Semi-automated and interactive segmentation of contrast-enhancing masses on breast DCE-MRI using spatial fuzzy clustering

Carmelo Militello[1], *E-mail*: carmelo.militello@ibfm.cnr.it

Institute of Molecular Bioimaging and Physiology, Italian National Research Council (IBFM-CNR), 90015 Cefalù (PA), Italy

Leonardo Rundo, *E-mail*: lr495@cam.ac.uk

Department of Radiology, University of Cambridge, CB2 0QQ Cambridge, United Kingdom

Cancer Research UK Cambridge Centre, CB2 0RE Cambridge, United Kingdom

Mariangela Dimarco, *E-mail*: maridimarco33@gmail.com

Section of Radiology – Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND), University Hospital "Paolo Giaccone", 90127 Palermo, Italy

Alessia Orlando, *E-mail*: orlandoalessiamed@gmail.com

Section of Radiology – Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND), University Hospital "Paolo Giaccone", 90127 Palermo, Italy

Vincenzo Conti, *E-mail*: vincenzo.conti@unikore.it

Faculty of Engineering and Architecture, University of Enna KORE, 94100 Enna, Italy

Ramona Woitek, *E-mail*: rw585@cam.ac.uk

Department of Radiology, University of Cambridge, CB2 0QQ Cambridge, United Kingdom

Cancer Research UK Cambridge Centre, CB2 0RE Cambridge, United Kingdom

Department of Biomedical Imaging and Image-guided Therapy, Medical University Vienna, 1090 Vienna, Austria

[1]Corresponding author. *Email*: carmelo.militello@ibfm.cnr.it

Address: Institute of Molecular Bioimaging and Physiology, Italian National Research Council (IBFM-CNR), 90015 Cefalù (PA), Italy.

Ildebrando D'Angelo, *E-mail*: ildebrando.dangelo@hsrgiglio.it

Department of Radiology, Fondazione Istituto "G.Giglio", 90015 Cefalù (PA), Italy

Tommaso Vincenzo Bartolotta, *E-mail*: tommasovincenzo.bartolotta@unipa.it

Section of Radiology – Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND), University Hospital "Paolo Giaccone", 90127 Palermo, Italy

Department of Radiology, Fondazione Istituto "G. Giglio", Cefalù (PA) 90015, Italy

Giorgio Russo, *E-mail*: giorgio.russo@ibfm.cnr.it

Institute of Molecular Bioimaging and Physiology, Italian National Research Council (IBFM-CNR), 90015 Cefalù (PA), Italy

Semi-automated and interactive segmentation of contrast-enhancing masses on breast DCE-MRI using spatial fuzzy clustering

Carmelo Militello[a,*], Leonardo Rundo[b,c], Mariangela Dimarco[d], Alessia Orlando[d], Vincenzo Conti[e], Ramona Woitek[b,c,f], Ildebrando D'Angelo[g], Tommaso Vincenzo Bartolotta[d,g], Giorgio Russo[a]

[a] *Institute of Molecular Bioimaging and Physiology, Italian National Research Council (IBFM-CNR), 90015 Cefalù (PA), Italy*
[b] *Department of Radiology, University of Cambridge, CB2 0QQ Cambridge, UK*
[c] *Cancer Research UK Cambridge Centre, CB2 0RE Cambridge, UK*
[d] *Section of Radiology – Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND), University Hospital "Paolo Giaccone", 90127 Palermo, Italy*
[e] *Faculty of Engineering and Architecture, University of Enna KORE, 94100 Enna, Italy*
[f] *Department of Biomedical Imaging and Image-guided Therapy, Medical University Vienna, 1090 Vienna, Austria*
[g] *Department of Radiology, Fondazione Istituto "G.Giglio", 90015 Cefalù (PA), Italy*

Abstract

Multiparametric Magnetic Resonance Imaging (MRI) is the most sensitive imaging modality for breast cancer detection and is increasingly playing a key role in lesion characterization. In this context, accurate and reliable quantification of the shape and extent of breast cancer is crucial in clinical research environments. Since conventional lesion delineation procedures are still mostly manual, automated segmentation approaches can improve this time-consuming and operator-dependent task by annotating the regions of interest in a reproducible manner. In this work, a semi-automated and in-

*Corresponding author.
E-mail address: carmelo.militello@ibfm.cnr.it (C. Militello)

teractive approach based on the spatial Fuzzy C-Means (sFCM) algorithm is proposed, used to segment masses on dynamic contrast-enhanced (DCE) MRI of the breast. Our method was compared against existing approaches based on classic image processing, namely (*i*) Otsu's method for thresholding-based segmentation, and (*ii*) the traditional FCM algorithm. A further comparison was performed against state-of-the-art Convolutional Neural Networks (CNNs) for medical image segmentation, namely SegNet and U-Net, in a 5-fold cross-validation scheme. The results showed the validity of the proposed approach, by significantly outperforming the competing methods in terms of the Dice similarity coefficient (84.47 ± 4.75). Moreover, a Pearson's coefficient of $\rho = 0.993$ showed a high correlation between segmented volume and the gold standard provided by clinicians. Overall, the proposed method was confirmed to outperform the competing literature methods. The proposed computer-assisted approach could be deployed into clinical research environments by providing a reliable tool for volumetric and radiomics analyses.

Keywords: `Semi-automated segmentation`, `Breast cancer`, `Unsupervised fuzzy clustering`, `Spatial information`, `Computer-assisted lesion detection`, `Magnetic Resonance Imaging`

1. Introduction

Breast cancer is the most common cancer among women worldwide [1] and persists as the second leading cause of cancer death [2]. The importance of screening programs based on imaging, such as mammography, has been widely acknowledged for early cancer detection [3]. Indeed, early breast cancer diagnosis showed to be associated with better outcomes [4].

Along with mammography and ultrasound (US), which are the most commonly used modalities for breast cancer detection and diagnosis, Magnetic Resonance Imaging (MRI) is playing an increasing role in the detection and characterization of breast cancer [5], irrespective of breast density. The main indications for a breast MRI examination are screening to detect possible occult breast cancer in women at increased risk, pre-operative assessment of the disease extent in women with confirmed breast cancer, and assessment of treatment response to neoadjuvant chemotherapy [6]. Especially, multiparametric MRI is the most sensitive imaging modality to detect breast cancer, including ductal carcinoma *in situ* (DCIS) [5]. In recent years, the trend of implementing Computer-Assisted Diagnosis (CADx) systems for breast lesions embraces all imaging modalities, with a variety of approaches that show the interest from the research community [7]. In fact, CADx systems can potentially speed up the reading/interpretation process and reduce the amount of lesions that might be misinterpreted or overlooked by radiologists during breast MRI screening, particularly in the case of Dynamic Contrast-Enhanced (DCE-MRI) [8, 9]. Moreover, Diffusion-Weighted Imaging (DWI) has also been increasingly used on breast lesion detection and characterization, with particular interest in the Apparent Diffusion Coefficient (ADC) maps by exploiting the restricted diffusion in regions of high cellular density, generally caused by the proliferation of glandular tissue [10].

In clinical routine, the radiologists' workflow is based on descriptors derived from medical images to identify and characterize the lesions. In particular, the American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) [11] aims at standardizing the categories of imag-

ing features that commonly characterize breast lesions on imaging, such as calcifications and lesion morphology on mammography [12] or kinetic curve type on DCE-MRI. By doing so, radiologists can provide a single malignancy descriptor for each lesion [13].

Accurate and reliable quantification of the shape and extent of the breast cancer has a vital role in clinical applications [13]. In this scenario, lesion detection and segmentation can affect predictive and classification models. In the case of radiomics analyses [14, 15], the quality of the extracted features and, consequently, the performance and generalization abilities of the developed models. As a matter of fact, manual breast lesion segmentation is a labor-intensive task and, even more importantly, is highly affected by operator-dependence [16].

In this work, a semi-automated approach based on spatial Fuzzy C-Means (sFCM) clustering algorithm to segment masses on breast DCE-MRI images is proposed. The main contributions are related to:

- **medical image analysis:** the availability of an accurate segmentation procedure reduces the radiologists' time-consuming labor, as well as the operator dependence, thus increasing result repeatability;
- **DCE-MRI:** only the time-point with the strongest enhancement phase is analyzed to provide an approach less dependent on acquisition protocols;
- **precision oncology:** automated and accurate segmentation methods are beneficial to the downstream quantitative imaging analyses for developing reliable diagnostic and prognostic biomarkers [13].

This manuscript is organized as follows: Section 2 outlines the literature concerning breast lesion segmentation on DCE-MRI. Section 3 introduces the theoretical background of spatial FCM clustering. Section 4 presents the characteristics of the processed DCE-MRI dataset, as well as the proposed semi-automated segmentation method. Section 5 shows and discusses the achieved experimental results. Finally, Section 6 provides conclusive remarks and future directions.

2. Related work

Automated segmentation tools could facilitate diagnosis and treatment planning tasks, with reduced reporting times compared to the analogous manual procedure. Medical image analysis software, tailored to specific applications, enables the extraction and analysis of quantitative imaging features, with the goal of gaining new insights into the disease under consideration. This could optimize the whole clinical workflow for the development of personalized therapies [17].

Quantitative imaging methods, such as radiomics [18], require segmentation procedures that are reliable and repeatable, and thus operator-independent. In the literature, there are several approaches for the assisted segmentation of breast lesions: traditional techniques [19], machine learning [16, 20, 21], and deep learning [9, 12, 22].

In [16], the authors presented a method based on the Fuzzy C-Means (FCM) clustering algorithm for breast lesion segmentation on DCE-MRI. The approach consists of six consecutive steps: (*i*) manual Region of Interest (ROI) selection; (*ii*) lesion ROI preprocessing; (*iii*) FCM clustering; lesion

extraction; (*iv*) postprocessing; (*v*) object selection; (*vi*) hole filling. The algorithm was tested on an MRI dataset consisting of 121 primary lesions and compared against a manual gold standard. Also in [20], the authors proposed a two-stage approach exploiting the FCM clustering and Gradient Vector Flow (GVF) snake algorithms for breast lesion segmentation on MR images. Manual delineations, provided by expert MR radiologists, represented the reference to evaluate the computerized segmentation method. The approach was also compared with the traditional FCM algorithm. Overall, a dataset of 60 mass-like lesions was used in the experimental trials. The morphological and textural radiomic features, extracted from the segmented ROIs, were used to classify the benign and malignant lesions.

Among computational approaches, Deep learning approaches are gaining ground in breast imaging by exhibiting a great potential [23]. Nevertheless, several of them did not generalize well on unseen test data. On the other hand, the conventional algorithms achieved robust and stable results, even though they often require human interaction. Wang *et al.* [9] proposed a 2D/3D mixed convolution module able to exploit the contexts between adjacent slices on 90 DCE-MRI studies. A multi-scale context extractor block—consisting of convolutions with different sampling rates—was introduced to extract multi-scale image features, essential to consider the diversity of shape and size of breast lesions. Piantadosi *et al.* [22] proposed a multi-planar combination of U-Nets by fusing the projections and enabling multi-protocol applications. The experiments were performed on 109 DCE-MRI studies with histopathologically proven lesions and two different acquisition protocols.

Classic machine learning represents a well-established class of algorithms,

which provide the best compromise since (*i*) it is able to better manage the intrinsic variability of the signal, typical of biomedical images (for instance, in the case of global thresholding [19, 24]), and (*ii*) does not require a large amount of labeled data for training, such as in CNN-based architectures [25].

The current challenge for clinicians is two-fold: (*i*) early diagnosis of the disease and (*ii*) personalized medicine, targeted to the specific clinical case. In order to guarantee these objectives, the support of computer-assisted systems is valuable because, alongside the knowledge provided by clinicians, it allows us to overcome the limitations of manual segmentation approaches.

3. The spatial FCM clustering algorithm

This section focuses on the spatial versions of the FCM clustering algorithm. For a detailed theoretical discussion, see Section S1 'Unsupervised fuzzy C-Means clustering techniques' of the Supplementary Materials.

The traditional FCM clustering algorithm does not take into account any spatial relationship among pixels since all the samples are analyzed as independent points, making it sensitive to noise and other imaging artifacts [26, 27]. Accordingly, the integration of spatial information might be beneficial.

The initial version of the sFCM algorithm was formulated as a regularization term to penalize the traditional FCM objective function—which relies on pixel values alone regardless of their location—by conveying spatial information and constraining the behavior of the membership functions [28], similarly to methods used in the regularization and Markov Random Field (MRF) theory [29]. Thus, a fuzzy local similarity measure is integrated into

its objective function, aiming at decreasing the sensitivity to noise, as well as preserving image details [30]. Accordingly, the objective function of the robust Fuzzy Local Information C-Means (FLICM) algorithm aims at estimating spatially smooth membership functions [31]:

$$\mathcal{J}_{\mathrm{FLICM}}(\mathbf{U}, \mathcal{V}; \mathbf{X}) = \sum_{i=1}^{C} \sum_{k=1}^{N} (u_{ik})^m \cdot d_{ik} + \mathcal{H}_{ik}, \tag{1}$$

where the fuzzy factor $\mathcal{H}_{ik}$ acts as a regularizer:

$$\mathcal{H}_{ik} = \sum_{j \in \mathcal{N}(\mathbf{x}_k)} \frac{1}{1 + d_{jk}} (1 - u_{ij})^m \cdot \|\mathbf{x}_j - \mathbf{v}_i\|^2. \tag{2}$$

For instance, a 3D median

The sFCM, alternatively introduced by Chuang *et al.* [32], allows maintaining the same formulation and objective function as the traditional FCM algorithm (described in Section S1 'Unsupervised fuzzy C-Means clustering techniques' of Supplementary Materials) by just modifying the update rules with the local spatial content in the image. Aiming at exploiting the contextual information, a spatial function is defined as [33, 32]:

$$h_{ik} = \sum_{j \in \mathcal{N}(\mathbf{x}_k)} \hat{u}_{ij}, \tag{3}$$

where $\mathcal{N}$ is the neighborhood centered on the pixel x_k.

Basically, just like the membership function, the spatial function h_{ij} represents the membership degree of the pixel $\mathbf{x}_i$ belonging to the j-th cluster: the higher its values, the larger the number of neighbors belonging to the same clusters. The incorporation of the spatial component considerably improves the performance when segmenting noisy regions. In a homogeneous

region, the spatial functions emphasize the original membership, so the clustering results are not affected. On the other hand, for noisy images, the weighting of a noisy cluster by the labels of its neighboring pixels and the misclassified pixels from noisy regions or spurious blobs may be corrected. Formally, the spatial function modifies the membership function of a pixel according to the membership statistics of its neighbors as follows:

$$\hat{u}'_{ik} = \frac{\hat{u}^p_{ik} \cdot h^q_{ik}}{\sum_{j=1}^{C} \hat{u}^p_{jk} \cdot h^q_{jk}}, \tag{4}$$

where $\hat{u}_{ik}$ is computed by the the traditional FCM clustering algorithm; finally, as in the case of the traditional FCM algorithm, the centroid vector is updated. All the details are provided in Supplementary Section S1 'Unsupervised fuzzy C-Means clustering techniques'.

4. Materials and methods

This section presents the analyzed DCE-MRI series, as well as the proposed segmentation method based on unsupervised fuzzy clustering. Lastly, the evaluation methodology, along with the competing methods, are described.

4.1. Patient dataset composition

The study used MR images of patients recruited at the Breast Unit at the Fondazione Istituto "G. Giglio" of Cefalù. In particular, 111 breast DCE-MRI studies were acquired and then reported by consultant breast radiologists. This retrospective study was approved by the local ethical review board.

With the goal of achieving clinical feasibility, we analyzed the subtracted images (obtained from two DCE-MRI time-points), which currently represent the routine examination on breast cancer patients [34, 35]. Aiming at clinical and radiomic applications, to avoid considering lesions without sufficient informative content, only contrast-enhancing masses with a minimum diameter of 0.5 cm were considered. Manual segmentations were performed by a radiologist (by using a MatLab-coded custom tool), with more than 5-year experience on breast MRI, in consensus with a consultant breast radiologist (with more than 30-year experience on breast imaging). Table 1 shows the main characteristics of the used sequences. Figure 1 shows the histopathologically proven sub-type distribution of the contrast-enhancing masses that compose the analyzed breast DCE-MRI dataset.

Table 1: Characteristics of MR sequences used in this study.

DCE-MRI Characteristic	**Value**
Sequence type	DCE-MRI
Series description	Ax VIBRANT mphase
Scanner model	GE Signa HDxt (1.5T)
Repetition time (TR)	37.720 – 56.920 ms
Echo time (TE)	17.640 – 26.800 ms
Flip angle	10°
Matrix size	512 × 512 pixels
Slice thickness	2–3 mm
Spacing between slices	1 – 1.5 mm
Pixel spacing	0.6875 – 0.7422 mm

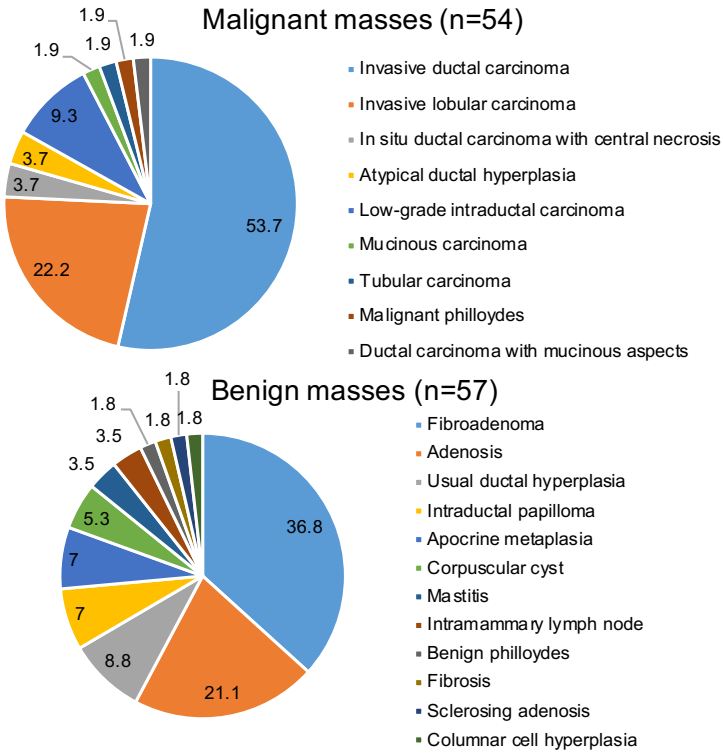


Figure 1: Histopathologically proven sub-type distribution of malignant ($n = 54$) and benign ($n = 57$) contrast-enhancing breast masses. The numbers denote percentage values.

4.2. The proposed contrast-enhancing mass segmentation method in DCE-MRI

Fig. 2 shows the overall processing pipeline for contrast-enhancing mass segmentation in breast DCE-MRI, which consists of several steps. Only the initial step (namely, 'ROI selection') requires user interaction, while the subsequent steps are fully automated; moreover, the 'ROI Refinements' step is optional.

The proposed approach and the competing methods were entirely developed using the MatLab® R2018b (64-bit version) environment (The Math-

Works, Natick, MA, USA). The tests were conducted on a Windows® 10 (Microsoft Corporation, Redmond, Washington, USA) computational platform equipped with an Intel® Core™ i7@2.7 GHz CPU and 16 GB of RAM. Fig. 3 depicts the Graphical User Interface (GUI) of the implemented tool based on the proposed approach.

4.2.1. DCE-MRI data preparation

The acquisition of DCE-MRI involves the administration of a Gadolinium-based contrast agent, which aids depicting morphological/physiological characteristics of breast lesions. Since the examination includes various acquisitions in well-defined time intervals, a DCE-MRI sequence can be analyzed as a 4D volume, wherein each voxel can be characterized by a tuple (x, y, z, t), consisting of three spatial and one temporal dimensions, respectively. Considering a specific position (x_s, y_s, z_s) within the volume, each voxel has a Time Intensity Curve (TIC(t)), function of time, reflecting the signal intensity variations due to the absorption/release of the contrast agent. The time course of the TIC curves can help clinicians infer the type of disease (e.g., benign/malignant). More recently, a CNN exploiting three time-points (3TPs) was introduced in [36], which allowed for leveraging the contrast agent effects on breast DCE-MRI by relying on the 3TP approach proposed in [37]. Unfortunately, such a choice is not straightforward and sometimes calibration maps have to be used to normalize the input data. In our work, relying upon clinicians' domain knowledge and previous literature [18], the strongest enhancement phase was analyzed since it best reflects the tumor heterogeneity and invasiveness [38]: this choice consistently resulted in a median phase (i.e., time-point) $\phi_m = 3$ and an interquartile range IQR = 1. The average

acquisition delay (i.e., between two consecutive time-points) for the analyzed DCE-MRI studies is 74 ± 7 s.

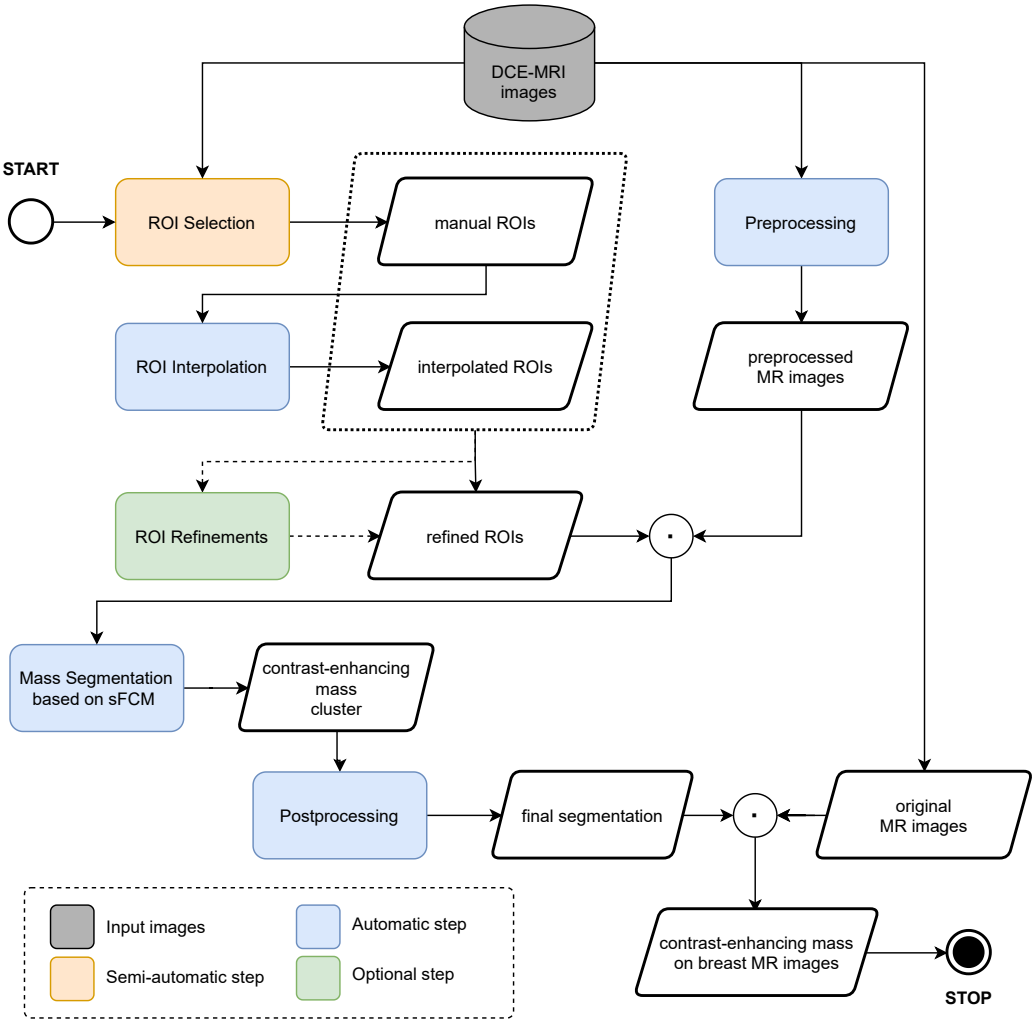


Figure 2: Overall flow diagram of the proposed semi-automated approach for contrast-enhancing mass segmentation on breast DCE-MRI. The color legend is reported in the left bottom corner.

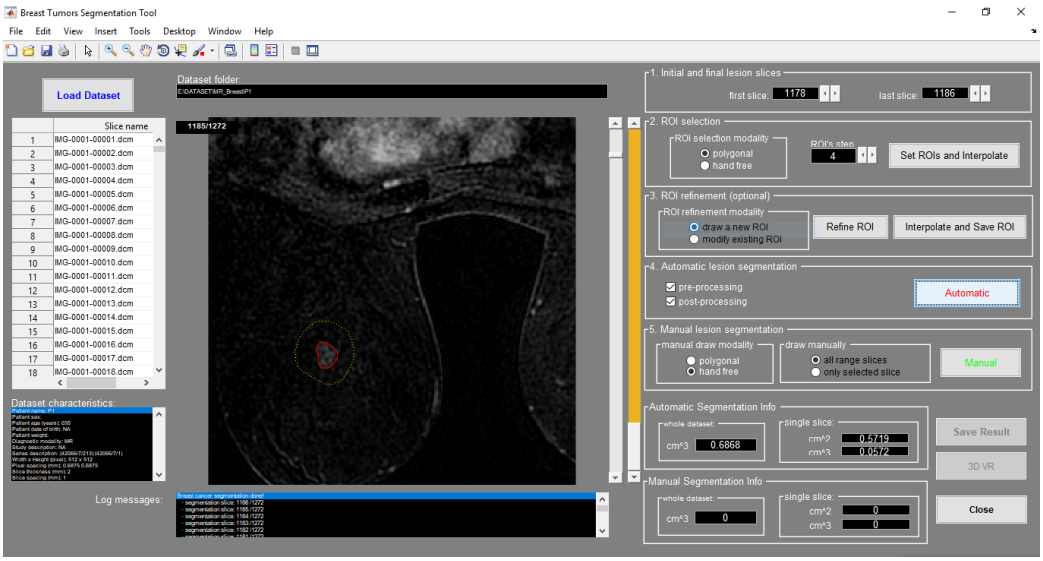


Figure 3: The GUI of the implemented tool based on the proposed semi-automated approach for contrast-enhancing mass segmentation in DCE-MRI of patients with breast cancer.

4.2.2. ROI selection

The ROIs were delineated on the whole tumor on the DCE-MRI phase with the strongest enhancement phase [18], where the ROI selection process is inspired by our previous successful strategy [39].

After loading the DCE-MRI series to be processed, the operator must select the bounding region (i.e., ROI) containing the contrast-enhancing mass. Considering that only slices including the lesion have to be segmented, the next step involves the specification of the $[S_I, S_F]$ range, which includes only the slices containing the lesion. Assuming that the lesions are generally round-shaped, it was decided to reduce the number of manually inserted ROIs. In fact, ROIs are drawn manually every ΔROI slices, where ΔROI is a parameter that indicates the offset between the slices wherein the ROI

has to be manually delineated. S_I and S_F represent the initial and the final slices, respectively (i.e., where the lesion starts and ends). The selection of the range $[S_I, S_F]$ is interactively performed by the radiologist. After the choice of S_I and S_F, the implemented tool supports the radiologist to manually delineate some rough ROIs (reference ROIs), one every ΔROI slices and, then, automatically interpolates each pair of adjacent reference ROIs, to obtain an ROI on each slice within the range $[S_I, S_F]$.

If necessary, the operator can modify the value of ΔROI. In our experiments, we always used $\Delta\text{ROI} = 4$ (the default value chosen jointly with clinicians). Fig. 4 schematizes the selection strategy used to set a ROI every ΔROI slices.

4.2.3. ROI interpolation and refinement

At the end of the previous step, reference ROIs were manually drawn, and the missing ROIs are automatically inferred by means of interpolation (see Fig. 5).

In particular, relying on the manually segmented ROIs, drawn every ΔROI slices, all the ROIs for the intermediate slices are computed by interpolating the ROIs delineated on the slices at positions i^{th} and $(i + \Delta\text{ROI})^{\text{th}}$. By doing so, the workload of the operator is sensibly reduced, because the number of slices manually set is reduced by a ΔROI factor. Considering that the points P_i and $P_{i+\Delta\text{ROI}}$ belong to the ROI on the i^{th} and $(i + \Delta\text{ROI})^{\text{th}}$ slices, respectively, it is possible to calculate the points $P_j, \forall j \in (i, i+\Delta\text{ROI})$,

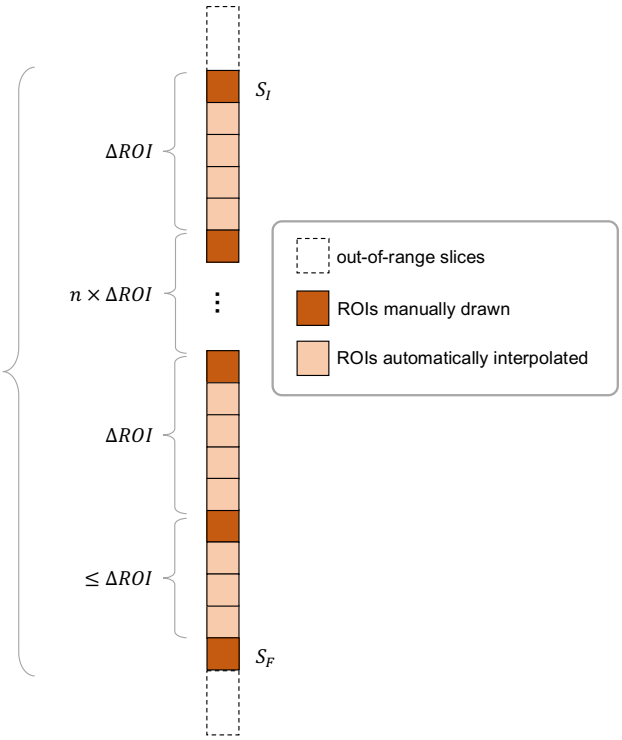


Figure 4: ROI selection strategy: within the range $[S_I, S_F]$, the user roughly delineates the ROI manually every ΔROI (dark orange squares). After delineating these ROIs, the remaining ROIs are interpolated automatically (light orange squares).

of the points of the j^{th} interpolated ROI according to Eq. (5):

$$\begin{aligned} P_j = \big\{ \big(x\text{ROI}_{i,n} + j \cdot \Delta x_n, y\text{ROI}_{i,n} + j \cdot \Delta y_n \quad | \quad n \in \mathcal{P}_{ROI} \big) \big\}, \\ \forall i \in [S_I, S_F] \wedge \text{mod}(i - S_I, \Delta\text{ROI}), \quad \forall j \in (i, i + \Delta\text{ROI}), \end{aligned} \tag{5}$$

where $\mathcal{P}_{\text{ROI}}$ represents the set of points composing the ROI.

As illustrated in Fig. 5e, the pair $(\Delta x_n, \Delta y_n)$—defined by Eq. (6)—represents the increments along the x and y directions for the n^{th} point of the ROI that must be added j times to obtain the points interpolating the

$(x\mathrm{ROI}_{i,n}, y\mathrm{ROI}_{i,n})$ and $(x\mathrm{ROI}_{i+\Delta\mathrm{ROI},n}, y\mathrm{ROI}_{i+\Delta\mathrm{ROI},n})$ point pairs, belonging to the i^{th} and $(i+\Delta\mathrm{ROI})^{\mathrm{th}}$ ROIs, respectively. This allows us to calculate the missing ROIs.

$$\left(\Delta x_n, \Delta y_n\right) = \left(\frac{x\mathrm{ROI}_{i+\Delta\mathrm{ROI},n} - x\mathrm{ROI}_{i,n}}{\Delta\mathrm{ROI}-1}, \frac{y\mathrm{ROI}_{i+\Delta\mathrm{ROI},n} - y\mathrm{ROI}_{i,n}}{\Delta\mathrm{ROI}-1} \right) \quad (6)$$

Because this interpolation strategy has been applied here to a clinical scenario different from the original (i.e., chest CT angiography) [39], before its use in the segmentation, an analysis and validation phase was accurately performed together with the clinicians, who endorsed its validity in the new breast DCE-MRI segmentation task and the effectiveness in remarkably reducing the workload of the human operator.

If deemed necessary, when the manual ROI is not sufficiently accurate or the result of the interpolation is not satisfactory, the operator can replace the ROI by redrawing it. The adjacent ROIs are automatically updated by performing a new interpolation that takes into account the newly drawn ROI. To endorse the quality of the interpolation results obtained by our procedure, this event never occurred.

4.2.4. Image preprocessing and postprocessing

Preprocessing. Before the sFCM clustering, a 2D median filter with a 5×5 pixel kernel was applied for image smoothing. The use of a median filter is a common choice in DCE-MRI image analysis. For instance, a 3D median filtering was applied to facilitate the ROI identification by removing (or at least reducing) the outlier introduced by anatomical peculiarities [21] or also deal with small patient shifts [40]. With specific reference to impulsive

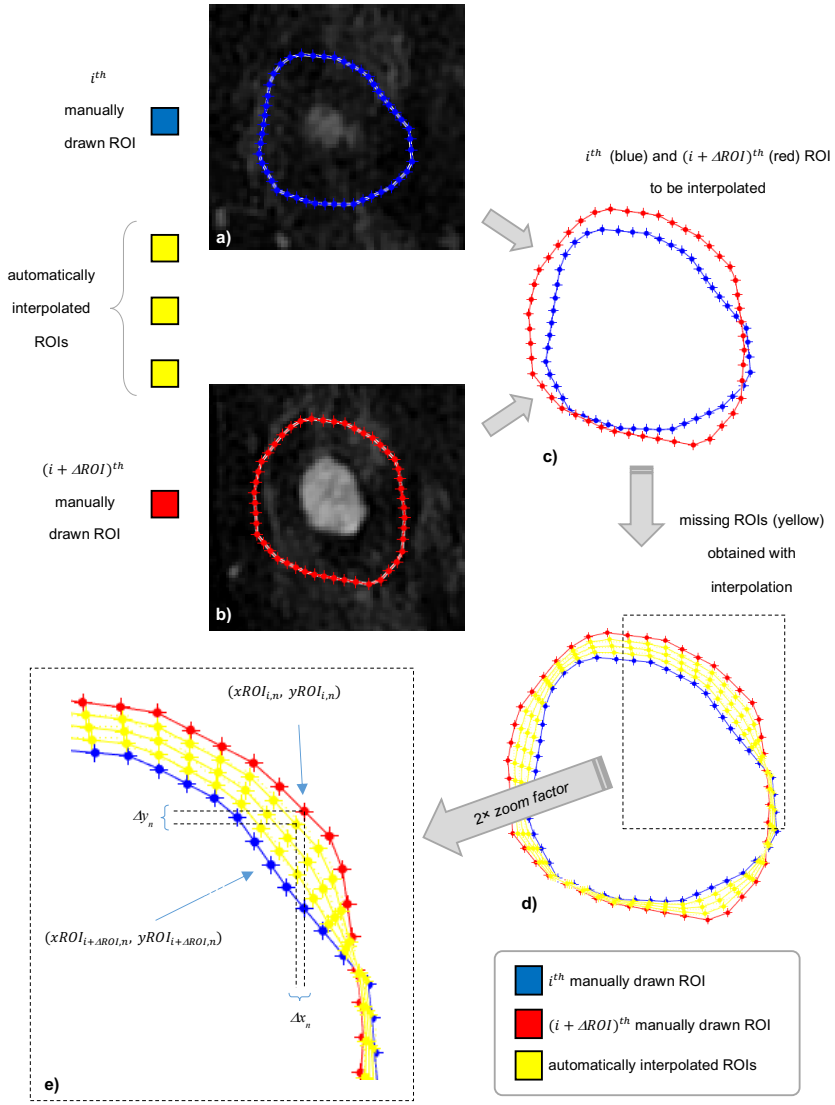


Figure 5: ROIs interpolation steps: a) and b) i^{th} and $(i + \Delta\text{ROI})^{\text{th}}$ slices with the overimposed manually delineated contours; c) i^{th} and the $(i + \Delta\text{ROI})^{\text{th}}$ contours (in blue and red, respectively) to be interpolated; d) the 3 ROIs (in this case $\Delta\text{ROI} = 4$) obtained after the interpolation process (in yellow); e) zoomed area showing some interesting details concerning interpolation algorithm.

noise models, a 3D median filter was also systematically applied in [41] to pharmacokinetic parameter maps before masking to suppress spatial noise.

Postprocessing. Although preprocessing reduces intensity discontinuities, it cannot always manage to completely eliminate non-uniformities that may cause holes within the segmented masses. This problem was overcome by applying a hole-filling algorithm, based on morphological reconstruction [42]. Moreover, considering that sFCM could detect, in addition to the lesion, also small connected-components, a flood-fill algorithm was executed for the identification of the largest connected-components, representing the contrast-enhancing mass.

4.2.5. Contrast-enhancing mass segmentation

Breast lesions generally appear brighter than the surrounding adipose and muscle tissues on DCE-MRI. Furthermore, anatomical atlases and radiology reports refer that breast lesions are typically spherical [43]. Relying on these morphological features, we decided to use the sFCM, a variant of the traditional FCM, which allowed us to improve the segmentation when the 'pixels of interest' are spatially close to each other [32].

In fact, one of the important characteristics of an image is that neighboring pixels are highly correlated. In other words, these neighboring pixels have similar feature values, and the probability that they belong to the same cluster is high [32, 44]. This spatial relationship is important in clustering, but it is not utilized in a traditional FCM algorithm.

The segmentation *via* sFCM analyzed the images after the application of the preprocessing, in order to obtain images with similar characteristics. The

masks yielded by clustering are used to extract the lesion from the original images (which did not undergo preprocessing). In fact, in some contexts—such as radiomics—it is essential to keep original data, without any alteration, so that they may reflect the real anatomical/metabolic characteristics of the breast masses.

4.3. Segmentation evaluation procedure

The quantitative evaluation of the proposed computational method was performed by comparing the automatically segmented DCE-MRI image (S) against the corresponding gold standard manual segmentation (G) using spatial overlap- and distance-based metrics [45, 46, 47]. Since our method analyses 2D MR images, we calculated slice-wise metrics that were then averaged per lesion.

The used overlap-based evaluation metrics were: Dice similarity coefficient (DSC), Jaccard index (JI), sensitivity (Sen), specificity (Spc). We also computed the following distance-based metrics: Hausdorff distance (HD), mean absolute distance (MAD), maximum absolute distance (MaxD). The detailed definitions are provided in Section S2 'Segmentation evaluation metrics' of Supplementary Materials.

4.3.1. Competing methods

With the goal of providing a fair and careful comparison with the classic image processing approaches for breast lesion segmentation on DCE-MRI proposed in the literature approaches, a direct evaluation on the same dataset was performed. In particular, the proposed sFCM-based approach was compared against the following segmentation approaches:

- Otsu's method [48], inspired by [19];
- traditional FCM algorithm, similar to [16].

To strengthen the experimental analysis, we present also a proper experimental comparison on the same dataset between the proposed method and other literature approaches. With more details, we compared our implemented approach against CNN-based approaches, for the following reasons:

- the only breast dataset, named QIN DCE-MRI [49, 50], was not collected for segmentation purposes (it is a longitudinal dataset created for assessing breast cancer response to neoadjuvant chemotherapy);
- even by re-implementing the approaches in Table 6, the full reproducibility cannot be ensured;
- no other public breast DCE-MRI datasets are currently available, therefore we used our curated DCE-MRI dataset to obtain a common database for appropriate and fair comparisons with the state-of-the-art.

In particular, we implemented and tested the most popular CNN-based approaches for semantic segmentation of medical images [51], namely SegNet [52] and U-Net [53], to directly compare them on the same data. To allow other researchers to reproduce the same results, we maintained the default architectural characteristics (provided by the MatLab Deep Learning Toolbox). Table 2 provides the most relevant details about implemented CNN-based architectures. In both cases, the used optimizer was the Stochastic Gradient Descent with Momentum (SGDM).

For an accurate and fair comparison, 2D slices were considered also for the CNN-based segmentation. In particular, the cropped images were derived from the manually segmented breast masses. The size of the input matrices (96×96 pixels) was determined considering the bounding-box (centered on the lesion centroid) that contains the largest lesion in the whole dataset. By doing so, we optimized the processing pipeline, avoiding to feed the whole image to the CNNs. Moreover, the use of the bounding-box for the CNN segmentation is equivalent to the ROI selection adopted in our processing pipeline.

Table 2: The most relevant architectural and training details of the SegNet and U-Net CNN-based architectures.

Characteristic	**SegNet**	**U-Net**
Input size	96×96 pixels	
Output classes	2	
CNN layers	32	58
Optimizer	SGDM	
Initial learning rate	1×10^{-3}	
Maximum number of epochs	100	

A 5-fold cross-validation (CV) was adopted for the CNN training. Therefore, for each CV round, the whole dataset was split into 80%/20%, where the 80% and the 20% were used for CNN training and test, respectively. The mask obtained by CNN segmentation was elaborated by applying the same postprocessing steps exploited for the proposed approach (*i*) hole filling, (*ii*) largest connected-component selection, and (*iii*) morphological opening). For an accurate result comparison, this CV scheme was adopted as a test dataset for the proposed approach, despite our sFCM is unsupervised and does not require any training phase (Table 7).

5. Experimental results

This section presents the experimental results achieved by the proposed computer-assisted method. Segmentation performance, as well as correlation between gold standard (manual) and automated volumes, are reported and discussed. Furthermore, the analysis performed for sFCM parameter tuning, necessary to optimize the final clustering performance, is shown.

5.1. Parameter settings

The spatial function—defined in Eq. (3)—and the membership function—defined in Eq. (4)—are controlled by means of p and q weighting parameters. The tuning of these two parameters was performed on a calibration subset composed of 30 randomly selected cases. We assessed the performance in terms of DSC. Table 3 illustrates the average results obtained in this preliminary setting phase, which justifies the final choice to use the sFCM with $\langle p, q\rangle = \langle 1, 2\rangle$. The configuration with $\langle p, q\rangle = \langle 1, 0\rangle$, corresponding to the traditional FCM, obtained a lower mean value of DSC, demonstrating that the clustering with spatial constrains more accurately identified the tumor. Supplementary Figures S1 and S2 show a comparison of manual delineations against the automated segmentation results obtained by sFCM during the tuning phase by varying the parameters p and q.

5.2. Segmentation performance

Along with the quantitative segmentation assessment, we computed the correlation between the volumes calculated from the radiologist's gold standard (V_{manual}) and the volumes obtained by the proposed method ($V_{\text{automated}}$).

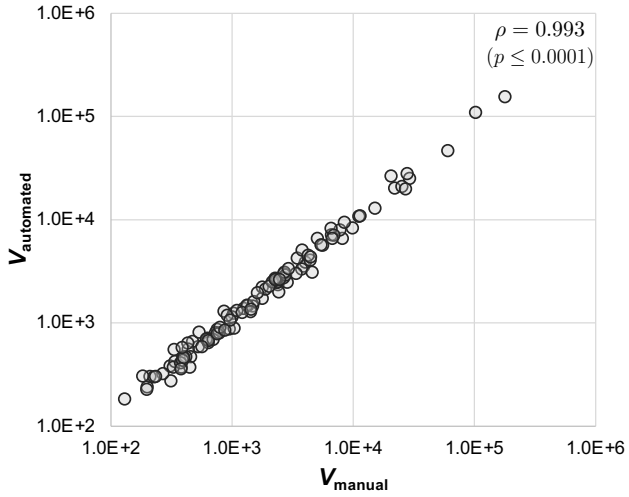
Table 3: DSC values (mean ± standard deviation) obtained during tuning of p and q parameters with different values of neighborhood $\mathcal{N}$.

$\langle p, q \rangle$	$\mathcal{N}_{3\times3}$	$\mathcal{N}_{5\times5}$
$\langle 1, 0 \rangle$	76.26 ± 18.19	74.59 ± 18.44
$\langle 1, 1 \rangle$	75.24 ± 14.78	69.67 ± 23.02
$\langle 1, 2 \rangle$	**77.66 ± 15.09**	**78.2 ± 15.07**
$\langle 2, 0 \rangle$	69.7 ± 21.46	71.2 ± 22.44
$\langle 2, 1 \rangle$	77.73 ± 15.13	75.86 ± 15.25
$\langle 2, 2 \rangle$	71.31 ± 22.3	77.97 ± 15.04

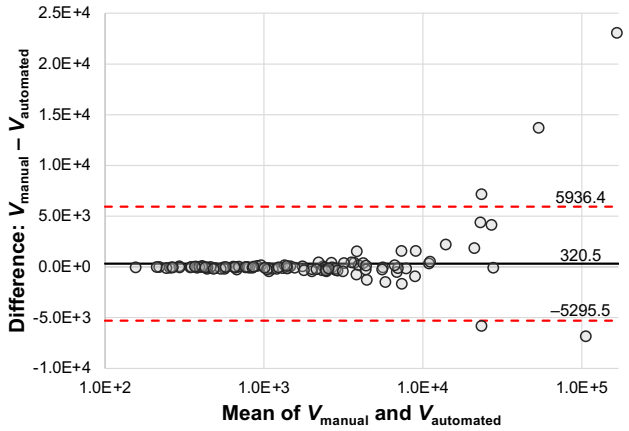
Fig. 6a shows the corresponding scatter diagram: the achieved Pearson's correlation coefficient ρ confirms the high accuracy of the method. Along with correlation coefficients, Fig. 6b depicts the corresponding Bland-Altman plot. As a matter of fact, the Bland–Altman analysis better shows the data distribution by plotting the pairwise difference between the volumes measured by the two methods against their mean. It can be observed that, except for a few outlier cases with very large volumes, there is no systematic bias.

Fig. 7 plots the mean DSC values achieved by the implemented and tested segmentation methods. For statistical validation of the results achieved by the compared methods, the two-sided Wilcoxon signed rank test on paired DSC results [54] was performed with the null hypothesis that the samples come from continuous distributions with equal medians. In all the tests, a significance level of 0.05 was considered. The p-values, corrected by the Bonferroni–Holm method [55] for multiple comparisons, are shown at the top of Fig. 7. According to Fig. 7a, both FCM and sFCM achieved signifi-

cantly higher DSC values compared to adaptive thresholding ($p \ll 0.001$ in



(a)



(b)

Figure 6: (a) Scatter diagram showing the correlation $V_{\text{automated}}$ *versus* V_{manual}. Mass volumes are expressed as 'number of pixels' (a logarithmic scale is used). The Pearson's correlation coefficient is $\rho = 0.993$ ($p \leq 0.0001$). (b) Bland–Altman plots of the automated volumes ($V_{\text{automated}}$) *versus* the manual measurements (V_{manual}). Solid black horizontal and dashed red lines denote the mean and ± 1.96 standard deviation values, respectively.

both cases). Importantly, sFCM significantly outperformed the traditional FCM algorithm ($p = 3.511 \times 10^{-4}$), thus confirming the importance of the incorporation of the spatial information.

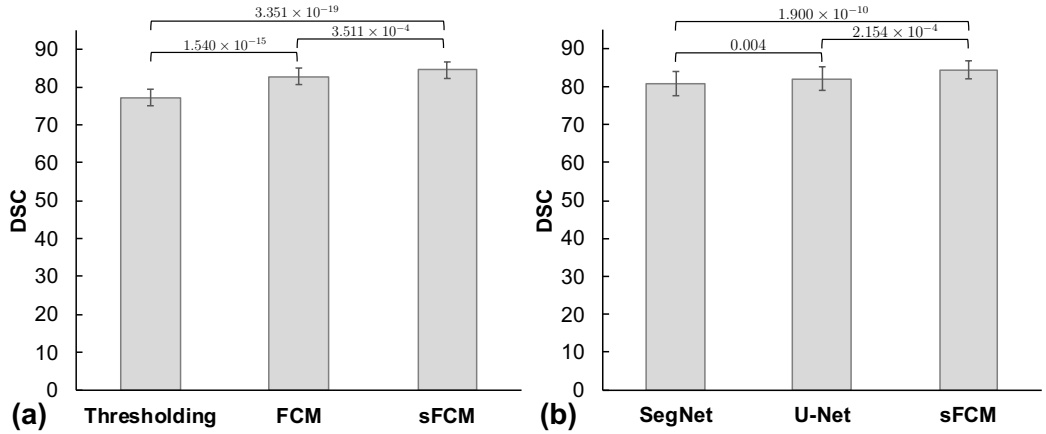


Figure 7: DSC values achieved by: (a) interactive methods based on classic image processing techniques (namely, thresholding, FCM and sFCM); (b) CNN-based approaches compared against the proposed sFCM method (in a 5-fold CV scheme). The bar graph and error bars denote the average value and the standard deviation DSC values, respectively. The p-values, obtained from the statistical validation procedure, are shown at the top of the bars as brackets. The Wilcoxon rank-sum test for pairwise result comparison was used with the alternative hypothesis that the samples do not have equal medians of DSC values. A significance level of $\alpha = 0.05$ with a correction using the Bonferroni–Holm method for multiple comparisons.

Tables 4 and 5 depict all the overlap-based and distance-based metrics, respectively, obtained by the implemented and tested approaches. Fig. 8 shows a segmentation example obtained using thresholding, FCM and sFCM, respectively. From a qualitative assessment, it is appreciable how the performances of FCM and sFCM are very similar (although the sFCM achieves overall better results). The segmentation based on thresholding obtained

the worst performances, due to many false misclassified pixels. Analogously, Fig. 9 shows two distinct examples of segmentation results yielded by SegNet, U-Net and the proposed sFCM-based method, respectively. The CNN-based approaches achieved accurate segmentation results, although our approach integrating the sFCM allowed us to obtain results superior to SegNet and U-Net without requiring any training phase.

Table 4: Segmentation results on the analyzed breast DCE-MRI dataset. The overlap-based metrics are expressed as the mean value ± standard deviation. Boldface indicates the best value for each metric.

Approach	JI	DSC	Sen	Spc
Thresholding	64.81 ± 5.36	77.27 ± 6.11	81.19 ± 5.32	75.20 ± 7.23
FCM	72.15 ± 6.23	82.86 ± 5.87	86.39 ± 4.56	81.09 ± 5.15
sFCM	**73.95 ± 5.8**	**84.47 ± 4.75**	**89.23 ± 5.71**	**82.42 ± 6.98**

Table 5: Segmentation results on the analyzed breast DCE-MRI dataset. The distance-based metrics are expressed as the mean value ± standard deviation. Boldface indicates the best value for each metric.

Approach	HD	MAD	MaxD
Thresholding	3.25 ± 0.82	2.21 ± 1.2	4.41 ± 3.16
FCM	2.71 ± 0.43	**1.13 ± 0.88**	3.62 ± 2.34
sFCM	**2.19 ± 0.51**	1.15 ± 0.71	**3.28 ± 2.56**

5.3. Comparison against existing work

For the sake of completeness, this section compares existing literature works addressing the problem of segmentation of breast lesions in DCE-MRI

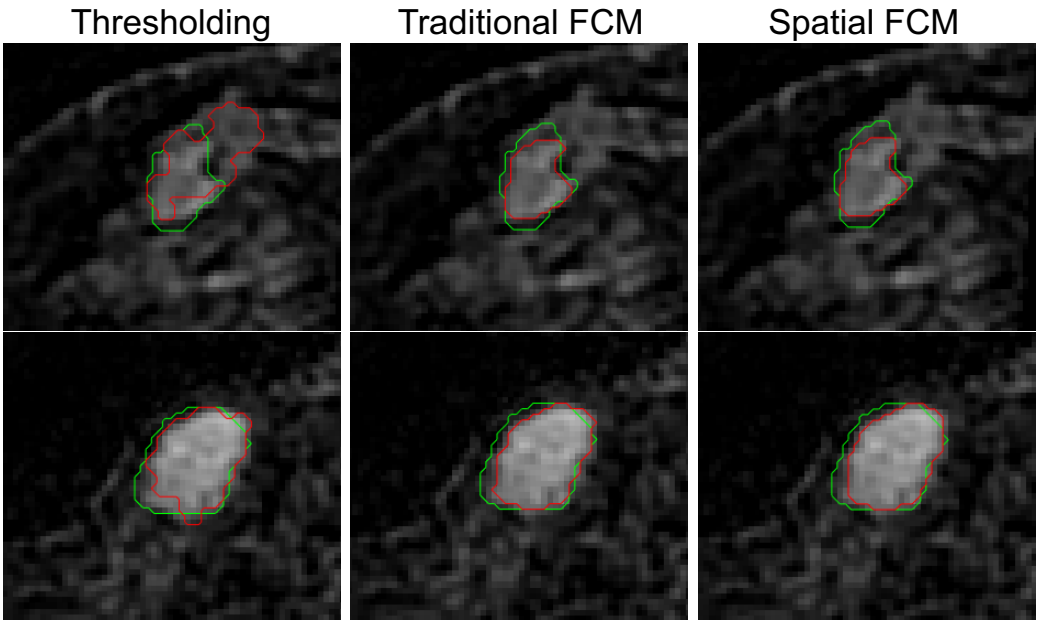


Figure 8: Segmentation results obtained using thresholding, FCM and sFCM integrated into the proposed semi-automated pipeline. The automated and the gold standard contours are denoted in red and green, respectively. Two distinct examples (in the rows) are shown and a 4× zoom factor was used.

sequences. Table 6, which merely reports the numerical values reported in the original publications, has the sole purpose of providing an indication of the performance that can be obtained by the various approaches. The main aim of this comparison is just to provide evidence of the performance that can be achieved by existing work; indeed, since different datasets were used in these studies, a direct comparison among them cannot be carried out. We showed that the achieved performance was in line with existing work based on either classic machine learning or deep learning techniques. Along with the introduction of the spatial FCM clustering algorithm, our novelty is represented by the smart and reliable interpolation strategy that allows us

to offer a user-friendly tool for radiologists.

The work in [19]—based on thresholding—obtained the highest sensitivity (98%), although it should be noted that only 65 lesions were processed and this method was designed to be highly sensitive at the cost of false positives. Unfortunately, no other evaluation metric was provided. Both the approaches proposed in [16] and [20] used FCM. Considering that [16] had twice as many patients/lesions with respect to [20], we could argue that they have comparable performance. Especially, the JI values reported in [16] and [20] were characterized by a higher standard deviation compared to the proposed method. Also the CNN-based method presented in [9] did not report superior performance in terms of DSC and sensitivity. In these scenarios, our approach—coupling a smart interactive strategy and an sFCM-based segmentation method—achieved comparable, or even the highest, performance on a dataset of 111 lesions. This evidence was confirmed also by the highest Pearson's correlation coefficient for ($V_{\text{automated}}$) *versus* (V_{manual}).

Table 6: Comparison of the proposed method against existing literature approaches (*N/A*: Not available).

Reference	Approach	Patients/Lesions	DSC	JI	Sen	Correlation
Vignati *et al.* (2011) [19]	Thresholding	48/65	*N/A*	*N/A*	98	*N/A*
Chen *et al.* (2006) [16]	FCM	121/121	*N/A*	64±12	*N/A*	0.98
Pang *et al.* (2012) [20]	FCM+GVF	60/60	*N/A*	78±8	*N/A*	0.976
Wang *et al.* (2021) [9]	CNN	90/90	76.48	*N/A*	75.93	*N/A*
Proposed approach	sFCM	111/111	84.47±4.75	73.95±5.8	89.23±5.71	0.993

Concerning the comparisons against CNN-based approaches, the results in Table 7 represent the evaluation metrics obtained in the 5-fold CV (expressed as mean ± standard deviation). In particular, as shown in Fig. 7b,

the proposed sFCM method achieved significantly higher DSC values ($p = 1.900 \times 10^{-10}$ and $p = 2.154 \times 10^{-4}$ compared to SegNet and U-Net, respectively), while U-Net outperformed SegNet ($p = 0.004$).

Unlike the approach proposed by Wang *et al.* [9], which combined three phases of the DCE-MRI series (namely, the subtraction image, the initial uptake image and the delayed response image) as input for a mixed 2D-3D architecture, in the sFCM segmentation we used only one phase (with the strongest enhancement, according to the clinicians' routine). This choice was motivated by keeping the existing clinical protocol for an immediate clinical feasibility of the implemented approach.

Table 7: Evaluation metrics (in terms of mean value ± standard deviation) obtained in 5-fold cross-validation. The standard deviation denotes the variability over the five CV rounds. Boldface indicates the best value for each metric.

Approach	JI	DSC	Sen	Spc	HD	MAD	MaxD
SegNet	69.89 ± 2.22	80.79 ± 1.83	85.13 ± 2.17	77.46 ± 4.86	3.2 ± 0.91	1.87 ± 0.71	4.18 ± 3.22
U-Net	72.74 ± 2.67	82.11 ± 2.28	86.21 ± 3.24	81.7 ± 6.19	2.78 ± 0.55	1.24 ± 0.67	3.4 ± 2.93
Proposed approach	**74.67 ± 4.11**	**84.42 ± 1.70**	**92.57 ± 2.36**	**84.93 ± 5.21**	**2.42 ± 0.62**	**1.03 ± 0.78**	**3.11 ± 2.85**

6. Discussion and conclusion

In clinical routine, the quantification of lesion extent on DCE-MRI plays a vital role must be as accurate as possible [13]. Along with volumetric analyses, radiomics enables non-invasive diagnosis by means of predictive models that rely on features extracted from images [14, 23]. While texture features obtained promising performance in distinguishing benign and malignant lesions [15], they can be combined with the current BI-RADS descriptors of internal tumor enhancement pattern and lesion margin [5]. For this reason,

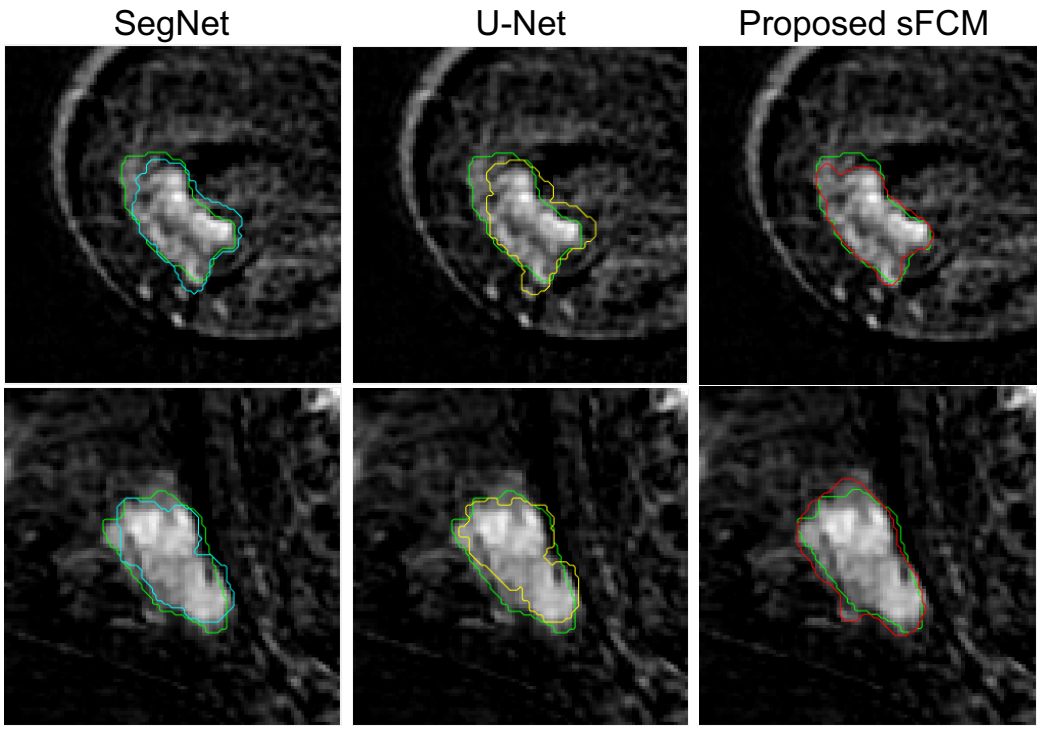


Figure 9: Segmentation results obtained comparing the proposed approach based on sFCM (red contour), SegNet segmentation (cyan contour) and U-Net segmentation (yellow contour) against the ground-truth (green contour). Examples are shown with a 2× zoom factor.

accurate quantification of the shape-based and texture features of the lesion is fundamental for precision oncology, targeted to the specific clinical case. Computer-assisted segmentation approaches could facilitate tumor annotation and, successively, diagnosis and treatment planning tasks, with reduced reporting-time and operator dependence compared to the analogous manual procedures.

In this work, we proposed an unsupervised machine learning approach based on sFCM used to segment masses on DCE-MRI images of patients

with breast lesions recruited by a single institution. In particular, a semi-automated segmentation tool was implemented to reliably help radiologists in accurate detection and segmentation. The obtained results, in terms of overlap-based and distance-based metrics, demonstrated the superiority of the sFCM, compared against global thresholding [19] and the traditional FCM clustering [16].

The proposed segmentation tool requires the user input for the rough selection of the mass region that is then propagated over all the slices of interest *via* a smart interpolation strategy. In the near future, the authors aim at further improving this computer-assisted segmentation approach, by automatically detecting the breast lesion [56]; in such a case, a CNN-based detection approach could be exploited [57]. Despite the analyzed DCE-MRI dataset having been prepared and curated by a single institution, our method could offer good generalization abilities also on external datasets collected at other institutions, using different MRI scanners and protocols. Indeed, our approach currently processes only the time-point with the strongest enhancement phase, and should therefore be less affected by different MRI acquisition protocols [18] which might facilitate its integration into the clinical practice compared to deep learning based methods.

The availability of publicly accessible datasets is a necessary resource for fair comparisons between a new method and existing approaches. Unfortunately, however, there are not always 'useful' resources available to the research area in question. In particular, our study used DCE-MRI images alone to enable the clinical feasibility of the approach. In fact, DCE-MRI is the routine exam in the clinical practice for patients with breast cancer [34, 35].

From this point of view, the publicly available resources [58, 59, 60, 49, 50] which do not comprise DCE-MRI for segmentation purposes, are not suitable for our purposes; however, the adoption of unenhanced MRI sequences, such as quantitative DWI, is highly relevant for future development [61, 10]. QIN Breast DCE-MRI dataset [49, 50] contains MRI images from a longitudinal study to assess breast cancer response to neoadjuvant chemotherapy of 20 patients. The only work we have found that uses this breast dataset is [62], where a two steps approach for breast lesion segmentation is proposed. To reduce the search area to a specific ROI, in the first step a raw breast mask is given by simply subtracting the first image and subsequent MRI images. In the second step, the breast lesions within the previously obtained breast mask are segmented. Three classic image processing methods were tested: (*i*) thresholding based on the Otsu's method, (*ii*) seeded region-growing, and (*iii*) K-means clustering. A postprocessing based on morphological operations was applied to remove some segmentation artifacts. Two area-based evaluation metrics (namely, DSC and JI) were considered, calculated only on 20 DCE-MRI 2D slices of patient *BC01*. The best result was obtained by K-means clustering (number of clusters $K = 4$), with DSC=0.93 and JI=0.87. The tests were performed on an arbitrary subset of the QIN Breast DCE-MRI dataset (one patient only), this therefore hampers a comprehensive comparison. Nevertheless, the most important experimental finding is that unsupervised K-means clustering outperformed the other classic image processing techniques in breast lesion segmentation on DCE-MRI.

Regarding the possibility of making the dataset publicly available for the research community, it should be noted that the approval of this retrospective

study, by the local ethical review board, currently allows for the data storage and analysis only inside the institution, despite the anonymization process. We plan to address this important issue for result reproducibility.

As a further development, with the goal of fully leveraging the several time-points provided by DCE-MRI, we plan to develop a multistep unsupervised clustering approach based on time-series analysis [63, 64], which exploits the sFCM for each single time-point, as well as the pharmacokinetics modeling derived from DCE-MRI [36, 65].

Acknowledgements

This study has received funding by GeSeToN project, funded by Italian MISE Grant No. 489 of 21/02/2018.

This study has also been partially supported by The Mark Foundation for Cancer Research and Cancer Research UK Cambridge Centre [C9685/A25177] and by the Royal Society for the International Exchanges 2020 Cost Share with the Italian CNR (project No. IEC/R2/202313).

Additional support was also provided by the National Institute of Health Research (NIHR) Cambridge Biomedical Research Centre [BRC-1215-20014]. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

References

[1] R. L. Siegel, K. D. Miller, H. E. Fuchs, A. Jemal, Cancer statistics, 2021, CA Cancer J. Clin. 71 (1) (2021) 7–33. doi:10.3322/caac.21654.

[2] C. E. DeSantis, J. Ma, M. M. Gaudet, L. A. Newman, K. D. Miller, A. Goding Sauer, A. Jemal, R. L. Siegel, Breast cancer statistics, 2019, CA Cancer J. Clin. 69 (6) (2019) 438–451. doi:10.3322/caac.21583.

[3] E. Le, Y. Wang, Y. Huang, S. Hickman, F. Gilbert, Artificial intelligence in breast imaging, Clin. Radiol. 74 (5) (2019) 357–366. doi:10.1016/j.crad.2019.02.006.

[4] B. Lauby-Secretan, C. Scoccianti, D. Loomis, L. Benbrahim-Tallaa, V. Bouvard, F. Bianchini, K. Straif, Breast-cancer screening—viewpoint of the IARC Working Group, N. Engl. J. Med. 372 (24) (2015) 2353–2358. doi:10.1056/NEJMsr1504363.

[5] B. Reig, L. Heacock, K. J. Geras, L. Moy, Machine learning in breast MRI, J. Magn. Reson. Imaging 52 (4) (2020) 998–1018. doi:10.1002/jmri.26852.

[6] F. Sardanelli, C. Boetes, B. Borisch, T. Decker, M. Federico, F. J. Gilbert, T. Helbich, S. H. Heywang-Köbrunner, W. A. Kaiser, M. J. Kerin, et al., Magnetic resonance imaging of the breast: recommendations from the eusoma working group, Eur. J. Cancer 46 (8) (2010) 1296–1316. doi:10.1016/j.ejca.2010.02.015.

[7] Y. Liu, L. Ren, X. Cao, Y. Tong, Breast tumors recognition based on edge feature extraction using support vector machine, Biomed. Signal Process. Control 58 (2020) 101825. doi:10.1016/j.bspc.2019.101825.

[8] D. Sheth, M. L. Giger, Artificial intelligence in the interpretation of

breast cancer on MRI, J. Magn. Reson. Imaging 51 (5) (2020) 1310–1324. doi:10.1002/jmri.26878.

[9] H. Wang, J. Cao, J. Feng, Y. Xie, D. Yang, B. Chen, Mixed 2D and 3D convolutional network with multi-scale context for lesion segmentation in breast DCE-MRI, Biomed. Signal Process. Control 68 (2021) 102607. doi:10.1016/j.bspc.2021.102607.

[10] G. C. Baxter, M. J. Graves, F. J. Gilbert, A. J. Patterson, A meta-analysis of the diagnostic performance of diffusion MRI for breast lesion characterization, Radiology 291 (3) (2019) 632–641. doi:10.1148/radiol.2019182510.

[11] S. G. Orel, N. Kay, C. Reynolds, D. C. Sullivan, BI-RADS categorization as a predictor of malignancy, Radiology 211 (3) (1999) 845–850. doi:10.1148/radiology.211.3.r99jn31845.

[12] C. Xu, M. Lou, Y. Qi, Y. Wang, J. Pi, Y. Ma, Multi-scale attention-guided network for mammograms classification, Biomed. Signal Process. Control 68 (2021) 102730. doi:10.1016/j.bspc.2021.102730.

[13] H. M. Whitney, H. Li, Y. Ji, P. Liu, M. L. Giger, Comparison of breast MRI tumor classification using human-engineered radiomics, transfer learning from deep convolutional neural networks, and fusion methods, Proc. IEEE 108 (1) (2019) 163–177. doi:10.1109/JPROC.2019.2950187.

[14] R. J. Gillies, P. E. Kinahan, H. Hricak, Radiomics: images are more

than pictures, they are data, Radiology 278 (2) (2015) 563–577. doi:10.1148/radiol.2015151169.

[15] W. Chen, M. L. Giger, G. M. Newstead, U. Bick, S. A. Jansen, H. Li, L. Lan, Computerized assessment of breast lesion malignancy using DCE-MRI: robustness study on two independent clinical datasets from two manufacturers, Acad. Radiol. 17 (7) (2010) 822–829. doi:10.1016/j.acra.2010.03.007.

[16] W. Chen, M. L. Giger, U. Bick, A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images, Acad. Radiol. 13 (1) (2006) 63–72. doi:10.1016/j.acra.2005.08.035.

[17] L. Rundo, R. Pirrone, S. Vitabile, E. Sala, O. Gambino, Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine, J. Biomed. Inform. 108 (2020) 103479. doi:10.1016/j.jbi.2020.103479.

[18] J. Liu, D. Sun, L. Chen, Z. Fang, W. Song, D. Guo, T. Ni, C. Liu, L. Feng, Y. Xia, et al., Radiomics analysis of dynamic contrast-enhanced magnetic resonance imaging for the prediction of sentinel lymph node metastasis in breast cancer, Front. Oncol. 9 (2019) 980. doi:10.3389/fonc.2019.00980.

[19] A. Vignati, V. Giannini, M. De Luca, L. Morra, D. Persano, L. A. Carbonaro, I. Bertotto, L. Martincich, D. Regge, A. Bert, et al., Perfor-

mance of a fully automatic lesion detection system for breast DCE-MRI, J Magn. Reson. Imaging 34 (6) (2011) 1341–1351.

[20] Y. Pang, L. Li, W. Hu, Y. Peng, L. Liu, Y. Shao, Computerized segmentation and characterization of breast lesions in dynamic contrast-enhanced mr images using fuzzy c-means clustering and snake algorithm, Comput. Math. Methods Med. 2012. doi:10.1155/2012/634907.

[21] S. Marrone, G. Piantadosi, R. Fusco, A. Petrillo, M. Sansone, C. Sansone, Breast segmentation using fuzzy c-means and anatomical priors in DCE-MRI, in: Proc. 23rd International Conference on Pattern Recognition (ICPR), IEEE, 2016, pp. 1472–1477. doi:10.1109/ICPR.2016.7899845.

[22] G. Piantadosi, M. Sansone, R. Fusco, C. Sansone, Multi-planar 3D breast segmentation in MRI via deep convolutional neural networks, Artif. Intell. Med. 103 (2020) 101781. doi:10.1016/j.artmed.2019.101781.

[23] I. Castiglioni, L. Rundo, M. Codari, G. Di Leo, C. Salvatore, M. Interlenghi, et al., AI applications to medical images: From machine learning to deep learning, Phys. Med. 83 (2021) 9–24. doi:10.1016/j.ejmp.2021.02.006.

[24] S. Wu, S. P. Weinstein, E. F. Conant, M. D. Schnall, D. Kontos, Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images, Med. Phys. 40 (4) (2013) 042301. doi:10.1118/1.4793255.

[25] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, G.-Z. Yang, Deep learning for health informatics, IEEE J. Biomed. Health Inform. 21 (1) (2017) 4–21. doi:10.1109/JBHI.2016.2636665.

[26] L. Caponetti, G. Castellano, V. Corsini, MR brain image segmentation: a framework to compare different clustering techniques, Information 8 (4) (2017) 138. doi:10.3390/info8040138.

[27] L. Rundo, C. Militello, A. Tangherloni, G. Russo, S. Vitabile, M. C. Gilardi, G. Mauri, NeXt for neuro-radiosurgery: A fully automatic approach for necrosis extraction in brain tumor mri using an unsupervised machine learning technique, Int. J. Imaging Syst. Technol. 28 (1) (2018) 21–37. doi:10.1002/ima.22253.

[28] D.-Q. Zhang, S.-C. Chen, A novel kernelized fuzzy c-means algorithm with application in medical image segmentation, Artif. Intell. Med. 32 (1) (2004) 37–50. doi:10.1016/j.artmed.2004.01.012.

[29] D. L. Pham, Spatial models for fuzzy clustering, Comput. Vis. Image Underst. 84 (2) (2001) 285–297. doi:10.1006/cviu.2001.0951.

[30] M. Gong, Y. Liang, J. Shi, W. Ma, J. Ma, Fuzzy c-means clustering with local information and kernel metric for image segmentation, IEEE Trans. Image Process. 22 (2) (2012) 573–584. doi:10.1109/TIP.2012.2219547.

[31] S. Krinidis, V. Chatzis, A robust fuzzy local information c-means clustering algorithm, IEEE Trans. Image Process, 19 (5) (2010) 1328–1337. doi:10.1109/TIP.2010.2040763.

[32] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, T.-J. Chen, Fuzzy c-means clustering with spatial information for image segmentation, Comput. Med. Imaging Graph. 30 (1) (2006) 9–15. doi:10.1016/j.compmedimag.2005.10.001.

[33] B. N. Li, C. K. Chui, S. Chang, S. H. Ong, Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation, Comput. Biol. Med. 41 (1) (2011) 1–10. doi:10.1016/j.compbiomed.2010.10.007.

[34] C. Kuhl, The current status of breast MR imaging. Part I. choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice, Radiology 244 (2) (2007) 356—-378. doi:10.1148/radiol.2442051620.

[35] E. Yeh, Breast magnetic resonance imaging: Current clinical indications, Obstet. Gynecol. Clin. North Am. 38 (1) (2011) 159–177. doi:10.1016/j.ogc.2011.02.008.

[36] M. Gravina, S. Marrone, M. Sansone, C. Sansone, DAE-CNN: exploiting and disentangling contrast agent effects for breast lesions classification in DCE-MRI, Pattern Recogn. Lett. 145 (2021) 67–73. doi:10.1016/j.patrec.2021.01.023.

[37] H. Degani, V. Gusis, D. Weinstein, S. Fields, S. Strano, Mapping pathophysiological features of breast tumors by mri at high spatial resolution, Nature Medicine 3 (1997) 780–782. doi:10.1038/nm0797-780.

[38] J. R. Teruel, M. G. Heldahl, P. E. Goa, M. Pickles, S. Lundgren, T. F. Bathen, P. Gibbs, Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer, NMR Biomed. 27 (8) (2014) 887–896. doi:10.1002/nbm.3132.

[39] C. Militello, L. Rundo, P. Toia, V. Conti, G. Russo, C. Filorizzo, E. Maffei, F. Cademartiri, L. La Grutta, M. Midiri, et al., A semi-automatic approach for epicardial adipose tissue segmentation and quantification on cardiac CT scans, Comput. Biol. Med. 114 (2019) 103424. doi:10.1016/j.compbiomed.2019.103424.

[40] G. Piantadosi, R. Fusco, A. Petrillo, M. Sansone, C. Sansone, LBP-TOP for volume lesion classification in breast DCE-MRI, Lecture Notes in Computer Science 9279 (2015) 647–657. doi:10.1007/978-3-319-23231-7_58.

[41] M. C. Schabel, A unified impulse response model for DCE-MRI, Journal of Magnetic Resonance Imaging 68 (5) (2012) 1632–1646. doi:10.1002/mrm.24162.

[42] P. Soille, Morphological image analysis: principles and applications, 2nd Edition, Principles and Applications, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.

[43] J. M. Feder, E. S. de Paredes, J. P. Hogge, J. J. Wilken, Unusual breast lesions: radiologic-pathologic correlation, Radiographics 19 (Spe-

cial Issue) (2019) S11–S26. doi:10.1148/radiographics.19.suppl_1.g99oc07s11.

[44] L. Rundo, L. Beer, S. Ursprung, P. Martin-Gonzalez, F. Markowetz, J. D. Brenton, M. Crispin-Ortuzar, E. Sala, R. Woitek, Tissue-specific and interpretable sub-segmentation of whole tumour burden on CT images by unsupervised fuzzy clustering, Comput. Biol. Med. 120 (2020) 103751. doi:10.1016/j.compbiomed.2020.103751.

[45] A. Taha, A. Hanbury, Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, BMC Med. Imaging 15 (1) (2015) 29. doi:10.1186/s12880-015-0068-x.

[46] A. Fenster, B. Chiu, Evaluation of segmentation algorithms for medical imaging, in: Proc. Annual International Conference of the Engineering in Medicine and Biology Society, IEEE, 2005, pp. 7186–7189. doi:10.1109/IEMBS.2005.1616166.

[47] Y. Zhang, A review of recent evaluation methods for image segmentation, in: Proc. IEEE International Symposium on Signal Processing and its Applications (ISSPA), Vol. 1, IEEE, 2001, pp. 148–151. doi:10.1109/ISSPA.2001.949797.

[48] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 11 (285-296) (1975) 23–27. doi:10.1109/TSMC.1979.4310076.

[49] W. Huang, A. Tudorica, S. Chui, K. Kemmer, A. Naik, M. Troxell, K. Oh, N. Roy, A. Afzal, M. Holtorf, Variations of dynamic contrast-

enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge, The Cancer Imaging Archive (TCIA)doi:10.7937/k9/tcia.2014.a2n1ixox.

[50] W. Huang, X. Li, Y. Chen, X. Li, M. Chang, M. Oborski, et al., Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge, Trans. Oncol. 7 (2014) 153–166. doi:10.1593/tlo.13838.

[51] L. Rundo, C. Han, J. Zhang, R. Hataya, Y. Nagano, C. Militello, C. Ferretti, A. Nobile, M.S. Tangherloni, M. Gilardi, S. Vitabile, H. Nakayama, G. Mauri, CNN-based prostate zonal segmentation on T2-weighted MR images: a cross-dataset study, in: Neural Approaches to Dynamics of Signal Exchanges, Smart Innovation, Systems and Technologies, Springer, 2020, pp. 269–280. doi:10.1007/978-981-13-8950-4_25.

[52] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Patt. Anal. Mach. Intell. 39 (2017) 2481–2495. doi:10.1109/TPAMI.2016.2644615.

[53] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, Cham., 2015, pp. 234–241. doi:10.1007/978-3-319-24574-4_28.

[54] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics Bull. 1 (6) (80–83) 196–202. doi:10.2307/3001968.

[55] S. Holm, A simple sequentially rejective multiple test procedure, Scand. J. Statist. 6 (2) (1979) 65–70. doi:10.2307/4615733.

[56] A. Gubern-Mérida, R. Martí, J. Melendez, J. L. Hauth, R. M. Mann, N. Karssemeijer, B. Platel, Automated localization of breast cancer in DCE-MRI, Med. Image Anal. 20 (1) (2015) 265–274. doi:10.1016/j.media.2014.12.001.

[57] Y. Zhang, S. Chan, V. Y. Park, K.-T. Chang, S. Mehta, M. J. Kim, et al., Automatic detection and segmentation of breast cancer on MRI using Mask R-CNN trained on non-fat-sat images and tested on fat-sat images, Acad. Radiol. In press. doi:10.1016/j.acra.2020.12.001.

[58] C. Meyer, T. Chenevert, C. Galbán, T. Johnson, D. Hamstra, A. Rehemtulla, B. Ross, Data from RIDER breast MRI, The Cancer Imaging Archive (TCIA)doi:10.7937/K9/TCIA.2015.H1SXNUXL.

[59] M. Heath, K. Bowyer, D. Kopans, W. P. Kegelmeyer, R. Moore, K. Chang, S. Munishkumaran, Current status of the digital database for screening mammography, in: Proc. Fourth International Workshop on Digital Mammography, Kluwer Academic Publishers, 1998, pp. 457–460. doi:10.1007/978-94-011-5318-8_75.

[60] M. Heath, K. Bowyer, D. Kopans, R. Moore, W. P. Kegelmeyer, The digital database for screening mammography, in: Proc. Fifth Interna-

tional Workshop on Digital Mammography, Medical Physics Publishing, M.J. Yaffe, ed, 2001, pp. 212–218.

[61] X. Chen, W. L. Li, Y. L. Zhang, Q. Wu, Y. M. Guo, Z. L. Bai, Meta-analysis of quantitative diffusion-weighted mr imaging in the differential diagnosis of breast lesions, BMC Cancer 10 (1) (2010) 693. doi:10.1186/1471-2407-10-693.

[62] M. Frackiewicz, Z. Koper, H. Palus, D. Borys, K. Psiuk-Maksymowicz, Breast lesion segmentation in DCE-MRI imaging, in: Proc. 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), IEEE, 2018, pp. 308–313. doi:10.1109/SITIS.2018.00054.

[63] M. Fadili, S. Ruan, D. Bloyet, B. Mazoyer, A multistep unsupervised fuzzy clustering analysis of fMRI time series, Human Brain Mapping 10 (4) (2000) 160–178. doi:10.1002/1097-0193(200008)10:4<160::AID-HBM20>3.0.CO;2-U.

[64] H. Izakian, W. Pedrycz, I. Jamal, Fuzzy clustering of time series data using dynamic time warping distance, Eng. Appl. Artif. Intell. 39 (2015) 235–244. doi:10.1016/j.engappai.2014.12.015.

[65] W. Chen, M. L. Giger, U. Bick, G. M. Newstead, Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI, Med. Phys, 33 (8) (2006) 2878–2887. doi:10.1118/1.2210568.