Article

Prediction of Breast Cancer Histological Outcome by Radiomics and Artificial Intelligence Analysis in Contrast-Enhanced Mammography

Antonella Petrillo [1,*], Roberta Fusco [2], Elio Di Bernardo [2], Teresa Petrosino [1], Maria Luisa Barretta [1], Annamaria Porto [1], Vincenza Granata [1], Maurizio Di Bonito [3], Annarita Fanizzi [4], Raffaella Massafra [5], Nicole Petruzzellis [5], Francesca Arezzo [6], Luca Boldrini [7] and Daniele La Forgia [8]

[1] Radiology Division, Istituto Nazionale Tumori—IRCCS—Fondazione G. Pascale, 80131 Naples, Italy; t.petrosino@istitutotumori.na.it (T.P.); m.barretta@istitutotumori.na.it (M.L.B.); a.porto@istitutotumori.na.it (A.P.); v.granata@istitutotumori.na.it (V.G.)

[2] Medical Oncology Division, Igea SpA, 80013 Naples, Italy; r.fusco@igeamedical.com (R.F.); e.dibernardo@igeamedical.com (E.D.B.)

[3] Pathology Division, Istituto Nazionale Tumori—IRCCS—Fondazione G. Pascale, 80131 Naples, Italy; m.dibonito@istitutotumori.na.it

[4] Direzione Scientifica—IRCCS Istituto Tumori Giovanni Paolo II, Via Orazio Flacco 65, 70124 Bari, Italy; a.fanizzi@oncologico.bari.it

[5] SSD Fisica Sanitaria—IRCCS Istituto Tumori Giovanni Paolo II, Via Orazio Flacco 65, 70124 Bari, Italy; r.massafra@oncologico.bari.it (R.M.); n.petruzzellis@oncologico.bari.it (N.P.)

[6] Obstetrics and Gynecology Unit, Department of Biomedical Sciences and Human Oncology, University of Bari Aldo Moro, Piazza Giulio Cesare 11, 70124 Bari, Italy; francescaarezzo@libero.it

[7] Dipartimento di Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, 00168 Roma, Italy; luca.boldrini@policlinicogemelli.it

[8] Struttura Semplice Dipartimentale di Radiodiagnostica Senologica—IRCCS Istituto Tumori Giovanni Paolo II, Via Orazio Flacco 65, 70124 Bari, Italy; d.laforgia@oncologico.bari.it

* Correspondence: a.petrillo@istitutotumori.na.it

Simple Summary: The assessment of breast lesions through mammographic images is currently challenging, especially in dense breasts. Contrast-enhanced mammography has been shown to overcome the limitations of standard mammography but it greatly depends on the interpretative skills of the physician. The aim of this study was to evaluate the potentialities of statistical and artificial intelligence algorithms as a tool for helping the radiologists in the interpretation of images. The most remarkable results were achieved in discriminating benign from malignant lesions and in the identification of the presence of the hormone receptor. A tool to support the physician's decision-making process may be designed starting from simple logistic regression and tree-based algorithms. This type of tool may help the radiologist in assessing the investigated breast and in choosing the appropriate follow-up without resorting to histology.

Abstract: Purpose: To evaluate radiomics features in order to: differentiate malignant versus benign lesions; predict low versus moderate and high grading; identify positive or negative hormone receptors; and discriminate positive versus negative human epidermal growth factor receptor 2 related to breast cancer. Methods: A total of 182 patients with known breast lesions and that underwent Contrast-Enhanced Mammography were enrolled in this retrospective study. The reference standard was pathology (118 malignant lesions and 64 benign lesions). A total of 837 textural metrics were extracted by manually segmenting the region of interest from both craniocaudally (CC) and mediolateral oblique (MLO) views. Non-parametric Wilcoxon–Mann–Whitney test, receiver operating characteristic, logistic regression and tree-based machine learning algorithms were used. The Adaptive Synthetic Sampling balancing approach was used and a feature selection process was implemented. Results: In univariate analysis, the classification of malignant versus benign lesions achieved the best performance when considering the original_gldm_DependenceNonUniformity feature extracted on CC view (accuracy of 88.98%). An accuracy of 83.65% was reached in the classification of grading, whereas a slightly lower value of accuracy (81.65%) was found in the classification of the presence

of the hormone receptor; the features extracted were the original_glrlm_RunEntropy and the original_gldm_DependenceNonUniformity, respectively. The results of multivariate analysis achieved the best performances when using two or more features as predictors for classifying malignant versus benign lesions from CC view images (max test accuracy of 95.83% with a non-regularized logistic regression). Considering the features extracted from MLO view images, the best test accuracy (91.67%) was obtained when predicting the grading using a classification-tree algorithm. Combinations of only two features, extracted from both CC and MLO views, always showed test accuracy values greater than or equal to 90.00%, with the only exception being the prediction of the human epidermal growth factor receptor 2, where the best performance (test accuracy of 89.29%) was obtained with the random forest algorithm. Conclusions: The results confirm that the identification of malignant breast lesions and the differentiation of histological outcomes and some molecular subtypes of tumors (mainly positive hormone receptor tumors) can be obtained with satisfactory accuracy through both univariate and multivariate analysis of textural features extracted from Contrast-Enhanced Mammography images.

1. Introduction

Mammography is one of the main techniques in the diagnosis of breast cancer, showing a key role in both screening and follow-up [1,2]. Mammographic screening has been shown to be highly accurate in detection of breast lesions; however, it suffers from some limitations, especially in the case of dense breasts. In fact, dense breasts show a hyper-intense signal over the mammary parenchyma, resulting in very little contrast between the latter and the lesions. For the mammographic screening of patients with dense breasts, other techniques, such as Magnetic Resonance Imaging (MRI), are commonly preferred [3,4]. In particular, one of the most recent novel approaches is Contrast-Enhanced Mammography (CEM). Combining the potential and benefits of Full-Field Digital Mammography (FFDM), CEM has been shown to be highly effective for the detection and the correct staging of cancer, particularly in dense breasts [4–10]. More specifically, CEM combines the enhancing properties of the intravenous administration of an iodinated contrast medium with the high precision of digital imaging from FFDM; therefore, the neo-vascularity associated with actively growing malignancy is remarkably emphasized. Due to this property, CEM is not only able to detect cancer with high accuracy, but it is also a powerful technique for the identification of cancers that are obscure at mammography; furthermore, it allows a more accurate evaluation of the disease extent and offers guidance in the planning of surgery and treatment [4–10]. However, as in all imaging techniques, the evaluation of CEM images depends on the experience and skills of the radiologist, making the identification of automated or semi-automated techniques, which can provide decision support, a considerable challenge.

Recent significant advancements in this sense rely on the application of artificial intelligence and radiomics for the processing of large quantities of data by different imaging modalities [11,12].

Radiomics is the process of extracting quantitative properties, named features, from medical images. This feature extraction generally includes pattern recognition algorithms and provides, as a result, a set of numbers, each representing a quantitative description of a specific either geometrical or physical property of the image portion under consideration. In the context of tumor characterization, the radiomics features typically considered are those that describe properties related to size, shape, intensity, and texture of the tumor [13–27].

Biological and molecular features related to breast cancer are commonly extracted by biopsy, which is invasive and not always able to detect tumor heterogeneity [28]. In recent years, there has been growing interest in non-invasive methods to directly derive

insights from radiologic images. In this context, the radiomics analysis of tumor features extracted from CEM represents an important tool for breast tumor characterization. As several authors suggest, radiomics analysis combined with artificial intelligence techniques can be used to create a tool to support the physician's decision-making process in the classification of breast cancer [29–44]. In fact, through an appropriate tool, the physician would be able to discriminate the tumor nature and/or grading, identifying the adequate treatment for a single patient (e.g., neoadjuvant therapy) or even a more conservative approach (e.g., wait-and-see or conservative surgery). However, based on our knowledge, only some recent studies have used CEM in the prediction of histological grading and receptor status of breast cancer [45,46].

This work aimed to evaluate radiomics features to differentiate malignant versus benign lesions, to predict low versus moderate and high grading, to identify positive or negative hormone receptors, and to discriminate positive versus negative human epidermal growth factor receptor 2 related to breast cancer.

2. Methods

2.1. Patient Selection

From October 2017 to December 2021, according to regulations issued by the local Institutional Review Board, 182 patients (mean age ± standard deviation of 55.3 ± 10.9 years (range 31–80)) with known breast lesions were enrolled retrospectively. All women signed informed consent.

Inclusion criteria: patients with known breast lesions (from radiological or clinical screening, symptom of palpable lesions), histologically proven, and that underwent dual-energy CEM. CEM images of patients were acquired at Istituto Nazionale Tumori-IRCCS-Fondazione G. Pascale (Naples, Italy) and at Istituto Tumori "Giovanni Paolo II" of Bari (Bari, Italy).

Exclusion criteria: patient with breast implants, presence of non-removable drilling at the nipple, pacemakers, clips or other metal implants, pregnancy or possible pregnancy, inability to keep upright immobility during the examination, renal disease, or chemotherapy treatment at the time of imaging [41].

Overall, 118 malignant lesions and 64 benign lesions were analyzed.

2.2. Imaging Protocol

A total of 136 CEM examinations were performed using the Selenia® Dimensions® Unit dual-energy mammography system (Hologic, Bedford, MA, USA), whereas the remaining 46 CEM image were acquired with the Senographe Essential dual-energy mammography system (GE Healthcare, Princeton, NJ, USA).

The same acquisition protocol was implemented for all the images using both scanners. Specifically, two minutes after the intravenous injection of 1.5 mL/(kg bw) of iodinated contrast medium (Visipaque 320; GE Healthcare, Inc., Princeton, NJ, USA) at a rate of 2–3 mL/s, a set of images was acquired in quick succession, in both CC and MLO views. The CEM examination obtained two images: a low-energy (LE) acquisition at 26–30 kVp and a high-energy (HE) acquisition at 45–49 kVp, depending on breast density and thickness. CEM acquisition details were reported in previous studies [41,42,45].

2.3. Image Processing

Two expert radiologists, with 25 and 20 years of experience in breast imaging, manually segmented images by drawing slice-by-slice the contours of the lesions where contrast uptake was emphasized both in CC and MLO views.

MRI Post-Processing with PyRadiomics Tool

For each region of interest, 837 radiomics features were extracted as median values using the PyRadiomics Python package [47] including: First Order Statistics, Grey Level Co-occurrence Matrix, Grey Level Run Length Matrix, Grey Level Size Zone Matrix, Neigh-

boring Grey Tone Difference Matrix, and Grey Level Dependence Matrix features before and after the wavelet filtering. The extracted features comply with feature definitions as described by the Imaging Biomarker Standardization Initiative (IBSI) [48] and as reported in (https://readthedocs.org/projects/pyradiomics/downloads/, accessed on 20 January 2017).

We used wavelet filtering, with all possible combinations of both high-pass (H) and low-pass (L) filters along the three axes (X, Y, and Z axes), to derive six different matrices:

- First Order (FIRST ORDER): Describes the individual values of voxels obtained as a result of ROI cropping. These are generally histogram-based properties (energy, entropy, kurtosis, skewness).
- Gray Level Co-occurrence Matrix (GLCM): Calculates how often the same and similar pixel values come together in an image and records statistical measurements according to this matrix. These resulting values numerically characterize the texture of the image.
- Gray Level Run Length Matrix (GLRLM): Defined as the number of homogeneous consecutive pixels with the same gray tone and quantifies the gray-level values.
- Gray Level Size Zone Matrix (GLSZM): Describes voxel counts according to the logic of measuring gray-level regions in an image.
- Neighboring Gray Tone Difference Matrix (NGTDM): Digitization of textures obtained from filtered images and their fractal properties.
- Gray Level Dependence Matrix (GLDM): Number of bound voxels at a fidex distance from the central voxel.

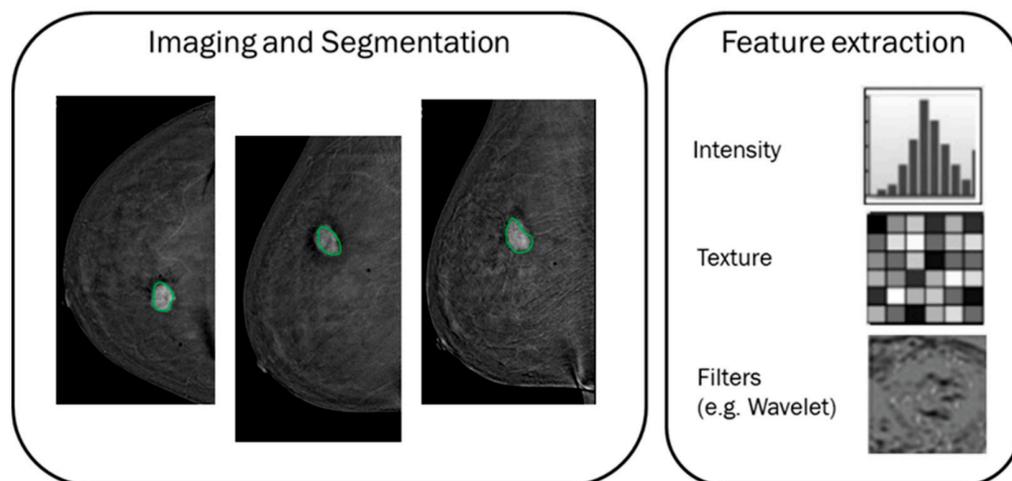A graphical representation of the process for features extraction in a radiomics context is reported in Figure 1.

Figure 1. A graphical representation of the features extraction process in a radiomics context. The green lines represent the segmentation of lesion contours.

2.4. Histopathological Analysis

The reference standard (ground truth) was the histopathologic examination of tissue as reported in [41]. Breast lesions were categorized based on the American Joint Committee on Cancer staging. The histological grade and the expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67 antigen associated with cell proliferation were determined by immune-histochemical analysis.

The tumor grade G was defined on a three-grade scale according to the Elston–Ellis modification of the Scar–Bloom–Richardson grading system.

The hormone receptor (HR) was also considered; a breast cancer is classified as HR-positive if its cells have receptors for the hormones estrogen and progesterone.

2.5. Statistical Analysis

The statistical analysis was performed using the R programming language (version 4.0.2) with the RStudio software, version 1.3.959 (https://www.rstudio.com/, accessed on 20 January 2017) [49].

Considering the histologic results as ground truth, four different types of outcomes were used in both univariate and multivariate analysis: (1) nature of tumor (benign versus malignant); (2) grading (G1 versus G2 + G3); (3) presence of human epidermal growth factor receptor 2 (HER2+ versus HER2−); (4) presence of hormone receptor (HR+ versus HR−).

Before proceeding with statistical analysis, the dataset was balanced with respect of each outcome. The balancing was performed through the synthetization of samples for the less-represented classes using the Adaptive Synthetic Sampling (ADASYN) approach [50,51].

In the context of univariate analysis, the non-parametric Wilcoxon–Mann–Whitney test for continuous variables was used. Receiver operating characteristic (ROC) analysis and the Youden index were considered to obtain the optimal cut-off value for each feature; then, the area under ROC curve (AUC), sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), negative predictive value (NPV), and accuracy (ACC) were computed. Bonferroni correction was used to adjust for multiple comparison.

In the context of multivariate analysis, logistic regression and tree-based algorithms were appropriately designed to predict each outcome individually; the main predictive features were also extracted. Before proceeding with the analysis, three pre-processing steps were performed.

Firstly, the dataset was randomly split into a training set and a test set, using the createDataPartition R function. Specifically, 90% of the entire dataset was used to train the algorithms, designing a cross-validated procedure; the remaining 10% of samples was used to estimate the accuracy of algorithms on 'new' samples, which are samples not used to train the algorithms themselves. Successively (and before running algorithms), a variable selection procedure was designed to remove redundant features from the training set. To achieve this aim, the cross-correlation between each predictor was calculated and all the features with a correlation higher than 0.7 (as an absolute value) with each single predictor were discarded. Finally, the input predictors were centered and scaled before running the logistic regression algorithm.

The machine learning approaches designed for the aim of this paper are described in the following. For each approach, the performance (accuracy) was assessed on both the training and test sets, also considering the values of sensitivity and specificity.

Logistic Regression. Considering the dichotomic nature of each outcome, a logistic regression was executed using all non-redundant features. The method was run using the glm R function.

Logistic Regression with least absolute shrinkage and selection operator (LASSO) method. In a different approach, the logistic regression model was fitted on training data, performing a further variable selection with the LASSO regularization method [52,53]. The LASSO was designed using the glmnet R function and the hyperparameter was tuned through a 10-fold cross validation procedure. The variables selected were saved to train the logistic regression algorithm.

Logistic Regression with two predictors. An additional variation of the logistic regression was considered predicting each outcome with all possible couples of features. All combinations that reached a test accuracy higher than 0.9 were saved and analyzed.

Tree-based algorithms. Among all tree-based algorithms, Classification and Regression Trees (CART) and Random Forest (RF) algorithms were chosen and designed. The CART algorithm was trained taking into account the possibility of obtaining a decision chart, whereas the RF method was used for a more robust evaluation of performances. Tuning of functions' hyperparameters was performed through a 10-fold cross validation procedure.

3. Results

Table 1 shows the distribution of characteristic of analyzed patients.

Table 1. Distribution of analyzed patients.

Characteristic	Distribution	
Age	Min value	25
	Max value	82
	Median value	52
Tumor nature	benign	64
	malignant	118
Tumor grading	G1	78
	G2 + G3	104
Human epidermal growth factor receptor 2	HER2+	135
	HER2−	47
Hormone receptor	HR+	93
	HR−	89
Histotype	0	16
	1	2
	2	80
	3	19
	4	14
	5	51

Table 2 reports the diagnostic accuracy of significant textural parameters for dual-energy CEM, in both CC and MLO views, obtained in the context of univariate analysis.

Table 2. Performance results of univariate analysis both on CC and MLO view.

Performance Results at Univariate Analysis	Benign Versus Malignant Lesions by CC-View	Benign Versus Malignant Lesions by MLO-View	G1 Versus G2 + G3 by CC-View	G1 Versus G2 + G3 by MLO-View	Identification of HER2+ by CC-View	Identification of HER2+ by MLO-View	Identification of HR+ by CC-View	Identification of HR+ by MLO-View
	original_gldm_ DependenceNonUniformity	wavelet_LLL_gldm_ DependenceNonUniformity	original_glrlm_ RunEntropy	wavelet_LLL_glrlm_ RunEntropy	wavelet_HLL_ glcm_Idn	wavelet_HLH_ glcm_Idm	original_gldm_ DependenceNonUniformity	wavelet_LLL_gldm_ DependenceNonUniformity
AUC	0.8587	0.8406	0.8237	0.7643	0.7150	0.7081	0.7500	0.7334
SENS	0.9237	0.8220	0.9038	0.7981	0.5481	0.5704	0.9699	0.8495
SPEC	0.8559	0.8814	0.7692	0.7692	0.8148	0.8148	0.6559	0.6882
PPV	0.8651	0.8739	0.7966	0.7757	0.7475	0.7549	0.7355	0.7315
NPV	0.9182	0.8320	0.8889	0.7921	0.6433	0.6548	0.9385	0.8205
ACC	0.8983	0.8517	0.8365	0.7837	0.6815	0.6926	0.8165	0.7688
Cut-off	2.3093	4.1147	0.8023	0.8732	0.8866	0.7384	2.5524	4.2121

As the table shows, in the classification of malignant versus benign lesions, the best performance was reached by the original_gldm_DependenceNonUniformity feature, extracted on CC view, with an accuracy of 89.83%, a sensitivity of 92.37%, and a specificity of 85.59%, and with a cut-off of 2.31.

In the classification of grading, the best performance was reached by the original_glrlm_RunEntropy feature, extracted on CC view, with an accuracy of 83.65%, a sensitivity of 90.38%, and a specificity of 76.92%, and with a cut-off of 0.80.

In the identification of HER2+, the best performance was reached by the wavelet_HLH_gldm_LargeDependenceHighGrayLevelEmphasis feature, extracted on MLO view, with an accuracy of 69.63%, a sensitivity of 62.22%, and a specificity of 77.04%, and with a cut-off of 0.74.

In the identification of HR+, the best performance was reached by the original_gldm_DependenceNonUniformity feature, extracted on CC view, with an accuracy of 81.65%, a sensitivity of 96.99%, and a specificity of 65.59%, and with a cut-off of 2.55.

Tables 3 and 4 show the results obtained with logistic regression-based and tree-based methods, respectively.

Table 3. Results for logistic regression with and without LASSO regularization.

Results for Single Outcome	Logistic Regression				Logistic Regression with LASSO			
	Trainset	Test Set			Trainset	Test Set		
	ACC	ACC	SENS	SPEC	ACC	ACC	SENS	SPEC
CC—Tumor nature	0.9583	0.9583	1.0000	0.9286	0.9167	0.9167	0.9000	0.9286
MLO—Tumor nature	0.7500	0.7500	0.8333	0.6667	0.8750	0.8750	1.0000	0.7500
CC—Grading	0.8333	0.8333	0.8571	0.8000	0.7917	0.7917	0.9286	0.6000
MLO—Grading	0.7083	0.7083	0.8462	0.5455	0.7917	0.7917	0.7692	0.8182
CC—HER2	0.7143	0.7143	0.7778	0.6000	0.7857	0.7857	1.0000	0.4000
MLO—HER2	0.6786	0.6786	0.5333	0.8462	0.8214	0.8214	0.8000	0.8462
CC—HR	0.8500	0.8500	0.8182	0.8889	0.8500	0.8500	0.7273	1.0000
MLO—HR	0.7500	0.7500	0.7500	0.7500	0.7000	0.7000	0.5000	1.0000

Table 4. Results for CART and RF methods.

Results for Single Outcome	CART				Random Forest			
	Trainset	Test Set			Trainset	Test Set		
	ACC	ACC	SENS	SPEC	ACC	ACC	SENS	SPEC
CC—Tumor nature	0.9122	0.9167	0.9000	0.9286	0.9259	0. 9167	0.9000	0.9286
MLO—Tumor nature	0.8825	0.8333	1.0000	0.6667	0.8968	0.8750	1.0000	0.7500
CC—Grading	0.8073	0.9167	0.9286	0.9000	0.8265	0.8750	0.9286	0.8000
MLO—Grading	0.7660	0.8333	0.8462	0.8182	0.8021	0.8750	0.9231	0.8182
CC—HER2	0.6992	0.6071	0.4444	0.9000	0.7463	0.7143	0.6111	0.9000
MLO—HER2	0.7084	0.8214	0.8667	0.7692	0.8289	0.8929	0.8667	0.9231
CC—HR	0.8045	0.8000	0.6364	1.0000	0.8125	0.8500	0.7273	1.0000
MLO—HR	0.7331	0.7000	0.5000	1.0000	0.7756	0.8000	0.6667	1.0000

Considering the CC view, the best performances were obtained when predicting the tumor nature (malignant versus benign). Logistic regression proved to be the best performing model (test accuracy of 95.83%) when using an approach without LASSO regularization. The goodness of the logistic regression method was also observed with LASSO regularization (test accuracy of 91.67%). Almost comparable results were obtained when using the tree-based algorithms (Table 5), with a test accuracy of 91.67% in the prediction of tumor nature. The decisional chart obtained with the CART method is shown in Figure 2 and the goodness of training procedure on 500 trees with the RF method is shown in Figure 3.

Table 5. Examples of results for logistic regression methods run using all possible combinations of two predictors.

| Results for Single Outcome	ACC	SENS	SPEC	Var 1	Var 2
CC—Tumor nature	0.9583	1.0000	0.9286	original_gldm_SmallDependenceEmphasis	original_firstorder_TotalEnergy
MLO—Tumor nature	0.9167	1.0000	0.8333	original_gldm_LargeDependenceHighGrayLevelEmphasis	wavelet_LHL_glcm_MaximumProbability
CC—Grading	0.9167	0.9286	0.9000	original_gldm_SmallDependenceEmphasis	wavelet_HLL_firstorder_Energy
MLO—Grading	0.9167	1.0000	0.8182	original_glrlm_RunPercentage	original_glszm_LargeAreaLowGrayLevelEmphasis
CC—HR	0.9000	0.8182	1.0000	original_glcm_InverseVariance	original_glcm_DifferenceVariance
MLO—HR	0.9500	0.9167	1.0000	original_firstorder_Maximum	wavelet_LHL_glrlm_RunPercentage
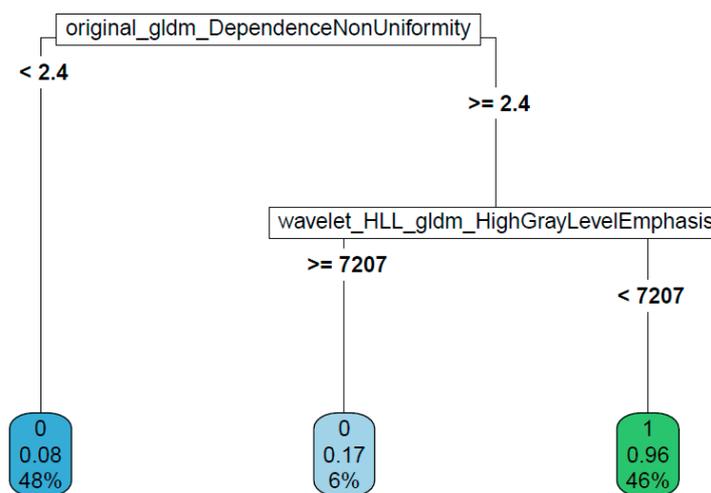
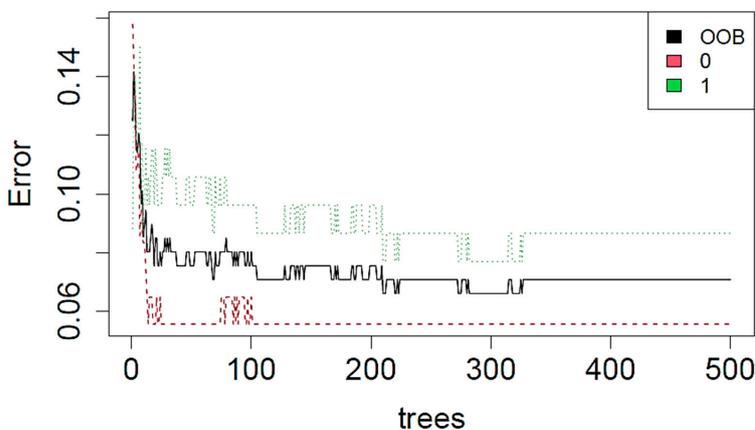**Figure 2.** Decisional chart for the prediction of tumor nature from CC images.

Figure 3. Error evolution during the training procedure of the RF method for the prediction of tumor nature from CC images.

Considering the features extracted from MLO view images, the best test accuracy (91.67%) was obtained when predicting the grading using a CART algorithm, while the

use of all non-redundant features (44 predictors) in a logistic regression model significantly reduces the accuracy value (max test accuracy of 75.00%). The goodness of tree-based algorithms is confirmed by the error evolution plot of RF, reaching an accuracy value of 87.50% on the test set. The decision chart and error evolution are shown in Figure 4.
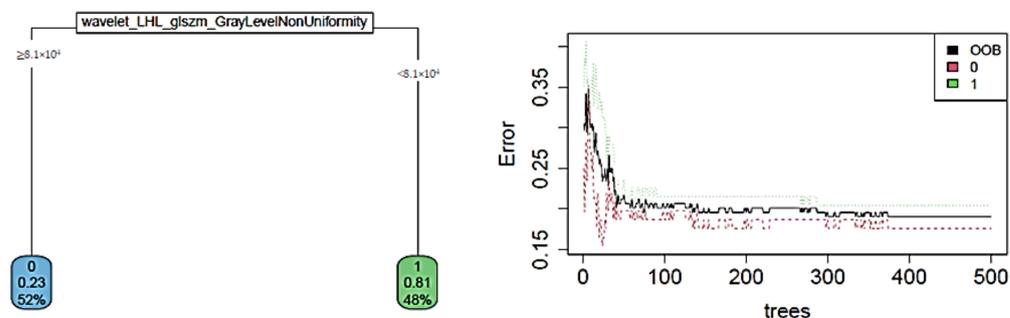

Figure 4. The decision chart and error evolution for the prediction of grading from MLO images.

Combinations of only two features, extracted from both CC and MLO views, always showed test accuracy values greater than or equal to 90.00% (Table 5), with the only exception being the HER2 outcome, where the best performance (test accuracy of 89.29%) was obtained with the RF algorithm.

4. Discussions

The radiomics analysis of tumor features extracted from CEM images represents an important tool for breast cancer characterization.

In this study, we aimed to perform radiomics analysis with texture features extracted by dual-energy CEM, evaluating its ability to classify malignant and benign breast lesions, and to predict grading and breast cancer receptors status (HER2+ and HR+).

In recent years, many studies have addressed the problem of breast lesion classification using several feature categories, such as morphological and textural features, in combination with different machine learning approaches, based on CEM and Dynamic Contrast-Enhanced MRI image analysis [29–40,54–58], whereas other studies used CEM to predict histological outcomes [45,46].

La Forgia et al. [45] assessed the discrimination power of the statistical features extracted from CEM images to predict histological outcomes and two particular subtypes of tumors, HER2-positive and triple-negative. In their work, they showed encouraging results for the differentiation between ER+/ER−, PR+/PR−, HER2+/HER2−, Ki67+/Ki67−, and High-Grade/Low-Grade. In particular, the highest performances were obtained for discriminating HER2+/HER2− (90.87%), ER+/ER− (83.79%), and Ki67+/Ki67− (84.80%).

In a retrospective study, Marino et al. [46] examined the potential of radiomics analysis using features from both CEM and MRI. In particular, they assessed the tumor invasiveness, the hormone receptor status, and the tumor grade in patients with primary breast cancer through common radiomics parameters. In their results, they showed remarkable accuracies when performing CEM radiomics analysis for discriminating HR+ versus HR− breast cancers (95.6%) and invasive versus non-invasive breast cancers (92.0%); slightly lower results were obtained, instead, in the classification of G1 + G2 versus G3 invasive cancers (77.8%).

The results of the univariate analysis of the present study show that the classification of malignant versus benign lesions achieved the best performance when considering the original_gldm_DependenceNonUniformity feature extracted on CC view (accuracy of 88.98%). The features extracted on CC view appeared to perform better, as the results in the classification of both grading and HR suggest. In fact, an accuracy of 83.65% was reached in the classification of grading, whereas a slightly lower value of accuracy (81.65%) was found in the classification of HR+; the features extracted were the original_glrlm_RunEntropy and the original_gldm_DependenceNonUniformity, respectively. In the identification of HER2+,

the best performance, however low (accuracy of 69.63%), was reached when considering the wavelet_HLH_gldm_LargeDependenceHighGrayLevelEmphasis feature, extracted on MLO view images.

The results of multivariate analysis showed that better performances could be achieved when using two or more features as predictors for the classification of malignant and benign lesions and for the prediction of HR positive status. The best performance was achieved when predicting the tumor nature from the CC images through a logistic regression model, where the test accuracy reached a value of 95.83% without LASSO regulation. Nevertheless, the LASSO regularization selected 12 out of 27 predictors, significantly reducing the model complexity, at the price of an imperceptible reduction in performance (91.67%).

The same performance of logistic regression was not observed when predicting the same outcome (malignant versus benign classification) using MLO images, confirming the tendency of results in the univariate analysis context.

Features extracted from MLO images were shown to be useful in the prediction of grading with a CART algorithm. However, the results obtained with a logistic regression approach using only two predictors (minimum test accuracy of 90.00%, maximum test accuracy of 95.83%) suggest that simpler models are preferred, with the only exception of the HER2 outcome, where the best performance (test accuracy of 89.29%) was obtained with the RF algorithm.

Remarkable results were also obtained in the prediction of both the grading and the HR+, from both CC and MLO views; however, the performance of logistic regression (regularized or not) and of tree-based algorithms is surpassed by the accuracies obtained when using only two predictive features. Therefore, it can be stated that the prediction of all the outcomes is preferable with less complex models (that is, logistic regression with only two predictors or with a regularized approach). It is furthermore useful to note that the results of univariate analysis are less performant of those of the multivariate approach, suggesting that artificial intelligence can be powerfully used to extract insights from CEM images analysis.

The main limitation of this study is the need for manual segmentation of the images, which is time consuming and operator dependent. The problem of biased results due to this weakness was addressed by having two radiologists perform the segmentation. A foreseeable solution may be the use of automatic or semi-automatic segmentation; however, this may be difficult to implement, especially in the cases of multicentric lesions or background parenchymal enhancement. A further limit is that the interpretation of machine learning algorithm results is not always intuitive and may require specific expertise from the clinician.

5. Conclusions

The results of this study confirm that radiomics textural features extracted from CEM images can be highly informative about both the tumor nature and grading, and some molecular subtypes of tumors. Therefore, the results suggest that the combination of artificial intelligence algorithms with the concept of radiomics analysis can be successfully used to create a tool for supporting the physician's decision-making process in the classification of breast cancer. In particular, the identification of malignant breast lesions and HR positive status can be performed with a high predictive power, even using simpler models.

Author Contributions: Conceptualization, L.B.; Data curation, A.P. (Antonella Petrillo), R.F.; F.A. and D.L.F.; Formal analysis, R.F., E.D.B. and T.P.; Investigation, A.P. (Antonella Petrillo), M.L.B., A.P. (Annamaria Porto), V.G., M.D.B., A.F., R.M., N.P. and D.L.F.; Methodology, R.F., E.D.B., T.P., M.L.B., V.G., M.D.B., A.F., R.M., N.P., F.A. and D.L.F.; Project administration, A.P.(Annamaria Porto); Writing—original draft, A.P. (Antonella Petrillo); R.F.; E.D.B.; Writing—review and editing, A.P. (Antonella Petrillo), A.F., R.F., E.D.B.; R.M. and D.L.F. All authors have read and agreed to the published version of the manuscript.

References

1. Patel, B.K.; Lobbes, M.; Lewin, J. Contrast Enhanced Spectral Mammography: A Review. *Semin. Ultrasound CT MRI* **2018**, *39*, 70–79. [CrossRef] [PubMed]
2. Heywang-Köbrunner, S.; Viehweg, P.; Heinig, A.; Küchler, C. Contrast-enhanced MRI of the breast: Accuracy, value, controversies, solutions. *Eur. J. Radiol.* **1997**, *24*, 94–108. [CrossRef]
3. Satake, H.; Ishigaki, S.; Ito, R.; Naganawa, S. Radiomics in breast MRI: Current progress toward clinical application in the era of artificial intelligence. *Radiol. Med.* **2022**, *127*, 39–56. [CrossRef]
4. Dromain, C.; Balleyguier, C.; Muller, S.; Mathieu, M.C.; Rochard, F.; Opolon, P.; Sigal, R. Evaluation of tumor angiogenesis of breast carcinoma using contrast-enhanced digital mammography. *AJR Am. J. Roentgenol.* **2006**, *187*, 528–537. [CrossRef] [PubMed]
5. Dromain, C.; Balleyguier, C.; Adler, G.; Garbay, J.R.; Delalogeet, S. Contrast-enhanced digital mammography. *Eur. J. Radiol.* **2009**, *69*, 34–42. [CrossRef]
6. Li, L.; Roth, R.; Germaine, P.; Ren, S.; Lee, M.; Hunter, K.; Tinney, E.; Liao, L. Contrast-enhanced spectral mammography (CESM) versus breast magnetic resonance imaging (MRI): A retrospective comparison in 66 breast lesions. *Diagn. Interv. Imaging* **2017**, *98*, 113–123. [CrossRef] [PubMed]
7. Fallenberg, E.M.; Dromain, C.; Diekmann, F.; Engelken, F.; Krohn, M.; Singh, J.M.; Ingold-Heppner, B.; Winzer, K.J.; Bick, U.; Renz, D.M. Contrast-enhanced spectral mammography versus MRI: Initial results in the detection of breast cancer and assessment of tumour size. *Eur. Radiol.* **2014**, *24*, 256–264. [CrossRef]
8. Lewin, J.M.; Isaacs, P.K.; Vance, V.; Larke, F.J. Dual-energy contrast-enhanced digital subtraction mammography: Feasibility. *Radiology* **2003**, *229*, 261–268. [CrossRef]
9. Jochelson, M.S.; Dershaw, D.D.; Sung, J.S.; Heerdt, A.S.; Thornton, C.; Moskowitz, C.S.; Ferrara, J.; Morris, E.A. Bilateral contrast-enhanced dual-energy digital mammography: Feasibility and comparison with conventional digital mammography and MR imaging in women with known breast carcinoma. *Radiology* **2013**, *266*, 743–751. [CrossRef]
10. Tagliafico, A.S.; Bignotti, B.; Rossi, F.; Signori, A.; Sormani, M.P.; Valdora, F.; Calabrese, M.; Houssami, N. Diagnostic performance of contrast-enhanced spectral mammography: Systematic review and meta-analysis. *Breast* **2016**, *28*, 13–19. [CrossRef]
11. Liney, G.P.; Sreenivas, M.; Gibbs, P.; Garcia-Alvarez, R.; Turnbull, L.W. Breast lesion analysis of shape technique: Semiautomated vs. manual morphological description. *J. Magn. Reason. Imaging* **2006**, *23*, 493–498. [CrossRef] [PubMed]
12. Fusco, R.; Sansone, M.; Filice, S.; Carone, G.; Amato, D.M.; Sansone, C.; Petrillo, A. Pattern Recognition Approaches for Breast Cancer DCE-MRI Classification: A Systematic Review. *J. Med. Biol. Eng.* **2016**, *36*, 449–459. [CrossRef] [PubMed]
13. Nie, K.; Chen, J.H.; Yu, H.J.; Chu, Y.; Nalcioglu, O.; Su, M.Y. Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI. *Acad. Radiol.* **2008**, *15*, 1513–1525. [CrossRef] [PubMed]
14. Hu, H.T.; Shan, Q.Y.; Chen, S.L.; Li, B.; Feng, S.T.; Xu, E.J.; Li, X.; Long, J.Y.; Xie, X.Y.; Lu, M.D.; et al. CT-based radiomics for preoperative prediction of early recurrent hepatocellular carcinoma: Technical reproducibility of acquisition and scanners. *Radiol. Med.* **2020**, *125*, 697–705. [CrossRef]
15. Rossi, F.; Bignotti, B.; Bianchi, L.; Picasso, R.; Martinoli, C.; Tagliafico, A.S. Radiomics of peripheral nerves MRI in mild carpal and cubital tunnel syndrome. *Radiol. Med.* **2020**, *125*, 197–203. [CrossRef]
16. Santone, A.; Brunese, M.C.; Donnarumma, F.; Guerriero, P.; Mercaldo, F.; Reginelli, A.; Miele, V.; Giovagnoni, A.; Brunese, L. Radiomic features for prostate cancer grade detection through formal verification. *Radiol. Med.* **2021**, *126*, 688–697. [CrossRef]
17. Fusco, R.; Sansone, M.; Filice, S.; Granata, V.; Catalano, O.; Amato, D.M.; Di Bonito, M.; D'Aiuto, M.; Capasso, I.; Rinaldo, M.; et al. Integration of DCE-MRI and DW-MRI Quantitative Parameters for Breast Lesion Classification. *BioMed. Res. Int.* **2015**, *2015*, 237863. [CrossRef]
18. Zhang, Y.; Zhu, Y.; Zhang, K.; Liu, Y.; Cui, J.; Tao, J.; Wang, Y.; Wang, S. Invasive ductal breast cancer: Preoperative predict Ki-67 index based on radiomics of ADC maps. *Radiol. Med.* **2020**, *125*, 109–116. [CrossRef]

19. Chianca, V.; Albano, D.; Messina, C.; Vincenzo, G.; Rizzo, S.; Del Grande, F.; Sconfienza, L.M. An update in musculoskeletal tumors: From quantitative imaging to radiomics. *Radiol. Med.* **2021**, *126*, 1095–1105. [CrossRef]

20. Kirienko, M.; Ninatti, G.; Cozzi, L.; Voulaz, E.; Gennaro, N.; Barajon, I.; Ricci, F.; Carlo-Stella, C.; Zucali, P.; Sollini, M.; et al. Computed tomography (CT)-derived radiomic features differentiate prevascular mediastinum masses as thymic neoplasms versus lymphomas. *Radiol. Med.* **2020**, *125*, 951–960. [CrossRef]

21. Karmazanovsky, G.; Gruzdev, I.; Tikhonova, V.; Kondratyev, E.; Revishvili, A. Computed tomography-based radiomics approach in pancreatic tumors characterization. *Radiol. Med.* **2021**, *126*, 1388–1395. [CrossRef] [PubMed]

22. Cellina, M.; Pirovano, M.; Ciocca, M.; Gibelli, D.; Floridi, C.; Oliva, G. Radiomic analysis of the optic nerve at the first episode of acute optic neuritis: An indicator of optic nerve pathology and a predictor of visual recovery? *Radiol. Med.* **2021**, *126*, 698–706. [CrossRef] [PubMed]

23. Benedetti, G.; Mori, M.; Panzeri, M.M.; Barbera, M.; Palumbo, D.; Sini, C.; Muffatti, F.; Andreasi, V.; Steidler, S.; Doglioni, C.; et al. CT-derived radiomic features to discriminate histologic characteristics of pancreatic neuroendocrine tumors. *Radiol. Med.* **2021**, *126*, 745–760. [CrossRef]

24. Nazari, M.; Shiri, I.; Hajianfar, G.; Oveisi, N.; Abdollahi, H.; Deevband, M.R.; Oveisi, M.; Zaidi, H. Noninvasive Fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning. *Radiol. Med.* **2020**, *125*, 754–762. [CrossRef] [PubMed]

25. Crivelli, P.; Ledda, R.E.; Parascandolo, N.; Fara, A.; Soro, D.; Conti, M. A New Challenge for Radiologists: Radiomics in Breast Cancer. *Biomed. Res. Int.* **2018**, *2018*, 6120703. [CrossRef]

26. Maglogiannis, I.; Zafiropoulos, E.; Anagnostopoulos, I. An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. *Appl. Intell.* **2007**, *30*, 24–36. [CrossRef]

27. Zheng, Y.; Englander, S.; Baloch, S.; Zacharaki, E.I.; Fan, Y.; Schnall, M.D.; Shen, D. STEP: Spatiotemporal enhancement pattern for MR-based breast tumor diagnosis. *Med. Phys.* **2009**, *36*, 3192–3204. [CrossRef]

28. Turashvili, G.; Brogi, E. Tumor Heterogeneity in Breast Cancer. *Front. Med.* **2017**, *8*, 227. [CrossRef]

29. Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; van Stiphout, R.G.; Granton, P.; Zegers, C.M.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **2012**, *48*, 441–446. [CrossRef]

30. Sinha, S.; Lucas-Quesada, F.A.; DeBruhl, N.D.; Sayre, J.; Farria, D.; Gorczyca, D.P.; Bassett, L.W. Multifeature analysis of Gd-enhanced MR images of breast lesions. *J. Magn. Reson. Imaging* **1997**, *7*, 1016–1026. [CrossRef]

31. Vomweg, T.W.; Buscema, P.M.; Kauczor, H.U.; Teifke, A.; Intraligi, M.; Terzi, S.; Heussel, C.P.; Achenbach, T.; Rieker, O.; Mayer, D.; et al. Improved artificial neural networks in prediction of malignancy of lesions in contrast-enhanced MR-mammography. *Med. Phys.* **2003**, *30*, 2350–2359. [CrossRef] [PubMed]

32. Sathya, D.J.; Geetha, K. Mass classification in breast DCE-MR images using an artificial neural network trained via a bee colony optimization algorithm. *Science* **2013**, *39*, 294. [CrossRef]

33. Sathya, J.; Geetha, K. Experimental Investigation of Classification Algorithms for Predicting Lesion Type on Breast DCE-MR Images. *Int. J. Comput. Appl.* **2013**, *82*, 1–8. [CrossRef]

34. Fusco, R.; Sansone, M.; Petrillo, A.; Sansone, C. A Multiple Classifier System for Classification of Breast Lesions Using Dynamic and Morphological Features in DCE-MRI. *Comput. Vis.* **2012**, *7626*, 684–692. [CrossRef]

35. Degenhard, A.; Tanner, C.; Hayes, C.; Hawkes, D.J.O.; Leach, M. The UK MRI Breast Screening Study Comparison between radiological and artificial neural network diagnosis in clinical screening. *Physiol. Meas.* **2002**, *23*, 727–739. [CrossRef] [PubMed]

36. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man. Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]

37. Fusco, R.; Sansone, M.; Sansone, C.; Petrillo, A. Segmentation and classification of breast lesions using dynamic and textural features in dynamic contrast enhanced-magnetic resonance imaging. In Proceedings of the 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS), Rome, Italy, 20–22 June 2012; pp. 1–4.

38. Abdolmaleki, P.; Buadu, L.D.; Naderimansh, H. Feature extraction and classification of breast cancer on dynamic magnetic resonance imaging using artificial neural network. *Cancer Lett.* **2001**, *171*, 183–191. [CrossRef]

39. Agner, S.C.; Soman, S.; Libfeld, E.; McDonald, M.; Thomas, K.; Englander, S.; Rosen, M.A.; Chin, D.; Nosher, J.; Madabhushi, A. Textural Kinetics: A Novel Dynamic Contrast-Enhanced (DCE)-MRI Feature for Breast Lesion Classification. *J. Digit. Imaging* **2010**, *24*, 446–463. [CrossRef]

40. Levman, J.; Leung, T.; Causer, P.; Plewes, D.; Martel, A.L. Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines. *IEEE Trans. Med. Imaging* **2008**, *27*, 688–696. [CrossRef]

41. Fusco, R.; Piccirillo, A.; Sansone, M.; Granata, V.; Rubulotta, M.R.; Petrosino, T.; Barretta, M.L.; Vallone, P.; Di Giacomo, R.; Esposito, E.; et al. Radiomics and Artificial Intelligence Analysis with Textural Metrics Extracted by Contrast-Enhanced Mammography in the Breast Lesions Classification. *Diagnostics* **2021**, *30*, 815. [CrossRef]

42. Fusco, R.; Piccirillo, A.; Sansone, M.; Granata, V.; Vallone, P.; Barretta, M.L.; Petrosino, T.; Siani, C.; Di Giacomo, R.; Petrillo, A.; et al. Radiomic and Artificial Intelligence Analysis with Textural Metrics, Morphological and Dynamic Perfusion Features Extracted by Dynamic Contrast-Enhanced Magnetic Resonance Imaging in the Classification of Breast Lesions. *Appl. Sci.* **2021**, *11*, 1880. [CrossRef]

43. Fanizzi, A.; Losurdo, L.; Basile, T.M.A.; Bellotti, R.; Bottigli, U.; Delogu, P.; Diacono, D.; Didonna, V.; Fausto, A.; Lombardi, A.; et al. Fully Automated Support System for Diagnosis of Breast Cancer in Contrast-Enhanced Spectral Mammography Images. *J. Clin. Med.* **2019**, *8*, 891. [CrossRef] [PubMed]

44. Massafra, R.; Bove, S.; Lorusso, V.; Biafora, A.; Comes, M.C.; Didonna, V.; Diotaiuti, S.; Fanizzi, A.; Nardone, A.; Nolasco, A.; et al. Radiomic Feature Reduction Approach to Predict Breast Cancer by Contrast-Enhanced Spectral Mammography Images. *Diagnostics* **2021**, *11*, 684. [CrossRef] [PubMed]

45. La Forgia, D.; Fanizzi, A.; Campobasso, F.; Bellotti, R.; Didonna, V.; Lorusso, V.; Moschetta, M.; Massafra, R.; Tamborra, P.; Tangaro, S.; et al. Radiomic Analysis in Contrast-Enhanced Spectral Mammography for Predicting Breast Cancer Histological Outcome. *Diagnostics* **2020**, *10*, 708. [CrossRef] [PubMed]

46. Marino, M.A.; Leithner, D.; Sung, J.; Avendano, D.; Morris, E.A.; Pinker, K.; Jochelson, M.S. Radiomics for Tumor Characterization in Breast Cancer Patients: A Feasibility Study Comparing Contrast-Enhanced Mammography and Magnetic Resonance Imaging. *Diagnostics* **2020**, *10*, 492. [CrossRef]

47. Radiomic Features. Available online: https://pyradiomics.readthedocs.io/en/latest/features.html (accessed on 20 January 2017).

48. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* **2020**, *295*, 328–338. [CrossRef]

49. R-Tools Technology Inc. Available online: https://www.r-tt.com/ (accessed on 15 October 2020).

50. He, H.; Bai, Y.; Garcia, E.A.; Li, S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–6 June 2008; pp. 1322–1328.

51. Gu, X.; Angelov, P.P.; Soares, E.A. A self-adaptive synthetic over-sampling technique for imbalanced classification. *Int. J. Intell. Syst.* **2020**, *35*, 923–943. [CrossRef]

52. Tibshirani, R. The lasso Method for Variable Selection in the Cox Model. *Statist. Med.* **1997**, *16*, 385–395. [CrossRef]

53. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Statist. Methodol.* **1996**, *58*, 267–288. [CrossRef]

54. Bernardi, D.; Belli, P.; Benelli, E.; Brancato, B.; Bucchi, L.; Calabrese, M.; Carbonaro, L.A.; Caumo, F.; Cavallo-Marincola, B.; Clauser, P.; et al. Digital breast tomosynthesis (DBT): Recommendations from the Italian College of Breast Radiologists (ICBR) by the Italian Society of Medical Radiology (SIRM) and the Italian Group for Mammography Screening (GISMa). *Radiol. Med.* **2017**, *122*, 723–730. [CrossRef]

55. Bucchi, L.; Belli, P.; Benelli, E.; Bernardi, D.; Brancato, B.; Calabrese, M.; Carbonaro, L.A.; Caumo, F.; Cavallo-Marincola, B.; Clauser, P.; et al. Recommendations for breast imaging follow-up of women with a previous history of breast cancer: Position paper from the Italian Group for Mammography Screening (GISMa) and the Italian College of Breast Radiologists (ICBR) by SIRM. *Radiol. Med.* **2016**, *121*, 891–896. [CrossRef] [PubMed]

56. Losurdo, L.; Fanizzi, A.; Basile, T.M.A.; Bellotti, R.; Bottigli, U.; Dentamaro, R.; Didonna, V.; Lorusso, V.; Massafra, R.; Tamborra, P.; et al. Radiomics Analysis on Contrast-Enhanced Spectral Mammography Images for Breast Cancer Diagnosis: A Pilot Study. *Entropy* **2019**, *21*, 1110. [CrossRef]

57. Ahmed, S.A.; Samy, M.; Ali, A.M.; Hassan, R.A. Architectural distortion outcome: Digital breast tomosynthesis-detected versus digital mammography-detected. *Radiol Med.* **2022**, *127*, 30–38. [CrossRef] [PubMed]

58. D'Angelo, A.; Orlandi, A.; Bufi, E.; Mercogliano, S.; Belli, P.; Manfredi, R. Automated breast volume scanner (ABVS) compared to handheld ultrasound (HHUS) and contrast-enhanced magnetic resonance imaging (CE-MRI) in the early assessment of breast cancer during neoadjuvant chemotherapy: An emerging role to monitoring tumor response? *Radiol Med.* **2021**, *126*, 517–526. [CrossRef] [PubMed]