Article

Breast Cancer Mass Detection in DCE–MRI Using Deep-Learning Features Followed by Discrimination of Infiltrative vs. In Situ Carcinoma through a Machine-Learning Approach

Luana Conte [1,2] **, Benedetta Tafuri** [1,2,*] **, Maurizio Portaluri** [3] **, Alessandro Galiano** [4] **, Eleonora Maggiulli** [5] **and Giorgio De Nunzio** [1,2]

[1] Laboratory of Biomedical Physics and Environment, Department of Mathematics and Physics "E. De Giorgi", University of Salento, 73100 Lecce, Italy; luana.conte@unisalento.it (L.C.); giorgio.denunzio@unisalento.it (G.D.N.)

[2] Advanced Data Analysis in Medicine (ADAM), Laboratory of Interdisciplinary Research Applied to Medicine (DReAM), University of Salento and ASL (Local Health Authority), 73100 Lecce, Italy

[3] Operative Unit of Radiotherapy, ASL (Local Health Authority), Brindisi, and 'Di Summa-Perrino' Hospital, 72100 Brindisi, Italy; m.portaluri@asl.brindisi.it

[4] Operative Unit of Radiodiagnostics, ASL (Local Health Authority), Brindisi, and 'Di Summa-Perrino' Hospital, 72100 Brindisi, Italy; alessgaliano@yahoo.it

[5] Operative Unit of Medical Physics, ASL (Local Health Authority), Brindisi, and 'Di Summa-Perrino' Hospital, 72100 Brindisi, Italy; eleonora.maggiulli@asl.brindisi.it

* Correspondence: benedetta.tafuri@gmail.com

check for updates

Featured Application: **Automatic CAD system in detecting tumors as a valuable support to radiologists for detection and characterization of breast cancers in dynamic contrast-enhanced magnetic resonance imaging (DCE–MRI) images.**

Abstract: Breast cancer is the leading cause of cancer deaths worldwide in women. This aggressive tumor can be categorized into two main groups—in situ and infiltrative, with the latter being the most common malignant lesions. The current use of magnetic resonance imaging (MRI) was shown to provide the highest sensitivity in the detection and discrimination between benign vs. malignant lesions, when interpreted by expert radiologists. In this article, we present the prototype of a computer-aided detection/diagnosis (CAD) system that could provide valuable assistance to radiologists for discrimination between in situ and infiltrating tumors. The system consists of two main processing levels—(1) localization of possibly tumoral regions of interest (ROIs) through an iterative procedure based on intensity values (ROI Hunter), followed by a deep-feature extraction and classification method for false-positive rejection; and (2) characterization of the selected ROIs and discrimination between in situ and invasive tumor, consisting of Radiomics feature extraction and classification through a machine-learning algorithm. The CAD system was developed and evaluated using a DCE–MRI image database, containing at least one confirmed mass per image, as diagnosed by an expert radiologist. When evaluating the accuracy of the ROI Hunter procedure with respect to the radiologist-drawn boundaries, sensitivity to mass detection was found to be 75%. The AUC of the ROC curve for discrimination between in situ and infiltrative tumors was 0.70.

1. Introduction

Breast cancer (BC) is one of the most common malignant tumors as well as the leading cause of mortality among women worldwide. In Italy, BC affected about 53,000 new cases out of a total of 175,000 cases of all female cancers in 2019 [1]. BC can be classified into two main types—in situ and invasive. Based on cytological characteristics and growth patterns, the in situ type is further subdivided into ductal and lobular, located within the ductal or lobular epithelium, respectively. Ductal carcinoma in situ (DCIS) is more common than lobular carcinoma in situ (LCIS), accounting for 30–50% of all detected BCs [2,3], and normally does not infiltrate through the basal membrane. On the other hand, the invasive ductal carcinoma (IDC) is the most common malignant lesion, accounting for approximately 70% of all malignant cases [4,5]. Treatment choice is different between in situ and IDC (e.g., for the latter, the identification of the sentinel lymph node is required, which is not for the former, see the "National Comprehensive Cancer Network Guidelines—Breast Cancer" at https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf). Moreover, clinical outcomes are worse for the invasive disease. Consequently, women might need to undergo additional surgical options if an invasive disease is missed.

In the current clinical imaging practice, the use of Magnetic Resonance Imaging (MRI) has a high sensitivity and strongly improves tumor mass detection and discrimination between benign and malignant lesions [6–10]. Naturally, breast MRI scans must be interpreted by experienced radiologists, as these examinations are often used to improve the outcome of the surgical practice, by reducing the number of re-excisions, allowing patient selection for neoadjuvant chemotherapy or therapy modification, as well as representing a technique of choice for pre-surgical assessment of residual tumor size, to determine breast conservation surgery candidacy [6].

In this scenario, a new field of research called Radiomics is becoming increasingly popular, with the general aim being a conversion of all information contained in digital medical images into quantifiable features. The latter are normally related to tumor size, shape, pixel intensity, and texture associated with clinical outcomes and prognosis, defining a proper tumor Radiomics signature [11]. The usage of Radiomics signatures can lead to a remarkable improvement of detection rate [12,13].

Starting from the above considerations, the aim of this study was to develop a software system that is able to differentiate in situ infiltration of BT in dynamic contrast-enhanced (DCE–MRI) images, based on lesion Radiomics signature. Preliminary results of this work, on a smaller dataset, with a partially different approach, and without the segmentation step, are reported in [14].

The problem of distinguishing invasive from in situ BC is debated in a few papers in the specific literature. In [15], Radiomics features are extracted from DCE–MRI scans (190 IDC and 58 DCIS) and are used to train a random forest classifier in a leave-one-out cross-validation scheme. AUC of the ROC curve was 0.90. A Radiomics signature of 569 features was tested by Li et al. [16] in mammographic images; the dataset was composed of 161 DCIS and 89 IDC and their best result was AUC = 0.72. In [17], the apparent diffusion coefficient (ADC) computed from diffusion-weighted MRI (DWI) was used to distinguish invasive from in situ DCIS. DWI characterizes tissue diffusivity, therefore it provides a description of tissue micro-structure [18,19]. The rationale was that invasive breast cancer spreads by degrading the tissue structure through proteolytic activity—the chronic inflammatory reaction to proteolysis causes the reduction of extra-cell water content, with a consequent reduction of ADC if compared to in situ tumors. In order to test the hypothesis, in [17], a dataset of 21 DCIS and 155 IDC was employed and a significant difference in ADC values between the two groups was found ($p < 0.001$ and AUC = 0.89). A Radiomics approach in DCE–MR images, using combined computer-extracted MR imaging kinetic and morphologic features, was tested by Bhooshan et al. [20] (in a dataset containing 32 benign, 71 DCIS, and 150 IDC cases), obtaining AUC = 0.83. Finally, deep learning was tried in [21], with the purpose of predicting invasive cancer after DCIS diagnosis. They used a transfer learning approach, in which a pre-trained GoogleNet was used to calculate features in 131 MRI images, then training a support vector machine (SVM). The result was AUC = 0.70.

The Radiomics calculations to classify tumors as in situ or infiltrative must be performed in a region of interest (ROI) containing the tumor tissue [22]. For this reason, a necessary pre-processing step is manual or (semi)automatic segmentation (contouring) of the lesions, separating the tumor from normal tissue in the image. Breast tumor segmentation, especially in DCE–MRI images, is still a challenging task in the clinical setting, although it is necessary in some circumstances, e.g., when tumor-response prediction to chemotherapy is assessed [23–25]. Automating this procedure would help radiologists to reduce their manual workload on image analysis, as they normally perform tumor diagnosis by locating lesions layer-by-layer, and that is an arduous and time-consuming task [26].

Different image segmentation methods in MRIs were proposed in the past decades but no optimal method exists yet. The simplest pixel-based approaches generally rely on thresholding the image intensity and grouping individual pixels through appropriate classifiers. For example, Tzacheva et al. [27] determined the boundary of the suspected tumor on the assumption that the lesion intensity range was 110–140 on the 0–255 scale, so they simply applied a threshold for obtaining a binary image. The use of thresholding for breast tumor segmentation was also used by Fusco et al. [28] through the exploitation of intensity differences between the pixels, before and after providing contrast, followed by morphological post-processing steps. Fuzzy C Means (FCM) clustering [26] and its various versions [29,30] is also one of the prevailing methods, due to its simplicity [26] in isolating suspicious lesions. Another popular method is the classic k-means used for segmenting the lesion [31,32].

Other typical techniques used for lesion segmentation are region-based methods. Adams and Bischof [33] proposed the algorithm of seeded region growing (SRG) and its advancement [34], which begins by determining the seed (or set of seeds) from which growth starts. Then, SRG grows these seeds into regions by successively adding surrounding pixels, until all pixels are assigned to one region. Other region-based methods exploit the watershed algorithm, followed by post-processing steps [35,36].

Contour-based methods are also much used in the task of breast lesion segmentation, especially in case of active-contours of the lesion boundary. A recent work [37] describes an interactive segmentation method for BC lesions in DCE–MRI images, based on the active contour without edges (ACWE) algorithm and using parallel programming with general purpose computing on graphics processing units (GPGPU). The ACWE was able to segment objects with low gradient information in their boundaries. The performance of this algorithm was evaluated on a set of 32 breast DCE–MRI cases in terms of speed-up, and compared to non-GPU based approach. A high speed-up (40 or more) was obtained for high-resolution images, providing real-time outputs.

Sun et al. [38] proposed a semi-supervised method for breast tumor segmentation. After image segmentation with advanced clustering techniques, they performed a supervised learning step, based on texture features and mean intensity levels, to classify the tumor and non-tumor patches, in order to automatically locate the tumor regions in an MRI image.

These manual or semi-automatic tumor annotation techniques are generally the most used [26,39], although these approaches are often time-consuming and can drive user variability. In addition, they often need manual delineation of ROIs as a first step, requiring expert knowledge in advance. On the contrary, breast tumor segmentation using deep learning approaches was recently used in some medical imaging applications [40–42] and showed promise in automatic lesion segmentation. El Adoui et al. [40], used two deep learning architectures, SegNet and U-Net [43,44] for the detection and segmentation of 86 breast DCE–MRI images. These two CNN architectures were successfully applied to biomedical imaging segmentation and could even be used with relatively small datasets [45]. A 2D U-Net [43] CNN architecture was also used by Dalmis et al. on 66 breast T1-MRI post-contrast images [41], with promising results. At the same time, Moeskops et al. [42] used a deep learning approach to segment the pectoral muscle in 34 T1-MRI breast images.

The next sections describe the software system developed in this work, composed of a segmentation step followed by classification. Technical details on the database employed and on the code structure are given in the Materials and Methods section, while the preliminary results are summed up and commented in the Results and Discussion section.

2. Materials and Methods

The dataset consists of 55 anonymized DCE–MRI scans of BC (11 DCIS + LCIS and 44 IDC). The MRI sequence was dynamic eTHRIVE with fat suppression, on a Philips Achieva 1.5 T MRI equipment. We considered images containing at least one tumor mass, as diagnosed by an expert radiologist and confirmed by biopsy. An ROI of the tumor mass was manually delimited for each slice by an expert radiologist in post-contrast images. The MRI volumes were resampled to isometric 1-mm pixel size, before processing.

The software system developed consisted of two main steps (see Figure 1)—tumor detection/segmentation (paragraphs 2.1 to 2.3) and tumor classification of in situ vs. invasive tumor (2.4 and 2.5). The former found the tumor and performed contouring by (a) automatic localization of candidate tumor ROIs (suspicious regions likely to contain a tumor mass) based on a dynamically changed threshold on the intensity values (ROI hunting); (b) feature extraction from candidate ROIs, through a pre-trained Deep-Learning Convolutional Neural Network (CNN); (c) false-positive ROI rejection through the training of a feed-forward multi-layer perceptron Artificial Neural Network (ANN), with the aim of preserving only the tumors (positive class) for subsequent processing. The second step concerned the discrimination between in situ and infiltrating tumors and was subdivided into—(d) Radiomics signature extraction from the detected ROIs; and (e) binary classification. The code was written partially in python 3.7 and pyradiomics (https://pyradiomics.readthedocs.io/en/latest/index.html), and partially in the Matlab environment. In the following sections, each of the above-mentioned processing steps is reviewed in detail.
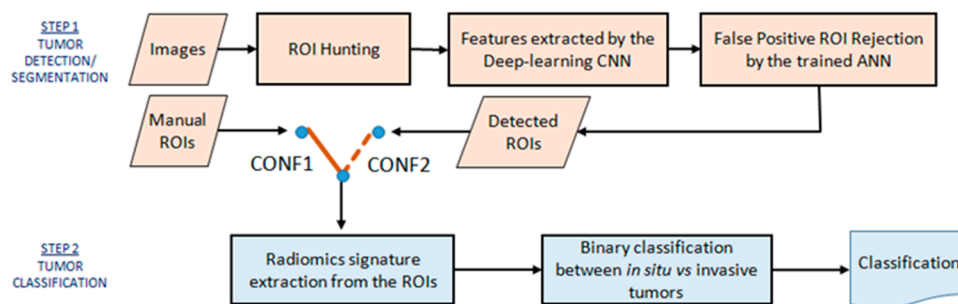
Figure 1. A flowchart of the software. CONF1 and CONF2 denote the two configurations in which the tumor classification (Step 2) was tested, i.e., respectively with manual ROIs as input, or fed up with the automatically detected ROIs (coming from Step 1).

2.1. ROI Hunter Procedure

In our particular application, accurate tumor borders were not fundamental, so we used a simple detection/segmentation method based on the application of thresholds followed by region classification.

Prior to be processed and in view to minimize false positive (FP) ROIs, the mammary area, containing the breast, was semi-automatically selected in all slices by a bounding box (working volume) and the tissue outside the box removed. Figure 2 shows an example of breast area selection.

Then the candidate tumors inside the working volume were detected. Since the tumor mass normally appears as a bright area, an iterative 2D ROI Hunter procedure, based on a dynamically changing threshold, was implemented. The number of ROIs detected from each slice was not set a priori, rather it was related to the intensity properties of the image.

First, the images were converted to pixel values in the range 0 to 1, where the 99.9th percentile of the gray values of the whole image was used for normalization, in order to exclude outliers. The following iterative procedure was then performed on a per-slice basis, giving a small number of 2D ROIs, per image section. An initial threshold (T) was set to 0.9 and only pixels with value $\geq$T were extracted, considering the found objects as tumor candidates. If no objects were detected, the threshold

was iteratively lowered by 5% from the current value, until at least one object was identified in the current slice. Tumor lesions were normally fairly round, thus, elongated and threadlike objects were excluded by thresholding on their geometrical features. In particular, for each ROI, the length of the major and minor axes of the ellipse that had the same second moments as the ROI were calculated (as the eigenvalues of the covariance matrix of the ROI point coordinates), and their ratio R was derived. In all examined cases, R was lower than about 1.5 or just slightly larger, so a conservative threshold was set at R = 2, discarding ROIs with larger R, which were always artefacts.
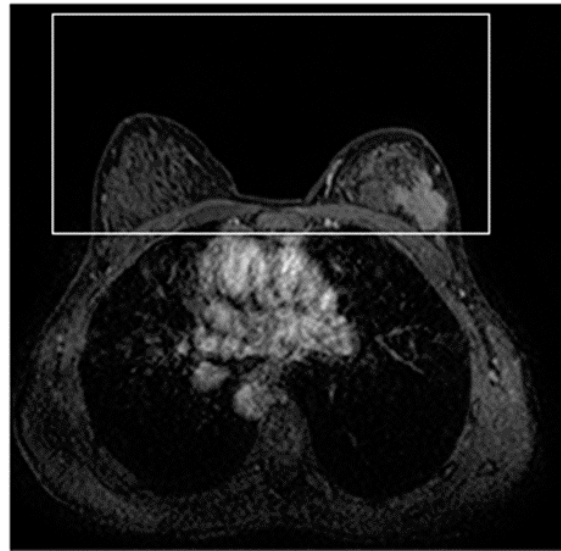
Figure 2. A single slice of the original image with the selected working area.

Each object gray-value median was calculated and the ROIs were labeled from 1 onwards in gray-value median descending order (the highest median value possibly selecting the most plausible tumor). Borders of the ROIs were also extracted.

2.2. Deep-Learning Feature Extraction

Many different approaches were experimented with the purpose of obtaining a set of features able to distinguish tumor regions from FP, the most successful being the one described hereafter. In the calculation of the features and subsequent classification (training and validation), we adopted a sliding window approach. Initially, in order to set some procedure parameters, the variability of tumor size was investigated, as the latter differed among different patients and of course in different slices. According to the statistics of our dataset, the longest edge of the lesion-bounding box was at most 120 mm. This was in accordance with previous studies, e.g., [21]. After some tests, we then chose 30×30 pixels as the size of the sliding window for ROI scanning. During operation, the bounding box containing each lesion section was enlarged if necessary (when smaller than 30×30) and the sliding window moved with a step of two pixels (on each axis) to explore the ROIs. The features were calculated for each position of the sliding window. As to the choice of the feature vector to employ, several tests were conducted, starting from the direct usage of the 900 patch pixel values, which however gave poor results. Finally, features were extracted using a GoogleNet model pre-trained on the ImageNet dataset [46], which is one the most representative networks in image classification. GoogleNet consists of 2 convolution layers, 9 inception layers, and 1 fully connected layer, which was used for feature calculation. The output size of the last fully connected layer was 1000, thus, the same number of features were extracted for each sliding window position. To fit the input image size of GoogleNet, all extracted patches were resized to 224×224 pixels using bilinear interpolation and converted to RGB images through replication of the image bitplane. The cardinality of the feature

set was then reduced through recursive feature elimination to only retain the most representative variables. After various tests with different cardinality, we reduced the feature set to only 200 variables, where the quality roughly saturated.

2.3. FP ROI Rejection through Binary Classification

In order to preserve only positive ROIs (true tumors) for further processing, so excluding FPs, the obtained features were used to train a binary classifier.

Tumor patches whose area was occupied by the lesion at least by 10% were considered to be positive, while the remaining ones, together with the supplementary patches randomly extracted from outside the lesion-bounding box, contributed to the negative samples.

To increase the size of the dataset and to favor generalization, data augmentation was obtained by random image rotations, taking care of finally obtaining a roughly balanced dataset. Several classifiers were tested (e.g., XGBoost, svm ...) and the best results were obtained with a feed-forward, backprop, multi-layer perceptron ANN, with one hidden layer composed of five neurons.

In order to clarify the terminology, hereafter, we will use the term "training set" for the "seen data" (of which a larger part was used by the classifier for actual training and a small part was used for early stop in case of overfitting), and the term "validation set" for the "unseen data" (for validating the model, i.e., for accuracy/ROC/etc. calculation and hyperparameter tuning). Training/validation was performed in a leave-one-patient-out (LOPO) cross-validation scheme [47]. Data was split by patient, ensuring that the ROIs related to each patient were totally contained in the training or in the validation set, and never in both, to avoid bias and consequent misleading figures of merit. For the same reason, at each iteration, feature value normalization to range 0–1 was performed using min–max normalization on the training set and, subsequently, validation set features were normalized by the parameters used for the training set. Fifty-four out of 55 patients were used for training the network, while the last one was used for validation, and a cyclical permutation of the patients was carried out. Statistics were calculated after a full LOPO cycle—an ROC curve was used to judge classification quality and to deduce an optimal threshold value on the ANN output, thus obtaining the binary classifier. Figure 3 shows an example of output of the whole process from ROI Hunting to classification. As our detection/segmentation code had a tendency to slightly underestimate the tumor area compared to the manually segmented ROIs, we performed a dilation operation to be sure to cover the lesion tissue.
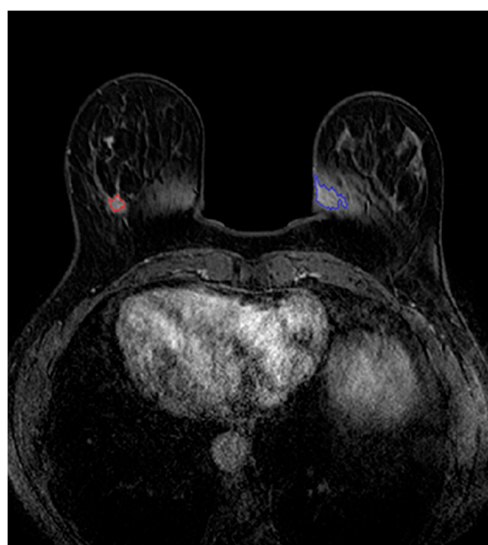
Figure 3. A typical result obtained by the CAD system. The objects found after the iterative procedure based on intensity values (ROI hunting) and the results of the classification prediction (red for tumor and blue for FP).

2.4. Tumor Characterization by Radiomics Signature

The ROI Hunter locates lesions without giving further information. The second and last part of the process concerns the characterization of the found ROIs, so that a decision-making system can correctly discriminate in situ from infiltrative lesions. This step consists of Radiomics feature extraction from the selected ROIs and classification. As the calculation was performed in 3D, before proceeding, 2D ROIs were grouped on the basis of continuity from slice to slice, so as to form 3D ROIs.

In order to discriminate the tumor volumes so obtained, we investigated a large dataset of Radiomics features. Overall, 1820 features comprising shape, first order, and higher order features were generated for each detected ROI, with original and filtered intensity. We computed 18 first-order statistic features describing the distribution of voxel intensities within the image region defined, and 68 textural features quantifying intra-tumor heterogeneity (22 from gray-level co-occurrence matrices (GLCM), 16 from gray-level run length matrices (GLRLM), 14 from gray level dependence matrices (GLDM), and 16 from gray level size zone matrices (GLSZM)) [48]. In addition to calculating the features on the original ROI volumes, we applied several preprocessing filters to each ROI, before computing the Radiomics signatures—Laplacian of Gaussian filter for edge enhancement, Wavelet filters yielding 8 subfilters (all possible combinations of applying either a High or a Low pass filter in each of the three dimensions), Square and SquareRoot filters that take the square and the square root of the image intensities, and linearly scale them back to the original range, Logarithm, Exponential, Gradient filters, the Local Binary Pattern filter (in a by-slice operation, i.e., 2D, and using spherical harmonics, in 3D). After this step, we applied recursive feature elimination to remove redundant and irrelevant features.

2.5. Classification to Discriminate In Situ vs. Invasive BC

Three different classifiers (Naive Bayes, random forests, and XGBoost) were employed and the best results were obtained with the Extreme Gradient Boosting (XGBoost) classifier (an implementation of gradient boosted decision trees) [49] in a LOPO cross-validation scheme. At each iteration, the features were normalized to [0, 1] using min–max normalization on the training subjects and subsequently applying the calculated normalization parameters to each test patient feature set. To overcome the severe class imbalance, we oversampled the minority class (in situ BC), using the Synthetic Minority Oversampling Technique (SMOTE) [50]. Performance for our imbalanced classification task was assessed using different metrics, such as balanced accuracy instead of accuracy, average precision-recall, confusion matrix, Matthews correlation coefficient, and AUC from ROC curve. All hyperparameters of the XGBoost classifier were optimized for our imbalanced dataset.

3. Results and Discussion

The sensitivity of the detection/segmentation procedure of our prototype, computed as the percentage of tumor masses correctly detected, was 75% (n = 41 out of a total of 55 samples). The Jaccard coefficient on the found masses was 0.7. As the system showed sub-optimal sensitivity, an interactive part that allowed the manual inclusion of regions missed by the automatic procedure was added for completeness. Four FPs were suggested by the ROI Hunter but excluded by the trained ANN, which thus showed an excellent specificity. With regards to the code for the discrimination between in situ and infiltrative lesions, two different configurations were explored (graphically shown in Figure 1). In the first (CONF1), the discrimination step was tested on its own—all masses visually detected and manually segmented by the radiologist (our ground truth) were used as input. In this way, discrimination quality figures (e.g., accuracy) were not influenced by the errors that the automatic detection/segmentation step introduced (in terms of false negatives, i.e., missed masses). In the second configuration (CONF2), the discrimination code was fed with only the masses found by the detection/segmentation code, giving a totally automatic standalone chain composed of detection/segmentation + discrimination. In CONF1, the evaluation of the classification performance of the trained XGBoost classifier reported a ROC curve with an AUC of 0.70. After choosing the

classifier threshold associated with the ROC curve point closest to the [0, 1] ROC space coordinates as the optimal threshold, the model correctly classified 47 subjects out of 55 (the confusion matrix was (6, 5; 3, 41); sensitivity 0.93, specificity 0.54, with an F1 score of 0.90, a balanced accuracy score of 0.74, a Matthews correlation coefficient of 0.44). In CONF2, where only the masses found by the detection/segmentation step were considered (which, as said, led to missing a non-negligible number of lesions), these values were slightly better, as most of classification errors actually came from masses not detected by the first CAD step. This suggests that the lesions missed by the detection step were also more difficult to characterize and assign to their class.

Our hypothesis was that the limit in our results might be justified, at least in part, by the small size of our monocentric dataset and its imbalanced nature. This conjecture might be supported by the observation that in our dataset there is a very small number of in situ tumors (i.e., 11, far less numerous than the infiltrative lesions) and only about half of in situ were correctly classified, giving a poor specificity value. This hypothesis might also be suggested by a check of the dataset size of the three reviewed articles working in (conventional) MRI, i.e., [15,20,21], against the corresponding AUC found by the respective authors. While our small database consisted of only 55 patients, the number of images employed in the mentioned papers were, respectively, 248, 221 (if only the malignant cases are considered), and 131. The AUC values were 0.90, 0.83, and 0.70, which evidently correlated with sample cardinality. In particular, the last cited study of this group [21] had an AUC comparable to ours, with a dataset that was more than double in size. These considerations encouraged us to go on with our tests, increasing the dataset size in the near future. A deeper test of our approach would require a larger sample size for each class, so as to guarantee generalization and result quality, avoiding overfitting. In perspective, we point to increase the size of the dataset by involving different hospitals, thus creating a multicenter study. In this way, after solving the well-known problem of image normalization between different scanners, we might possibly build a CAD system with better quality and larger applicability. Subsequently, to the few groups that are currently active on the specific subject of in situ vs. infiltrative BC discrimination, we plan to propose a project on an ensemble classification system built by merging (with various approaches) the classifiers developed by each group.

From the algorithmic point of view, it is our intention to soon explore the inclusion, in the calculation of features, of the peritumoral area, which is known to be informative in certain Radiomics applications (see e.g., [51], also working in DCE–MRI for BC).

4. Conclusions

The automatic pre-operative, non-invasive distinction between infiltrative and in situ breast cancer represents an important challenge in the biomedical field.

In this work, a two-step CAD system was developed and tested on DCE–MRI scans, with the aim of discriminating infiltrating from in situ breast tumors. The first step initially performed an ROI Hunting procedure to automatically extract 2D ROIs exploiting intensity values. This level consisted of a dynamical threshold algorithm that allowed us to select suspicious regions that were likely to contain a tumor mass. From the candidate ROIs, 1000 features were extracted through a deep learning method (starting from a pre-trained GoogleNet), followed by a classical machine-learning classifier (ANN) in the task of excluding FP regions. The second step performed the classification of in situ vs. invasive breast cancer of the previously detected ROIs (merged into 3D regions), through a Radiomics-based analysis. The results showed that the ROI Hunter segmentation procedure correctly identified 75% of tumor volumes, but the software contained an interactive part that allowed manual inclusion of the regions missed by the automatic detection/segmentation procedure. The infiltrative vs. in situ classification task achieved a final F1 score of 0.90 on all masses and a slightly better score on the masses automatically identified by the detection/segmentation step.

Our preliminary results on tumor type classification were still worse than those reported in the few specific studies existing in the literature, which could be partly explained by the small size of the dataset we used, which was moreover quite imbalanced.

Our future efforts would be towards the enrichment of the database employed, considering a multicentric development of the research. From the algorithmic point of view, we shall pursue an increase in sensitivity in the detection/segmentation step, with alternative approaches, and an increment of the accuracy of the classification step. In particular, the latter would also explore the effectiveness of including the peritumoral area in the ROIs. The aim would be the complete automatization of our CAD system in detecting the tumors and then distinguishing the two classes, so that in perspective this could be used as a valuable support to radiologists for detection and characterization of breast cancers in DCE–MRI images. As a long-term project, we plan to later propose the building of an ensemble classification system merging the classifiers developed by the few groups currently active on the specific subject of in situ vs. infiltrative BC discrimination.

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

Conflicts of Interest: The authors report no conflict of interest. The authors alone are responsible for the content and writing of this article.

References

1. Various Authors. *I Numeri del Cancro in Italia*; Intermedia Editore: Brescia, Italy, 2019.
2. Bijker, N.; Meijnen, P.; Peterse, J.L.; Bogaerts, J.; Van Hoorebeeck, I.; Julien, J.-P.; Gennaro, M.; Rouanet, P.; Avril, A.; Fentiman, I.S.; et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma-in-situ: Ten-year results of randomized phase III trial. *J. Clin. Oncol.* **2006**, *24*, 3381–3387. [CrossRef] [PubMed]
3. Ernster, V.L.; Ballard-Barbash, R.; Barlow, W.E.; Zheng, Y.; Weaver, D.L.; Cutter, G.; Yankaskas, B.C.; Rosenberg, R.; Carney, P.A.; Kerlikowske, K.; et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J. Natl. Cancer Inst.* **2002**, *94*, 1546–1554. [CrossRef] [PubMed]
4. Northridge, M.E.; Rhoads, G.G.; Wartenberg, D.; Koffman, D. The importance of histologic type on breast cancer survival. *J. Clin. Epidemiol.* **1997**, *50*, 283–290. [CrossRef]
5. Gamel, J.W.; Meyer, J.S.; Feuer, E.; Miller, B.A. The impact of stage and histology on the long-term clinical course of 163,808 patients with breast carcinoma. *Cancer* **1996**, *77*, 1459–1464. [CrossRef]
6. Mann, R.M.; Cho, N.; Moy, L. Breast MRI: State of the art. *Radiology* **2019**, *292*, 520–536. [CrossRef]
7. Kuhl, C.K.; Schild, H.H. Dynamic image interpretation of MRI of the breast. *J. Magn. Reson. Imaging* **2000**, *12*, 965–974. [CrossRef]
8. Kuhl, C.K.; Mielcareck, P.; Klaschik, S.; Leutner, C.; Wardelmann, E.; Gieseke, J.; Schild, H.H. Dynamic breast MR imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology* **1999**, *211*, 101–110. [CrossRef]
9. Schnall, M.D. Breast MR imaging. *Radiol. Clin. N. Am.* **2003**, *41*, 43–50. [CrossRef]
10. Morris, E.A. Breast cancer imaging with MRI. *Radiol. Clin. N. Am.* **2002**, *40*, 443–466. [CrossRef]
11. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [CrossRef]
12. Brown, M.S.; Goldin, J.G.; Rogers, S.; Kim, H.J.; Suh, R.D.; McNitt-Gray, M.F.; Shah, S.K.; Truong, D.; Brown, K.; Sayre, J.W.; et al. Computer-aided lung nodule detection in CT: Results of large-scale observer test. *Acad. Radiol.* **2005**, *12*, 681–686. [CrossRef] [PubMed]
13. Peldschus, K.; Herzog, P.; Wood, S.A.; Cheema, J.I.; Costello, P.; Schoepf, U.J. Computer-aided diagnosis as a second reader: Spectrum of findings in CT studies of the chest interpreted as normal. *Chest* **2005**, *128*, 1517–1523. [CrossRef] [PubMed]
14. Tafuri, B.; Conte, L.; Portaluri, M.; Galiano, A.; Maggiulli, E.; De Nunzio, G. Radiomics for the Discrimination of Infiltrative vs In Situ Breast Cancer. *Biomed. J. Sci. Tech. Res.* **2019**, *24*, 17890–17893.

15. Karen, D.; Schram, J.; Burda, S.; Li, H.; Lan, L. Radiomics Investigation in the distinction between in situ and invasive breast cancers. *Med. Phys.* **2015**, *42*, 3602–3603.

16. Li, J.; Song, Y.; Xu, S.; Wang, J.; Huang, H.; Ma, W.; Jiang, X.; Wu, Y.; Cai, H.; Li, L. Predicting underestimation of ductal carcinoma in situ: A comparison between radiomics and conventional approaches. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 709–721. [CrossRef]

17. Bickel, H.; Pinker-Domenig, K.; Bogner, W.; Spick, C.; Bagó-Horváth, Z.; Weber, M.; Helbich, T.; Baltzer, P. Quantitative apparent diffusion coefficient as a noninvasive imaging biomarker for the differentiation of invasive breast cancer and ductal carcinoma in situ. *Investgative Radiol.* **2015**, *50*, 95–100. [CrossRef]

18. Pinker, K.; Bickel, H.; Helbich, T.H.; Gruber, S.; Dubsky, P.; Pluschnig, U.; Rudas, M.; Bago-Horvath, Z.; Weber, M.; Trattnig, S.; et al. Combined contrast-enhanced magnetic resonance and diffusion-weighted imaging reading adapted to the 'Breast Imaging Reporting and Data System' for multiparametric 3-T imaging of breast lesions. *Eur. Radiol.* **2013**, *23*, 1791–1802. [CrossRef]

19. Spick, C.; Pinker-Domenig, K.; Rudas, M.; Helbich, T.H.; Baltzer, P.A. MRI-only lesions: Application of diffusion-weighted imaging obviates unnecessary MR-guided breast biopsies. *Eur. Radiol.* **2014**, *24*, 1204–1210. [CrossRef]

20. Bhooshan, N.; Giger, M.L.; Jansen, S.A.; Li, H.; Lan, L.; Newstead, G.M. Cancerous breast lesions on dynamic contrast-enhanced MR images: Computerized characterization for image-based prognostic markers. *Radiology* **2010**, *254*, 680–690. [CrossRef]

21. Zhu, Z.; Harowicz, M.; Zhang, J.; Saha, A.; Grimm, L.J.; Hwang, E.S.; Mazurowski, M.A. Deep learning analysis of breast MRIs for prediction of occult invasive disease in ductal carcinoma in situ. *Comput. Biol. Med.* **2019**, *115*, 103498. [CrossRef]

22. Halalli, B.; Makandar, A. Computer Aided Diagnosis—Medical Image Analysis Techniques. In *Breast Imaging*; InTechOpen Limited: London, UK, 2018.

23. El Adoui, M.; Drisis, S.; Larhmam, M.A.; Lemort, M.B.M. Breast Cancer Heterogeneity Analysis as Index of Response to Treatment Using MRI Images: A Review. *Imaging Med.* **2017**, *9*, 109–119.

24. El Adoui, M.; Drisis, S.; Benjelloun, M. A PRM approach for early prediction of breast cancer response to chemotherapy based on registered MR images. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *13*, 1233–1243. [CrossRef] [PubMed]

25. Benjelloun, M.; El Adoui, M.; Larhmam, M.A.; Mahmoudi, S.A. Automated Breast Tumor Segmentation in DCE-MRI Using Deep Learning. In Proceedings of the 2018 4th International Conference on Cloud Computing Technologies and Applications, Cloudtech 2018, Brussels, Belgium, 26–28 November 2018.

26. Chen, W.; Giger, M.L.; Bick, U. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Acad. Radiol.* **2006**, *13*, 63–72. [CrossRef] [PubMed]

27. Tzacheva, A.A.; Najarian, K.; Brockway, J.P. Breast cancer detection in gadolinium-enhanced MR images by static region descriptors and neural networks. *J. Magn. Reson. Imaging* **2003**, *17*, 337–342. [CrossRef] [PubMed]

28. Fusco, R.; Sansone, M.; Sansone, C.; Petrillo, A. Selection of suspicious ROIs in breast DCE-MRI. In Proceedings of the International Conference on Image Analysis and Processing, Ravenna, Italy, 14–16 September 2011; Volume 6978 LNCS, pp. 48–57.

29. Kannan, S.R.; Ramathilagam, S.; Devi, R.; Sathya, A. Robust kernel FCM in segmentation of breast medical images. *Expert Syst. Appl.* **2011**, *38*, 4382–4389. [CrossRef]

30. Kannan, S.R.; Sathya, A.; Ramathilagam, S. Effective fuzzy clustering techniques for segmentation of breast MRI. *Soft Comput.* **2011**, *15*, 483–491. [CrossRef]

31. Gnonnou, C.; Smaoui, N. Segmentation and 3D reconstruction of MRI images for breast cancer detection. In Proceedings of the International Image Processing, Applications and Systems Conference, IPAS 2014, Sfax, Tunisia, 5–7 November 2014.

32. Moftah, H.M.; Azar, A.T.; Al-Shammari, E.T.; Ghali, N.I.; Hassanien, A.E.; Shoman, M. Adaptive k-means clustering algorithm for MR breast image segmentation. *Neural Comput. Appl.* **2014**, *24*, 1917–1928. [CrossRef]

33. Adams, R.; Bischof, L. Seeded Region Growing. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 641–647. [CrossRef]

34. Al-Faris, A.Q.; Ngah, U.K.; Isa, N.A.M.; Shuaib, I.L. Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering. In *Soft Computing in Industrial Applications*; Springer: Cham, Switzerland, 2014; Volume 223, pp. 49–60.

35. Cui, Y.; Tan, Y.; Zhao, B.; Liberman, L.; Parbhu, R.; Kaplan, J.; Theodoulou, M.; Hudis, C.; Schwartz, L.H. Malignant lesion segmentation in contrast-enhanced breast MR images based on the marker-controlled watershed. *Med. Phys.* **2009**, *36*, 4359–4369. [CrossRef]

36. Saaresranta, T.; Hedner, J.; Bonsignore, M.R.; Riha, R.L.; McNicholas, W.T.; Penzel, T.; Anttalainen, U.; Kvamme, J.A.; Pretl, M.; Sliwinski, P.; et al. Clinical Phenotypes and Comorbidity in European Sleep Apnoea Patients. *PLoS ONE* **2016**, *11*, e0163439. [CrossRef]

37. Zavala-Romero, O.; Meyer-Baese, A.; Lobbes, M.B.I. Breast lesion segmentation software for DCE-MRI: An open source GPGPU based optimization. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 211–215.

38. Sun, L.; He, J.; Yin, X.; Zhang, Y.; Chen, J.H.; Kron, T.; Su, M.Y. An image segmentation framework for extracting tumors from breast magnetic resonance images. *J. Innov. Opt. Health Sci.* **2018**, *11*, 1850014. [CrossRef]

39. 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]

40. El Adoui, M.; Mahmoudi, S.A.; Larhmam, M.A.; Benjelloun, M. MRI breast tumor segmentation using different encoder and decoder CNN architectures. *Computers* **2019**, *8*, 52. [CrossRef]

41. Dalmiş, M.U.; Litjens, G.; Holland, K.; Setio, A.; Mann, R.; Karssemeijer, N.; Gubern-Mérida, A. Using deep learning to segment breast and fibroglandular tissue in MRI volumes. *Med. Phys.* **2017**, *44*, 533–546.

42. Moeskops, P.; Wolterink, J.M.; van der Velden, B.H.M.; Gilhuijs, K.G.A.; Leiner, T.; Viergever, M.A.; Išgum, I. Deep learning for multi-task medical image segmentation in multiple modalities. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; Volume 9901, pp. 478–486.

43. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Volume 9351, pp. 234–241.

44. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]

45. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651.

46. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.

47. Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions. *J. R. Stat. Soc. Ser. B* **1974**, *36*, 111–133. [CrossRef]

48. Parmar, C.; Grossmann, P.; Bussink, J.; Lambin, P.; Aerts, H.J.W.L. Machine Learning methods for Quantitative Radiomic Biomarkers. *Sci. Rep.* **2015**, *5*, 1–11. [CrossRef]

49. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

50. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]

51. Braman, N.M.; Etesami, M.; Prasanna, P.; Dubchuk, C.; Gilmore, H.; Tiwari, P.; Plecha, D.; Madabhushi, A. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res.* **2017**, *19*, 57. [CrossRef]