the iteration.

E. EXPERIMENTS

The server used in this study was equipped with Intel(R) Xeon(R) E5-2650 v4 CPUs @ 2.20 GHz (2 CPUs, 24 cores, 2 threads/core, 128 GB of memory) and an NVIDIA-SMI 384.81. The KERAS deep learning framework was used. In addition, the learning rate used was 0.01, the momentum was 0.9, and the decay rate was 1.0×10^{-6}. The total number of parameters was approximately 243 M and 96 M in our model, which was based on VGG19 and Inception V3, respectively.

F. PERFORMANCE EVALUATION

The data were randomly split into training (75%) (invasive ductal carcinoma: 74 patients, 79 lesions; fibroadenoma: 54 patients, 63 lesions) and testing (25%) (invasive ductal carcinoma: 34 patients, 36 lesions; fibroadenoma: 22 patients, 22 lesions) subsets of patients. During the training process, the training dataset was randomly split into five groups in order to perform a five-fold cross validation. Model parameter exploration was performed by five-fold cross validation on the training dataset. This approach was chosen on account of the small size of the dataset. We used our subset by partitioning as previously described in [24] to prevent contamination of the training, validation and testing sets with images of the same patient. The results were reported in terms of accuracy (ACC), sensitivity (SEN), specificity (SPE), Matthews correlation coefficient (MCC) and AUC. During the validation
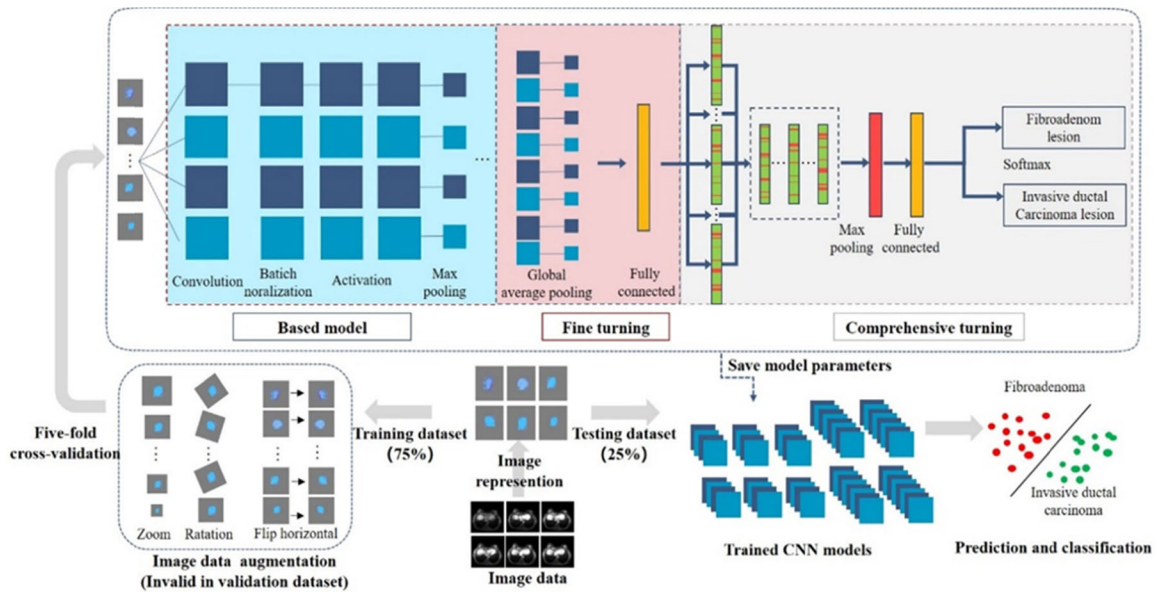
FIGURE 2. Flowchart of the study. First, the DCE-MRI images were preprocessed and represented. Second, five-fold cross-validation was used to select the fine turning model with high predictive performance. In the end, comprehensive model was built to distinguish the fibroadenoma and I.

TABLE 1. Five-fold cross-validation result on these models with different depth weight of CNN.

	Inception V3					VGG 19			
	ACC	SEN	SPEC	MCC		ACC	SEN	SPEC	MCC
Layer100	0.79	0.71	0.86	0.59	**Layer9**	0.83	0.70	0.98	0.69
Layer196	0.79	0.76	0.83	0.61	**Layer13**	*0.84*	*0.77*	*0.95*	*0.71*
Layer276	*0.83*	*0.75*	*0.93*	*0.69*	**Layer17**	0.80	0.77	0.85	0.63
Layer310	0.79	0.75	0.84	0.61	**Layer21**	0.81	0.78	0.81	0.61

ACC, Accuracy; SEN, Sensitivity; SPEC, Specificity; MCC, Matthews correlation coefficient. The results of optimal transfer learning depth were underlined with bold and italics text

phase, all the metrics were calculated based on the average five-fold cross validation results. Then, a better model structure was chosen, the model was retrained with all the training data, and the model parameters were saved.

III. RESULTS

A. PERFORMANCES WITH DIFFERENT DEPTHS OF TRANSFER LEARNING

The models were trained with the iteration stopping criteria, which were determined by monitoring the convergence of the ACC and the loss of the validation and training datasets. The convergence ranges of validation loss were 0.5-0.7 and 0.5-0.7, and the averaged validation accuracies were 89-94% and 81-91%, in Inception V3 and VGG19 respectively. Combined with the ACC and loss curves in Fig. 3 and a series of metrics in Table 1, it was found that the trend of fine-turning results on both lesion-level models (based on Inception V3 and VGG19) was consistent. For the specific

target domain in this study, the classification results were barely satisfactory when the number of freezing layers was too large or too small.

For the lesion-level model based on Inception V3 and VGG19, the validation results of AUCs from freezing layer-276 (AUC = 0.83) and layer-13 (AUC = 0.87) were larger in Fig. 4(A) and Fig. 4(B), respectively. Other metrics such as the SEN, SPEC, MCC and ACC in Table 1 presented the same trends as the AUCs in Fig. 4. As shown in Table 1, for the lesion-level model based on Inception V3, the fine-turning result of freezing the 100 layers was the worst. The model had many parameters, but with less data, possible reasons. For the lesion-level model based on VGG19, the fine-turning result of freezing the 21 layers was the worst. Thus, the natural image training model is not fully suitable for medical images. Both of them are difficult to converge and fluctuate greatly.

The appropriate fine-turning depth and the related training parameters were selected, yielding a better performance
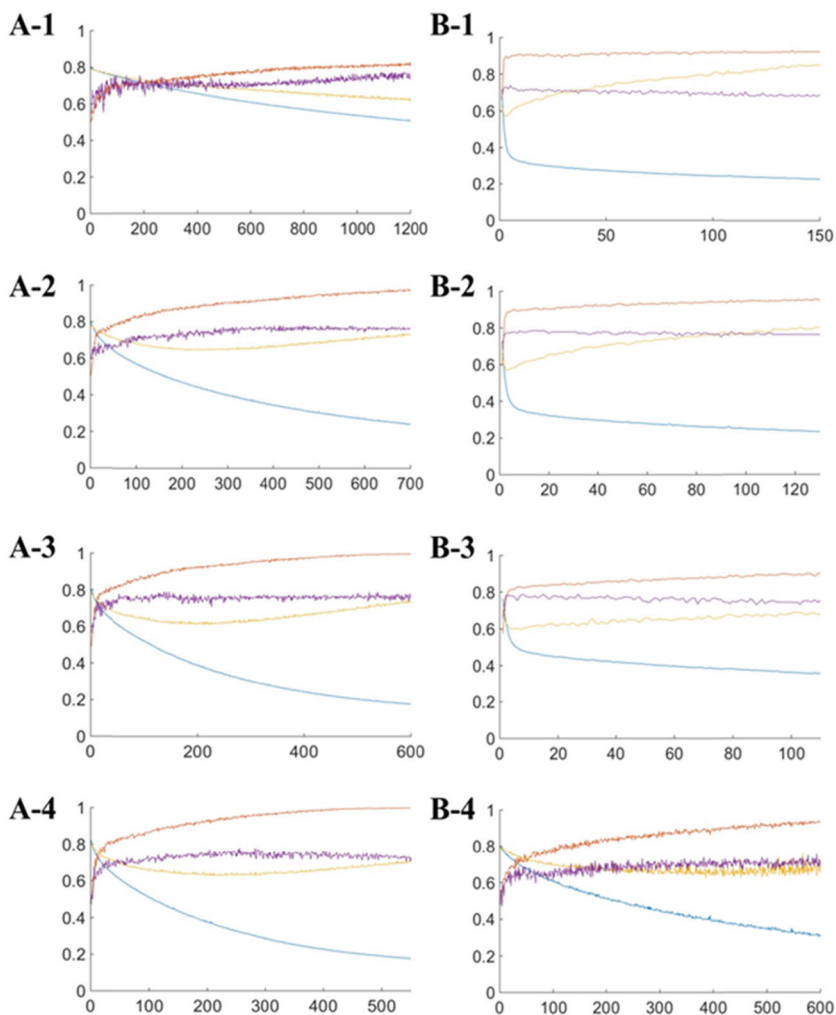
FIGURE 3. Traces of training loss and validation loss (blue and yellow lines) and training ACC and validation accuracy (orange and purple lines). A column was trained on the Inception V3, and −1, −2, −3 and −4 denoted freezing the weights of CNN before 100, 197, 276 and 310 layers. B column was trained on the VGG19, and −1, −2, −3 and −4 denoted freezing the weights of CNN before 9, 13, 17 and 21 layer.

**TABLE 2.** Test result on these models with different depth weight of CNN.

		ACC	SEN	SPEC	MCC	AUC
Inception V3(layer276)						
	Image-level	0.76	0.74	0.83	0.49	0.81
	Lesion-level	*0.90*	*0.94*	*0.82*	*0.78*	*0.93*
VGG19(layer13)						
	Image-level	0.78	0.75	0.89	0.56	0.88
	Lesion-level	*0.91*	*0.97*	*0.82*	*0.82*	*0.92*

ACC Accuracy; SEN Sensitivity; SPEC Specificity; MCC Matthews correlation coefficient; AUC area under the receiver operating characteristic curve. The lesion-level model notably improved the classification capability, and the corresponding results were underlined with bold and italics text

in five-fold cross validation; then, the selected model was retrained with all the training datasets. We employed the corresponding model parameters on our testing dataset. The results are listed in Table 2.

B. COMPARISONS OF LESION-LEVEL MODELS

To classify fibroadenoma and IDC from slice images, we removed the comprehensive part and set two nodes (fibroadenoma/invasive ductal carcinoma) in the Softmax
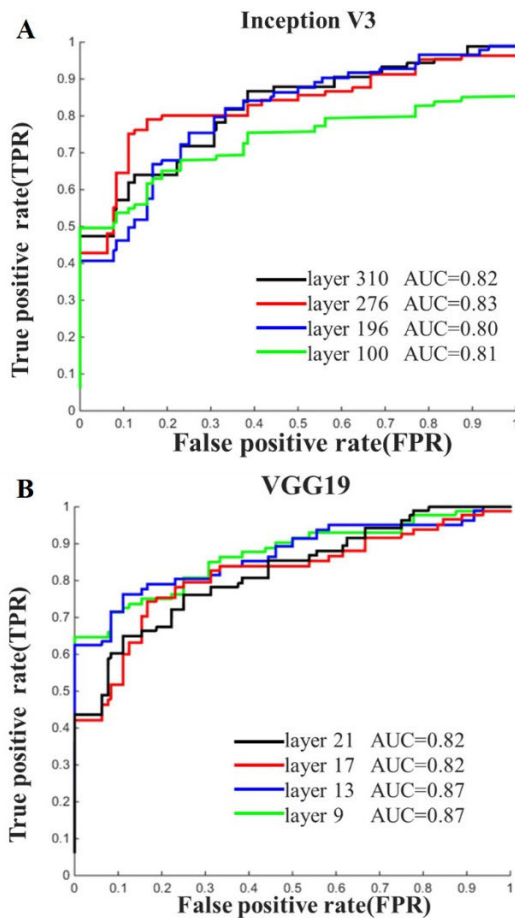
FIGURE 4. ROC averaged on five-fold cross validation of the transfer learning with freezing different layers. The plot of AUC calculated from ROC with freezing different layers.

layer of Inception V3 and VGG19. The other structures of the models and the depths of transfer learning remained the same.

Compared with the image-level model, our two lesion-level models obviously improved and were similar (the ACC, SEN, MCC, and AUC increased by 13%, 22%, 26% and 4% in the lesion-level model based on VGG19; the ACC, SEN, MCC, and AUC increased by 14%, 20%, 29% and 12% in the lesion-level model based on Inception V3), as shown in in Table 2

IV. DISCUSSION
Our study demonstrates that transfer learning can be used to classify fibroadenoma and IDC in DCE-MRI images. Different depths of transfer learning resulted in different performances, and our lesion-level model notably improved the classification accuracy. Computerized algorithms can assist radiologists by automatically characterizing the detected lesions. Pang *et al.* [25] used texture and morphology features with an SVM classifier to achieve an accuracy of 0.9 for discriminating benign and malignant breast lesions. Chen *et al.* [26] extracted the statistical and Haralick texture features in DCE-MRI based on which predictive models were

built to predict histological grade in breast cancer. Their classifier achieved good performance with an AUC of 0.859. Those methods based on feature engineering need defined features artificially, which might limit the accuracy and scalability of the radiomics model. In this study, we proposed a deep learning model to mine MRI image information related to breast lesion status. While our results are comparable with these reports, our classification approach is more general, as it requires only the manually selected tumor region in MRI images without human-defined features, which is different from conventional radiomic methods based on feature engineering and can be naturally extended to other imaging modalities or other MR sequences.

The applicability of a CNN classifier is often coupled with the availability of large training datasets. Some researchers have demonstrated that the classification of small sample size image data can be achieved by transfer learning [27]. Anderson *et al.* [28] fine-tuned the final FC layers of the pretrained VGG19 to classify the DCE-MRI breast images as malignant or benign (AUC = 0.86). However, they only trained the last FC layers. Zhe *et al.* [29] used the deep features approach with the GoogleNet model pretrained on ImageNet as the feature extractor and a polynomial kernel SVM used as the classifier (AUC = 0.70) to identify occult invasive disease in patients diagnosed with ductal carcinoma in situ. However, they only trained the classifier while not training the feature extractor. Thus, regulating the trainable layers in transfer learning is worthwhile to investigate. In addition, they only took care of the central slice image of each lesion. Previous studies have shown that feature extraction by nonmedical models is feasible for the classification task at hand [21], [28]. To improve the transfer learning performance, we experimented with varying the number of trainable layers. As Table 1 shows, the classification ability for breast tumors remained at a similarly low level when the weights of layers that were too large or too small were frozen. Because the features extracted directly from the pretrained model were unsuitable and insufficient for breast tumor classification, the fewer layers that were fine turned, the less classification accuracy could be obtained. When training a complex model with one small dataset, the more layers that were fine turned, the poorer classification accuracy could be obtained.

Our proposed multichannel image representation naturally fits the three-channel input architecture of the pretrained convolutional network while effectively encoding both spatial and temporal characteristics of the imaging contrast. In addition to its contribution to improving classification accuracy, it may also be used effectively for the simultaneous visualization of the morphology and kinetics of a lesion to a human interpreter. For clinical diagnoses, an experienced radiologist usually observes and detects tumors based on many slices along the Z-axis. We merged all slice features of each lesion together to achieve diagnosis at the lesion level. By fusing the learning of intra-slice and inter-slice features, the detection performance for breast tumors was obviously improved, as shown in Table 2.

Despite the encouraging performance of the deep learning model, this study has several limitations. First, we only examined patients in one hospital. In future work, populations from multiple sources are necessary to test whether our model can be generalized. Second, although our model showed better performance than clinical and radiomics models, the effects of a combination of these models are unclear. The predictive performance may be improved if we combine these models together. Third, our study only focused on two special categories of lesions. Collecting more different subtype data and further investigating fine-grained recognition can be explored in future work.

V. CONCLUSION

In this paper, we have proposed a transfer learning-based automatic detection and differentiation method in DCE-MRI images. In our trained dataset, training layers from approximately half the initially trained models performed better than previous transfer-learning studies in which only the last layer was trained. In addition, our lesion-level model notably improved the classification accuracy. Such recognition ability could eliminate the need for patients identified with benign tumors to undergo invasive procedures. Deep learning technology bears the potential to revolutionize the capabilities of computer-aided diagnosis tools. We believe that as the available datasets expand and models are further optimized, this method will be able to facilitate the incorporation of cognitive technologies into the radiology workflow.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA, Cancer J. Clin.*, vol. 69, no. 1, pp. 7–34, Jan/Feb. 2019.

[2] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D. Parkin, M. Piñeros, A. Znaor, and F. Bray, "Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods," *Int. J. Cancer*, vol. 144, no. 8, pp. 1941–1953, Apr. 2019.

[3] A. Noone, N. Howlader, M. Krapcho, D. Miller, A. Brest, and M. Yu, "SEER cancer statistics review, 1975–2015," Nat. Cancer Inst., Bethesda, MD, USA, Tech. Rep., 2018. [Online]. Available: https://seer.cancer.gov/archive/csr/1975_2015/

[4] A. E. Giuliano, J. L. Connolly, S. B. Edge, E. A. Mittendorf, H. S. Rugo, L. J. Solin, D. L. Weaver, D. J. Winchester, and G. N. Hortobagyi, "Breast cancer-major changes in the American joint committee on cancer eighth edition cancer staging manual," *CA, Cancer J. Clin.*, vol. 67, no. 4, pp. 290–303, Jul. 2017.

[5] T. Stephens, "Breast cancer screening with imaging: Recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer," *Yearbook Diagnostic Radiol.*, vol. 2011, pp. 46–47, Jan. 2011.

[6] N. Qu, Y. Luo, T. Yu, and H. Yu, "Differentiation between pure mucinous breast carcinomas and fibroadenomas with strong high-signal intensity on T2-weighted images from dynamic contrast-enhanced magnetic resonance imaging," *Breast Care*, vol. 13, no. 1, pp. 32–37, Mar. 2018.

[7] C. D. Lehman, C. Gatsonis, C. K. Kuhl, R. E. Hendrick, E. D. Pisano, and L. Hanna, "MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer," *New England J. Med.*, vol. 356, no. 13, pp. 1295–1303, Mar. 2007.

[8] L. Neal, C. L. Tortorelli, and A. Nassar, "Clinician's guide to imaging and pathologic findings in benign breast disease," *Mayo Clin. Proc.*, vol. 85, no. 3, pp. 274–279, Mar. 2010.

[9] Q. Zhu, A. Ricci, P. Hegde, M. Kane, E. Cronin, A. Merkulov, Y. Xu, B. Tavakoli, and S. Tannenbaum, "Assessment of functional differences in malignant and benign breast lesions and improvement of diagnostic accuracy by using US-guided diffuse optical tomography in conjunction with conventional US," *Radiology*, vol. 280, no. 2, pp. 387–397, Aug. 2016.

[10] T. Kinoshita, N. Yashiro, N. Ihara, H. Funatu, E. Fukuma, and M. Narita, "Diffusion-weighted half-Fourier single-shot turbo spin echo imaging in breast tumors: Differentiation of invasive ductal carcinoma from fibroadenoma," *J. Comput. Assist. Tomogr.*, vol. 26, no. 6, pp. 1042–1046, Nov. 2002.

[11] American College of Radiology. *ACR BI-RADS Atlas Fifth Edition Quick Reference.* [Online]. Available: https://www.acr.org/-/media/ACR/Files/RADS/BI-RADS/BIRADS-Reference-Card.pdf

[12] Office of the Assistant Secretary for Planning and Evaluation. *The Importance of Radiology and Pathology Communication in the Diagnosis and Staging of Cancer: Mammography as a Case Study.* [Online]. Available: https://aspe.hhs.gov/pdf-report/importance-radiology-and-pathology-communication-diagnosis-and-staging-cancer-mammography-case-study

[13] K. Kasireddy, K. Manjula, and C. S. B. R. Prasad, "Clinicopathological study of rare invasive epithelial tumors of breast: An institutional study," *BLDE Univ. J. Health Sci.*, vol. 1, no. 2, p. 125, 2016.

[14] H.-L. Liu, M. Zong, H. Wei, J.-J. Lou, S.-Q. Wang, Q.-G. Zou, H.-B. Shi, and Y.-N. Jiang, "Preoperative predicting malignancy in breast mass-like lesions: Value of adding histogram analysis of apparent diffusion coefficient maps to dynamic contrast-enhanced magnetic resonance imaging for improving confidence level," *Brit. J. Radiol.*, vol. 90, no. 1079, Nov. 2017, Art. no. 20170394.

[15] R. Fusco, M. Sansone, C. Sansone, and A. Petrillo, "Segmentation and classification of breast lesions using dynamic and textural features in dynamic contrast enhanced-magnetic resonance imaging," in *Proc. 25th IEEE Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2012, pp. 1–4.

[16] R. Fusco, M. Sansone, A. Petrillo, and C. Sansone, "A multiple classifier system for classification of breast lesions using dynamic and morphological features in DCE-MRI," in *Proc. Joint IAPR Int. Workshops Stat. Techn. Pattern Recognit. Struct. Syntactic Pattern Recognit. (SPR&SSPR)*, 2012, pp. 684–692.

[17] S. Glasser, U. Niemann, B. Preim, and M. Spiliopoulou, "Can we distinguish between benign and malignant breast tumors in DCE-MRI by studying a tumor's most suspect region only?" in *Proc. 26th IEEE Int. Symp. Comput.-Based Med. Syst.*, Jun. 2013, pp. 77–82.

[18] N. Antropova, B. Huynh, and M. Giger, "SU-D-207B-06: Predicting breast cancer malignancy on DCE-MRI data using pre-trained convolutional neural networks," *Med. Phys.*, vol. 43, no. 6, pp. 3349–3350, Jun. 2016.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[20] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[21] G. Amit, R. Ben-Ari, O. Hadad, E. Monovich, N. Granot, and S. Hashoul, "Classification of breast MRI lesions using small-size training sets: Comparison of deep learning approaches," in *Proc. Med. Imag., Comput.-Aided Diagnosis*, Mar. 2017, Art. no. 101341H.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[23] H. Ye, Z. Wu, R.-W. Zhao, X. Wang, Y.-G. Jiang, and X. Xue, "Evaluating two-stream CNN for video classification," in *Proc. 5th ACM Int. Conf. Multimedia Retr. (ICMR)*, 2015, pp. 435–442.

[24] L. Zhou, Z. Zhang, Y.-C. Chen, Z.-Y. Zhao, X.-D. Yin, and H.-B. Jiang, "A deep learning-based radiomics model for differentiating benign and malignant renal tumors," *Transl. Oncol.*, vol. 12, no. 2, pp. 292–300, Feb. 2019.

[25] Z. Pang, D. Zhu, D. Chen, L. Li, and Y. Shao, "A computer-aided diagnosis system for dynamic contrast-enhanced MR images based on level set segmentation and ReliefF feature selection," *Comput. Math. Methods Med.*, vol. 2015, pp. 1–10, Jan. 2015.

[26] Q. Chen, M. Fan, P. Zhang, L. Li, and M. Xu, "Heterogeneity of tumor and its surrounding stroma on DCE-MRI and diffusion weighted imaging in predicting histological grade and lymph node status of breast cancer," in *Proc. Med. Imag., Imag. Informat. Healthcare, Res., Appl.*, Mar. 2019, Art. no. 109541G.

[27] S. Wang, J. Shi, Z. Ye, D. Dong, D. Yu, M. Zhou, Y. Liu, O. Gevaert, K. Wang, Y. Zhu, H. Zhou, Z. Liu, and J. Tian, "Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning," *Eur. Respiratory J.*, vol. 53, no. 3, Mar. 2019, Art. no. 1800986.

[28] R. Anderson, H. Li, Y. Ji, P. Liu, and M. L. Giger, "Evaluating deep learning techniques for dynamic contrast-enhanced MRI in the diagnosis of breast cancer," *Med. Imag., Comput.-Aided Diagnosis*, vol. 10950, Mar. 2019, Art. no. 1095006.

[29] Z. Zhu, M. Harowicz, J. Zhang, A. Saha, L. J. Grimm, E. S. Hwang, and M. A. Mazurowski, "Deep learning analysis of breast MRIs for prediction of occult invasive disease in ductal carcinoma *in situ*," *Comput. Biol. Med.*, vol. 115, Dec. 2019, Art. no. 103498.

LEILEI ZHOU received the master's degree in biomedical engineering from Nanjing Medical University, Nanjing, China, in 2019. Her research interest includes deep learning and its applications.
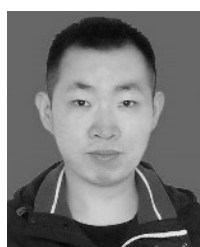
ZUOHENG ZHANG is currently pursuing the doctor degree in molecular dynamics simulation for nucleation and growth of iron oxide with the School of Biological Science and Medical Engineering, Southeast University. His research interest includes parameter optimization and deep learning and its applications.
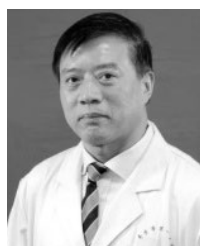
XINDAO YIN is currently the Director of the Department of Radiology, Nanjing First Hospital, Nanjing Medical University, Nanjing, China. His research interests include MRI data processing, machine learning, and deep learning and its applications.

HONG-BING JIANG received the master's degree from Southeast University, Nanjing, China, in 2010. Since 2010, he has been a Professor at Southeast University, Nanjing. Since 2012, he has been a Professor with Nanjing Medical University, Nanjing. His current research interests include mobile robot, computer vision, medical image processing, and machine learning.

JIE WANG is currently pursuing the master's degree in radiology with Nanjing Medical University. She is also working with the Affiliated Obstetrics and Gynecology Hospital of Nanjing Medical University, Nanjing, China. Her research interest includes breast MRI scanning and its applications.

GUAN GUI (Senior Member, IEEE) received the Dr.Eng. degree in information and communication engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2012.

From 2009 to 2014, he was a Research Assistant as well as a Postdoctoral Research Fellow with the Wireless Signal Processing and Network Laboratory (Prof. Adachi Laboratory), Department of Communications Engineering, Graduate School of Engineering, Tohoku University. From 2014 to 2015, he was an Assistant Professor with the Department of Electronics and Information System, Akita Prefectural University. Since 2015, he has been a Professor at the Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China. He is currently engaged in the research of deep learning, compressive sensing, and advanced wireless techniques. He has published more than 200 international peer-reviewed journal/conference papers. He received the Member and Global Activities Contributions Award from the IEEE ComSoc and seven best paper awards, i.e., ICEICT 2019, ADHIP 2018, CSPS 2018, ICNC 2018, ICC 2017, ICC 2014, and VTC 2014-Spring. He was also selected as a Jiangsu Specially-Appointed Professor, in 2016, the Jiangsu High-level Innovation and Entrepreneurial Talent, in 2016, the Jiangsu Six Top Talent, in 2018, the Nanjing Youth Award, in 2018. He was an Editor of *Security and Communication Networks,* from 2012 to 2016. He has been an Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, since 2017, IEEE ACCESS, since 2018, *KSII Transactions on Internet and Information Systems*, since 2017, and *Journal of Communications*, since 2019, and the Editor-in-Chief of *EAI Transactions on Artificial Intelligence*, since 2018.

YU-CHEN CHEN received the Ph.D. degree from Southeast University, Nanjing, China, in 2016. He is currently a Radiologist working with the Radiology Department, Nanjing First Hospital, Nanjing Medical University, Nanjing. His research interests include functional MRI data processing, machine learning, and transfer learning and its applications.

JIN-XIA ZHENG is currently an Associate Chief Physician with the Radiology Department, Affiliated Obstetrics and Gynecology Hospital of Nanjing Medical University, Nanjing, China. Her research interest includes breast DCE-MRI scanning data processing and its applications.

• • •