

University of Groningen

Breast Tumor Identification in Ultrafast MRI Using Temporal and Spatial Information

Jing, Xueping; Dorrius, Monique D.; Wielema, Mirjam; Sijens, Paul E.; Oudkerk, Matthijs; van Ooijen, Peter

Published in:
Cancers

DOI:
[10.3390/cancers14082042](https://doi.org/10.3390/cancers14082042)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Jing, X., Dorrius, M. D., Wielema, M., Sijens, P. E., Oudkerk, M., & van Ooijen, P. (2022). Breast Tumor Identification in Ultrafast MRI Using Temporal and Spatial Information. *Cancers*, 14(8), [2042]. <https://doi.org/10.3390/cancers14082042>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Article

Breast Tumor Identification in Ultrafast MRI Using Temporal and Spatial Information

Xueping Jing ¹, Monique D. Dorrius ^{2,*}, Mirjam Wielema ², Paul E. Sijens ², Matthijs Oudkerk ^{3,4} and Peter van Ooijen ^{1,*}

¹ Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands; x.jing@umcg.nl

² Department of Radiology, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands; m.wielema@umcg.nl (M.W.); p.e.sijens@umcg.nl (P.E.S.)

³ Faculty of Medical Sciences, University of Groningen, 9700 RB Groningen, The Netherlands; m.oudkerk@umcg.nl

⁴ Institute for Diagnostic Accuracy, 9713 GH Groningen, The Netherlands

* Correspondence: m.d.dorrius@umcg.nl (M.D.D.); p.m.a.van.ooijen@umcg.nl (P.v.O.)

Simple Summary: The diagnosis of breast cancer with MRI is based on both morphological evaluation and kinetic curve assessment. Current computer-aided diagnosis methods for malignancy determination mainly focus on morphology features but ignored the temporal information in dynamic contrast-enhanced MRI sequences. Malignant and benign lesions usually have different enhancement patterns during the wash-in phase. Ultrafast breast MRI with high temporal resolution can capture the inflow of contrast in breast lesions. This advantage of ultrafast MRI enables the combination of both temporal and spatial information for automatic breast lesion analysis model development. We found that temporal information helps to significantly improve the performance of breast lesion classification. This suggests that ultrafast MRI provides useful information for malignancy identification and temporal information, which is indispensable for similar model development.

Abstract: Purpose: To investigate the feasibility of using deep learning methods to differentiate benign from malignant breast lesions in ultrafast MRI with both temporal and spatial information. Methods: A total of 173 single breasts of 122 women (151 examinations) with lesions above 5 mm were retrospectively included. A total of 109 out of 173 lesions were benign. Maximum intensity projection (MIP) images were generated from each of the 14 contrast-enhanced T1-weighted acquisitions in the ultrafast MRI scan. A 2D convolutional neural network (CNN) and a long short-term memory (LSTM) network were employed to extract morphological and temporal features, respectively. The 2D CNN model was trained with the MIPs from the last four acquisitions to ensure the visibility of the lesions, while the LSTM model took MIPs of an entire scan as input. The performance of each model and their combination were evaluated with 100-times repeated stratified four-fold cross-validation. Those models were then compared with models developed with standard DCE-MRI which followed the same data split. Results: In the differentiation between benign and malignant lesions, the ultrafast MRI-based 2D CNN achieved a mean AUC of 0.81 ± 0.06 , and the LSTM network achieved a mean AUC of 0.78 ± 0.07 ; their combination showed a mean AUC of 0.83 ± 0.06 in the cross-validation. The mean AUC values were significantly higher for ultrafast MRI-based models than standard DCE-MRI-based models. Conclusion: Deep learning models developed with ultrafast breast MRI achieved higher performances than standard DCE-MRI for malignancy discrimination. The improved AUC values of the combined models indicate an added value of temporal information extracted by the LSTM model in breast lesion characterization.

Keywords: ultrafast breast MRI; lesion classification; deep learning



Citation: Jing, X.; Dorrius, M.D.; Wielema, M.; Sijens, P.E.; Oudkerk, M.; van Ooijen, P. Breast Tumor Identification in Ultrafast MRI Using Temporal and Spatial Information. *Cancers* **2022**, *14*, 2042. <https://doi.org/10.3390/cancers14082042>

Academic Editor: Katja Pinker-Domenig

Received: 24 January 2022

Accepted: 11 April 2022

Published: 18 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has a higher sensitivity compared with other modalities for breast cancer detection [1]. It is widely used for breast cancer staging, screening in high-risk women, and chemotherapy evaluation. For breast cancer screening, recent research reveals that supplemental MRI screening in a high-risk population, especially in women with extremely dense breasts, could significantly reduce the interval cancer rate compared with mammography alone [2]. For women with a lifetime breast cancer risk above 20%, annual MRI screening helps to detect cancers at an earlier stage than mammography [3]. However, the advantages of breast MRI are accompanied by higher cost and a higher false positive rate.

During DCE-MRI scanning, the concentration of gadolinium contrast material in the permeable blood vessels of tumor tissue causes a decrease in the relaxation time of water protons and results in a high signal in T1-weighted images. The current DCE-MRI protocol usually contains multiple T1-weighted acquisitions after contrast. The diagnosis of breast lesions is highly dependent on the high spatial resolution and kinetic pattern in the wash-out phase, making MRI scanning tedious. To enable a wider use of breast MRI, protocols should be shorter to improve cost effectiveness [4,5]. Different abbreviated protocols have been proposed to shorten the scanning time [6–8]. Studies also showed that the diagnostic performance of abbreviated protocols matches the full diagnostic breast MRI protocol [9–13].

Most abbreviated protocols are based on T1-weighted imaging before and after contrast agent administration. T1-weighted acquisitions may be performed in different time resolutions. Conventional T1-weighted MRI has high spatial resolution, but is time consuming, while the time-resolved angiography with stochastic trajectories (TWIST) technique enables ultrafast acquisition at the cost of a lower spatial resolution [14,15]. Compared with the use of only one high spatial resolution post-contrast T1-weighted series for image interpretation [7], repeated acquisition of a T1-weighted sequence in TWIST enables not only morphological evaluation but also pharmacokinetic analysis of the wash-in phase (Figure 1). The current literature recognizes the critical role of kinetic analysis in the differentiation of benign and malignant lesions in ultrafast MRI. The ultrafast T1-weighted sequence might change the future of breast DCE-MRI and has proven useful for the characterization of breast lesions. Abe et al. [16] concluded that there was significant difference between benign and malignant lesions in enhancement rate and kinetic area under the curve (AUC) in ultrafast MRI, and that the differential utility of ultrafast imaging is comparable to standard kinetic assessment. Other studies [17] found that features (time to enhancement and maximum slope) derived from ultrafast MRI during initial enhancement could help improve the performance of differentiating between malignant and benign lesions, especially in case of non-mass enhancement. These studies revealed the potential of kinetic analysis in ultrafast MRI for breast lesions [7,17–19].

Recently, artificial intelligence (AI) was introduced to help discriminate benign from malignant lesions in breast MRI [20,21]. Most AI studies focus on 2D feature extraction with, either, radiomics methods [22,23] or convolutional neural networks (CNN) [24]. Even though it has proven to be a crucial factor and is widely used in clinical practice, temporal information (kinetic features) has been ignored in the development of automated breast lesion analysis systems. Long short-term memory (LSTM) networks [25], which are widely used for time series analysis, have been employed for medical image analysis recently [26–28]. For example, Zou et al. [29] used an LSTM model for pharmacokinetic parameter estimation in head and neck cancer using a TWIST sequence and achieved promising results. These successful applications indicate that temporal information extracted by LSTM models may further improve the performance of automatic breast lesion diagnosis in DCE-MRI.

We hypothesized that fully automatic deep learning-based methods, with only ultrafast DCE-MRI as input, could be used for malignancy identification. Therefore, the aim of this study is to evaluate the feasibility of this automatic deep learning-based method and investigate the usefulness of temporal information extracted with an LSTM model.

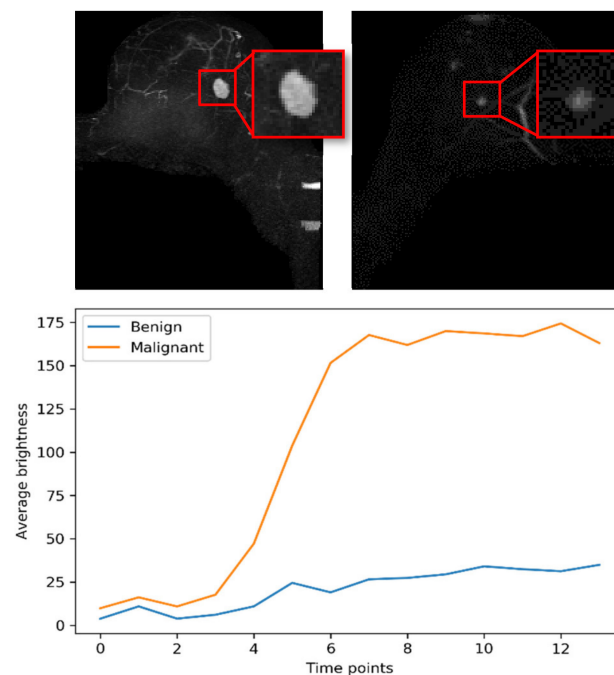


Figure 1. Maximum intensity projection images (MIP) of a malignant (**left**) and a benign (**right**) lesion and their enhancement curves during ultrafast MRI. Plotted curves are the average brightness of 10×10 pixels in the center of each lesion.

2. Materials and Methods

2.1. Study Population

This retrospective study was approved by the local Medical Ethics Review Board (METc Nr: 2018/652) with a waiver of the requirement for informed consent from each patient. A total of 1447 breast MRIs, including both screening and diagnostic examinations, from 809 patients who underwent breast MRI scans at our hospital between 2016 and 2019 were retrospectively included. Among the 1447 consecutive breast MRIs, 443 were excluded because no ultrafast DCE-MRI sequences had been acquired. Furthermore, 694 MRI scans from 374 patients with only focus/foci (enhancing dots with diameters smaller than 5 mm) or with no lesion reported were excluded. A total of 34 MRI scans for chemotherapy evaluation, 31 for breast implants, and 94 post-operative scans were also excluded, resulting in 151 MRI scans from 122 patients. For the remaining 151 MRI scans, left and right breasts were assessed separately for the presence of breast lesions. A total of 129 breasts with no suspicious lesion were excluded, and 173 breasts with one or more lesions were included in the final model development.

The label of each breast was determined by the primary lesion contained. For breasts with no malignant lesions, the benign lesion with the biggest size was analyzed. For breasts with malignant lesions, the possible additional present benign lesions were not included in the analysis. The ground truth of those included lesions was obtained from pathology reports (surgery or core needle biopsy) or follow-up examinations (2 years follow up, verified by a senior radiologist). The final study dataset contained 109 benign and 64 malignant lesions (Figure 2).

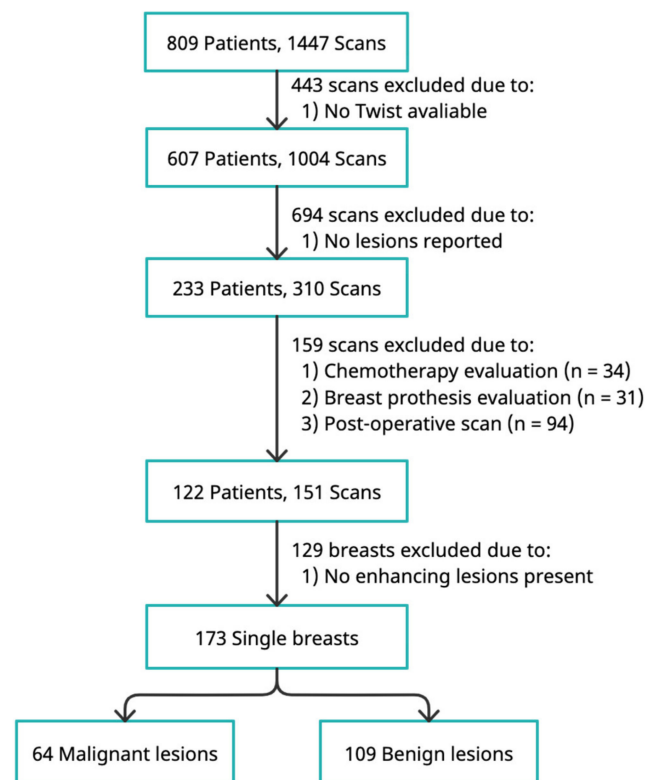


Figure 2. Flowchart of the patient inclusion.

2.2. MRI Protocol

All DCE-MRI scans were acquired on a 3 T (Siemens Magnetom Skyra, Siemens Medical Solutions, Erlangen, Germany) or 1.5 T (Siemens Magnetom Avanto_fit, Siemens Medical Solutions, Erlangen, Germany) scanner with a circularly polarized bilateral breast coil (Siemens). The protocol consisted of a pre-contrast T1 and T2 and the following sequences after contrast: DWI, TWIST, and four DCE-T1-weighted series. A measurement of 15 mL of contrast agent (Dotarem, Guerbet, Villepinte, France) was injected using a power injector. Acquisition parameters of the T1-weighted and TWIST sequence are list in Table 1.

Table 1. Acquisition Parameters for ultrafast and standard DCE-MRI.

Parameter	TWIST		T1-Weighted	
	1.5 T	3.0 T	1.5 T	3.0 T
TR/TE (ms)	2.50/0.90	4.12/2.08	5.27/2.39	4.50/1.60
Flip angle (°)	20	20	10	10
Phase oversampling (%)	26	20	N/A	N/A
Slice oversampling (%)	20	0	N/A	N/A
Voxel size (mm ³)	0.68 × 0.68 × 3.0	0.91 × 0.91 × 3.0	0.84 × 0.84 × 1.2	0.89 × 0.89 × 1.2
Temporal resolution (s)	5.2	4.3	120	120
Field of view (mm)	350	350	350	370
Fat suppression	None	None	SPAIR	SPAIR

2.3. Development of the Deep-Learning-Based Classification System

Breast segmentations were first performed by using a 3D U-Net [30] to remove the redundant background; the left and right breasts were then separated by splitting the breast regions in the middle. Maximum intensity projection (MIP) images were then generated by applying MIP operation on the subtraction volume of each sequence and the sequence before contrast agent injection.

To take advantage of both morphological and temporal information provided by ultrafast breast MRI, a CNN model and a long short-term memory LSTM network were employed to extract morphological and temporal features, respectively. (Figure 3) The CNN model, modified from the pre-trained ResNet-18 [31] by replacing the last output layer, were trained with MIP images from the last four acquisitions in each ultrafast MRI series. The LSTM model, with the ResNet-18 model as feature extractor, was trained with the MIP clip (14 timed acquisitions of the ultrafast DCE sequence yielding 14 MIP images) of each ultrafast MRI series. For standard DCE-MRI, the same method was used for the training, where MIPs of the four T1-weighted acquisitions after contrast injection were used as the input for both the 2D CNN and LSTM model.

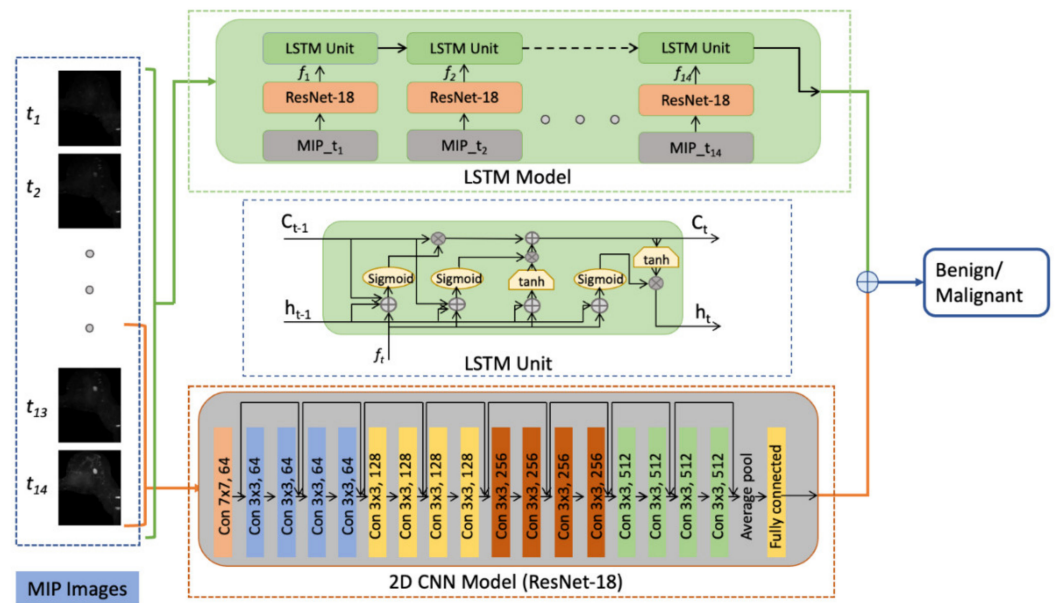


Figure 3. Pipeline of the proposed system. The 2D CNN model takes a maximum intensity projection (MIP) image of the last four acquisitions as input, while the LSTM model takes all 14 MIP images in a TWIST sequence as a single input. In the LSTM model, a ResNet-18 model was used for feature extraction. The extracted feature vector f_t was then inputted to the LSTM unit, in which C_{t-1} represents the memory from last unit, h_{t-1} represents the output of the previous unit, C_t represents the memory of the current unit, and h_t represents the output of the current unit. The output probability of each model was added up to generate a combined prediction.

The output of each model was a numerical value between zero and one, which indicates the probability of malignancy. The performance of each model and their combination was evaluated with 100-times repeated stratified 4-fold cross-validation. During inference, the CNN models assessed each of the four MIP images from a single validation sample, and the highest output value was then determined as the malignancy probability of the sample. The output of the proposed system was then determined by averaging the predicted scores from both models. The receiver operating characteristic (ROC) curve was plotted for each model and the entire system based on the predicted probability. Sensitivity and specificity were used to evaluate the performance of each model with different thresholds applied.

Models were developed with Pytorch framework (v1.9.1) and trained with an NVIDIA Quadro P4000 GPU. For the 2D CNN (ResNet-18) model, transfer learning was adopted with pre-trained weights on ImageNet; convolution layers were frozen while only fully connected layers' parameters were finetuned. The training was run for 30 epochs with a learning rate of $1e-4$, a batch size of 16, and a momentum of 0.9. The Adam optimizer and a categorical cross-entropy loss function were used. For the LSTM model, a ResNet-18 model was used as feature extractor. The same hyperparameters were used for training, except the batch size was reduced to four and training epochs were extended to 40. During

training, MIP images were first resized to 224×224 , and then sent into the models; random horizontal flip (probability of 0.5) and random rotation (within 10°) were performed for data augmentation. The training (30 epochs) and validation of the 2D CNN model took 28.8 and 1.5 min, respectively, while the training (40 epochs) and validation of the LSTM model took 34.3 and 2.2 min, respectively, in each fold.

3. Results

3.1. Patient and Lesion Characteristics

A total of 122 patients with 173 breasts were included in this study. The patient and lesion characteristics are provided in Table 2. The median age was 47 years (range, 24–80 years). The median size of all lesions was 13.0 mm (range, 5.0–110.0 mm), while for benign lesions it was 9.0 mm (range, 5.0–81.0 mm) and malignant lesions it was 22.0 mm (range, 6–110 mm).

Table 2. Lesion characteristics.

Characteristics	Value (Proportion)
Benign lesions	109 (0.63)
Adenosis	24 (0.14)
Fibroadenoma	19 (0.11)
Hyperplasia	6 (0.03)
Glandular tissue	4 (0.02)
Cyst	3 (0.02)
Inflammation	1 (0.01)
Other ¹	51 (0.29)
Malignant lesions	64 (0.37)
Invasive ductal carcinoma	51 (0.29)
Invasive lobular carcinoma	4 (0.02)
Ductal carcinoma in situ	4 (0.02)
Micropapillary carcinoma	2 (0.01)
Apocrine carcinoma	1 (0.01)
Mucinous carcinoma	2 (0.01)
Lesion size (mm) ²	
Overall	19.9 ± 18.4
Malignant	28.6 ± 20.8
Benign	13.9 ± 13.6

¹ The “Other” category included enhancement around fat necrosis, scar tissue, hyperplasia, atheroma cyst, regional background enhancement, and other benign-appearing enhancements not specified. ² Data are ± standard deviation.

3.2. Performance of Models

The ROC curves of each model for malignancy differentiation with TWIST sequences are shown in Figure 4. During the repeated stratified cross-validation, the 2D CNN showed a mean AUC of 0.81 ± 0.06 , the LSTM network showed a mean AUC of 0.78 ± 0.07 , while their combination showed a mean AUC of 0.83 ± 0.06 . For standard DCE-MRI, the 2D CNN model showed a mean AUC of 0.67 ± 0.09 , the LSTM network showed a mean AUC of 0.66 ± 0.08 , and their combination showed a mean AUC of 0.70 ± 0.07 .

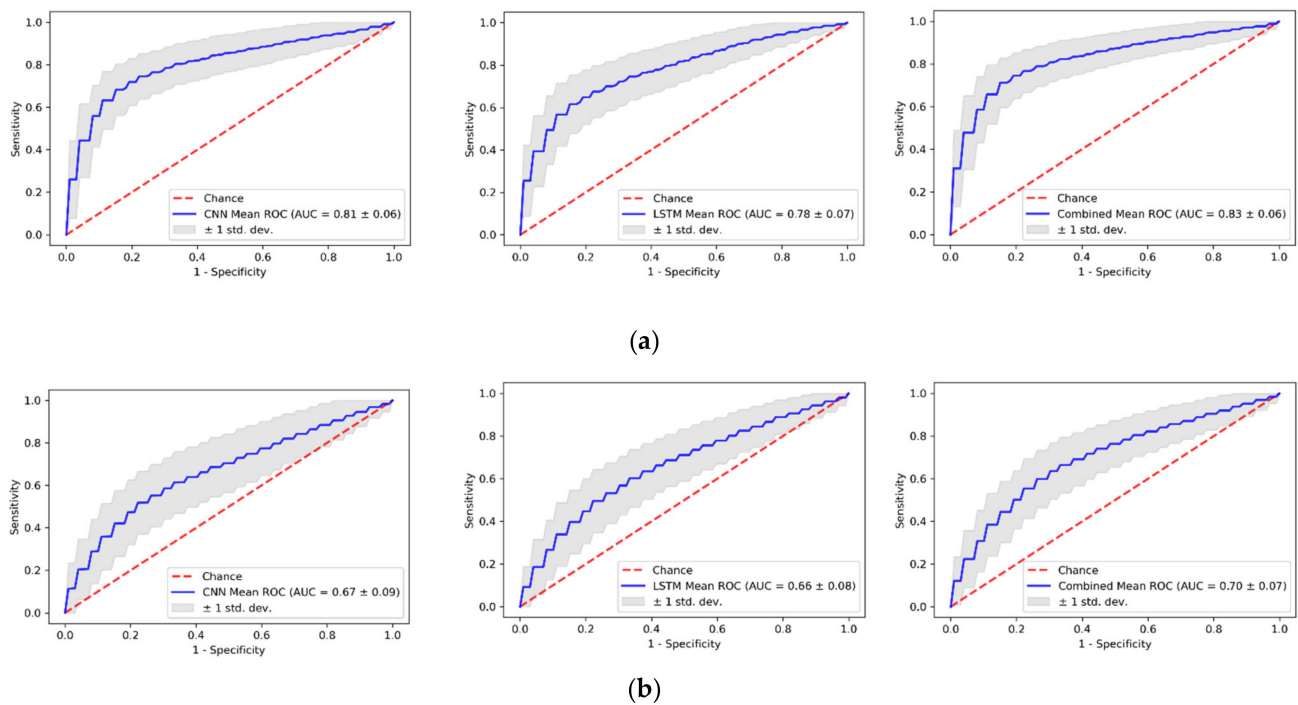


Figure 4. Receiver operating characteristic (ROC) curves for CNN, LSTM model, and their combination in the (a) TWIST and (b) standard DCE-MRI validation sets. The plots show the mean ROC of each model over 100-times repeated cross-validation, reflecting the variance of the curves when the dataset is split into different training and validation sets.

The 2D CNN achieved a higher mean AUC than LSTM ($p < 0.01$), but the combined prediction could further improve the performance (all $p < 0.01$). There is clearly an added value in using the temporal information extracted by the LSTM model. A detailed analysis of the diagnosis performance with different cutoff values is listed in Table 3.

Table 3. Diagnostic performance of each model with ultrafast MRI under different threshold settings.

Threshold	2D CNN		LSTM		Combined	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
0.1	0.92 (0.91, 0.93)	0.25 (0.23, 0.27)	0.92 (0.91, 0.93)	0.22 (0.20, 0.24)	0.96 (0.95, 0.96)	0.14 (0.13, 0.16)
0.2	0.82 (0.81, 0.84)	0.56 (0.54, 0.58)	0.81 (0.80, 0.83)	0.48 (0.46, 0.50)	0.87 (0.86, 0.88)	0.48 (0.46, 0.50)
0.3	0.73 (0.72, 0.74)	0.76 (0.75, 0.77)	0.73 (0.72, 0.75)	0.65 (0.63, 0.67)	0.78 (0.77, 0.79)	0.72 (0.71, 0.74)
0.4	0.64 (0.63, 0.66)	0.88 (0.87, 0.88)	0.66 (0.64, 0.68)	0.78 (0.77, 0.79)	0.68 (0.67, 0.70)	0.86 (0.85, 0.86)
0.5	0.56 (0.55, 0.57)	0.93 (0.92, 0.93)	0.59 (0.57, 0.61)	0.88 (0.87, 0.88)	0.57 (0.55, 0.58)	0.92 (0.91, 0.93)
0.6	0.44 (0.43, 0.46)	0.95 (0.94, 0.95)	0.46 (0.44, 0.48)	0.92 (0.91, 0.93)	0.44 (0.43, 0.46)	0.96 (0.95, 0.96)
0.7	0.34 (0.33, 0.35)	0.96 (0.96, 0.97)	0.34 (0.33, 0.36)	0.94 (0.94, 0.95)	0.32 (0.31, 0.33)	0.98 (0.97, 0.98)
0.8	0.25 (0.23, 0.26)	0.98 (0.97, 0.98)	0.24 (0.22, 0.25)	0.97 (0.96, 0.97)	0.18 (0.17, 0.19)	0.99 (0.99, 0.99)
0.9	0.15 (0.14, 0.16)	0.99 (0.99, 0.99)	0.12 (0.10, 0.13)	0.99 (0.98, 0.99)	0.07 (0.06, 0.08)	1.0 N/A

To better illustrate the classification pattern of the models developed with ultrafast MRI, boxplots of the mean AUC values of each model during cross-validation are illustrated in

Figure 5a. Compared with the 2D CNN and LSTM model, less dispersed AUC values and fewer outliers indicate an improved performance of the combination. The distribution of predicted risk scores of the combined model in relation to malignancy is shown in Figure 5b. Malignant lesions were shown to likely be predicted with a higher risk score compared with benign lesions. The mean risk scores for malignant and benign lesions were 0.588 (95% CI: 0.582, 0.594) and 0.294 (95% CI: 0.291, 0.297), respectively.

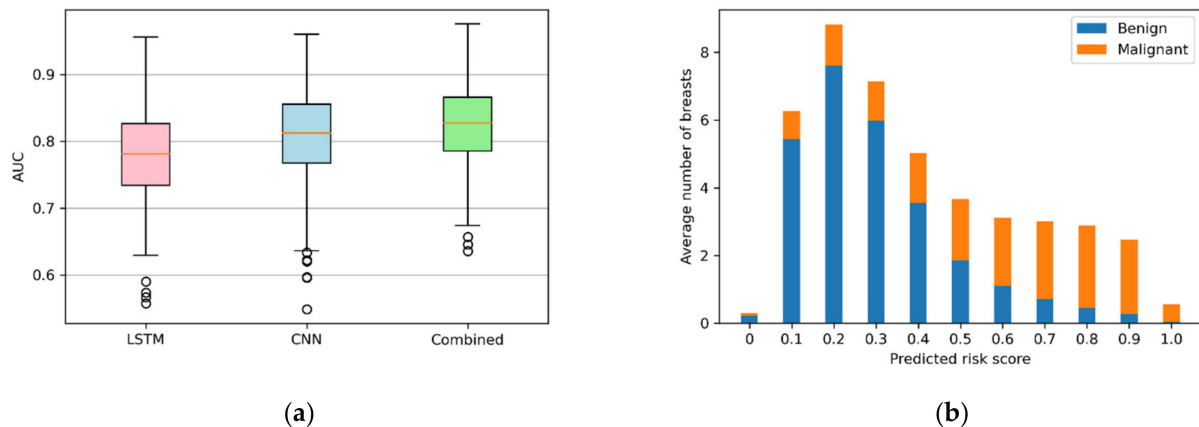


Figure 5. The boxplots of the mean AUC values of each model on the TWIST validation set (a) and the distribution of AI risk scores predicted by the combined model for benign and malignant lesions (b).

4. Discussion

In this study, we investigated the possibility of using a deep learning model for breast lesion discrimination with only ultrafast MRI as input. Motivated by recently discovered kinetic features derived from ultrafast MRI, we looked at the effectiveness of deep learning-based automatic pharmacokinetic analysis of the wash-in phase and possible improvements to morphological analysis. For this purpose, a 2D CNN model and an LSTM model were first trained and evaluated separately for the classification of benign and malignant lesions, and then combined by the fusion of the predicted scores of each model.

The results showed that deep learning models could achieve high performance for lesion differentiation with only ultrafast breast MRI. The models developed with ultrafast MRI had a significantly higher mean AUC values compared with the model developed with standard DCE-MRI. Meanwhile, the combined model achieved the highest mean AUC value during repeated stratified cross-validation. Even though the mean AUC value of the LSTM model was lower than the 2D CNN model, after integrating them together, the temporal information extracted by the LSTM model turned out to have added value for diagnostic performance. This upward trend was observed with both ultrafast and standard MRI.

Machine learning was also adopted for malignancy prediction in previous studies with different sequences. Platel et al. [32] used 5 morphological features and 18 uptake-curve features extracted from ultrafast MRI to train support vector machine (SVM) models. The AUC of the SVM models trained with only kinetic features or only morphological features were 0.80 and 0.79, respectively, while the SVM model trained with all the 23 handcraft features together showed an AUC of 0.85. These results are consistent with our findings, indicating that both temporal and morphological features are indispensable for the development of breast lesion classification models in ultrafast MRI. In another study, Dalmış et al. [24] developed a deep-learning-based classifier which takes image patches from TWIST, T2w, DWI sequences, and patient information as input. When only image patches were used, an AUC of 0.83 was achieved, while with combining all imaging and patient information, the AUC was 0.85. Pötsch et al. [22] extracted 86 pharmacokinetic enhancement features from TWIST sequences to train a multilayer perceptron (MLP) artificial

neural network. The MLP model achieved an AUC of 0.83 on the test dataset and could help avoid 14.5% biopsies without false negatives. Radiomics has also been used for breast lesion differentiation [33–35].

Compared with previous studies developing classifiers with handcraft or radiomics features, in this work, we used a 2D CNN model and an LSTM model to automatically extract morphological and kinetic features separately. Instead of highly depending on the manually delineated lesion area for feature extraction, a big advantage of the proposed method is that neither manual annotation nor dedicated segmentation of the lesions are required before analysis. The models take raw images as input and provide predictions based on the content only; therefore, it could be easily integrated in the clinic workflow and works fully automated.

What should also be noticed is that there is no associate relationship of identification ability between machine learning models and radiologists. Radiologists still rely more on high-resolution T1-weighted sequences for malignancy discrimination, which intuitively contradicts the finds of this paper. A possible explanation for the inferior performance of the standard DCE-MRI-based models might be that temporal analysis of the TWIST sequence with the LSTM model made up for the defect of morphological analysis caused by insufficient spatial resolution.

Despite these promising results achieved with ultrafast MRI sequences, we can easily envision the improvements when taking other sequences into consideration. For accurate diagnosis, delayed sequences are indispensable in clinical practice, especially the wash-out phase in DCE-MRI which provides irreplaceable functionality of indicating malignancy. However, in the trend of shortening MRI protocol for screening, identifying the malignancy of a tumor from an easily accessible sequence with a fully automatic model could provide support for a flexible scanning strategy, in which additional MRI sequences could be added from the protocol as needed. The flexibility of MRI could save time while maintaining the performance of breast cancer screening.

There are still several limitations of our study. This retrospective study had a relatively small dataset from a single medical center, and not all the benign lesions were pathologically confirmed. Even though the dataset has been randomly split 100 times to illustrate the robustness of the models, a large multivendor multicenter dataset is still needed for further validation. Another limitation of the work is that the study was conducted in a high-risk screening population and included pre-operation examinations; therefore, the prevalence of malignant lesions in our study is higher compared with the general screening population, which may weaken the model's generalizability. Additionally, patient information has been proven effective to improve the decision-making of classification models, and properly integrating patient information may further improve the performance. Furthermore, we focused on the validation of the usefulness of temporal information extracted with an LSTM model; all the models were trained with the same hyperparameter setting and training strategy, which may lead to an underestimated performance of the system. Finally, models were trained and validated with the output of the primary lesion in each breast. Therefore, the predictions only revealed the status of primary lesions but overlooked associated lesions. Object detection models, which could locate each single lesion, may help address this problem.

5. Conclusions

In conclusion, deep learning models achieved promising results for the fully automatic differentiation of benign and malignant lesions. Kinetic information extracted by LSTM model has added value in the 2D CNN-based morphology analysis.

Author Contributions: Conceptualization, X.J. and P.v.O.; methodology, X.J.; software, X.J.; validation, X.J. and M.D.D.; investigation, X.J.; resources, M.W. and M.D.D.; data curation, M.D.D.; writing—original draft preparation, X.J.; writing—review and editing, P.E.S., M.W., P.v.O., and M.D.D.; supervision, M.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of University Medical Centre Groningen on 4 December 2018 (METc 2018/652).

Informed Consent Statement: Patient consent was waived due to the Medical Ethics Review Board conclude this research is not a clinical research with human subjects as meant in the Medical Research Involving Human Subjects Act (WMO).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy concerns, in accordance with GDPR.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vandermeer, F.Q.; Bluemke, D.A. Breast MRI: State of the art. *Cancer Investig.* **2007**, *25*, 384–392. [[CrossRef](#)] [[PubMed](#)]
2. Bakker, M.F.; de Lange, S.V.; Pijnappel, R.M.; Mann, R.M.; Peeters, P.H.M.; Monninkhof, E.M.; Emaus, M.J.; Loo, C.E.; Bisschops, R.H.C.; Lobbes, M.B.I.; et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N. Engl. J. Med.* **2019**, *381*, 2091–2102. [[CrossRef](#)] [[PubMed](#)]
3. Saadatmand, S.; Geuzinge, H.A.; Rutgers, E.J.T.; Mann, R.M.; de Roy van Zuidewijn, D.B.W.; Zonderland, H.M.; Tollenaar, R.A.E.M.; Lobbes, M.B.I.; Ausems, M.G.E.M.; van 't Riet, M.; et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): A multicentre, randomised, controlled trial. *Lancet Oncol.* **2019**, *2045*, 1136–1147. [[CrossRef](#)]
4. Kuhl, C.K. A Call for Improved Breast Cancer Screening Strategies, Not Only for Women with Dense Breasts. *JAMA Netw. Open* **2021**, *4*, e2121492. [[CrossRef](#)]
5. Mann, R.M.; Kuhl, C.K.; Moy, L. Contrast-enhanced MRI for breast cancer screening. *J. Magn. Reson. Imaging* **2019**, *50*, 377–390. [[CrossRef](#)]
6. Chhor, C.M.; Mercado, C.L. Abbreviated MRI protocols: Wave of the future for breast cancer screening. *Am. J. Roentgenol.* **2017**, *208*, 284–289. [[CrossRef](#)]
7. Kuhl, C.K.; Schrading, S.; Strobel, K.; Schild, H.H.; Hilgers, R.D.; Bieling, H.B. Abbreviated breast Magnetic Resonance Imaging (MRI): First postcontrast subtracted images and maximum-intensity projection—A novel approach to breast cancer screening with MRI. *J. Clin. Oncol.* **2014**, *32*, 2304–2310. [[CrossRef](#)]
8. Pinker, K.; Moy, L.; Sutton, E.J.; Mann, R.M.; Weber, M.; Thakur, S.B.; Jochelson, M.S.; Bago-Horvath, Z.; Morris, E.A.; Baltzer, P.A.; et al. Diffusion-Weighted Imaging with Apparent Diffusion Coefficient Mapping for Breast Cancer Detection as a Stand-Alone Parameter: Comparison with Dynamic Contrast-Enhanced and Multiparametric Magnetic Resonance Imaging. *Investig. Radiol.* **2018**, *53*, 587–595. [[CrossRef](#)]
9. Van Zelst, J.C.M.; Vreemann, S.; Witt, H.J.; Gubern-Merida, A.; Dorrius, M.D.; Duvivier, K.; Lardenoije-Broker, S.; Lobbes, M.B.I.; Loo, C.; Veldhuis, W.; et al. Multireader Study on the Diagnostic Accuracy of Ultrafast Breast Magnetic Resonance Imaging for Breast Cancer Screening. *Investig. Radiol.* **2018**, *53*, 579–586. [[CrossRef](#)]
10. Partovi, S.; Sin, D.; Lu, Z.; Sieck, L.; Marshall, H.; Pham, R.; Plecha, D. Fast MRI breast cancer screening—Ready for prime time. *Clin. Imaging* **2020**, *60*, 160–168. [[CrossRef](#)]
11. Kim, E.S.; Cho, N.; Kim, S.Y.; Kwon, B.R.; Yi, A.; Ha, S.M.; Lee, S.H.; Chang, J.M.; Moon, W.K. Comparison of abbreviated mri and full diagnostic mri in distinguishing between benign and malignant lesions detected by breast mri: A multireader study. *Korean J. Radiol.* **2021**, *22*, 297–307. [[CrossRef](#)] [[PubMed](#)]
12. Grimm, L.J.; Soo, M.S.; Yoon, S.; Kim, C.; Ghate, S.V.; Johnson, K.S. Abbreviated Screening Protocol for Breast MRI. A Feasibility Study. *Acad. Radiol.* **2015**, *22*, 1157–1162. [[CrossRef](#)] [[PubMed](#)]
13. Harvey, S.C.; Di Carlo, P.A.; Lee, B.; Obadina, E.; Sippo, D.; Mullen, L. An Abbreviated Protocol for High-Risk Screening Breast MRI Saves Time and Resources. *J. Am. Coll. Radiol.* **2016**, *13*, R74–R80. [[CrossRef](#)] [[PubMed](#)]
14. Herrmann, K.H.; Baltzer, P.A.; Dietzel, M.; Krumbein, I.; Geppert, C.; Kaiser, W.A.; Reichenbach, J.R. Resolving arterial phase and temporal enhancement characteristics in DCE MRM at high spatial resolution with TWIST acquisition. *J. Magn. Reson. Imaging* **2011**, *34*, 973–982. [[CrossRef](#)] [[PubMed](#)]
15. Le, Y.; Kipfer, H.; Majidi, S.; Holz, S.; Dale, B.; Geppert, C.; Kroeker, R.; Lin, C. Application of time-resolved angiography with stochastic trajectories (twist)-dixon in dynamic contrast-enhanced (dce) breast mri. *J. Magn. Reson. Imaging* **2013**, *38*, 1033–1042. [[CrossRef](#)] [[PubMed](#)]
16. Abe, H.; Mori, N.; Tsuchiya, K.; Schacht, D.V.; Pineda, F.D.; Jiang, Y.; Karczmar, G.S. Kinetic analysis of benign and malignant breast lesions with ultrafast dynamic contrast-enhanced MRI: Comparison with standard kinetic assessment. *Am. J. Roentgenol.* **2016**, *207*, 1159–1166. [[CrossRef](#)]
17. Goto, M.; Sakai, K.; Yokota, H.; Kiba, M.; Yoshida, M.; Imai, H.; Weiland, E.; Yokota, I.; Yamada, K. Diagnostic performance of initial enhancement analysis using ultra-fast dynamic contrast-enhanced MRI for breast lesions. *Eur. Radiol.* **2019**, *29*, 1164–1174. [[CrossRef](#)]

18. Onishi, N.; Sadinski, M.; Gibbs, P.; Gallagher, K.M.; Hughes, M.C.; Ko, E.S.; Dashevsky, B.Z.; Shanbhag, D.D.; Fung, M.M.; Hunt, T.M.; et al. Differentiation between subcentimeter carcinomas and benign lesions using kinetic parameters derived from ultrafast dynamic contrast-enhanced breast MRI. *Eur. Radiol.* **2020**, *30*, 756–766. [[CrossRef](#)]
19. Mus, R.D.; Borelli, C.; Bult, P.; Weiland, E.; Karssemeijer, N.; Barentsz, J.O.; Gubern-Mérida, A.; Platel, B.; Mann, R.M. Time to enhancement derived from ultrafast breast MRI as a novel parameter to discriminate benign from malignant breast lesions. *Eur. J. Radiol.* **2017**, *89*, 90–96. [[CrossRef](#)]
20. Codari, M.; Schiaffino, S.; Sardanelli, F.; Trimboli, R.M. Artificial intelligence for breast MRI in 2008–2018: A systematic mapping review. *Am. J. Roentgenol.* **2019**, *212*, 280–292. [[CrossRef](#)]
21. Sheth, D.; Giger, M.L. Artificial intelligence in the interpretation of breast cancer on MRI. *J. Magn. Reson. Imaging* **2020**, *51*, 1310–1324. [[CrossRef](#)] [[PubMed](#)]
22. Pötsch, N.; Dietzel, M.; Kapetas, P.; Clauser, P.; Pinker, K.; Ellmann, S.; Uder, M.; Helbich, T.; Baltzer, P.A.T.; An, A.I. classifier derived from 4D radiomics of dynamic contrast-enhanced breast MRI data: Potential to avoid unnecessary breast biopsies. *Eur. Radiol.* **2021**, *31*, 5866–5876. [[CrossRef](#)] [[PubMed](#)]
23. Fan, M.; Wu, G.; Cheng, H.; Zhang, J.; Shao, G.; Li, L. Radiomic analysis of DCE-MRI for prediction of response to neoadjuvant chemotherapy in breast cancer patients. *Eur. J. Radiol.* **2017**, *94*, 140–147. [[CrossRef](#)] [[PubMed](#)]
24. Dalmış, M.U.; Gubern-Mérida, A.; Vreemann, S.; Bult, P.; Karssemeijer, N.; Mann, R.; Teuwen, J.; Dalmış, M.U.; Gubern-Mérida, A.; Vreemann, S.; et al. Artificial Intelligence Based Classification of Breast Lesions Imaged with a Multi-Parametric Breast MRI Protocol With ultrafast DCE-MRI, T2 and DWI. *Investig. Radiol.* **2019**, *54*, 325–332. [[CrossRef](#)]
25. Greff, K.; Srivastava, R.K.; Koutnik, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A Search Space Odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [[CrossRef](#)]
26. Benou, A.; Veksler, R.; Friedman, A.; Riklin Raviv, T. Ensemble of expert deep neural networks for spatio-temporal denoising of contrast-enhanced MRI sequences. *Med. Image Anal.* **2017**, *42*, 145–159. [[CrossRef](#)]
27. Huynh, B.; Giger, M.L. Breast lesion classification based on dynamic contrast-enhanced magnetic resonance images sequences with long short-term memory networks. *J. Med. Imaging* **2018**, *6*, 1. [[CrossRef](#)]
28. Feng, C.; Elazab, A.; Yang, P.; Wang, T.; Zhou, F.; Hu, H.; Xiao, X.; Lei, B. Deep Learning Framework for Alzheimer’s Disease Diagnosis via 3D-CNN and FSBi-LSTM. *IEEE Access* **2019**, *7*, 63605–63618. [[CrossRef](#)]
29. Zou, J.; Balter, J.M.; Cao, Y. Estimation of pharmacokinetic parameters from DCE-MRI by extracting long and short time-dependent features using an LSTM network. *Med. Phys.* **2020**, *47*, 3447–3457. [[CrossRef](#)]
30. Zhang, J.; Saha, A.; Zhu, Z.; Mazurowski, M.A. Hierarchical Convolutional Neural Networks for Segmentation of Breast Tumors in MRI with Application to Radiogenomics. *IEEE Trans. Med. Imaging* **2018**, *38*, 435–447. [[CrossRef](#)]
31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016*; IEEE: Manhattan, NY, USA, 2016; pp. 770–778.
32. Platel, B.; Mus, R.; Welte, T.; Karssemeijer, N.; Mann, R. Automated characterization of breast lesions imaged with an ultrafast DCE-MR protocol. *IEEE Trans. Med. Imaging* **2014**, *33*, 225–232. [[CrossRef](#)] [[PubMed](#)]
33. Boca, I.; Ciurea, A.I.; Ciortea, C.A.; Ștefan, P.A.; Lisencu, L.A.; Dudea, S.M. Differentiating breast tumors from background parenchymal enhancement at contrast-enhanced mammography: The role of radiomics—A pilot reader study. *Diagnostics* **2021**, *11*, 1248. [[CrossRef](#)] [[PubMed](#)]
34. Conti, A.; Duggento, A.; Indovina, I.; Guerrisi, M.; Toschi, N. Radiomics in breast cancer classification and prediction. *Semin. Cancer Biol.* **2021**, *72*, 238–250. [[CrossRef](#)] [[PubMed](#)]
35. Tagliafico, A.S.; Piana, M.; Schenone, D.; Lai, R.; Massone, A.M.; Houssami, N. Overview of radiomics in breast cancer diagnosis and prognostication. *Breast* **2020**, *49*, 74–80. [[CrossRef](#)]