SEGMENTATION AND CLASSIFICATION OF TRIPLE NEGATIVE BREAST CANCERS USING DCE-MRI

Shannon C. Agner, Jun Xu, Hussain Fatakdawala, Shridar Ganesan, Anant Madabhushi *

Rutgers, The State Univeristy of New Jersey
Department of Biomedical Engineering
Piscataway, NJ 08854

Sarah Englander, Mark Rosen, Kathleen Thomas, Mitchell Schnall, Michael Feldman, John Tomaszewski

University of Pennsylvania
Philadelphia, PA 19104

ABSTRACT

Triple-negative (TN) breast cancer has gained much interest recently due to its lack of response to receptor-targeted therapies and its aggressive clinical nature. In this study, we evaluate the ability of a computer-aided diagnosis (CAD) system to not only distinguish benign from malignant lesions on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), but also to quantitatively distinguish triple negative breast cancers from other molecular subtypes of breast cancer. 41 breast lesions (24 malignant, 17 benign) as imaged on DCE-MRI were included in the dataset. Of the 24 malignant cases, 13 were of the TN phenotype. Using the dynamic signal intensity information from the DCE-MRIs, an Expectation Maximization-driven active contours scheme is used to automatically segment the breast lesions. Following quantitative morphological, textural, and kinetic feature extraction, a support vector machine classifier was employed to distinguish (a) benign from malignant lesions and (b) TN from non-TN cancers. In the former case, the classifier yielded an accuracy of 83%, sensitivity of 79%, and specificity of 88%. In distinguishing TN from non-TN cases, the classifier had an accuracy of 92%, sensitivity of 92%, and specificity of 91%. The results suggest that the TN phenotype has distinct and quantifiable signatures on DCE-MRI that will be instrumental in the early detection of this aggressive breast cancer subtype.

Index Terms— triple negative, CAD, breast cancer, kinetic texture curves, image analysis, classification, molecular subtypes

1. INTRODUCTION

Triple negative (TN) breast cancer is a molecular subtype that lacks expression of the estrogen, progesterone, and the HER2 receptors. Because of a lack of receptor expression, targeted therapies are ineffective, and chemotherapy is currently the only treatment available. TN breast cancer is also particularly aggressive and accounts for 12% to 26% of all breast cancers, most often occurring in young and also in African-American patients [1]. Studies examining the pathological phenotype have revealed a heterogeneous group of breast cancers that often present as interval cancers, emerging in the months in between annual breast cancers screenings [2]. Because of its particularly aggressive clinical course and lack of targeted therapies, TN breast cancer is of particular interest to the research community.

*Thanks to the Coulter Foundation, Aresty Research Center, New Jersey Commission on Cancer Research, the National Cancer Institute and the Society for Imaging Informatics in Medicine, and the Life Science Commercialization Fund for providing funding to make this work possible.

Studies have shown that x-ray mammography is likely suboptimal for TN breast cancer screening [1, 3]. Chen et al.[1] concluded that DCE-MRI was significantly more sensitive in detecting TN and basal-like breast cancer phenotypes, and a similar pattern of enhancement, described as rim enhancement, was observed in both subtypes. Although qualitative radiologic descriptors have been developed for small pilot studies in TN breast cancer cases, we are unaware of any attempts to quantitatively describe the TN subtype.

CAD has been used extensively to distinguish malignant from benign breast lesions on MRI, driven by quantitative morphologic, textural, and kinetic descriptors [4, 5]. While several breast MRI CAD systems rely on manual lesion detection and segmentation, semi-automated segmentation schemes have recently been developed. These include approaches based on fuzzy c-means[4], backpropagation neural networks[5], and support vector machines (SVMs)[6]. These methods are based on the assumption that the time-series signal intensities within the lesion are homogeneous in their enhancement patterns. The segmentation algorithms then group the voxels that have similar intensity enhancement characteristics and delineate the regions with uniform malignant-type enhancement as described by Kuhl et al [7].

In this paper, we present a comprehensive CAD system for the discrimination of (a) benign from malignant breast lesions, and (b) TN from non-TN breast lesions. In our study, we work from a similar understanding of the typical lesion enhancement patterns to create an Expectation Maximization (EM)-driven active contour scheme to automatically extract the lesion contour. Quantitative features are then automatically obtained for each lesion. Although we consider previously documented features[4], such as morphology and texture, we also calculate a new DCE-MRI feature called *kinetic texture*, which characterizes spatio-temporal changes in lesion texture. An SVM classifer is used to quantitatively classify the breast lesions in the dataset. We also employ graph embedding, a nonlinear dimensionality reduction technique, to reduce data dimensionality and for visualization of relationships between different breast lesion classes.

The 3 novel components of our CAD system are: (1) Lesion detection and segmentation by combining the time-series signal intensity data with an EM-driven active contour scheme; (2) Automated extraction of over 500 different features for quantitatively characterizing the appearance of different breast cancer subtypes; and (3) Breast lesion classification performed in a hierarchical manner, first distinguishing malignant from benign lesions and then, within the group of malignant lesions, identifying those of the aggressive TN molecular phenotype.

2. LESION DETECTION AND SEGMENTATION

2.1. Data Collection

A total of 41 (24 malignant, 17 benign) breast DCE-MRI studies were collected at the University of Pennsylvania in clinical cases where a screening mammogram demonstrated a lesion suspicious for malignancy. All studies were collected under Institutional Review Board approval, and lesion diagnosis was confirmed by biopsy. Immunohistochemistry was performed on the 24 malignant lesions to measure estrogen receptor, progesterone receptor, and HER2 receptor status. Of these, 13 were determined to be TN cases and 11 non-TN. Sagittal T1 weighted, spoiled gradient echo sequences with fat suppression consisting of one series pre-contrast injection of Gd-DTPA and three to eight series post-contrast injection were acquired (Matrix 384×384, 512×512, or 896×896, slice thickness 3cm). Temporal resolution between post-contrast acquisitions was 45-90 seconds. The lesion region of interest (ROI) was then manually segmented in MRIcro [8] by an attending radiologist with expertise in MR mammography. The radiologist selected a lesion slice most representative of each lesion, and analyses were performed only for that slice of the lesion volume.

2.2. Expectation Maximization (EM)-based Lesion Detection Using Functional Data

For each lesion, the radiologist-selected slice at each timepoint pre- and post-contrast is compiled into a three-dimensional matrix. We define the image scene as $\mathcal{C} = (C, f^t)$, where C is a spatial grid of voxels $c \in C$ and f^t is the associated signal intensity at timepoint $t \in \{0, 1, 2, \ldots, T-1\}$, where T is the number of timepoints in the MRI time-series. A signal enhancement curve is then generated for each voxel based on the signal intensity values at each timepoint. The coefficients of a third order polynomial (model degree determined experimentally) are then obtained from fitting the time-series curve $[f^0(c), f^1(c), f^2(c), \ldots, f^{T-1}(c)]$ in a least-squares sense of each individual $c \in C$ and can be expressed as:

$$f^t(c) = a_3^c t^3 + a_2^c t^2 + a_1^c t + a_0^c \tag{1}$$

The EM algorithm groups all voxels based on a time-series coefficient matrix, $[a_c = (a_3^c, a_2^c, a_1^c, a_0^c), \forall c \in C]$.

Based on Bayes' theorem, the EM algorithm aims to compute the posterior probability, $P(D_k|a_c)$, of each voxel c belonging to class $D_k \in \{1, 2, \ldots, K\}$, given $p(D_k|a_c)$, the prior probability that voxel c belongs to class D_k [9]. In this case, K is the number of Gaussian mixtures, and we empirically set $K = 4$, the dimension of the row vector a_c. The algorithm is run iteratively, comprising two steps: the Expectation step (E-step) and the Maximization step (M-step). The E-step calculates the posterior probability $P(D_k|a_c)$ based on the current parameters of the Gaussian mixture model while the M-step recalculates or updates the model parameters, $\Sigma_k = \{\mu_k, \sigma_k, \tau_k\}$, where μ_k and σ_k are the mean and covariance of each Gaussian component, respectively, and τ_k are mixture coefficients in the Gaussian mixture model. After convergence of EM algorithm, voxel c is assigned to class D_k such that the k chosen corresponds to the class with the highest $P(D_k|a_c)$ for that voxel, c. This class assignment is used in the creation of the binarized scenes described in the next section.

2.3. Geodesic Active Contour-based Lesion Segmentation

After the EM is calculated, it is used as a specific initialization to an active contour model to obtain the final lesion segmentation. In this paper, the Magnetostatic Active Contour (MAC) model as described in [10] is employed to extract lesion boundaries. The MAC model implements a bidirectional force field F generated from a hypothetical magnetostatic interaction between the contour and the object boundary. The force field F drives the contour toward the boundary. Both the boundary and the contour are treated as current carrying loops and the magnetic field from the boundary is computed using the well known Biôt-Savart law. This magnetic field determines the force F acting on the contour, where F is defined over the entire scene $\mathcal{C}$. The level set implementation of the contour as proposed in [10] takes the form:

$$\frac{\partial \phi}{\partial n} = \alpha q(\mathcal{C}) \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) |\nabla \phi| - (1 - \alpha) F(\mathcal{C}) \cdot \nabla \phi, \tag{2}$$

where α is a real constant, $q(\mathcal{C}) = 1/(1+|\nabla \mathcal{C}|)$ and $\nabla(\cdot)$ represents the 2D image gradient $\left(\frac{\partial(\cdot)}{\partial X}, \frac{\partial(\cdot)}{\partial Y} \right)$ along the X and Y axes. We refer the reader to [10] for additional details on the MAC model.

For a given image $\mathcal{C}$, K class likelihood scenes $L_k = (C, \ell_k)$ are constructed where $\ell_k(c) = P(D_k|a_c)$ for each $c \in C$. For each L_k, we obtain the corresponding binarized scene $L_k^B = (C, \ell_k^B)$ where

$$\ell_k^B(c) = \begin{cases} 1, & \text{if } \ell_k = \max\limits_k [P(D_k|a_c)] \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The appropriate scene L_k^B representing the lesion class is manually selected by the user and is used to initialize the active contour. The initialization of the contour is defined as circles centered at centroids of the objects detected in L_k^B via connected component labeling. The lesion contour, $\mathcal{R}$, is then evolved until an empirically determined convergence threshold. Figure 1 illustrates the EM-driven, active contour-based lesion segmentation.

3. FEATURE EXTRACTION AND CLASSIFCATION

3.1. Morphologic Features

The set of boundary points, $\mathcal{R} \subset C$, correspond to the segmented lesion. 2D morphological features are calculated based on Green's theorem, where the centroid coordinates $\bar{c} = (\bar{x}, \bar{y})$ are calculated from $\mathcal{R}$. From the calculated centroid and area values of the lesion, the following features are computed [11]: Area overlap ratio, Average distance ratio, Standard deviation of distance ratio, Variance of distance ratio, Compactness, and Smoothness.

3.2. Pre-contrast (t=0) Textural Features

3.2.1. Gradient Features

Eleven non-steerable gradient features were obtained using Sobel, Kirsch and standard derivative operations. Gabor gradient operators [12] comprising the steerable class of gradient features were defined for every $c \in C$ where $c = (x, y)$,

$$h_u(c) = \frac{1}{\sqrt{2}\psi_X \psi_Y} e^{\frac{-1}{2}\left[\frac{x^2}{\psi_X{}^2} + \frac{y^2}{\psi_Y{}^2}\right]} \cos(2\pi\omega x), \tag{4}$$

where $u \in \{1, 2, \ldots, U_{gab}\}$, ω is the frequency of a sinusoidal plane wave along the $\mathbf{X}$-axis, ψ_X and ψ_Y are the space constraints of the Gaussian envelope along the $\mathbf{X}$ and $\mathbf{Y}$ directions, respectively, and U_{gab} is the number of Gabor filters. Filter orientation, θ, is affected by the coordinate transformations: $x' = z(x\cos\theta + y\sin\theta)$ and $y' = z(-x\sin\theta + y\cos\theta)$, where z is the scaling

(a) (b) (c) (d)

Fig. 1. Segmentation of a fibroadenoma: (a) orginal image; (b) binarized scene showing lesion segmentation obtained via EM; (c) MAC segmentation driven by EM scene (breast boundary in red and lesion in green); and (d) corresponding manual segmentation of the lesion.

factor. Gabor gradient features were calculated at 13 scales ($z \in \{-\frac{\pi}{16}, -\frac{\pi}{8\sqrt{2}}, \ldots, \frac{\pi}{16}\}$), 8 orientations ($\theta \in \{0, \frac{\pi}{8}, \ldots, \frac{7\pi}{8}\}$), and 3 window sizes ($s \in \{3, 5, 8\}$).

3.2.2. Statistical Features

Four first order statistical features (mean, median, standard deviation, and range) for 3 different window sizes were calculated for the gray values of pixels within the sliding window neighborhood $\mathcal{N}_s$, $s \in \{3, 5, 8\}$.

To calculate the second order statistical (Haralick) feature scenes [12], we compute a $G \times G$ co-occurrence matrix $O_{d,c,s}$ associated with $\mathcal{N}_s(c_i)$, where G is the maximum grayscale intensity in C. The value at any location $[g_1, g_2]$ in $O_{d,c,s}$, where $g_1, g_2 \in \{1, 2, \ldots, G\}$, represents the frequency with which two distinct voxels $c_i, c_j \in \mathcal{N}_s(c)$ where $i, j \in \{1, 2, \ldots, |C|\}$ with associated image intensities $f(c_i) = g_1$, $f(c_j) = g_2$ are separated by distance d. A total of 13 Haralick features including contrast energy, contrast inverse moment, contrast average, contrast variance, contast entropy, intensity average, intensity variance, intensity entropy, entropy, energy, correlation, and 2 information measures were extracted at every voxel $c \in C$, based on $O_{d,c,s}$, for $s \in \{3, 5, 7\}$, $d = 1$ and $G \in \{64, 128, 256\}$.

3.3. Kinetic Textural Features

Kinetic texture-based (Gabor, first order statistical, and second order statistical) features are computed at each time point for each voxel in the lesion. Each textural feature's mean value, ρ_u, is plotted over time such that a kinetic texture curve is created, which is analogous to the one created for signal intensity. A third order polynomial, chosen empirically to find the best fit to the texture curves, is fitted to this curve to characterize its shape, defining four associated coefficients. The equation in this case is:

$$\rho_u(t) = r_{u,3}t^3 + r_{u,2}t^2 + r_{u,1}t + r_{u,0}, \quad (5)$$

where u$\in \{1, \ldots, U_{ktext}\}$ and U_{ktext} is the number of kinetic textural attributes, and $[r_{u,3}, r_{u,2}, r_{u,1}, r_{u,0}]$ is the feature vector, describing the kinetic texture feature for each texture feature as a function of time.

3.4. Classification of Breast Lesions

A Support Vector Machine (SVM) classifier was applied to evaluate the ability of each feature class (morphology, texture, kinetic texture and kinetic signal intensity) to classify each lesion as benign or malignant. The SVM was also used to classify the malignant lesions as being TN or non-TN. The features corresponding to each of the feature classes were used as inputs to the classifier individually and in combination. Classifier evaluation was done via a leave-one-out validation scheme [9].

4. RESULTS AND DISCUSSION

4.1. Evaluation of Automated Lesion Segmentation

Our automated lesion segmentation scheme was compared against the expert segmentation via various quantitative performance measures (Table 1: HD-Hausdorff distance (voxels), MAD-Mean Absolute Distance (voxels), FAs-fibroadenomas, Sens.-sensitivity, Spec.-specificity). Figure 1 shows a qualitative comparison of our segmentation scheme against that of an expert.

Table 1: Evaluation of Automated vs. Manual Segmentation

Lesion Type	*HD*	*MAD*	*Sens.(%)*	*Spec.(%)*
Benign, FAs	11.23	8.85	73	99
Malignant, non-TN	9.01	4.41	82	99
Malignant, TN	14.49	7.65	47	99

4.2. Discriminating Benign from Malignant Lesions

Graph embedding (GE) is a nonlinear dimensionality reduction scheme that we apply to the feature vectors associated with each of the lesions and that allows us to visualize biomedical data on a 3-dimensional manifold. We use it here to visualize class separability. The 3 dominant embedding vectors obtained from first order kinetic texture features were able to separate the data into benign and malignant lesion classes via GE (see Figure 2(a)). Note that a linear dimensionality reduction scheme (PCA) was unable to separate the 2 classes as seen in Figure 2(b).

Results in Table 2 show that kinetic texture features outperform morphology, precontrast texture, and kinetic signal intensity features in distinguishing benign from malignant lesions.

Table 2: SVM Evaluation for Benign vs. Malignant Case

Feature Set	*Acc.(%)*	*Sens.(%)*	*Spec.(%)*
Morphological	73	88	53
Pre-contrast Texture	63	90	25
Kinetic Signal Intensity	63	67	59
Kinetic 1st Order Textural	83	79	88
Kinetic 2nd Order Statistical	73	88	53

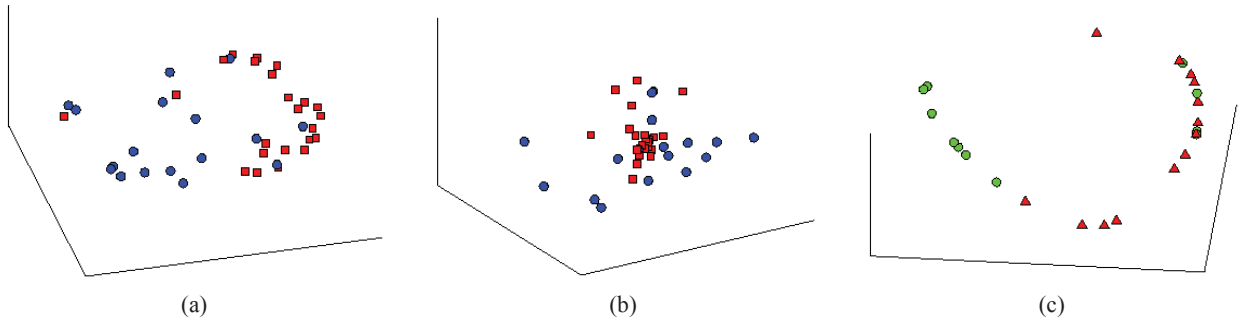

Fig. 2. Embedding plots showing: (a,b) all lesions from first order kinetic texture feature space (malignant: red squares, benign: blue circles) in (a) GE- reduced dimensional space and (b) PCA- reduced dimensional space; (c) malignant lesions only (TN: red triangles, non-TN: green circles) using GE from second order kinetic texture feature space. Note the good separation between lesion classes in the reduced GE space.

4.3. Discriminating Triple from Non-Triple Negative Lesions

Figure 2(c) shows excellent separation between TN and non-TN lesions based on a second order kinetic texture feature in the reduced GE space. Table 3 also shows that second order kinetic texture features performed best at TN versus non-TN lesion classification.

Table 3: SVM Evaluation for TN vs. non-TN Case

Feature Set	*Acc.(%)*	*Sens.(%)*	*Spec.(%)*
Morphological	58	56	9
Pre-contrast Texture	71	92	45
Kinetic Signal Intensity	58	77	36
Kinetic 1st Order Textural	83	77	91
Kinetic 2nd Order Statistical	92	92	91

5. CONCLUDING REMARKS

In this paper, we presented an automated segmentation and classification scheme to distinguish between (a) benign and malignant and (b) TN and non-TN breast lesions on DCE-MRI. The primary novel contributions of this work are:

- An EM-driven active contour lesion segmentation scheme driven by kinetic information; and
- Demonstration of quantitative radiologic phenotypic signatures of breast cancer that identify triple negative breast cancers on DCE-MRI with a relatively high degree of accuracy, sensitivity, and specificity.

For the classification of the TN datasets, high accuracy values indicate that these tumors may have a specific, quantitative radiologic phenotype, unreported thus far in the literature. Consistent with qualitative studies, morphology was not a discriminating feature for TN lesions [1], yet second order kinetic texture features performed with a high degree of accuracy, outperforming current standard methods of measuring kinetic signal intensity changes. While this result is limited by the size of the dataset, it is still a promising one, giving insight into the nature of the triple negative molecular phenotype, and further study with a larger dataset is warranted. Additional avenues for future research will include: (a) integrating kinetic texture and morphological features to construct a more accurate meta-classifier for identifying aggressive molecular subtypes of breast cancer on DCE-MRI and (b) further refining our annotated lesion segmentation scheme.

6. REFERENCES

[1] J.-H. Chen et al., "Triple-negative breast cancer: Mri features in 29 patients.," *Ann Oncol*, vol. 18, no. 12, pp. 2042–2043, Dec 2007.

[2] N.C. Turner and J.S. Reis-Filho, "Basal-like breast cancer and the brca1 phenotype," *Oncogene*, vol. 25, no. 43, pp. 5846–5853, 2006.

[3] J. Veltman et al., "Breast tumor characteristics of brca1 and brca2 gene mutation carriers on mri.," *Eur Radiol*, vol. 18, no. 5, pp. 931–938, May 2008.

[4] W. Chen et al., "Automatic identification and classification of characteristic kinetic curves of breast lesion on dce-mri," *Medical Physics*, vol. 33, no. 8, pp. 2878–2887, 2006.

[5] L.A. Meinel et al., "Breast mri lesion classification: Improving performance of human readers with a backpropagation neural network computer-aided (cad) system," *JMRI*, vol. 25, no. 1, pp. 89–95, 2007.

[6] J. Levman et al., "Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines," *IEEE Trans. on Med. Imag.*, vol. 27, no. 5, pp. 688–696, 2008.

[7] C.K. Kuhl et al., "Dynamic breast mr imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions," *Radiology*, vol. 211, pp. 101–110, 1999.

[8] C. Rorden and M. Brett, "Stereotaxic display of brain lesions.," *Behavioural Neurology*, vol. 12, pp. 191–200, 2000.

[9] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer–Verlag, 2006.

[10] X. Xie and M. Mirmehdi, "Mac: Magnetostatic active contour model," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 4, pp. 632–646, April 2008.

[11] S. Naik et al., "Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information," 2007, MIAAB.

[12] A. Madabhushi et al., "Automated detection of prostate adenocarcinoma from high-resolution ex vivo mri," *IEEE Trans. on Med. Imag.*, vol. 24, no. 12, pp. 1611–1625, 2005.