applied sciences

MDPI

Article

Radiomics and Machine Learning Approaches for the Preoperative Classification of In Situ vs. Invasive Breast Cancer Using Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE–MRI)

Luana Conte [1,2,†], Rocco Rizzo [2,3,4,†], Alessandra Sallustio [5], Eleonora Maggiulli [6], Mariangela Capodieci [7], Francesco Tramacere [5], Alessandra Castelluccia [5], Giuseppe Raso [1], Ugo De Giorgi [8], Raffaella Massafra [9], Maurizio Portaluri [10], Donato Cascio [1,*] and Giorgio De Nunzio [2,4,11]

1 Department of Physics and Chemistry "E. Segrè", University of Palermo, 90128 Palermo, Italy; luana.conte@unipa.it (L.C.); giuseppe.raso@unipa.it (G.R.)
2 Laboratory of Interdisciplinary Research Applied to Medicine (DReAM), University of Salento & ASL Lecce, 73100 Lecce, Italy; rocco.rizzo1@unisalento.it (R.R.); giorgio.denunzio@unisalento.it (G.D.N.)
3 Department of Engineering for Innovation, University of Salento, 73100 Lecce, Italy
4 Istituto Nazionale di Fisica Nucleare (INFN), 73100 Lecce, Italy
5 Radiotherapy Unit, 'Di Summa-Perrino' Hospital, ASL Brindisi, 72100 Brindisi, Italy; ale15.sally@gmail.com (A.S.); francesco.tramacere@asl.brindisi.it (F.T.); alessandra.castelluccia@asl.brindisi.it (A.C.)
6 Medical Physics Unit, 'Di Summa-Perrino' Hospital, ASL Brindisi, 72100 Brindisi, Italy; eleonora.maggiulli@asl.brindisi.it
7 Clinical Senology, 'Di Summa-Perrino' Hospital, ASL Brindisi, 72100 Brindisi, Italy; mariangela.capodieci@asl.brindisi.it
8 Department of Experimental Medicine, University of Salento, 73100 Lecce, Italy; ugo.degiorgi@unisalento.it
9 IRCCS Istituto Tumori "Giovanni Paolo II", 70124 Bari, Italy; r.massafra@oncologico.bari.it
10 Radiotherapy Unit, ASST 'Papa Giovanni XXIII', 24127 Bergamo, Italy; mportaluri@asst-pg23.it
11 Laboratory of Biomedical Physics and Environment, Department of Mathematics and Physics "E. De Giorgi", University of Salento, 73100 Lecce, Italy
* Correspondence: donato.cascio@unipa.it
† These authors contributed equally to this work.

Abstract

Accurate preoperative distinction between in situ and invasive Breast Cancer (BC) is critical for clinical decision-making and treatment planning. Radiomics and Machine Learning (ML) have shown promise in enhancing diagnostic performance from breast MRI, yet their application to this specific task remains underexplored. The aim of this study was to evaluate the performance of several ML classifiers, trained on radiomic features extracted from DCE–MRI and supported by basic clinical information, for the classification of in situ versus invasive BC lesions. In this study, we retrospectively analysed 71 post-contrast DCE–MRI scans (24 in situ, 47 invasive cases). Radiomic features were extracted from manually segmented tumour regions using the PyRadiomics library, and a limited set of basic clinical variables was also included. Several ML classifiers were evaluated in a Leave-One-Out Cross-Validation (LOOCV) scheme. Feature selection was performed using two different strategies: Minimum Redundancy Maximum Relevance (MRMR), mutual information. Axial 3D rotation was used for data augmentation. Support Vector Machine (SVM), K Nearest Neighbors (KNN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) were the best-performing models, with an Area Under the Curve (AUC) ranging from 0.77 to 0.81. Notably, KNN achieved the best balance between sensitivity and specificity without the need for data augmentation. Our findings confirm that radiomic features extracted from DCE–MRI, combined with well-validated ML models, can effectively support the differentiation of in situ vs. invasive breast cancer. This approach

check for updates

Academic Editor: Pedro Couto

Received: 29 May 2025
Revised: 16 July 2025
Accepted: 16 July 2025
Published: 18 July 2025

Citation: Conte, L.; Rizzo, R.; Sallustio, A.; Maggiulli, E.; Capodieci, M.; Tramacere, F.; Castelluccia, A.; Raso, G.; De Giorgi, U.; Massafra, R.; et al. Radiomics and Machine Learning Approaches for the Preoperative Classification of In Situ vs. Invasive Breast Cancer Using Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE–MRI). *Appl. Sci.* **2025**, *15*, 7999. https://doi.org/10.3390/app15147999

is quite robust even in small datasets and may aid in improving preoperative planning. Further validation on larger cohorts and integration with additional imaging or clinical data are recommended.

Keywords: breast cancer; radiomics; machine learning; artificial intelligence

1. Introduction

Breast cancer (BC) remains one of the most prevalent malignancies affecting women worldwide. In Italy alone, approximately 53,686 new cases and 15,500 deaths were reported in 2024 [1].

Malignant breast tumours can be classified as either lobular or ductal, depending on whether they originate in the lobules or the ducts of the breast [2,3]. These carcinomas may remain confined to their site of origin, in which case they are referred to as in situ, or they may spread to the surrounding tissues, in which case they are considered invasive. In situ lesions are characterised by malignant cells that respect the integrity of the basement membrane and do not penetrate the surrounding stroma. By contrast, invasive disease is defined by tumour cells that breach this membrane and infiltrate adjacent connective and adipose tissues [4,5].

The clinical management pathways for in situ versus invasive lesions differ significantly; for instance, sentinel lymph node biopsy, which is routinely recommended for invasive lesions, is generally not indicated in the treatment of in situ lesions. Moreover, invasive tumours are associated with a poorer prognosis, higher risk of metastasis, and increased likelihood of recurrence [6–10]. Consequently, misclassification of an invasive lesion as in situ could result in under-treatment, leading to inadequate surgical planning and compromised clinical outcomes. In this context, accurate preoperative differentiation between in situ and invasive breast lesions is crucial for optimising treatment decisions and improving patient care.

Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE–MRI) has emerged as a key imaging modality for breast cancer detection, staging, and characterisation, particularly in women with dense breast tissue or high familial risk [11–15]. Its superior sensitivity compared to other imaging techniques makes it an invaluable tool in identifying suspicious lesions. MRI is commonly employed to optimise surgical outcomes by minimising the likelihood of re-excisions, assisting in the identification of patients suitable for neoadjuvant chemotherapy or adjustments in therapeutic strategies. Moreover, it serves as a preferred modality for the preoperative evaluation of residual tumour burden, helping to assess eligibility for breast-conserving surgery [11]. However, interpreting breast MRI scans requires considerable expertise, as these examinations play a pivotal role in informing surgical decision-making. Despite its high diagnostic potential, conventional radiological evaluation of DCE–MRI is still largely qualitative and subject to inter-observer variability, which can limit its reliability in accurately distinguishing between in situ and invasive lesions. In this context, quantitative imaging approaches, named Radiomics, have garnered increasing interest [16]. Radiomics enables the extraction of a large number of features from medical images, offering a non-invasive window into tumour phenotype, heterogeneity, and underlying biology [17,18]. When combined with Artificial Intelligence (AI) and Machine Learning (ML) techniques, Radiomics has the potential to enhance diagnostic precision, improve lesion classification, and support clinical decision-making in breast cancer care [19–21]. Several studies have demonstrated that ML models trained on radiomic features extracted from DCE–MRI can outperform conventional assessment in distinguishing malignant subtypes [22–26].

In this study, we developed and evaluated an ML framework for the binary classification of breast tumours as either in situ or invasive, based on radiomic and clinical features extracted from a region of interest (ROI) containing the tumour tissue, in post-contrast DCE–MRI images. Our final aim was to investigate whether a Radiomics-based AI approach can aid in the pre-operative classification of breast lesions, potentially contributing to more personalised and effective patient management.

2. Methods

2.1. Dataset

The dataset included 71 anonymised DCE–MRI scans of BC patients, comprising 24 cases of in situ tumours and 47 cases of invasive tumours. All scans were acquired using a Philips Achieva 1.5 T MRI scanner, utilising a dynamic eTHRIVE sequence with fat suppression. Only MRI images containing at least one tumour mass, as identified by an experienced radiologist and subsequently confirmed through histopathological biopsy, were selected for analysis. For each case, an ROI was manually delineated around the tumour mass by two expert radiologists on the post-contrast images. The original spatial resolution of the MRI scans varied, with voxel sizes ranging from 1.6 mm to 2.0 mm, depending on acquisition parameters and slice thickness. To ensure consistency across the dataset and enable reliable feature extraction, all MRI volumes were resampled to an isotropic voxel size of 1 mm^3 prior to radiomic processing.

Figure 1 shows a single 2D slice of an image with the manually segmented tumour.

Figure 1. A typical single slice of an image with the manually segmented region of interest (ROI) in red.

2.2. Feature Extraction

For each case included in the dataset, a total of 109 features were extracted. Among these, 107 are radiomic features computed using the PyRadiomics open-source library (https://pyradiomics.readthedocs.io/en/latest/index.html, accessed on 15 June 2025), which enables standardised extraction of quantitative descriptors from medical images.

The radiomic features encompass shape descriptors, first-order intensity statistics, and higher-order texture features derived from grey-level co-occurrence matrices (GLCM), grey-level run length matrices (GLRLM), grey-level size zone matrices (GLSZM), and grey-level dependence matrices (GLDM) [27].

Prior to feature extraction, a visual and quantitative pre-check of grey-level histograms was performed to ensure consistency across the MRI scans. This step confirmed that the intensity distributions were comparable among images, allowing reliable application of uniform extraction parameters across the entire dataset.

In addition to radiomic descriptors, two clinical variables were included: patient age at diagnosis and tumour laterality, indicating whether the lesion was located in the right or left breast. Tumour laterality was numerically encoded.

2.3. Feature Selection

To evaluate the relevance of individual variables and reduce dimensionality, two feature selection strategies were tested: the mutual information-based selection, which scores features based on their statistical dependency with the classification label, and the Minimum Redundancy Maximum Relevance (MRMR) algorithm, which aims to optimise both relevance and feature diversity.

In particular, the first function, given the feature matrix and corresponding class labels, returns a vector of non-negative values, each representing an estimate of the mutual information between a single feature and the label variable. By keeping track of the corresponding features, we obtained a ranked list based on their estimated information gain.

The MRMR is a mutual information-based method that selects features by optimising both relevance and redundancy. It is designed to identify a subset of features that not only exhibit high statistical dependency with the target variable (maximum relevance), but also minimal redundancy among themselves. Although MRMR inherently considers feature redundancy, we also computed a correlation matrix across selected features to qualitatively verify the absence of highly correlated pairs.

2.4. Data Augmentation

To compensate for the relatively small number of available cases and to improve the robustness of the classification models, data augmentation was performed through geometric transformations of the original imaging data. Specifically, each MRI volume was rotated in three-dimensional space along the axial plane by six angles: -30°, -20°, -10°, $+10^\circ$, $+20^\circ$, and $+30^\circ$. These transformations aimed to simulate variations in patient positioning while preserving the anatomical structure of the tumour.

Radiomic features were extracted from each augmented volume using the same PyRadiomics pipeline applied to the original data. The features from the rotated images were then integrated into the training set, effectively increasing the sample size and variability. This augmentation strategy enabled the models to learn more robust and rotation-invariant patterns and helped reduce overfitting, particularly in the presence of class imbalance.

2.5. Classification Algorithms

Several classifiers were tested and comparatively evaluated, including K-Nearest Neighbors (KNN), Naive Bayes, Random Forest (RF), Support Vector Machine (SVM), Logistic Regression, Extreme Gradient Boosting (XGBoost), and a Multi-Layer Perceptron (MLP). All computations and model development steps were implemented using Python (version 3.13) and related open-source libraries (in particular, the scikit-learn library).

To address the class imbalance between in situ (n = 24) and invasive (n = 47) lesions, class weighting strategies were applied to selected models. Specifically, the SVM was configured with class_weight = 'balanced', and the XGBoost classifier was adjusted using the scale_pos_weight parameter to enhance learning from the minority class. The other classifiers were trained without explicit rebalancing.

Model validation was conducted using Leave-One-Out Cross-Validation (LOOCV), where each individual sample is used once as a test case while the rest serve as the training set. This process is repeated for every sample in the dataset. LOOCV is particularly well-suited for small and heterogeneous datasets. At each iteration of the LOOCV scheme, feature standardisation was performed using the training set only. Specifically, z-score normalisation was applied by fitting the scaler on the training data and subsequently applying the same transformation parameters to the test instance.

To prevent data leakage arising from the use of data augmentation techniques within the LOOCV process, the following strategy was adopted. Each original sample was identified by a unique suffix, distinguishing it from its corresponding augmented versions. During each LOOCV fold, the original sample of the current patient was selected as the test set, and all its augmented versions were completely excluded from the training data. The training set was composed exclusively of samples (original and augmented) belonging to the other patients, with the exclusion of any data related to the patient under test. Also, data standardisation (scaling) was performed only on the training data and then applied to the test sample.

For each classifier, the area under the ROC curve (AUC) was computed to comprehensively evaluate discriminative performance. In our analysis, we defined in situ lesions (the minority class) as the positive class. The decision threshold corresponding to the highest overall classification accuracy was then selected to calculate the final evaluation metrics, including precision, recall (sensitivity for in situ), F1-score, and to construct the confusion matrix. A flow chart is provided in Figure 2.

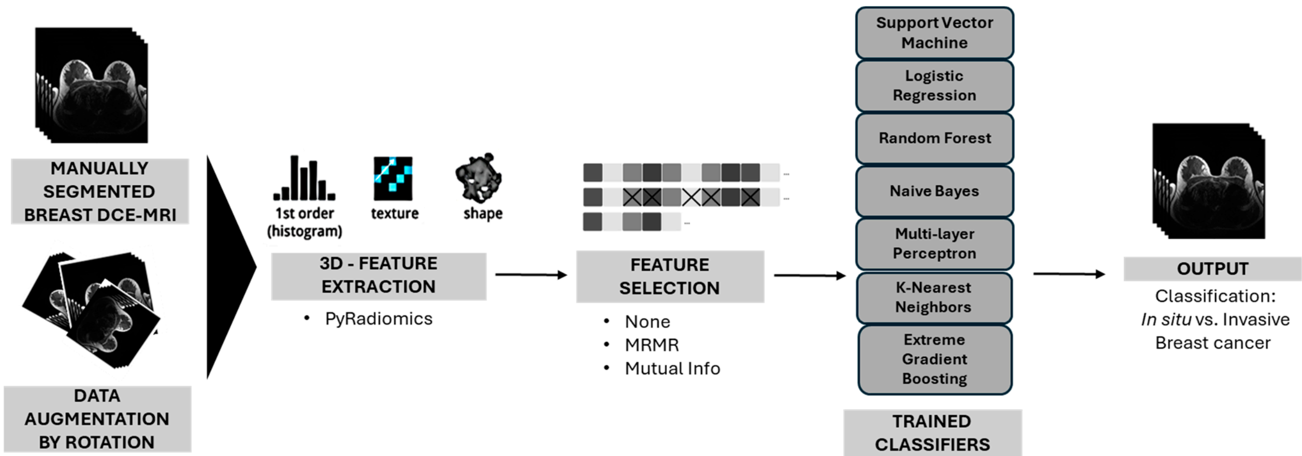

Figure 2. Flow chart of the study.

3. Results

The performance of multiple classification models was evaluated using an LOOCV scheme, in combination with different feature selection strategies and data augmentation. Among all tested configurations, the best-performing models were selected based on the maximum AUC of the ROC curve.

Table 1 summarises the results obtained with the best-performing classifiers: KNN, RF, SVM, and XGBoost. The ROC curves corresponding to the two best-performing configurations are shown in Figure 3, while Figure 4 shows an example distribution of the prediction scores obtained for each sample during LOOCV with the RF model (with the same parameters in Table 1), grouped by class. This plot allows for a visual comparison of the model output between the two classes and provides insight into the variability of the predicted scores across folds.

Table 1. Comparison of the best-performing Machine Learning classifiers in terms of classification performance for distinguishing in situ from invasive breast cancer lesions. Metrics include accuracy, precision, recall, and F1-score calculated at the best-accuracy point.

Model	Hyper Parameters	Augmentation	Feature Selection	Number of Selected Features	AUC	Accuracy (%)	Threshold (Best-Accuracy Point)	Precision (%)	Recall (%)	F1-Score (%)	Confusion Matrix
KNN	n_neighbors = 5 weights = uniform metric = cosine	No	MRMR	13	0.78	0.76	0.80	0.77	0.42	0.54	[44 3] [14 10]
XGBoost	objective = binary:logistic max_depth = 6 learning_rate = 0.01 n_estimators = 300 lambda = 0.5 alpha = 0.5 eval_metric = auc scale_pos_weight = 0.5435	No	Mutual Info	4	0.77	0.76	0.53	0.73	0.46	0.56	[43 4] [16 8]
RF	n_estimators = 200 criterion = gini max_depth = 7 min_samples_split = 5 min_samples_leaf = 2 random_state = 42	Yes	Mutual Info	3	0.81	0.79	0.55	0.76	0.54	0.63	[43 4] [11 13]
SVM	C = 5 kernel = rbf, gamma = scale shrinking = True probability = True class_weight = balanced random_state = 42	Yes	MRMR	3	0.80	0.79	0.80	0.76	0.54	0.63	[43 4] [11 13]
XGBoost	objective = binary:logistic max_depth = 6 learning_rate = 0.01 n_estimators = 300 lambda = 0.5 alpha = 0.5 eval_metric = auc scale_pos_weight = 0.5435	Yes	MRMR	6	0.79	0.77	0.80	0.67	0.67	0.67	[43 4] [12 12]

SVM = Support Vector Machine; KNN = K-Nearest Neighbors; RF = Random Forest; XGBoost = Extreme Gradient Boosting; MRMR = Minimum Redundancy Maximum Relevance.

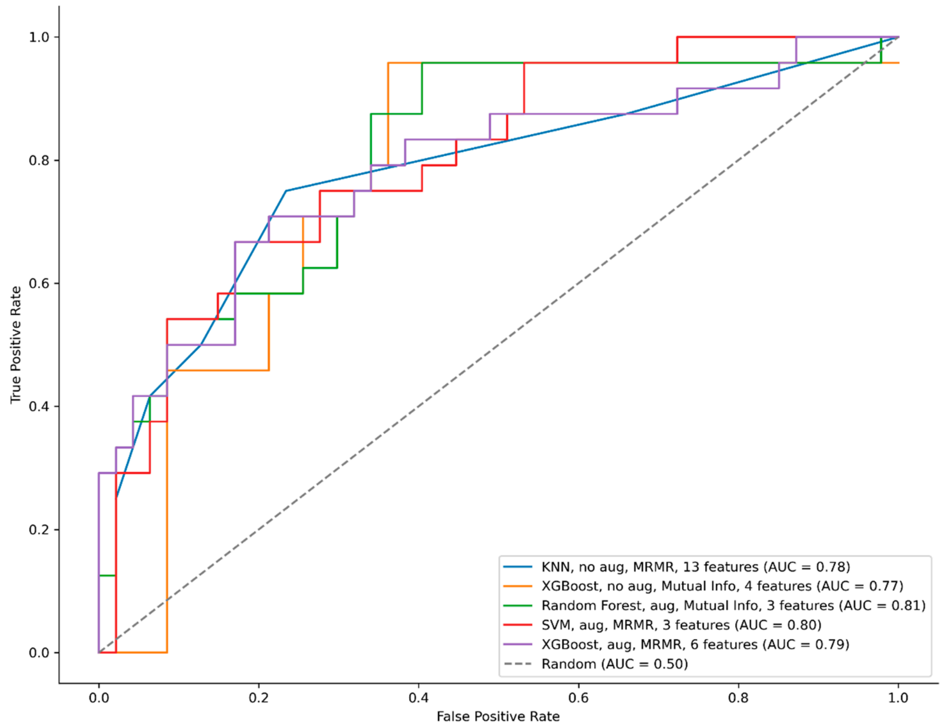

Figure 3. Receiver Operating Characteristic (ROC) curves for the best-performing classifiers. The blue curve represents the K-Nearest Neighbors (KNN) model trained on non-augmented data using the Minimum Redundancy Maximum Relevance (MRMR)-selected features. The orange curve corresponds to the XGBoost model trained on non-augmented data with mutual information-based feature selection. The green curve shows the Random Forest (RF) model trained on augmented data and using mutual information-selected features. The red curve refers to the Support Vector Machine (SVM) model trained on augmented data with MRMR feature selection. The purple curve illustrates the performance of the XGBoost model trained on augmented data with MRMR-selected features. The diagonal grey dashed line indicates the performance of a random classifier (AUC = 0.50).

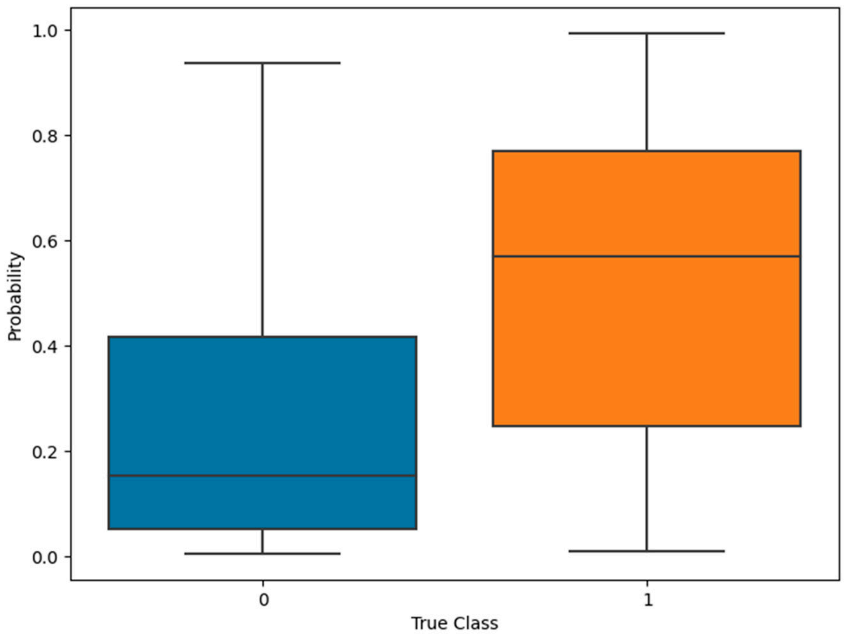

Figure 4. Boxplot of the prediction scores obtained in LOOCV with a Random Forest model, shown separately for the positive and negative classes. Each included value represents the score assigned to a sample when it was used as the validation case. This visualisation highlights the distribution of model outputs across classes.

The KNN classifier, trained on the non-augmented dataset with 13 MRMR-selected features, reached an AUC of 0.78 and a best accuracy of 76%. Specifically, it achieved a recall of 42% and a precision of 77%, resulting in an F1-score of 54%.

The XGBoost model, trained on the original dataset and evaluated using four features selected through mutual information, achieved an AUC of 0.77. At its optimal threshold, the model reached a best accuracy of 76%, precision of 73%, recall of 46%, and an F1-score of 56%.

The RF classifier, trained on the augmented dataset with three features selected via mutual information, yielded the best performance among all models, with an AUC of 0.81. It achieved a best accuracy of 79%, precision of 76%, recall of 54%, and an F1-score of 63%.

The SVM model, trained on the augmented dataset with the top 3 MRMR-selected features, achieved an AUC of 0.80. At the accuracy-maximising operating point, the model reached an accuracy of 79%, with a precision of 76%, a recall of 54%, and an F1-score of 63%.

The XGBoost model, trained on the augmented dataset with 6 MRMR-selected features, reported an AUC of 0.79. It reached a best accuracy of 77%, precision of 67%, recall of 67%, and an F1-score of 67%.

Table 2 summarises the radiomic and clinical features selected by the best-performing configurations of each ML classifier. The number of features retained by each model varied from 3 to 13, depending on the selection method and use of data augmentation. Notably, several features were selected across multiple classifiers, suggesting shared patterns of predictive relevance. For instance, shape-based descriptors such as Elongation, Flatness, and LeastAxisLength were frequently chosen by different models, as were texture-based metrics including GLCM Correlation, ClusterShade, MCC, and RunVariance. The recurrence of these features supports their potential role in capturing biologically meaningful differences between in situ and invasive lesions, as further discussed in the following section. To further support the feature selection process and assess potential redundancies among variables, a correlation matrix was computed over the union of all features selected by the different classifiers, as shown in Figure 5. This analysis allowed us to qualitatively verify the suitability of the selected subsets, especially in the context of redundancy-aware methods such as MRMR, by highlighting the presence or absence of strongly correlated feature pairs.

Table 2. Selected radiomic and clinical features for best-performing classifiers.

Model	Augmentation	Selection Method	Selected Features	Number of Selected Features
KNN	No	MRMR	original_glcm_MCC, original_glrlm_RunPercentage, original_ngtdm_Coarseness, original_glcm_ClusterShade, original_glcm_Correlation, original_shape_SurfaceVolumeRatio, original_glcm_Idn, original_glszm_ZoneEntropy, original_shape_Flatness, original_gldm_DependenceEntropy, original_gldm_DependenceVariance, original_gldm_SmallDependenceLowGrayLevelEmphasis, original_shape_LeastAxisLength	13
XGBoost	No	Mutual Info	age, original_shape_Elongation, original_shape_Flatness, original_shape_LeastAxisLength	4

Table 2. *Cont.*

Model	Augmentation	Selection Method	Selected Features	Number of Selected Features
RF	Yes	Mutual Info	original_shape_Elongation, original_shape_Flatness, original_shape_LeastAxisLength	3
SVM	Yes	MRMR	original_glcm_Correlation, original_glrlm_RunVariance, original_shape_Elongation	3
XGBoost	Yes	MRMR	original_glcm_Correlation, original_glrlm_RunVariance, original_shape_Elongation, original_glszm_ZoneEntropy, original_glcm_ClusterShade, original_glcm_MCC	6

SVM = Support Vector Machine; KNN = K-Nearest Neighbors; RF = Random Forest; XGBoost = Extreme Gradient Boosting; MRMR = Minimum Redundancy Maximum Relevance.

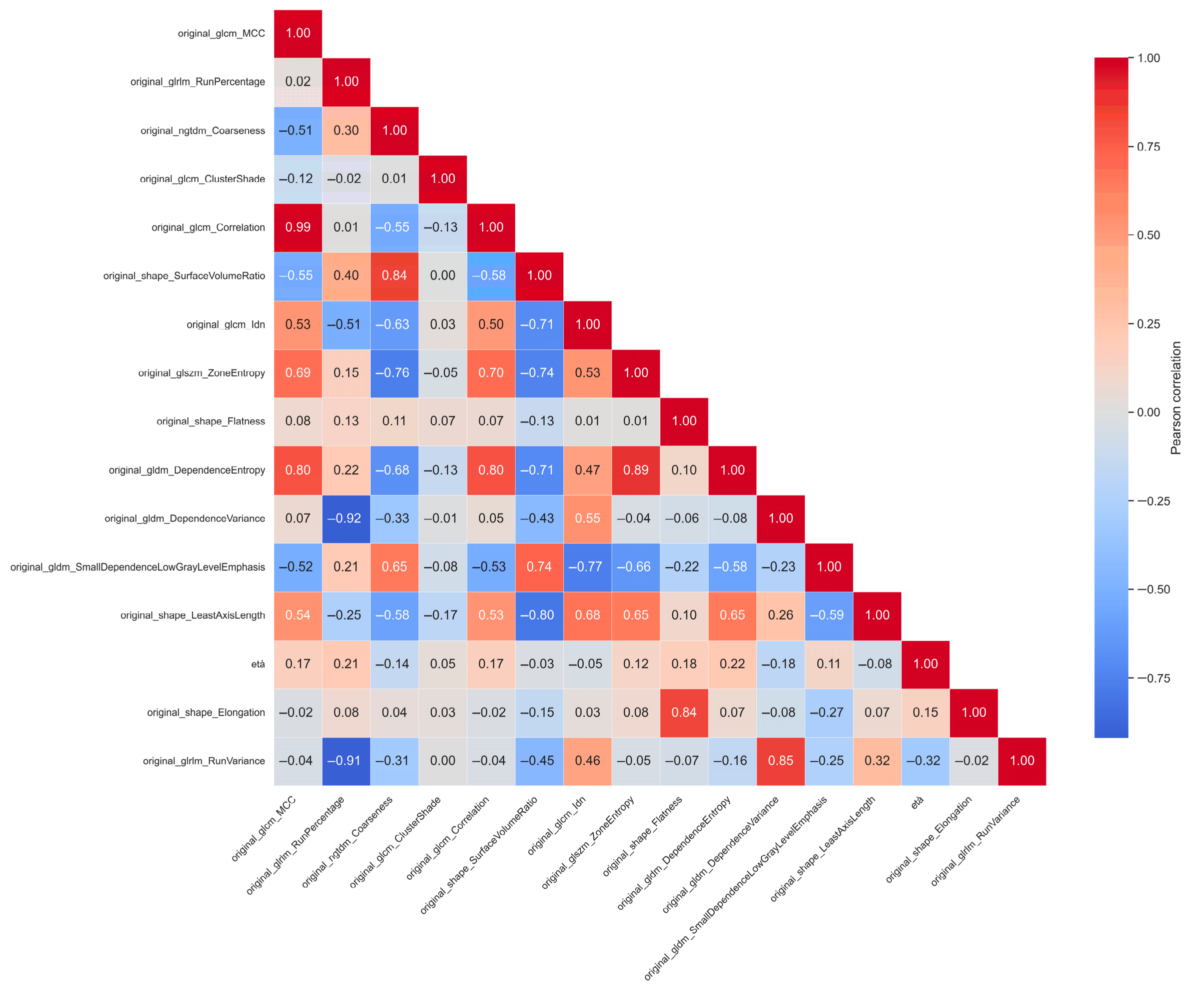

Figure 5. Correlation matrix computed across the union of all radiomic and clinical features selected by the best-performing classifiers.

To complement the analysis of selected features and improve interpretability, we generated a comparative feature importance plot (Figure 6), illustrating the most influential radiomic features across three high-performing models: XGBoost (with and without augmentation) and Random Forest. Feature importance was derived using the internal ranking metrics provided by each algorithm.

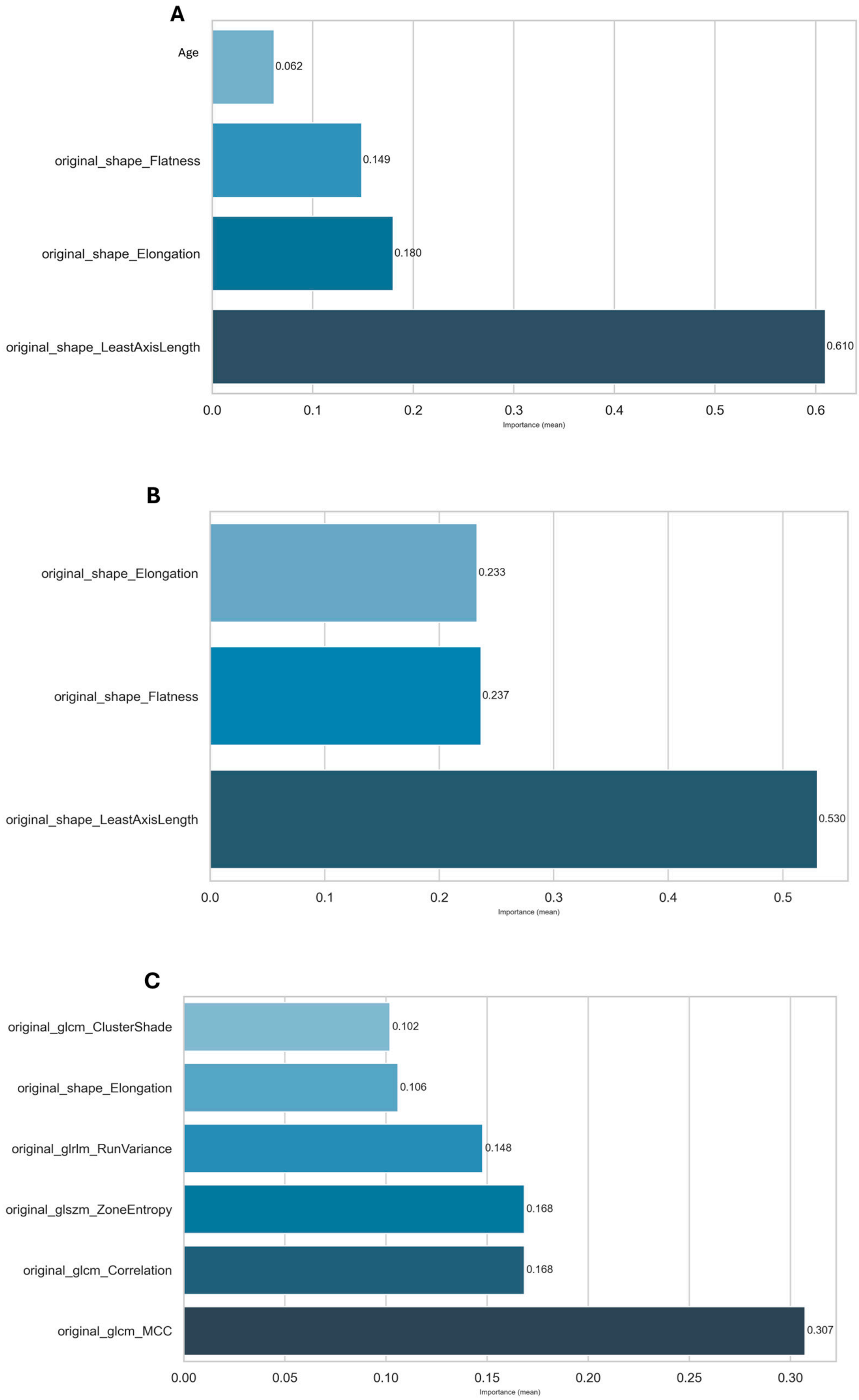

Figure 6. Mean feature importance obtained from ML models trained on sets of selected features. Panel (**A**): Mean importance calculated using an XGBoost model trained on four features. Panel (**B**): Mean importance derived from a Random Forest model trained on augmented images and three morphological features. Panel (**C**): Analysis performed with XGBoost on six features.

4. Discussion

The challenge of distinguishing between invasive and in situ BC has been addressed in a few studies, although it remains an open and evolving area of research. In this context, our study investigated the performance of several ML algorithms combined with radiomic features extracted from manually segmented ROIs on post-contrast DCE–MRI.

Among all classifiers, the highest performing models were SVM, KNN, XGBoost, and RF. While SVM and RF favoured higher precision and lower false-positive rates, KNN and XGBoost showed greater sensitivity toward in situ lesions. In particular, the best-performing model was the Random Forest classifier, which, when trained on the augmented dataset using 3 features selected through mutual information, achieved the highest AUC of 0.81 and an accuracy of 79%. These results demonstrate that traditional ML approaches, when properly optimised and validated, can yield robust performance even on relatively small datasets.

These findings are in line with prior work exploring the potential of Radiomics in the classification of BC subtypes. For instance, in one study, radiomic features extracted from DCE–MRI in a large cohort (190 IDC and 58 DCIS cases) were used to train an RF classifier under LOOCV, achieving a remarkably high AUC of 0.90 [28]. Similarly, Li et al. tested a radiomic signature composed of 569 features derived from mammographic images, applied to a dataset of 161 DCIS and 89 IDC cases, obtaining a maximum AUC of 0.72 [29].

Alternative imaging modalities have also been explored. Diffusion-weighted imaging (DWI) has been used to compute apparent diffusion coefficient (ADC) values, which reflect tissue diffusivity and microstructure. Since invasive tumours typically alter the extracellular matrix through proteolytic activity, lower ADC values are observed compared to in situ lesions [30,31]. One study confirmed this trend by demonstrating a statistically significant difference in ADC between 21 DCIS and 155 IDC cases ($p < 0.001$), with a resulting AUC of 0.89 [32].

In another approach, Bhooshan et al. combined kinetic and morphological MRI features to train a classifier on a dataset comprising 32 benign, 71 DCIS, and 150 IDC cases, achieving an AUC of 0.83 [33]. More recently, deep learning (DL) techniques have also been tested. In one study, a transfer learning strategy using a pre-trained GoogleNet was employed to extract features from 131 breast MRI scans, followed by classification with an SVM, resulting in an AUC of 0.70 [34].

Compared to similar studies, our approach yielded competitive results despite the limited dataset size. For instance, Drukker et al. [28] reported an AUC of 0.90 using a large cohort and Random Forest classifier under LOOCV, while Li et al. [29] obtained an AUC of 0.72 with mammographic radiomics. Bhooshan et al. [33] combined kinetic and morphological features, achieving an AUC of 0.83, and Zhu et al. [34] reported an AUC of 0.70 using deep learning. Our best model, a Random Forest trained on three shape-based features, reached an AUC of 0.81, demonstrating that a carefully designed classical ML pipeline with robust feature selection can achieve results comparable to more complex or data-hungry methods. These findings confirm that radiomics-based ML approaches hold promise even in small, heterogeneous datasets when supported by appropriate preprocessing and validation strategies.

To provide additional interpretability of our ML models, we analysed the specific features selected in the best-performing configurations for each classifier. The selected feature subsets reveal both recurring and unique variables, suggesting that different algorithms may prioritise distinct aspects of tumour morphology and texture in the classification of in situ versus invasive lesions.

For the KNN classifier (MRMR selection, original dataset), the best configuration included 13 features, many of which were texture-based, including original_glcm_MCC, original_glcm_Correlation, original_glcm_ClusterShade, and original_ngtdm_Coarseness, alongside shape descriptors like original_shape_SurfaceVolumeRatio, original_shape_Flatness, and original_shape_LeastAxisLength. This suggests that KNN leveraged both morphological and microtextural heterogeneity of the lesion to achieve robust classification.

The XGBoost model trained on the original dataset with Mutual Information selection retained only four features: patient age, original_shape_Elongation, original_shape_Flatness, and original_shape_LeastAxisLength, highlighting the model's reliance on shape descriptors and the clinical variable age, which may implicitly reflect biological aggressiveness and risk profiles.

The SVM classifier (MRMR selection, with augmentation) achieved an AUC of 0.80 using only three features: original_glcm_Correlation, original_glrlm_RunVariance, and original_shape_Elongation. This minimalistic set suggests that a combination of correlation-based texture and elongation-related morphology can be sufficient for accurate classification.

Similarly, XGBoost with MRMR selection and augmentation achieved high performance (AUC = 0.79) using a 6-feature set, which included both texture (e.g., original_glcm_Correlation, original_glrlm_RunVariance, original_glcm_ClusterShade, original_glcm_MCC) and shape descriptors (original_shape_Elongation), again underlining the relevance of both radiomic domains.

Finally, the RF classifier (with augmentation and Mutual Information selection) relied solely on three shape-related features: original_shape_Elongation, original_shape_Flatness, and original_shape_LeastAxisLength. This indicates that, for tree-based methods, morphological geometry may carry the strongest discriminative signal in this dataset.

Across models, a set of radiomic descriptors was consistently prioritised across different ML algorithms. These recurring features suggest robust discriminatory potential for distinguishing between in situ and invasive breast lesions, regardless of the specific classifier architecture or validation strategy.

Among the most frequently selected features, original_shape_Elongation appeared in four out of five models (XGBoost with and without augmentation, RF, and SVM). This metric captures the deviation of a lesion from a spherical form and may reflect infiltrative growth patterns characteristic of invasive cancers.

Similarly, original_glcm_Correlation, original_shape_Flatness, and original_shape_LeastAxisLength were each included in three models. These features describe different aspects of lesion morphology and internal texture organisation. GLCM Correlation quantifies the linear dependency of grey-level intensities in neighbouring pixels, potentially reflecting architectural complexity. Flatness and Least Axis Length are shape descriptors that provide insights into tumour geometry and spatial extension.

Other features like original_glcm_MCC, original_glcm_ClusterShade, and original_glrlm_RunVariance were selected in two models, confirming their secondary but notable contribution to classification. These texture metrics capture variations in intensity patterns and asymmetry within the lesion, both of which are associated with tumour heterogeneity and invasive potential.

The consistency of these features across models suggests that they capture systematic imaging differences between in situ and invasive tumours, likely related to structural and morphological heterogeneity observable on MRI. Although radiomic features are mathematical descriptors without direct biological interpretation, their recurrence across independent algorithms reinforces their potential value as robust imaging biomarkers. This consistency also supports the interpretability and possible generalizability of the models developed in this study, warranting further validation in larger and more diverse cohorts. From a clinical perspective, the ability to distinguish between in situ and invasive BC preoperatively has significant implications. A reliable, non-invasive radiomic classifier could assist radiologists by providing a second-read tool to support image interpretation, particularly in ambiguous cases or high-volume settings. For surgeons, early differentiation may inform surgical planning, such as the extent of resection or sentinel lymph node biopsy decisions, while potentially reducing overtreatment in patients with in situ disease. Moreover, this approach could facilitate more tailored patient counselling and streamline multidisciplinary team discussions in breast tumour boards.

Compared to previous studies, our results are competitive, particularly given the limited sample size. The use of 3D augmentation and feature selection enabled performance comparable to more complex approaches. Nonetheless, some limitations should be noted: the dataset included only 71 patients, with class imbalance between invasive and in situ lesions, which may affect generalizability despite mitigation strategies. Manual segmentation by an expert radiologist, while high-quality, introduces potential operator dependency and is inherently subject to intra- and inter-observer variability. This variability may impact the consistency of feature extraction and thus affect the model's reproducibility and generalizability to external datasets. Future work could benefit from automated methods to improve reproducibility, reduce observer bias, and enhance scalability. Moreover, only post-contrast DCE–MRI sequences were used—adding other imaging modalities could enhance accuracy. Finally, relying solely on radiomic features may not fully capture image complexity; integrating clinical, histopathological, or genomic data, along with DL techniques, could further improve performance and support personalised decision-making.

5. Conclusions

This study demonstrates the feasibility and effectiveness of combining radiomic features extracted from DCE–MRI with traditional ML algorithms to distinguish between in situ and invasive breast cancer. Despite the limited dataset size, models such as SVM, KNN, RF, and XGBoost achieved promising performance, with a maximum accuracy of 78% and AUC values up to 81%. The results confirm that classical ML models can offer competitive performance compared to more complex approaches. The proposed study could ultimately support radiologists and surgeons in refining diagnostic and therapeutic strategies, particularly in borderline or uncertain cases.

Future research should aim to validate these models on larger, multicentre cohorts, integrate additional imaging modalities and clinical variables, and move toward automated, standardised pipelines for segmentation, feature extraction, and model deployment to facilitate real-world clinical translation.

Author Contributions: Conceptualization, L.C., R.R. and M.P.; Methodology, L.C. and R.R.; Software, L.C.; Validation, L.C. and R.R.; Formal analysis, L.C. and R.R.; Investigation, L.C., R.R., A.S., E.M., M.C., F.T. and A.C.; Resources, L.C.; Data curation, L.C. and G.R.; Writing—original draft, L.C.; Writing—review & editing, L.C., D.C. and G.D.N.; Visualization, L.C., G.R., U.D.G., M.P., D.C. and G.D.N.; Supervision, G.R., U.D.G., R.M., M.P., D.C. and G.D.N. All authors have read and agreed to the published version of the manuscript.

Funding: Some of the techniques adopted in this study were developed within the framework of the PNRR—M4C2-I1.3 Project PE_00000019 "HEAL ITALIA", funded by the European Union—NextGenerationEU through the Italian Ministry of University and Research. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Institutional Review Board Statement: This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Bari, Italy, under protocol number 36, dated 16 January 2024. All patient data were anonymised prior to analysis to ensure compliance with privacy regulations and the General Data Protection Regulation (GDPR, EU 2016/679).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

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.

Conflicts of Interest: The authors have no relevant conflicts of interest to disclose.

References

1. AIOM-AIRTUM-Siapec-Iap I Numeri Del Cancro in Italia 2024. Available online: https://www.aiom.it/wp-content/uploads/2025/01/2024_NDC_web-def.pdf (accessed on 30 April 2025).
2. Bijker, N.; Meijnen, P.; Peterse, J.L.; Bogaerts, J.; Van Hoorebeeck, I.; Julien, J.-P.; Gennaro, M.; Rouanet, P.; Avril, A.; Fentiman, I.S.; et al. Breast-Conserving Treatment with or without Radiotherapy in Ductal Carcinoma-in-Situ: Ten-Year Results of Randomized Phase III Trial. *J. Clin. Oncol.* **2006**, *24*, 3381–3387. [CrossRef] [PubMed]
3. Ernster, V.L. Detection of Ductal Carcinoma In Situ in Women Undergoing Screening Mammography. *CancerSpectrum Knowl. Environ.* **2002**, *94*, 1546–1554. [CrossRef] [PubMed]
4. Northridge, M.E.; Rhoads, G.G.; Wartenberg, D.; Koffman, D. The Importance of Histologic Type on Breast Cancer Survival. *J. Clin. Epidemiol.* **1997**, *50*, 283–290. [CrossRef] [PubMed]
5. Gamel, J.W.; Meyer, J.S.; Feuer, E.; Miller, B.A. The Impact of Stage and Histology on the Long-Term Clinical Course of 163,808 Patients with Breast Carcinoma. *Cancer* **1996**, *77*, 1459–1464. [CrossRef]
6. Conte, L.; De Nunzio, G.; Lupo, R.; Mieli, M.; Lezzi, A.; Vitale, E.; Carriero, M.C.; Calabrò, A.; Carvello, M.; Rubbi, I.; et al. Breast Cancer Prevention: The Key Role of Population Screening, Breast Self-Examination (BSE) and Technological Tools. Survey of Italian Women. *J. Cancer Educ.* **2023**, *38*, 1728–1742. [CrossRef] [PubMed]
7. Conte, L.; Lupo, R.; Lezzi, A.; Paolo, V.; Rubbi, I.; Rizzo, E.; Carvello, M.; Calabrò, A.; Botti, S.; De Matteis, E.; et al. A Nationwide Cross-Sectional Study Investigating Adherence to the Mediterranean Diet, Smoking, Alcohol and Work Habits, Hormonal Dynamics between Breast Cancer Cases and Healthy Subjects. *Clin. Nutr. Open Sci.* **2024**, *55*, 1–19. [CrossRef]
8. Cortesi, L.; Galli, G.R.; Domati, F.; Conte, L.; Manca, L.; Berio, M.A.; Toss, A.; Iannone, A.; Federico, M. Obesity in Postmenopausal Breast Cancer Patients: It Is Time to Improve Actions for a Healthier Lifestyle. The Results of a Comparison Between Two Italian Regions with Different "Presumed" Lifestyles. *Front. Oncol.* **2021**, *11*, 769683. [CrossRef] [PubMed]
9. Conte, L.; Lupo, R.; Lezzi, A.; Sciolti, S.; Rubbi, I.; Carvello, M.; Calabrò, A.; Botti, S.; Fanizzi, A.; Massafra, R.; et al. Breast Cancer Prevention Practices and Knowledge in Italian and Chinese Women in Italy: Clinical Checkups, Free NHS Screening Adherence, and Breast Self-Examination (BSE). *J. Cancer Educ.* **2024**, *40*, 30–43. [CrossRef] [PubMed]
10. Conte, L.; Lupo, R.; Sciolti, S.; Lezzi, A.; Rubbi, I.; Botti, S.; Carvello, M.; Fanizzi, A.; Massafra, R.; Vitale, E.; et al. Exploring the Landscape of Breast Cancer Prevention among Chinese Residents in Italy: An In-Depth Analysis of Screening Adherence, Breast Self-Examination (BSE) Practices, the Role of Technological Tools, and Misconceptions Surrounding Risk Factors and Symptoms. *Int. J. Environ. Res. Public Health* **2024**, *21*, 308. [CrossRef]
11. Mann, R.M.; Cho, N.; Moy, L. Breast MRI: State of the Art. *Radiology* **2019**, *292*, 520–536. [CrossRef] [PubMed]
12. Kuhl, C.K.; Schild, H.H. Dynamic Image Interpretation of MRI of the Breast. *J. Magn. Reson. Imaging* **2000**, *12*, 965–974. [CrossRef] [PubMed]
13. Kuhl, C.K.; Mielcareck, P.; Klaschik, S.; Leutner, C.; Wardelmann, E.; Gieseke, J.; Schild, H.H. Dynamic Breast MR Imaging: Are Signal Intensity Time Course Data Useful for Differential Diagnosis of Enhancing Lesions? *Radiology* **1999**, *211*, 101–110. [CrossRef] [PubMed]
14. Schnall, M.D. Breast MR Imaging. *Radiol. Clin. N. Am.* **2003**, *41*, 43–50. [CrossRef] [PubMed]
15. Morris, E.A. Breast Cancer Imaging with MRI. *Radiol. Clin. N. Am.* **2002**, *40*, 443–466. [CrossRef] [PubMed]
16. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [CrossRef] [PubMed]
17. Brown, M.S.; Goldin, J.G.; Rogers, S.; Kim, H.J.; Suh, R.D.; McNitt-Gray, M.F.; Shah, S.K.; Truong, D.; Brown, K.; Sayre, J.W.; et al. Computer-Aided Lung Nodule Detection in CT: Results of Large-Scale Observer Test. *Acad. Radiol.* **2005**, *12*, 681–686. [CrossRef] [PubMed]
18. Peldschus, K.; Herzog, P.; Wood, S.A.; Cheema, J.I.; Costello, P.; Schoepf, U.J. Computer-Aided Diagnosis as a Second Reader: Spectrum of Findings in CT Studies of the Chest Interpreted as Normal. *Chest* **2005**, *128*, 1517–1523. [CrossRef] [PubMed]
19. Jalloul, R.; Chethan, H.K.; Alkhatib, R. A Review of Machine Learning Techniques for the Classification and Detection of Breast Cancer from Medical Images. *Diagnostics* **2023**, *13*, 2460. [CrossRef] [PubMed]
20. Radak, M.; Lafta, H.Y.; Fallahi, H. Machine Learning and Deep Learning Techniques for Breast Cancer Diagnosis and Classification: A Comprehensive Review of Medical Imaging Studies. *J. Cancer Res. Clin. Oncol.* **2023**, *149*, 10473–10491. [CrossRef] [PubMed]

21. Conte, L.; Rizzo, E.; Grassi, T.; Bagordo, F.; De Matteis, E.; De Nunzio, G. Artificial Intelligence Techniques and Pedigree Charts in Oncogenetics: Towards an Experimental Multioutput Software System for Digitization and Risk Prediction. *Computation* **2024**, *12*, 47. [CrossRef]
22. Adam, R.; Dell'Aquila, K.; Hodges, L.; Maldjian, T.; Duong, T.Q. Deep Learning Applications to Breast Cancer Detection by Magnetic Resonance Imaging: A Literature Review. *Breast Cancer Res.* **2023**, *25*, 87. [CrossRef] [PubMed]
23. Chen, Y.; Shao, X.; Shi, K.; Rominger, A.; Caobelli, F. AI in Breast Cancer Imaging: An Update and Future Trends. *Semin. Nucl. Med.* **2025**, *55*, 358–370. [CrossRef] [PubMed]
24. Conte, L.; Tafuri, B.; Portaluri, M.; Galiano, A.; Maggiulli, E.; De Nunzio, G. Breast Cancer Mass Detection in DCE–MRI Using Deep-Learning Features Followed by Discrimination of Infiltrative vs. In Situ Carcinoma through a Machine-Learning Approach. *Appl. Sci.* **2020**, *10*, 6109. [CrossRef]
25. Conte, L.; Rizzo, E.; Civino, E.; Tarantino, P.; De Nunzio, G.; De Matteis, E. Enhancing Breast Cancer Risk Prediction with Machine Learning: Integrating BMI, Smoking Habits, Hormonal Dynamics, and BRCA Gene Mutations—A Game-Changer Compared to Traditional Statistical Models? *Appl. Sci.* **2024**, *14*, 8474. [CrossRef]
26. Do, L.-N.; Lee, H.-J.; Im, C.; Park, J.H.; Lim, H.S.; Park, I. Predicting Underestimation of Invasive Cancer in Patients with Core-Needle-Biopsy-Diagnosed Ductal Carcinoma In Situ Using Deep Learning Algorithms. *Tomography* **2022**, *9*, 1–11. [CrossRef] [PubMed]
27. Parmar, C.; Grossmann, P.; Bussink, J.; Lambin, P.; Aerts, H.J.W.L. Machine Learning Methods for Quantitative Radiomic Biomarkers. *Sci. Rep.* **2015**, *5*, 13087. [CrossRef] [PubMed]
28. Drukker, K.; Schram, J.; Burda, S.; Li, H.; Lan, L.; Giger, M. Radiomics Investigation in the Distinction between in Situ and Invasive Breast Cancers. *Med. Phys.* **2015**, *42*, 3602–3603. [CrossRef]
29. Li, J.; Song, Y.; Xu, S.; Wang, J.; Huang, H.; Ma, W.; Jiang, X.; Wu, Y.; Cai, H.; Li, L. Predicting Underestimation of Ductal Carcinoma in Situ: A Comparison between Radiomics and Conventional Approaches. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 709–721. [CrossRef] [PubMed]
30. Pinker, K.; Bickel, H.; Helbich, T.H.; Gruber, S.; Dubsky, P.; Pluschnig, U.; Rudas, M.; Bago-Horvath, Z.; Weber, M.; Trattnig, S.; et al. Combined Contrast-Enhanced Magnetic Resonance and Diffusion-Weighted Imaging Reading Adapted to the "Breast Imaging Reporting and Data System" for Multiparametric 3-T Imaging of Breast Lesions. *Eur. Radiol.* **2013**, *23*, 1791–1802. [CrossRef] [PubMed]
31. Spick, C.; Pinker-Domenig, K.; Rudas, M.; Helbich, T.H.; Baltzer, P.A. MRI-Only Lesions: Application of Diffusion-Weighted Imaging Obviates Unnecessary MR-Guided Breast Biopsies. *Eur. Radiol.* **2014**, *24*, 1204–1210. [CrossRef] [PubMed]
32. Bickel, H.; Pinker-Domenig, K.; Bogner, W.; Spick, C.; Bagó-Horváth, Z.; Weber, M.; Helbich, T.; Baltzer, P. Quantitative Apparent Diffusion Coefficient as a Noninvasive Imaging Biomarker for the Differentiation of Invasive Breast Cancer and Ductal Carcinoma in Situ. *Investig. Radiol.* **2015**, *50*, 95–100. [CrossRef] [PubMed]
33. Bhooshan, N.; Giger, M.L.; Jansen, S.A.; Li, H.; Lan, L.; Newstead, G.M. Cancerous Breast Lesions on Dynamic Contrast-Enhanced MR Images: Computerized Characterization for Image-Based Prognostic Markers. *Radiology* **2010**, *254*, 680–690. [CrossRef] [PubMed]
34. Zhu, Z.; Harowicz, M.; Zhang, J.; Saha, A.; Grimm, L.J.; Hwang, E.S.; Mazurowski, M.A. Deep Learning Analysis of Breast MRIs for Prediction of Occult Invasive Disease in Ductal Carcinoma In Situ. *Comput. Biol. Med.* **2019**, *115*, 103498. [CrossRef] [PubMed]