Physica Medica 80 (2020) 101–110

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

Multi-parametric MRI lesion heterogeneity biomarkers for breast cancer diagnosis

Marialena I. Tsarouchi [a], Georgios F. Vlachopoulos [a], Anna N. Karahaliou [a], Katerina G. Vassiou [b], Lena I. Costaridou [a,*]

[a] Department of Medical Physics, School of Medicine, University of Patras, Patras 26500, Greece
[b] Radiology and Anatomy Department, Medical School, University of Thessaly, Larissa, Greece

ARTICLE INFO

Keywords:
Breast cancer diagnosis
Intra-lesion heterogeneity
Multi-parametric Magnetic Resonance Imaging
Radiomics analysis

ABSTRACT

Purpose: To identify intra-lesion imaging heterogeneity biomarkers in multi-parametric Magnetic Resonance Imaging (mpMRI) for breast lesion diagnosis.
Methods: Dynamic Contrast Enhanced (DCE) and Diffusion Weighted Imaging (DWI) of 73 female patients, with 85 histologically verified breast lesions were acquired. Non-rigid multi-resolution registration was utilized to spatially align sequences. Four (4) DCE (2nd post-contrast frame, Initial-Enhancement, Post-Initial-Enhancement and Signal-Enhancement-Ratio) and one (1) DWI (Apparent-Diffusion-Coefficient) representations were analyzed, considering a representative lesion slice. 11 1st-order-statistics and 16 texture features (Gray-Level-Co-occurrence-Matrix (GLCM) and Gray-Level-Run-Length-Matrix (GLRLM) based) were derived from lesion segments, provided by Fuzzy C-Means segmentation, across the 5 representations, resulting in 135 features. Least-Absolute-Shrinkage and Selection-Operator (LASSO) regression was utilized to select optimal feature subsets, subsequently fed into 3 classification schemes: Logistic-Regression (LR), Random-Forest (RF), Support-Vector-Machine-Sequential-Minimal-Optimization (SVM-SMO), assessed with Receiver-Operating-Characteristic (ROC) analysis.
Results: LASSO regression resulted in 7, 6 and 7 features subsets from DCE, DWI and mpMRI, respectively. Best classification performance was obtained by the RF multi-parametric scheme (Area-Under-ROC-Curve, (AUC) ± Standard-Error (SE), AUC ± SE = 0.984 ± 0.025), as compared to DCE (AUC ± SE = 0.961 ± 0.030) and DWI (AUC ± SE = 0.938 ± 0.032) and statistically significantly higher as compared to DWI. The selected mpMRI feature subset highlights the significance of entropy (1st-order-statistics and 2nd-order-statistics (GLCM)) and percentile features extracted from 2nd post-contrast frame, PIE, SER maps and ADC map.
Conclusion: Capturing breast intra-lesion heterogeneity, across mpMRI lesion segments with 1st-order-statistics and texture features (GLCM and GLRLM based), offers a valuable diagnostic tool for breast cancer.

1. Introduction

Breast tumor phenotype varies with respect to genetic and epigenetic properties of individual cancer cells (inter-lesion heterogeneity), as well as with respect to spatial lesion properties within lesion (intra-lesion heterogeneity) [1]. Biopsy-proven histopathologic biomarkers supporting clinical diagnosis, prognosis and treatment, encountering difficulties in localizing lesion area extent, are prone to heterogeneity under-sampling and may result in high rates of breast cancer misdiagnosis, suggesting the need for non-invasive whole lesion imaging biomarkers [2].

Advances in medical imaging and image analysis provide multilayered insights into capturing non-invasively, spatially, and temporally lesion heterogeneity. However, qualitative analysis of medical images in the clinical routine induces subjective bias (intra- and inter-observer variability) and labor-intensive interpretations [3]. Recent studies in the rapidly evolving field of image-based Computer Aided Detection/Diagnosis (CADe/CADx) have developed models for improving decision support, utilizing high throughput extraction of quantitative imaging features (radiomics) and advanced machine learning algorithms in order to provide objective and reproducible tools of capturing important intra-lesion imaging biomarkers [4,5].

* Corresponding author at: Department of Medical Physics, School of Medicine, University of Patras, Patras 26500, Greece.
E-mail address: costarid@upatras.gr (L.I. Costaridou).

https://doi.org/10.1016/j.ejmp.2020.10.007
Received 5 June 2020; Received in revised form 7 October 2020; Accepted 10 October 2020
Available online 1 November 2020

While digital mammography, breast tomosynthesis and sonography remain the clinical standard for breast screening, Magnetic Resonance Imaging (MRI) is the modality of choice for screening patients of high risk and preoperative imaging [6]. Anatomical and functional information derived from advanced sequences, such as Dynamic Contrast Enhanced (DCE) and Diffusion Weighted Imaging (DWI), makes MRI a dominant imaging tool in manifold clinical applications towards breast tumor diagnosis (detection and characterization), prognosis and treatment management, facilitating a deeper understanding of pathophysiological lesion properties [7].

DCE is a well-established MRI technique that allows assessment of tissue angiogenesis and vascular permeability through evaluation of signal intensity time curves of paramagnetic contrast agent uptake qualitatively (Breast Imaging Reporting and Data System, BI-RADS Lexicon [8]), semi-quantitatively (lesion enhancement parametric maps [9]) or fully quantitatively by pharmacokinetic (PK) modeling [10]. Image quantitative analysis in DCE has considered 1st, 2nd and higher order statistics as well as temporal texture, in addition to lesion morphology [11]. DCE holds great potential in classifying breast lesions with high sensitivity up to 99% [7], however, with low specificity resulting in increased false positive detection rates[2]. In addition, DCE-based imaging biomarkers have shown to be highly correlated with breast cancer molecular subtypes and other prognostic factors [12,13] and have also been introduced for monitoring response to treatment,risk of recurrence and overall survival scores [14–16].

DWI is a contrast agent free technique that depicts random movement of water molecules, reflecting tissue microstructural characteristics and cellularity [17]. Quantitative analysis of mono-exponential Apparent Diffusion Coefficient (ADC) has demonstrated high discriminating ability (Area Under Receiver Operating Characteristic Curve, AUC, ranging from 0.889 to 0.940) between benign and malignant breast lesions, based mainly on 1st order statistics intra-lesion analysis, resulting in an increment of the overall breast MRI specificity [18,19]. Beyond diagnosis, the prognostic value of ADC has been recently evaluated with promising results [20], while it may serve as an early predictive biomarker for treatment response [21].

Furthermore, quantitative multi-parametric MRI approaches have been introduced to investigate associations of imaging biomarkers with histological types [22], molecular subtypes [23] and predicting chemotherapy response of breast cancer patients [24]. The above state-of-the art studies support the usefulness of radiomics and especially texture analysis in capturing breast lesion heterogeneity regarding lesion characterization, prognosis, and response to therapy.

Classification of benign and malignant breast lesions by means of multi-parametric MRI approaches, combining lesion imaging appearance features from DCE and DWI sequences, [25,26], has demonstrated high diagnostic performance (AUC ranging from 0.90 to 0.93). The majority of informative features of the above schemes resulted from kinetic enhancement, lesion morphology, and 2nd order texture features (GLCMs) features from DCE, while the contribution of DWI was reflected by 1st order statistics features. Parekh *et al.* [27] proposed the quantification of local lesion heterogeneity represented by Radiomics Feature Maps depicting 1st order statistics and texture features (Gray-Level-Co-occurrence-Matrices, GLCMs, and Gray-Level-Run-Length-Matrices, GLRLMs, based), achieving high diagnostic performance (AUC = 0.91), with the1st order statistics-entropy feature demonstrating the highest discriminative power both on DCE and ADC map.

The aim of this study is to quantify breast intra-lesion heterogeneity, across multi-parametric MRI representation schemes, for classifying benign from malignant breast lesions. Four (4) representations are derived from the DCE sequence, acknowledging its diagnostic significance [28,29], while one (1) representation (ADC map) considers the added value of DWI sequence. To deal with variability introduced by lesion segmentation in different parametric representations, a registration step dealing with intra- and inter-sequence image alignment is introduced, an approach also considered by [27,30]. Eleven (11) 1st order statistics and sixteen (16) texture features (GLCMs and GLRLMs based) features are derived from lesion segments spatially aligned by image registration across the 5 multi-parametric representations, while their informative power is investigated with a machine learning approach.

2. Materials and methods

2.1. Patient cohort

This retrospective study, was approved by local Ethics Committee and informed consent was obtained from patients with suspicious breast findings in mammography and/or ultrasonography, according to BI-RADS lexicon [8], who underwent routine clinical breast MRI in a single 3.0 Tesla scanner. Study inclusion criteria considered: (i) mass like lesions, (ii) diagnosis histopathologically confirmed, (iii) no breast surgery or chemotherapy/radiotherapy prior to MRI examination and no breast implants. From an original database of 130 female patients (n = 130), twenty five (n = 25) were diagnosed with non-mass like breast lesions, eleven (n = 11) had no histopathological verification and eighteen (n = 18) patients were subjected to chemotherapy/radiotherapy procedures prior to MRI examination. Finally, three patients (n = 3) were excluded due to breast implants.

2.2. MRI protocol

Imaging was performed in a 3.0 T MR scanner (Signa HDx; GE Healthcare, Milwaukee, WI, USA) using a dedicated bilateral four-element two-channel, phased array breast coil, with the patients in prone position. Severe motion artifacts were reduced due to breast holders. Bilateral breast imaging in axial plane was obtained utilizing the following protocol: (1) Axial T_2-weighted fast spin echo (FSE) imaging sequence (T2–FSE, TR/TE, 3600/100 ms; slice thickness, 4 mm; spacing, 0 mm); (2) Axial short time inversion recovery sequence (STIR) (TR/TE, 3875/90 ms; slice thickness, 4 mm; spacing, 0 mm); (3) Axial conventional diffusion-weighted echo-planar imaging (DWI) sequence (DW-EPI, TR/TE, 6000/63.7 ms; slice thickness, 4 mm; spacing, 0 mm; matrix, 96x108; FOV, 360 mm × 360 mm). Sensitizing diffusion gradients were applied in three orthogonal directions (x, y, z) with b values of 0 and 900 s/mm^2, respectively. A three-dimensional (3D) fat-suppressed (FS) T_1-weighted VIBRANT dynamic sequence (flip angle, 10°; TR/TE, 5.8/2.1 ms; slice thickness, 1.2 mm; spacing, 1.2 mm; matrix, 512x512; FOV, 350 mm × 350 mm) was acquired once before and five times after intravenous injection of contrast medium (0.1 mmol/kg of gadopentate dimeglumine agent followed by a 20-mL flush of saline solution). Temporal resolution for each dynamic acquisition was 84 s.

2.3. Proposed radiomics scheme

The proposed radiomics analysis scheme (Fig. 1) exploits multi-parametric representations of intra-lesion heterogeneity from DCE and DWI sequences. DCE representations considered the 2nd post-contrast DCE frame [31] and three enhancement kinetic maps (Initial Enhancement (IE), Post Initial Enhancement (PIE), Signal Enhancement Ratio (SER)[29]) reflecting tissue angiogenesis and vascular permeability. DWI representation considered the ADC map, reflecting tissue microstructure and cellularity. Lesion region segmentation was performed on the 2nd post-contrast DCE frame (i.e. on a single slice representative of lesion largest diameter, selected by an experienced radiologist) and propagated to the remaining four (4) maps. Lesion contour propagation was enabled by intra-sequence and inter-sequence registration step. Intra-lesion heterogeneity, as depicted on a single representative slice of each multi-parametric representation scheme, is quantified by means of 1st order statistics and texture features (GLCMs and GLRLMs based) analysis. The discriminating ability of multi-parametric heterogeneity biomarkers is investigated with three machine learning models (Logistic

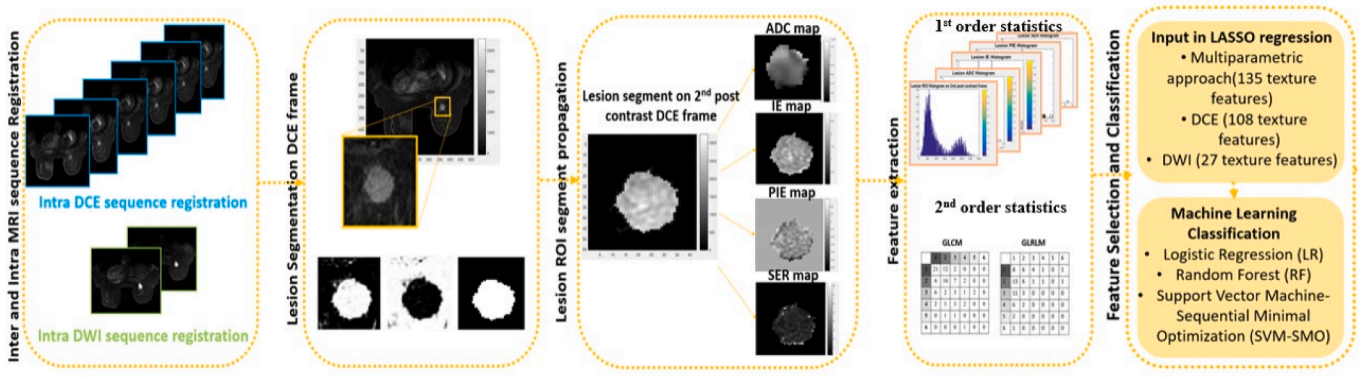

Fig. 1. Proposed radiomics analysis scheme.

Regression (LR) [32], Random Forest (RF) [33], Support Vector Machine-Sequential Minimal Optimization (SVM-SMO) [34]), following feature selection method Least-Absolute-Shrinkage and Selection-Operator (LASSO regression [35]).

2.3.1. Intra- and inter-sequence image registration

In order to capture intra-lesion pixel-wise heterogeneity properties across multi-parametric representations, all DCE and DWI images were aligned to a single post-contrast slice (2nd post-contrast DCE frame considered the reference image) exploiting multi-resolution non-rigid registration schemes [30,36]. The adopted registration schemes combine three levels of affine and a final resolution level of a 3rd order b-spline transform, utilizing the Mutual Information as a cost function [37], while different pyramids and optimizers have been selected for the two intra and the inter-sequence registration schemes.

2.3.1.1. Intra DCE sequence registration. The selected 2nd post-contrast DCE frame (slice representative of lesion diameter) served as the reference image, to which the pre-contrast and the remaining four (4) post-contrast frame slices (moving images) were registered to reduce motion and geometrical distortions [38]. In case of DCE intra-sequence registration the applied registration scheme employs a shrinking pyramid [39] and the regular step gradient descent optimizer [37].

2.3.1.2. Intra DWI sequence registration. The non-diffusion weighted image (named S_0, with corresponding value $b_0 = 0$ s/mm^2), proximal to the reference slice according to z axis slice location, was selected to be initially aligned to the reference image, as subjected to less gradient and motion correlated artifacts, compared to the high b value image (named S_{900}, with corresponding value $b_{900} = 900$ s/mm^2) [40]. The high b value image/slice (S_{900}) is then aligned to the registered S_{0_reg}, to remove intra DWI motion artifacts. Similar to DCE intra-sequence registration, DWI intra-sequence registration employs a shrinking pyramid [39] and the regular step gradient descentoptimizer [37].

2.3.1.3. Inter-sequence (DCE-DWI) registration. The adopted registration between DCE and DWI targets to propagate lesion segment, defined on 2nd post-contrast DCE frame, across the remaining four (4) parametric maps (IE, PIE, SER and ADC), allowing accurate texture sampling [36]. In case of inter-sequence registration, the adopted scheme utilized a Gaussian pyramid [39] and the adaptive stochastic gradient decent optimizer [37]. The Elastix 4.5 [39] software for intensity-based medical image registration based on the open source software Insight Toolkit (ITK) version 4.0 [41] was used for implementation of the registration scheme.

2.3.2. Lesion region segmentation

Breast lesion contouring was performed on the 2nd post-contrast DCE frame, characterized by improved lesion contour/border identification as compared to remaining post-contrast time frames (*e.g.* 3rd, 4th) and improved spatial resolution as compared to corresponding DWI slices [31,36]. Initially, an experienced radiologist defined a loose rectangular region of interest (ROI), containing the whole lesion, including necrotic or hemorrhagic/cystic portions of the lesion (2D analysis). A semi-automated segmentation algorithm based on Fuzzy C-Means clustering (FCM) [42,43] was adopted for lesion contour/border delineation. The FCM algorithm was applied on each defined ROI, building the lesion likelihood membership map (cluster number, 2; weighting exponent, 2; stop criteria, 0.0005; max iteration, 100), subsequently, followed by binarization (of the lesion membership map with empirically chosen threshold value), connected component labeling and morphological operations (i.e. object selection and hole filling) to derive the final lesion contour (MatLab R2017b Math Works, Natick, MA, www.mathworks.com). Evaluation of the adopted segmentation method accuracy was assessed with the Dice similarity coefficient (Sørensen index) [44,45]. The segmented lesion contour (defined on the 2nd post-contrast DCE frame) was propagated to the remaining four (4) four parametric maps (IE, PIE, SER and ADC) a process enabled by the previously described intra- and inter-sequence registration step. In order to evaluate the reliability of the adopted semi-automated segmentation process in the subsequent analysis, all radiomics features extracted from both FCM and manual ROIs (masks/segments) have been tested for robustness/reproducibility. The two-way mixed effect model, single measurement for absolute agreement Intra-class Correlation Coefficient (ICC) index [46], with a pre-defined threshold of ICC > 0.90, was used to identify highly stable features that are insensitive to segmentation variability (IBM SPSS Statistics software 24.0, 2016) [47].

2.3.3. DCE model-free parametric maps (IE, PIE and SER map)

The interpretation of contrast uptake signal intensity curves (related to tissue perfusion, vessel permeability and tumor vascularity) may be assessed qualitatively (BI-RADS lexicon), semi quantitatively (model-free approach) and quantitatively (pharmacokinetic modeling) [9,10]. In the model-free approach, signal intensity curves demonstrate two main phases of enhancement, early or initial enhancement and delayed or post initial enhancement [29]. IE demonstrates the rise of signal tissue intensity within the first 3 min after contrast injection, while PIE demonstrates the behavior of signal intensity curve after peak enhancement. The SER parameter is related with the post-initial appearance of signal intensity curve, incorporating both the signal change in the initial and the post initial phase relative to the pre-contrast signal measurement.

Adopting a model-free approach, the following three independent enhancement kinetic parameters were calculated in a pixel-wise basis maps, according to [9,29]:

$$IE = \frac{\max_{t=[1:3]} S(t) - S(0)}{S(0)} \quad (1)$$

$$PIE = \frac{S(5) - \max_{t=[1:3]} S(t)}{\max_{t=[1:3]} S(t)} \quad (2)$$

$$SER = \frac{\max_{t=[1:3]} S(t) - S(0)}{S(5) - S(0)} \quad (3)$$

where *S(t)* is the signal intensity of each pixel at registered DCE slices, at distinct time points: pre-contrast (t = 0) and five post-contrast time points (1st, 2nd, 3rd, 4th and 5th), respectively.

2.3.4. DWI parametric map (ADC map)

Apparent Diffusion Coefficient (ADC) has emerged as a valuable metric for quantifying changes in water mobility in vivo [17] which is directly correlated with tissue cellularity and structure. Restricted water mobility within tissues, due to increased cellularity, corresponds to high signal intensity on DWI but low ADC values.

In this study a mono-exponential model was fitted to the registered S_{0_reg} and S_{900_reg} lesion slices, allowing pixel-wise calculation ADC map generation:

$$ADC = -\frac{1}{b_{900}} \ln \frac{S_{900\_reg}}{S_{0\_reg}} \quad (4)$$

where b_{900} = 900 s/mm^2 and S_{0_reg}, S_{900_reg} the signal intensity of the registered S_0, S_{900} images, respectively. DCE and ADC parametric mapping was achieved with in-house code in MatLab R2017b (Math-Works, Natick, MA, www.mathworks.com).

Fig. 2 depicts intra-lesion heterogeneity of a benign (1st column) and a malignant (2nd column) case at 2nd post-contrast DCE frame at the four (4) parametric maps (IE, PIE, SER and ADC), respectively.

2.3.5. Feature extraction

Each lesion ROI segment, as depicted on a single representative slice of each parametric MRI representation (2nd post-contrast DCE frame, IE map, PIE map, SER map and ADC map), was subjected to 1st order statistics and texture features (GLCMs and GLRLMs based) feature extraction. A total of 27 quantitative imaging features was extracted, including eleven (11) 1st order statistics and sixteen (16) texture features (GLCMs and GLRLMs based) features. Specifically, eleven (11) were calculated from Gray Level Co-occurrence Matrices-GLCM [48] and five (5) from Gray Level Run Length Matrices-GLRLM [49]).

1st order statistics were analyzed on pixel raw values, in order to preserve the biophysical meaning of parametric MRI representations intensity values, while for the 1st order statistics entropy (Shannon's entropy) an intensity histogram with fixed bin number (8 bits/pixel) was considered as this feature can only be derived from a discretized distribution. 1st order statistics features were analyzed with in-house code (MatLab R2017b Math Works, Natick, MA, www.mathworks.com).

In case of texture feature extraction (GLCMs (2nd order statistics), and GLRLMs based features), each lesion ROI segment across the 5 parametric MRI representations was processed, considering the within ROI intensity range selection [μ ± 3σ], where μ is the mean of pixel intensity values within lesion ROI and σ is the standard deviation, (outlier correction) and discretized to 6 bits/pixel in order to minimize contrast and brightness variation and reduce computational cost of texture features [50]. The 11 features calculated from GLCM, considereda distance of one pixel (d = 1), while four directions (0°, 45°, 90° and 135°) were considered both for GLCM and GLRLM features. Features extracted from GLCM and GLRLM were averaged over the four directions to obtain rotationally invariant features. Texture features were extracted employing the publicly available MaZda package v4.6 [51].

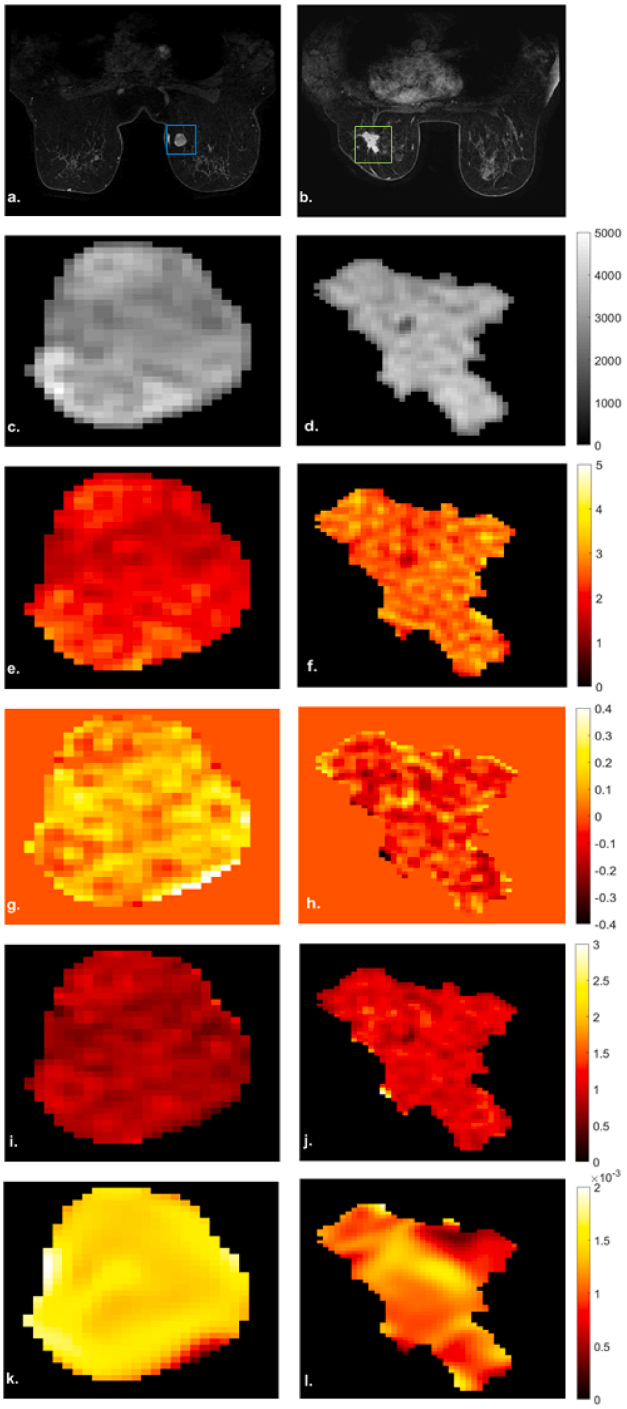

Fig. 2. Multi-parametric representation of one benign and one malignant breast lesion: a.2nd post-contrast breast DCE frame for the benign breast lesion, b. 2nd post-contrast breast DCE frame for the malignant breast lesion, **c.** benign breast lesion ROI segment on 2nd post-contrast DCE frame, d. malignant breast lesion ROI segment on 2nd post-contrast DCE frame, e. benign breast lesion ROI segment on Initial Enhancement (IE) map, f. malignant breast lesion ROI segment on Initial Enhancement (IE) map, g. benign breast lesion ROI segment on Post Initial Enhancement (PIE) map, h. malignant breast lesion ROI segment on Post Initial Enhancement (PIE) map, i. benign breast lesion ROI segment on Signal Enhancement Ratio (SER) map, j. malignant breast lesion ROI segment on Signal Enhancement Ratio (SER) map, k. benign breast lesion ROI segment on Apparent Diffusion Coefficient (ADC) map (mm^2/sec) and l. malignant breast lesion ROI segment on Apparent Diffusion Coefficient (ADC) map (mm^2/sec).

The mathematical definitions of the 27 quantitative imaging features extracted for each parametric MRI representation are provided in the Appendix (Supplementary material).

Regarding DCE parametric representations (2nd post-contrast DCE frame, IE, PIE and SER map), a total of 108 texture features was extracted (27 quantitative imaging features × 4 parametric DCE representations), while in case of DWI parametric representation, 27 features were extracted from the ADC map, for each breast lesion of the patient cohort. 135 features (27 features × 5 parametric MRI representations) were extracted from the multi-parametric scheme.

2.3.6. Statistical analysis

The Shapiro-Wilk test was utilized ($p < 0.05$) to test normality of feature values distribution in the dataset analyzed (IBM SPSS Statistics software 24.0, 2016) [47]. As the extracted features follow a non-normal distribution, the non-parametric Mann-Whitney *U* test was used (MatLabR2017b, Math Works, Natick, MA, www.mathworks.com), to assess existence of statistical significant difference of individual features between benign and malignant lesion status. Bonferroni correction for multiple tests was considered to adjust the level of significance [52].

2.4. Univariate classification schemes

Three efficient machine learning classifiers including Logistic Regression (LR), Random Forest (RF) and Support Vector Machine-Sequential Minimal Optimization (SVM- SMO), were exploited to further assess the univariate diagnostic performance of each feature. The above machine learning models were implemented in Weka [53], using the 5-fold cross validation splitting randomly into training/testing scheme. The performance of each classifier was evaluated by means of ROC analysis.

Specifically, the core of Logistic Regression (LR) classifier is the analysis of the relationship between multiple independent variables and a categorical dependent variable by fitting data to a logistic curve. LR establishes a cost function for a classification problem and then iteratively obtains the optimal model parameters through an optimization method [32]. Random Forest (RF) is based on the construction of a bunch of independent decision tree classifiers that are trained on various subsamples of the input data. RF outputs the class with the most occurrences among independent decision tree classifiers [33]. Finally, Support Vector Machine-Sequential Minimal Optimization (SVM-SMO) performs the classification by constructing a hyperplane for separating the classes of the input data and considers an optimization algorithm for solving the quadratic problem that arises during the SVM training [34].

2.5. Feature selection

In order to reduce feature dimensionality and avoid redundant information, prior to the multivariate classification the LASSO feature selection method [35] was implemented using the glmnet package in R programming language [54]. Specifically, the cohort was initially divided randomly into different subsets: The training dataset (66.6% of the dataset samples) used to develop the LASSO model, and the testing set (33.3% of the dataset samples) used for the evaluation of the obtained model. During the training process a nested 5-fold cross validation method was used to identify the optimal parameter λ of LASSO regression method. Since LASSO is a regularization method that shrinks to zero regression coefficients of features considered as attribute noise, the LASSO selection method was targeted to overcome overfitting and improve the generalization ability of the proposed machine learning models.

2.6. Multivariate classification schemes

Subsequently, the three machine learning classifiers were utilized for multivariate analysis in Weka [53], using the 5- fold cross validation training/testing scheme and the performance was evaluated with ROC analysis. The experimental study design was targeted to assess the contribution of individual DCE and DWI parametric representations as well as their combination in classifying malignant from benign breast lesions.

2.7. Z-score analysis

De Long non-parametric approach of z-score analysis was employed to investigate statistically significant classification performance differences ($p < 0.05$), comparing the AUC values of the ROC analysis [55].

Specifically, z-score was employed to compare classification performance amongthe DCE, DWI and multi-parametric selected features subsets. Additionally, z-score was employed to evaluate classification performance among the LR, RF and SVM-SMO classifiers, utilizing the same three selected feature subsets.

3. Results

In this retrospective study, a total of 73 female patients with 85 (n = 85) histologically verified breast lesions were analyzed. The study population included 33 female patients (mean age ± standard deviation, range; 46.8 ± 13.7, 22–74 years) with 41 (n = 41) benign breast lesions (mean lesion size ± standard deviation, range; 103.7 ± 95.0, 20.0–410.3 mm^2) and 40 female patients (mean age ± standard deviation, range; 58.6 ± 13.1, 25–79 years) with 44 (n = 44) malignant breast lesions (mean lesion size ± standard deviation, range; 251.4 ± 197.0, 89.2–1070.0 mm^2). Table 1 provides distribution of breast lesions with respect to histopathological subtype and corresponding lesion size (area in mm^2).

Accuracy evaluation of the adopted segmentation method (Section 2.3.2) was achieved by comparing the FCM segmented lesion regions with manual segments (ground truth), provided by an expert radiologist, resulting in high values (0.90 for benign lesions, 0.92 for malignant lesions) of the Dice similarity coefficient (Sørensen index) [44,45]. The segmented lesion contour (defined on the 2nd post-contrast DCE frame) was propagated to the remaining four (4) parametric maps (IE, PIE, SER and ADC), a process enabled by the previously described intra-and inter-sequence registration step.

Results of the ICC evaluation metric of reliability [46] indicate that the percentage of features that scored an ICC > 0.90 was 42.2% (57 out of the 135), while the median ICC for all the radiomics features (n = 135) extracted in all the 5 multi-parametric MRI representations, is 0.84.

Regarding the 27 features extracted from the 2nd post contrast DCE frame the median ICC is 0.80, while for the 3 parametric maps of DCE the median ICC for the 27 features extracted from each one is: median ICC for IE map is 0.87, median ICC for PIE map is 0.75 and median ICC for SER map is 0.83.

The median ICC for the 27 features extracted from the ADC map is 0.91.

Table 1
Distribution of breast lesions with respect to histopathological subtype and corresponding lesion size.

Histological Subtype	n (%)	Lesion size (mm^2) (Mean ± Standard Deviation)
Fibro adenomas	39 (45.8)	102.8 ± 83.5
Fibrocystic changes	1 (1.2)	401.4 ± 102.0
Sclerosing Adenosis	1 (1.2)	289.3 ± 105.3
Total Benign Breast Lesions	**41 (48.2)**	**103.7 ± 95.0**
Invasive Ductal Carcinoma	36 (42.3)	262.2 ± 211.0
Invasive Lobular Carcinoma	8 (9.5)	244.9 ± 133.6
Total Malignant Breast Lesions	**44 (51.8)**	**251.4 ± 197.0**
Total of Breast Lesions	**85 (100)**	**185.5 ± 177.5**

3.1. DCE intra-lesion heterogeneity biomarkers

Table 2 provides median and interquartile values of benign and malignant breast lesions for the 7 selected features (LASSO regression with optimal parameter λ of 0.092) extracted from DCErepresentations. Three entropy features, DCE 2nd order (GLCM)-entropy, SER 1st order -entropy and SER 2nd order (GLCM)-entropy, reflecting image complexity and randomness, demonstrate high classification performance for the three classifiers. In addition, descriptive parameters of signal intensity distributions, such as percentiles from both PIE (PIE_25thperc) and SER map (SER_50thperc), seem to provide additive information to lesion heterogeneity.

However, features extracted from IE map seem to have no contribution to the selected subset. The selected feature subset (2nd post-contrast DCE frame, PIE and SER map) achieved the highest classification performance with the RFclassifier (AUC ± SE = 0.961 ± 0.030).

Considering the reliability of the proposed 7 selected features (LASSO feature selection scheme) for the DCE representations (2nd post contrast frame, IE, PIE and SER map) the median ICC value is 0.91, indicating a repeatable scheme.

3.2. DWI intra-lesion heterogeneitybiomarkers

Table 3 provides median and interquartile values of benign and malignant breast lesions for the 6 selected features (LASSO regression with optimal parameter λ of 0.068) extracted from DWI representation. ADC_25thperc, ADC_50th perc and mean ADC value describing signal intensity, as well as 1st order and 2nd order (GLCM) entropy features seem to capture intra-lesion heterogeneity.

The highest classification performance was achieved with RF classifier (AUC ± SE = 0.938 ± 0.032).

Considering theproposed 6 selected features (LASSO feature selection scheme) from the DWI representation (ADC map) the median ICC value is 0.95, indicating a repeatable scheme.

Table 2
Median (interquartile) values and corresponding classification performance, in terms of AUC ± SE of features selected from DCE representations.

Selected Individual Features	Benign Breast Lesions (n = 41)	Malignant Breast Lesions (n = 44)	LR	RF	SVM-SMO
	Median (interquartile values)		AUC ± SE	AUC ± SE	AUC ± SE
2nd post contr. DCE Skewness (1st order)	0.0104 (−0.243, 0.414)	−0.2817 (−0.488, −0.081)	0.701 ± 0.051	0.575 ± 0.063	0.589 ± 0.068
2nd post contr. DCE Entropy (GLCM)	2.468 (2.316, 2.605)	2.673 (2.573, 2.751)	0.814 ± 0.045	0.707 ± 0.057	0.699 ± 0.060
PIE 25thPerc (1st order)	0.0261 (−0.0247, 0.0524)	−0.1235 (0.162, 0.0873)	0.915 ± 0.036	0.869 ± 0.045	0.771 ± 0.059
SER 50th Perc (1st order)	0.898 (0.829, 0.961)	1.104 (1.049, 1.228)	0.921 ± 0.035	0.873 ± 0.042	0.849 ± 0.042
SER Entropy (1st order)	5.991 (5.615, 6.397)	6.561 (6.410, 6.684)	0.813 ± 0.044	0.700 ± 0.060	0.781 ± 0.050
SER Correlation (GLCM)	0.4302 (0.336, 0.491)	0.5510 (0.444, 0.597)	0.756 ± 0.049	0.758 ± 0.052	0.704 ± 0.059
SER Entropy (GLCM)	2.360 (2.244, 2.565)	2.700 (2.559, 2.792)	0.817 ± 0.045	0.710 ± 0.057	0.714 ± 0.058
Selected Feature Subset			**0.899 ± 0.030**	**0.961 ± 0.030**	**0.899 ± 0.030**

Abbreviations: AUC: Area Under Receiver Operating Characteristic Curve; SE: Standard Error; LR: Logistic Regression; RF: Random Forest; SVM-SMO: Support Vector Machine-Sequential Minimal Optimization; 1st order statistics; GLCM: Gray Level Co-occurrenceMatrices; DCE: Dynamic Contrast Enhanced; PIE: Post Initial Enhancement and SER: Signal Enhancement Ratio.

Table 3
Median (interquartile) values and corresponding classification performance, in terms of AUC ± SE, of features selected from DWI representation.

Selected Individual Features	Benign Breast Lesions (n = 41)	Malignant Breast Lesions (n = 44)	LR	RF	SVM-SMO
	Median (interquartile values)		AUC ± SE	AUC ± SE	AUC ± SE
ADC 25th perc (1st order)	1.50 (1.30, 1.70) (×10⁻³ mm²/s)	0.83 (0.69, 0.72) (×10⁻³ mm²/s)	0.866 ± 0.040	0.884 ± 0.037	0.869 ± 0.039
ADC 50th perc (1st order)	1.6 (1.50, 1.80) (×10⁻³ mm²/s)	0.94 (0.80, 1.10) (×10⁻³ mm²/s)	0.861 ± 0.040	0.874 ± 0.041	0.869 ± 0.039
ADC mean (1storder)	1.7 (1.50, 1.80) (×10⁻³ mm²/s)	0.95 (0.83, 1.20) (×10⁻³ mm²/s)	0.847 ± 0.042	0.881 ± 0.041	0.735 ± 0.055
ADC Entropy (1st order)	5.967 (5.521, 6.429)	6.412 (6.214, 6.835)	0.802 ± 0.046	0.824 ± 0.048	0.574 ± 0.070
ADC Entropy (GLCM)	2.178 (2.010, 2.303)	2.378 (2.262, 2.480)	0.798 ± 0.046	0.723 ± 0.055	0.724 ± 0.056
RLNonUni (GLRLM)	142.51 (116.90, 221.45)	369.96 (250.98, 482.17)	0.838 ± 0.045	0.753 ± 0.053	0.700 ± 0.060
Selected Feature Subset			**0.907 ± 0.037**	**0.938 ± 0.032**	**0.881 ± 0.037**

Abbreviations: AUC: Area Under Receiver Operating Characteristic Curve; SE: Standard Error; LR: Logistic Regression; RF: Random Forest; SVM-SMO: Support Vector Machine-Sequential Minimal Optimization; 1st order statistics; GLCM: Gray Level Co-occurrence Matrices; GLRLM: Gray Level Run Length Matrices; RLNonUni: Run Length Non Uniformity; DWI: Diffusion Weighted Imaging and ADC: Apparent Diffusion Coefficient.

3.3. Multi-parametric MRI intra-lesion heterogeneity biomarkers

Table 4 provides the median and interquartile values of benign and malignant breast lesions for the 7 selected features (LASSO regression withoptimal parameter λ of 0.122) out of the 135 originally extracted features (27 features × 5 parametric MRI representations).

In the multi-parametric approach, the feature selection method highlighted the role of DCE parametric representations providing 6 out of 7 selected features, while DWI provides one feature. The role of percentiles and entropy features is in accordance to individual DCE and DWI parametric analysis. Combining DCE and DWI features resulted in increased classification performance for all three machine learning classifiers with the RF achieving the highest (AUC ± SE = 0.984 ± 0.025).

In addition, the 7 selected features (LASSO analysis) that compose the final proposed multi-parametric MRI feature subset, achieved ICC > 0.90 (median value of ICC 0.92), indicating a repeatable scheme.

3.4. Comparison of the classification performance of DCE, DWI and Multi-parametric MRI representations

Z-score was employed to investigate statistically significant classification performance differences ($p < 0.05$): (a) among the selected feature subsets for the two individual (DCE and DWI) and the multi-parametric scheme (mpMRI scheme) and (b) among the LR, RF and

Table 4
Median (interquartile) values and corresponding classification performance, in terms of AUC ± SE, of features selected from multi-parametric MRI representation scheme.

Selected Individual Features	Benign Breast Lesions (n = 41)	Malignant Breast Lesions (n = 44)	LR	RF	SVM-SMO
	Median (interquartile values)		AUC ± SE	AUC ± SE	AUC ± SE
2nd post contr. DCE Entropy (1storder)	6.236 (6.045, 6.467)	6.587 (6.435, 6.782)	0.835 ± 0.044	0.782 ± 0.049	0.795 ± 0.048
2nd post contr. DCE Entropy (GLCM)	2.468 (2.316, 2.605)	2.673 (2.573, 2.751)	0.814 ± 0.045	0.707 ± 0.057	0.699 ± 0.060
PIE 25thPerc (1st order)	0.0261 (−0.0247, 0.0524)	−0.1235 (0.162, 0.0873)	0.915 ± 0.036	0.869 ± 0.045	0.771 ± 0.059
SER 50th Perc (1st order)	0.898 (0.829, 0.961)	1.104 (1.049, 1.228)	0.921 ± 0.035	0.873 ± 0.042	0.849 ± 0.042
SER Entropy (1st order)	5.991 (5.615, 6.397)	6.561 (6.410, 6.684)	0.813 ± 0.044	0.700 ± 0.060	0.781 ± 0.050
SER Entropy (GLCM)	2.360 (2.244, 2.565)	2.700 (2.559, 2.792)	0.817 ± 0.045	0.710 ± 0.057	0.714 ± 0.058
ADC 25th Perc (1st order)	1.50 (1.30, 1.70) ($\times 10^{-3}$ mm^2/s)	0.83 (0.69, 0.72) ($\times 10^{-3}$ mm^2/s)	0.866 ± 0.040	0.884 ± 0.037	0.869 ± 0.039
Selected Feature Subset			**0.943 ± 0.028**	**0.984 ± 0.025**	**0.918 ± 0.032**

All the selected features differ statistically significantly between benign and malignant breast lesion status (Mann Whitney *U* test, $p < 0.001$)

Abbreviations: AUC: Area Under Receiver Operating Characteristic Curve; SE: Standard Error; LR: Logistic Regression; RF: Random Forest; SVM-SMO: Support Vector Machine-Sequential Minimal Optimization; 1st order statistics; GLCM: Gray Level Co-occurrence Matrices; DCE: Dynamic Contrast Enhanced; PIE: Post Initial Enhancement; SER: Signal Enhancement Ratio and ADC: Apparent Diffusion Coefficient.

SVM-SMO classifiers, utilizing the same three selected feature subsets (Fig. 3).

As it is demonstrated, the classification performance of the selected feature subset from DCE representations with the RF classifier was statistically significantly higher as compared to LR and SVM-SMO ($p = 0.0099$) (Fig. 3a). In case of DWI representation (ADC map), the RF had statistically significantly higher performance only with the respect to SVM-SMO ($p = 0.022$), while no statistically significantly difference was found between RF and LR (Fig. 3b). In the multi-parametric MRI scheme, the superior performance of the RF classifier (AUC ± SE = 0.984 ± 0.025) was statistically significant with respect to SVM-SMO and LR classification model ($p = 0.05$) (Fig. 3c).

The performance of the LR classifier in the selected multi-parametric feature subset, was statistically significantly higher as compared to the individual DCE representation ($p = 0.019$) (Fig. 3d). In case of the selected feature subsets classified with the RF model, statistically significantly higher performance was only demonstrated between DWI and the multi-parametric scheme ($p = 0.038$) (Fig. 3e). Finally, the multi-parametric SVM-SMO scheme did not demonstrate statistically significant different performance as compared to individual parametric representation (DCE and DWI) (Fig. 3f).

4. Discussion

Classifying breast lesions into benign or malignant pathological status in the clinical routine is mainly focused on imaging biomarkers such as average signal intensity values of the lesions, however, this may lead in misdiagnosis when suspicious non typical tumor phenotypes occur (BIRADS 3 or 4). Thus, novel approaches aiming to quantify imaging breast intra-lesion heterogeneity, including several quantitative gray level intensity biomarkers, are investigated towards more reliable diagnostic conclusions.

In the current study, intra-lesion heterogeneity, manifested at aligned positions of 5 parametric representations (4 DCE and ADC map), has been quantified by a total of 135 1st order statistics and texture features (GLCMs and GLRLMs based) features. The selected feature subset from DCE representations comprised of 7 features, demonstrated high classification performance ranging from 0.899 ± 0.030 to 0.961 ± 0.030 for the 3 classifiers, with the RF classifier performing best. The selected feature subset from DWI representation, composed of 6 features, also demonstrated high classification performance ranging from 0.881 ± 0.037 to 0.938 ± 0.037, again with the RF classifier performing best. The proposed multi-parametric scheme, composed of 7 selected features, has demonstrated the highest performance ranging from 0.918 ± 0.032 to 0.984 ± 0.028 for the 3 classifiers, also verifying best performance for the RF classifier. According to z-score analysis, the performance of RF multi-parametric scheme was statistically significantly higher as compared to DWI, while no statistically significant difference was observed between multi-parametric and DCE representation. Finally, the multi-parametricRF classification scheme performed statistically significantly higher as compared to LR and SVM-SMO schemes.

Results confirm the dominant role of the DCE parameter in the proposed multi-parametric scheme, agreeing with previously reported studies mainly focusing on DCE breast cancer diagnosis [11,12,28,29]. Specifically, Karahaliou *et al.* [29], highlighted the role of PIE and SER maps in the quantification of intra-lesion heterogeneity, using texture features (GLCMs and GLRMs based) obtained high diagnostic performance (AUC of 0.906 and 0.922 for features extracted from PIE and SER, respectively). Recently, three-dimensional semi-quantitative [11] or fully quantitative parametric maps (PK maps) [12] have been introduced exploiting 1st order statistics and texture features (GLCMs and GLRLMs based) features and reportingperformance of AUC of 0.91 and 0.85, respectively.

Regarding DWI, quantitative imaging analysis of lesion ADC map has mainly considered 1st order statistics for breast cancer diagnosis [18,19]. Suo *et al.* [18] suggested that minimum or percentiles as well as 1st order statistics entropy of ADC lesion value, may assist in differentiating benign from malignant status, achieving high performance (AUC = 0.938). Liu *et al.* [19] point to the same direction, exploiting a sophisticated exponential diffusion model, reporting a diagnostic performance of AUC = 0.893 for 10% percentage point of histogram ADC lesion values. Texture features (GLCMs and GLRLMs based) features in an ADC stand-alone approach have only been reported in the classification of breast lesions into histological or molecular subtypes [20] and not in the differentiation of benign from malignant breast lesions.

Multi-parametric MRI studies, combining DCE and DWI sequences, have been previously reported including both qualitative and quantitative variables in their analysis. Especially, Cai *et al.* [25] and Jiang *et al.* [26] have demonstrated high diagnostic performance (AUC of 0.93 and 0.90, respectively) with a feature subset including 7 enhancement kinetic indices, morphological features, GLCM texture features from DCE as well as the ADC mean value. GLCM features derived from DCE frames obtained AUC of 0.72 [25] and 0.67 [26], while the ADC mean lesion value has obtained AUC of 0.63 [25] and 0.86 [26]. Another multi-parametric approach proposed by Parekh *et al.* [27], focusing on quantification of local lesion heterogeneity of lesion Radiomic Feature Maps of 1st order statistics and texture features (GLCMs and GLRLMs based), has also reported high diagnostic performance (AUC = 0.91), with the 1st order statistics entropy feature demonstrating the highest discriminative power on both DCE and ADC maps. In addition, Xie *e tal.* [22], investigated 1st order statistics analysis on both DCE and DWI

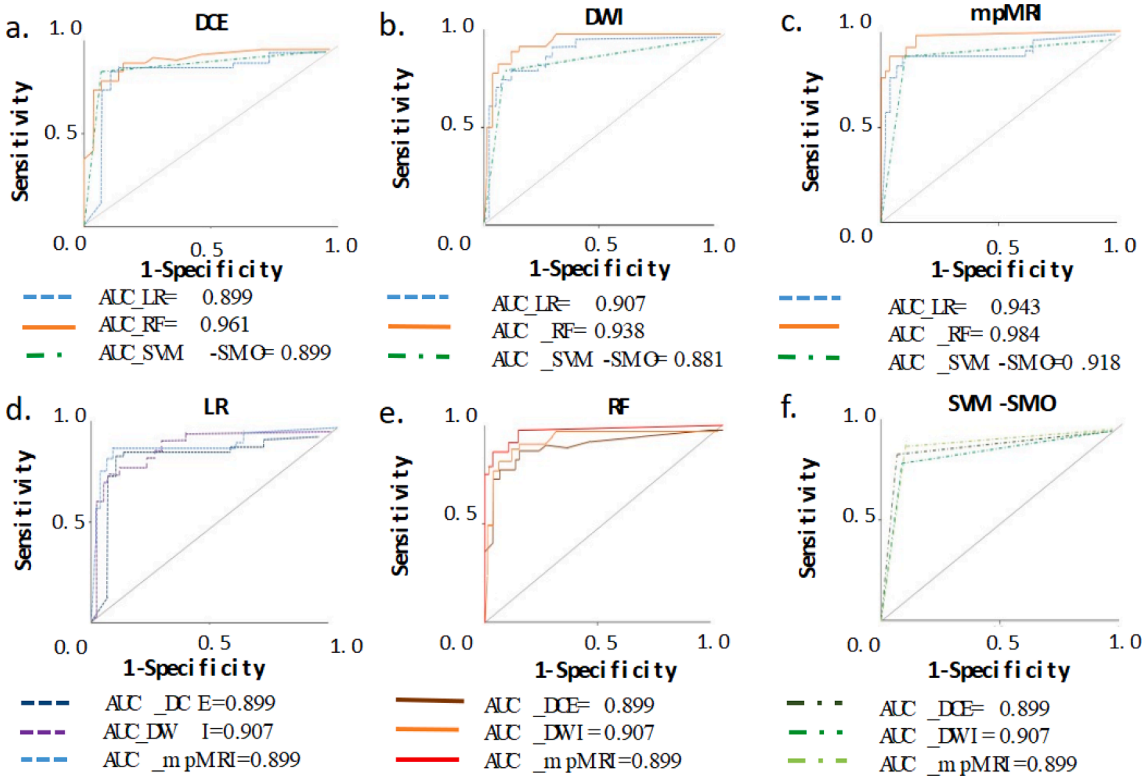

Fig. 3. ROC curves a. classification performance (LR, RF and SVM-SMO) of the selected feature subsets from DCE representations, b. classification performance (LR, RF and SVM-SMO) of the selected feature subsets from DWI and c. classification performance (LR, RF and SVM-SMO) of the selected feature subsets from multi-parametric MRI representations, d. LR classification performance of the selected subsets is investigated among MRI representations (DCE, DWI, multi-parametric), e. RF classification performance of the selected subsets is investigated among MRI representations (DCE, DWI, multi-parametric) and f. SVM-SMO classification performance of the selected subsets is investigated among MRI representations (DCE, DWI, multi-parametric).

parametric maps, for breast cancer subtype classification, reporting on the discriminating ability of histogram percentile features (AUC range: 0.683 to 0.763). Finally, Xie *et al.* [23], exploiting multiple machine learning approachesto classify breast cancer histological subtypes, reported an accuracy of 72.4% based on a feature subset of 20 selected features, including DCE and DWI texture features.

The current study agrees with [25–27] regarding the contribution of DCE in the mpMRI scheme, as in the selected feature subset, DCE representations provide 6 out of 7 selected features, while DWI provided one feature. Furthermore, four (4) entropy and three (3) percentile features have been identified as capable of capturing breast intra-lesion heterogeneity, in agreement to [18–20,25–27]. The value of percentiles and entropy features, as intra-lesion phenotypic heterogeneity biomarkers have also been identified across individual DCE and DWI representations. Finally, the contribution of SER map (model-free enhancement kinetic approach) in the multi-parametric scheme is also highlighted (3 out of 7 features of the best performing multi-parametric subset were extracted from SER map).

In addition, acknowledging the importance of pre-processing steps in the workflow of radiomics analysis [56,57], the current study has considered a non-rigid registration scheme towards establishing a mechanism for consistent lesion ROI propagation across the different MRI parametric representations. The affine registration model applied at the first three resolution levels followed by the B-spline transformation, as well as corresponding intensity interpolation at each resolution level, might introduce changes to image appearance and subsequently texture. However, results of this study demonstrate that despite possible alteration of the calculated values of image texture features, they retain their discriminating ability. To the best of our knowledge no study has so far thoroughly investigated the effect of the registration step on either radiomics or texture analysis workflow in breast MRI radiomics.

This study is limited by the relatively smallsize and composition of patient cohort, the latter comprised of mass like breast lesions (mainly benign fibroadenomas and malignant Invasive Ductal Carcinomas tumors), originating from a single institution. A dataset enriched with lesions of additional histology types will strengthen the generalization ability of the presented findings. When considering small size data sets, as the one used in our study, the use of cross validation [56] is the only option available in ensuring generalizability of classification results, which is an internal validation process and does not waive the need of additional external validation utilizing independent testing datasets. 2D lesion heterogeneity analysis remains at present a viable option in the radiomics field, offering conformance to diagnostic criteria of mass-like lesions in DCE MRI, according to BIRADS terminology [8], as well as lowering computational cost and thus facilitating clinical translatability of the proposed pipeline. 3D analysis is emerging to provide the exploitation of all available data from the entire tumor. Last but not least, exploitation of additional radiomics features, as well as comprehensive investigation of the effect of pre-processing steps such as registration, normalization and discretization, on the performance of the proposed analysis scheme should be considered in future efforts.

5. Conclusion

In conclusion, investigation of intra-lesion imaging heterogeneity biomarkers across multiple breast MRI representations highlighted the value of a multi-parametric feature subset, reflecting the significance of entropy (1^{st} order statistics and 2^{nd} order (GLCM)) and percentile features of 2^{nd}post-contrast DCE frame, PIE and SER maps, as well as the ADC map. The classification performance of the multi-parametric feature subset across three (3) classifiers ranged from 0.918 to 0.984, while the dominant role of the DCE parameter in the multi-parametric

scheme is confirmed. Capturing mpMRI breast intra-lesion phenotypic heterogeneity with1st order statistics and texture features (GLCMs and GLRLMs based), offers a high classification performance in breast cancer diagnosis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Special thanks to the Department of Radiology, University Hospital of Larissa, University of Thessaly, Greece, for contributing to this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmp.2020.10.007.

References

[1] O'Connor JPB, Rose CJ, Waterton JC, Carano RAD, Parker GJM, Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. Clin Cancer Res 2015;21(2):249–57.
[2] Chitalia RD, Kontos D. Role of texture analysis in breast MRI as a cancer biomarker: A review. J Magn Reson Imaging 2019;49(4):927–38.
[3] Saha A, Harowicz MR, Mazurowski MA. Breast cancer MRI radiomics: An overview of algorithmic features and impact of inter-reader variability in annotating tumors. Med Phys 2018;45(7):3076–85. https://doi.org/10.1002/mp.12925.
[4] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5(1). https://doi.org/10.1038/ncomms5006.
[5] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749–62.
[6] Mann RM, Balleyguier C, Baltzer PA, Bick U, Colin C, Cornford E, et al. Breast MRI: EUSOBI recommendations for women's information. Eur Radiol 2015;25(12):3669–78.
[7] Leithner D, Wengert GJ, Helbich TH, Thakur S, Ochoa-Albiztegui RE, Morris EA, et al. Clinical role of breast MRI now and going forward. Clin Radiol 2018;73(8):700–14.
[8] Morris EA, Comstock CE, Lee CH, et al. ACR BI-RADS® magnetic resonance imaging. In: ACR BI-RADS® atlas, breast imaging reporting and data system. Reston, VA: American College of Radiology; 2013. p. 125–43.
[9] Kuhl CK, Mielcareck P, Klaschik S, Leutner C, Wardelmann E, Gieseke J, Schild HH. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? Radiology 1999;211(1):101–10.
[10] Tofts PS, Brix G, Buckley DL, Evelhhoch JL, Henderson E, Knopp MV, et al. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols.JMagnReson. Imaging 1999;10(3):223–32. https://doi.org/10.1002/(sici)1522-2586(199909)10:3<223:aid-jmri2>3.0.co;2-s.
[11] Zhou J, Zhang Y, Chang K-T, Lee KE, Wang O, Li J, et al. Diagnosis of benign and malignant breast lesions on DCE-MRI by using radiomics and deep learning with consideration of peritumor tissue. J Magn Reson Imaging 2020;51(3):798–809. https://doi.org/10.1002/jmri.26981.
[12] Sun K, Zhu H, Chai W, Zhan Y, Nickel D, Grimm R, Fu C, Yan F. Whole-lesion histogram and texture analyses of breast lesions on inline quantitative DCE mapping with CAIPIRINHA-Dixon-TWIST-VIBE. Eur Radiol 2020;30(1):57–65. https://doi.org/10.1007/s00330-019-06365-8.
[13] Liu C, Ding J, Spuhler K, Gao Yi, Serrano Sosa M, Moriarty M, Hussain S, He X, Liang C, Huang C. Preoperative prediction of sentinel lymph node metastasis in breast cancer by radiomic signatures from dynamic contrast-enhanced MRI: Radiomic SLNM Prediction in Breast Cancer. J Magn Reson Imaging 2019;49(1):131–40. https://doi.org/10.1002/jmri.26224.
[14] Cain EH, Saha A, Harowicz MR, Marks JR, Marcom PK, Mazurowski MA. Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using an independent validation set. Breast Cancer Res Treat 2019;173(2):455–63. https://doi.org/10.1007/s10549-018-4990-9.
[15] Jahani N, Cohen E, Hsieh M-K, Weinstein SP, Pantalone L, Hylton N, et al. Prediction of treatment response to neoadjuvant chemotherapy for breast cancer via early changes in tumor heterogeneity captured by DCE-MRI registration. Sci Rep 2019;9(1). https://doi.org/10.1038/s41598-019-48465-x.
[16] Chitalia RD, Rowland J, McDonald ES, Pantalone L, Cohen EA, Gastounioti A, et al. Imaging phenotypes of breast cancer heterogeneity in preoperative breast dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) scans predict 10-year recurrence. Clin Cancer Res 2020;26(4):862–9. https://doi.org/10.1158/1078-0432.CCR-18-4067.
[17] Partridge SC, Nissan N, Rahbar H, Kitsch AE, Sigmund EE. Diffusion-weighted breast MRI: Clinical applications and emerging techniques: Diffusion-Weighted Breast MRI. J Magn Reson Imaging 2017;45(2):337–55. https://doi.org/10.1002/jmri.25479.
[18] Suo S, Zhang K, Cao M, Suo X, Hua J, Geng X, et al. Characterization of breast masses as benign or malignant at 3.0T MRI with whole-lesion histogram analysis of the apparent diffusion coefficient: Histogram Analysis of ADC in Breast Masses. J Magn Reson Imaging 2016;43(4):894–902. https://doi.org/10.1002/jmri.25043.
[19] Liu C, Wang K, Li X, Zhang J, Ding J, Spuhler K, Duong T, Liang C, Huang C. Breast lesion characterization using whole-lesion histogram analysis with stretched-exponential diffusion model: Stretched-Exponential Model in Breast DWI. J Magn Reson Imaging 2018;47(6):1701–10. https://doi.org/10.1002/jmri.25904.
[20] Leithner D, Bernard-Davila B, Martinez DF, Horvat JV, Jochelson MS, Marino MA, et al. Radiomic signatures derived from diffusion-weighted imaging for the assessment of breast cancer receptor status and molecular subtypes. Mol Imaging Biol 2020;22(2):453–61. https://doi.org/10.1007/s11307-019-01383-w.
[21] Kim Y, Kim SH, Lee HW, Song BJ, Kang BJ, Lee A, et al. Intravoxel incoherent motion diffusion weighted MRI for predicting response to neoadjuvant chemotherapy in breast cancer. Magn Reson Imaging 2018;48:27–33. https://doi.org/10.1016/j.mri.2017.12.
[22] Xie T, Wang Z, Zhao Q, Bai Q, Zhou X, Gu Y, et al. Machine learning-based analysis of MR multi-parametric radiomics for the subtype classification of breast cancer. Front Oncol 2019;9:505. https://doi.org/10.3389/fonc.2019.00505.
[23] Xie T, Zhao Q, Fu C, Bai Q, Zhou X, Li L, Grimm R, Liu Li, Gu Y, Peng W. Differentiation of triple-negative breast cancer from other subtypes through whole-tumor histogram analysis on multiparametric MR imaging. Eur Radiol 2019;29(5):2535–44. https://doi.org/10.1007/s00330-018-5804-5.
[24] Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, et al. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. Invest Radiol 2019;54(2):110–7. https://doi.org/10.1097/RLI.0000000000000518.
[25] Cai H, Peng Y, Ou C, Chen M, Li L. Diagnosis of breast masses from dynamic contrast-enhanced and diffusion weighted MR: a machine learning approach. PLoS One 2014;9(1). https://doi.org/10.1371/journal.pone.0087387.
[26] Jiang X, Xie F, Liu L, Peng Y, Cai H, Li L. Discrimination of malignant and benign breast masses using automatic segmentation and features extracted from dynamic contrast-enhanced and diffusion-weighted MRI. Oncol Lett 2018;16(2):1521–8. https://doi.org/10.3892/ol.2018.8805x.
[27] Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. npj Breast Cancer 2017;3(1). https://doi.org/10.1038/s41523-017-0045-3.
[28] Nie Ke, Chen J-H, Yu HJ, 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(12):1513–25.
[29] Karahaliou A, Vassiou K, Arikidis NS, Skiadopoulos S, Kanavou T, Costaridou L. Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis. BJR 2010;83(988):296–309.
[30] Vogl W-D, Pinker K, Helbich TH, Bickel H, Grabner G, Bogner W, Gruber S, Bago-Horvath Z, Dubsky P, Langs G. Automatic segmentation and classification of breast lesions through identification of informative multiparametric PET/MRI features. Eur Radiol Exp 2019;3(1). https://doi.org/10.1186/s41747-019-0096-3.
[31] Degani H, Gusis V, Weinstein D, Fields S, Strano S. Mapping pathophysiological features of breast tumors by MRI at high spatial resolution. Nat Med 1997;3(7):780–2. https://doi.org/10.1038/nm0797-780.
[32] Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 2002;35(5-6):352–9.
[33] Ho TK. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 1998;20(8):832–44. https://doi.org/10.1109/34.709601.
[34] Platt JC. Fast training of support vector machines using sequential minimal optimization, Advances in kernel methods: support vector learning. Cambridge MA: MIT Press; 1999.
[35] Tibshirani R. Regression shrinkage and selection via the Lasso. J Roy Stat Soc: Ser B (Methodol) 1996;58(1):267–88. https://www.jstor.org/stable/2346178.
[36] Méndez CA, Pizzorni Ferrarese F, Summers P, Petralia G, Menegaz G. DCE-MRI and DWI integration for breast lesions assessment and heterogeneity quantification. Int J Biomed Imaging 2012;2012:1–8. https://doi.org/10.1155/2012/676808.
[37] Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: a survey. IEEE Trans Med Imaging 2013;32(7):1153–90.
[38] Giannini V, Mazzetti S, Marmo A, Montemurro F, Regge D, Martincich L. A computer-aided diagnosis (CAD) scheme for pretreatment prediction of pathological response to neoadjuvant therapy using dynamic contrast-enhanced MRI texture features. BJR 2017;90(1077):20170269. https://doi.org/10.1259/bjr.20170269.
[39] Klein S, Staring M, Murphy K, Viergever MA, Pluim J. elastix: A toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010;29(1):196–205. https://doi.org/10.1109/TMI.2009.2035616.
[40] Arlinghaus LR, Welch EB, Chakravarthy AB, Xu L, Farley JS, Abramson VG, et al. Motion correction in diffusion-weighted MRI of the breast at 3.0T. J Magn Reson Imaging 2011;33(5):1063–70. https://doi.org/10.1002/jmri.22562.
[41] Ibanez L, Schroeder W, Ng L, Cates J. The ITK software guide. second ed. Clifton Park, NY): Kitware; 2005. ISBN: 1-930934-15-7, 2005.

[42] Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. Comput Geosci 1984;10(2-3):191–203.
[43] Chen W, Giger ML, Bick U. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. AcadRadiol 2006;13(1):63–72. https://doi.org/10.1016/j.acra.2005.08.035.
[44] Dice LR. Measures of the amount of ecologic association between species. ECOLOGY.1945;26(3):297-302. doi: 10.2307/1932409.
[45] Sørensen T. A Method of Establishing Groups of Equal Amplitude in plant Sociology Based on Similarity of Species Content and its Application to Analyses of the Vegatation on Danish Commons, serBiologiskeSkrifter// DetKongelige Danske VidanskabernesSelskab. I commission hos E. Munksgaard, 1948.
[46] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropractic Med 2016;15(2):155–63.
[47] Field AP. Discovering statistics using SPSS. Los Angeles, CA: Sage Publications; 2009.
[48] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst, Man, Cybern 1973;SMC-3(6):610–21.
[49] Galloway MM. Texture analysis using grey level run lengths. Comput Graph Image Process 1975;4:172–9. https://doi.org/10.1016/S0146-664X(75)80008-6.
[50] Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging 2004;22(1):81–91.
[51] Szczpinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda- a software package for image texture analysis. Comput Methods Programs Biomed 2009;94(1):66–76. https://doi.org/10.1016/j.cmpb.2008.08.
[52] Wright SP. Adjusted P-values for simultaneous inference. Biometrics 1992;48(4):1005. https://doi.org/10.2307/2532694.
[53] Frank E, Hall MA, and. Witten IH. The WEKA Workbench. Online Appendix for "DataMining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition.
[54] R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
[55] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837. https://doi.org/10.2307/2531595.
[56] Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015;162(1):W1. https://doi.org/10.7326/M14-0698.
[57] Shukla-Dave A, Obuchowski NA, Chenevert TL, Jambawalikar S, Schwartz LH, Malyarenko D, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. J Magn Reson Imaging 2019;49(7):e101–21.