Novel Kinetic Texture Features for Breast Lesion Classification on Dynamic Contrast Enhanced (DCE) MRI

Shannon C. Agner,[a] Salil Soman,[b] Edward Libfeld,[b] Margie McDonald,[b] Mark A. Rosen,[c] Mitchell D. Schnall,[c] Deanna Chin,[b] John Nosher,[b] Anant Madabhushi[a]

[a]Department of Biomedical Engineering, Rutgers University, 599 Taylor Road, Piscataway, NJ, USA 08854;
[b]Department of Radiology, Robert Wood Johnson Medical School, New Brunswick, NJ, USA 08903;
[c]Department of Radiology, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, USA, 19104

ABSTRACT

Dynamic contrast enhanced (DCE) MRI has emerged as a promising new imaging modality for breast cancer screening. Currently, radiologists evaluate breast lesions based on qualitative description of lesion morphology and contrast uptake profiles. However, the subjectivity associated with qualitative description of breast lesions on DCE-MRI introduces a high degree of inter-observer variability. In addition, the high sensitivity of MRI results in poor specificity and thus a high rate of biopsies on benign lesions. Computer aided diagnosis (CAD) methods have been previously proposed for breast MRI, but research in the field is far from comprehensive. Most previous work has focused on either quantifying morphological attributes used by radiologists, characterizing lesion intensity profiles which reflect uptake of contrast dye, or characterizing lesion texture. While there has been much debate on the relative importance of the different classes of features (e.g., morphological, textural, and kinetic), comprehensive quantitative comparisons between the different lesion attributes have been rare. In addition, although kinetic signal enhancement curves may give insight into the underlying physiology of the lesion, signal intensity is susceptible to MRI acquisition artifacts such as bias field and intensity non-standardness. In this paper, we introduce a novel lesion feature that we call the *kinetic texture feature*, which we demonstrate to be superior compared to the lesion intensity profile dynamics. Our hypothesis is that since lesion intensity is susceptible to artifacts, lesion texture changes better reflect lesion class (benign or malignant). In this paper, we quantitatively demonstrate the superiority of kinetic texture features for lesion classification on 18 breast DCE-MRI studies compared to over 500 different morphological, kinetic intensity, and lesion texture features. In conjunction with linear and non-linear dimensionality reduction methods, a support vector machine (SVM) classifier yielded classification accuracy and positive predictive values of 78% and 86% with kinetic texture features compared to 78% and 73% with morphological features and 72% and 83% with textural features, respectively.

Keywords: Methods: classification and classifier design, computer-aided diagnosis; Modalities: magnetic resonance; Diagnostic Task: diagnosis

1. INTRODUCTION

Breast cancer is the second leading cause of cancer-related deaths in women. According to the American Cancer Society, there were approximately 178,000 new cases of invasive breast cancer in 2007 in the United States alone. However, early detection, diagnosis, and treatment of breast cancer can reduce the chances of death by 15-30%. X-ray mammography continues to be the primary modality for breast cancer screening, but the American Cancer Society revised their MRI screening guidelines in 2007 to specifically include women with a 20 to 25 percent increased lifetime risk, including family history of breast and ovarian cancers or a history of prior treatment for Hodgkin's disease.[1] These changes have increased the frequency with which MRI is and will

Contact: Anant Madabhushi E-mail: anantm@rci.rutgers.edu, Telephone: 1 732 445 4500 x6213

Medical Imaging 2008: Computer-Aided Diagnosis, edited by Maryellen L. Giger, Nico Karssemeijer
Proc. of SPIE Vol. 6915, 69152C, (2008) · 1605-7422/08/$18 · doi: 10.1117/12.770920

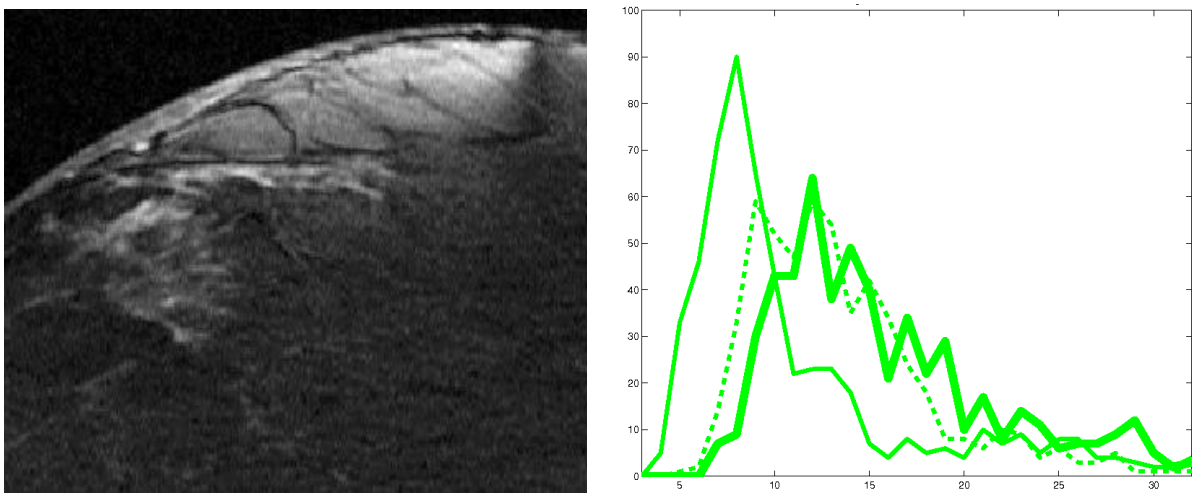
be used for screening of breast cancer. To meet the needs of increased rates of MR imaging for breast cancer screening and diagnosis, a breast MRI lexicon was developed by adapting the existing Breast Imaging and Data Reporting System (BI-RADS) lexicon that is used by radiologists for qualitative descriptions of breast lesions on mammography, but only moderate inter-observer agreement was found.[2] Clearly, quantifiable, accurate, and reliable methods for assessing breast lesions using MRI are necessary.

In field of breast cancer diagnosis as well as in other disease processes, CAD systems can be used to highlight regions of interest on the image, quantify lesion characteristics, and render a diagnosis regarding lesion type. In accordance with breast MRI being a relatively new component of clinical breast cancer diagnosis, the CAD literature in the field of breast MRI is sparse compared to that of mammography or breast ultrasound. Radiologists agree that both morphological and kinetic attributes are important for distinguishing benign from malignant breast lesions on MRI.[2–4] Morphologic attributes aim to quantify lesion margin characteristics typically described by radiologists using the breast MRI lexicon.[2] Although CAD schemes focus on morphology and kinetics in combination as well,[5,6] they rarely incorporate texture, which has been shown to be useful in classification of lesions on various imaging modalities.[7–9] Furthermore, Chen et al.[8] have also previously described the successful use of textural feature sets in breast MRI.

Recently, attempts have been made to quantify classify breast lesions in terms of dynamic contrast enhancement. The two most widely used methods for quantifying such data are the pharmacokinetic model approach[10] and the three time point empirical model.[11] The pharmacokinetic model is widely accepted in most contrast enhanced imaging disciplines and is based on a mathematical model that describes the velocity of blood flow as it traverses the arterial and capillary circulations and subsequent diffusion through the breast tissue.[10] The theory behind this is to physiologically describe the vascularity and tissue consistency of the patient's lesion and to compare those characteristics to other lesions, both benign and malignant. Signal intensity within the lesion is derived from the DCE-MRI images and transformed from a signal intensity space into a concentration space, which more accurately describes the physiology observed in the images. In the second method, semiquantitative measures are used to evaluate signal intensity changes, such as time to peak intensity, rate of washout, and signal enhancement ratio, which is usually defined as the ratio of contrast uptake to washout. These measures capture those characteristics of the kinetic enhancement curves that the radiologists qualitatively describe in the diagnostic radiology report.

Both methods of evaluating contrast enhancement are reasonably effective, and as a result, both methods continue to be subjects of much research.[4,11] However, the inherent nonlinearity of the MRI data and the image non-standardness created by intensity drift that occurs during image acquisition introduce inaccuracies to a numeric representation of any MRI based purely on signal intensity.[12,13] Nonlinearity of MRI images is manifest in the appearance of areas of hyperintensity in regions of the tissue proximal to the receiver coil used for image acquisition, and the signal hyperintensity tends to progressively decrease as the distance from the coil increases. This phenomenon can be observed in Figure 1(a). Note that the enhancement in the upper right corner of the image is not solely due to contrast dye enhancement, but is also a result of bias field inhomogeneity. Image nonstandardness occurs due to MRI scanner signal intensity optimization for each image acquisition. This results in varying signal intensity scales for image acquisition, even within the same MRI study. Figure 1(b) shows a plot of intensity histograms of non-enhancing breast tissue from the same MRI slice for the three postcontrast time points within the same DCE-MRI study. Drift is clearly observable in the lack of alignment between intensity histograms. Despite being well documented and even observed in this dataset, these artifacts are not typically corrected prior to analysis of kinetic curves for lesion classification. For these reasons, an alternate representation of the intensity space may be more robust to some of the artifacts to which signal intensity is typically sensitive, thereby yielding more consistent representations across the contrast enhancement time series and improved discriminability in lesion classification.

In addition to its advantages over signal intensity measures, texture dynamics (second order intensity statistics) may still be more discriminatory than first order intensity statistics (i.e., signal intensity), even if the aforementioned artifacts were corrected in the signal enhancement images. This theory is supported by the fact that previous work has shown texture features (e.g., steerable first order statistics such as Gabor filters, non-steerable first order statistics such as grey level features, and second order statistics such as Haralick features) to not only be useful in discrimination of lesion types in other pathologies such as prostate cancer,[14] but also in



(a) (b)

Figure 1. Illustrations of MRI artifacts. (a)Image showing the result of bias field. Note the high intensity at the top of the breast, where the coil interfaces with the skin. (b) Histogram illustrating the phenomenon of drift. This shows the intensity histograms of the same image during the postcontrast time points of a nonenhancing section of the tissue. One would expect the signal intensity to remain similar in distribution, but a shift is seen in the histograms.

image registration, where the above artifacts are often a great obstacle to overcome in MRI image registration.[15] As such, the primary motivations and contributions of this paper are the following:

- Novel kinetic texture features that are more robust to MR intensity artifacts and therefore have improved discriminability between breast lesion types compared to signal intensity profile features; and
- A quantitative comparison between kinetic texture features and over 500 morphological, signal intensity, and lesion texture features to determine the relative importance of individual feature types in breast lesion classification.

The remainder of this paper is organized as follows: In Section 2, we provide a description of the data used in this study and the notational convention adopted. In Section 3, we describe our feature extraction schemes, and in Section 4, we provide details on the use of dimensionality reduction and a classifier ensemble to determine the relative importance of the lesion attributes. Quantitative and qualitative results are presented in Section 5 with concluding remarks and future directions presented in Section 6.

2. DATA DESCRIPTION

T1 sagittal images were obtained at one time point pre-contrast and three time points postcontrast at both the Hospital at the University of Pennsylvania and Robert Wood Johnson University Hospital according to each hospitals' respective clinical breast cancer screening MRI protocol for 18 patients with mammograms showing lesions suspicious for cancer. Lesions were identified on MRI by expert radiologists and diagnosis was confirmed by histopathology. The benign lesion class contains 9 lesions (4 fibroadenomas, 5 fibrocystic changes), and the malignant lesion class contains 9 invasive ductal carcinomas.

3. FEATURE EXTRACTION

In this section, the methods for extracting the following feature types is described:

1. Kinetic Signal Intensity features and Kinetic Textural features (both based on the T1 DCE-MRI);
2. Precontrast Textural Features (based on the T1 image before contrast injection); and

3. Morphological Features (based on the contour created by expert manual segmentation).

Expert manual segmentations were obtained for each lesion on each slice of the MRI volume of a lesion. Over 350 3D texture feature scenes, corresponding to three different texture classes were extracted from each MRI scene. These feature representations were chosen since they have been demonstrated to be able to discriminate between the cancer and non-cancer classes.[16] We define a 3D MR scene $\mathcal{C} = (C, f)$ where C is a set of spatial locations $c \in C$, and $f(c)$ is a function that assigns an intensity value to every $c \in C$. The pre-contrast scene is denoted as $\mathcal{C}^0 = (C, f^0)$ and subsequent post-contrast scenes as $\mathcal{C}^1, \mathcal{C}^2, \mathcal{C}^3$. Coregistration of all scenes $\mathcal{C}^t, t \in \{0, \ldots, 3\}$ for each study were verified via the COFEMI scheme.[15]

3.1 Dynamic Contrast-Enhancement Features

3.1.1 Kinetic Signal Intensity Features

These features are computed as the coefficients $[a_3, a_2, a_1, a_0]$ of a third order polynomial obtained from fitting a curve in a least-squares sense to the signal intensity contrast enhancement curves. Hence, for each c in $\mathcal{C}^0$, $\mathcal{C}^1, \mathcal{C}^2, \mathcal{C}^3$, a third order curve is fitted as:

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

where $t \in \{0, \ldots, 3\}$ and $c \in \mathcal{C}$. Note that $f^t(c)$ represents the signal intensity at each spatial location, $c \in \mathcal{C}$, across the pre- and post-contrast MRI scenes.

3.1.2 Kinetic Texture Features

The Gabor, first order statistical, and second order statistical features are computed at each time point for each voxel in the lesion. Then, each modal value 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 is then fitted to this curve to characterize its shape, defining four associated coefficients, and the equation in this case is:

$$\gamma_u(t) = p_{u,3} t^3 + p_{u,2} t^2 + p_{u,1} t + p_{u,0}. \tag{2}$$

where $\gamma_u(t) = \underset{d \in L}{\text{MODE}} [f_u^t(d)]$ is the set of pixels corresponding to the lesion area and where MODE refers to the modal operator. $[p_{u,3}, p_{u,2}, p_{u,1}, p_{u,0}]$, is the feature vector, describing the texture feature as a function of time. The Precontrast Textural Features are defined as the modal texture values for each feature described above before contrast injection (t=0). Figure 2 shows examples of the feature maps that would be typically used to calculate the kinetic coefficients for a malignant lesion, a fibroadenoma, and fibrocystic changes. In Figure 2, one can visually appreciate the subtle differences in texture between each time point that would allow quantification of differences between various lesion types.

A. Gradient Features Gradient features are calculated using steerable and non-steerable linear gradient operators. Eleven non-steerable gradient features were obtained using Sobel, Kirsch and standard derivative operations. Gabor gradient operators[14] comprising the steerable class of gradient calculations were defined for every location $c \in C$ where $c = (x, y, z)$,

$$f_u(c) = \frac{1}{2^{\frac{3}{2}} \sigma_X \sigma_Y \sigma_Z} e^{\frac{-1}{2}\left[\frac{x^2}{{\sigma_X}^2} + \frac{y^2}{{\sigma_Y}^2} + \frac{z^2}{{\sigma_Z}^2}\right]} \cos(2\pi\omega x), \tag{3}$$

where ω is the frequency of a sinusoidal plane wave along the **X**-axis, and σ_X, σ_Y, and σ_Z are the space constraints of the Gaussian envelope along the **X**, **Y**, and **Z** directions respectively. The orientation of the filter,θ, is affected by the coordinate transformations: $x' = r(x\cos\theta + y\sin\theta)$, $y' = r(-x\sin\theta + y\cos\theta)$ and $z' = r(z)$, where r is the scaling factor. These were computed within the sliding square window neighborhood, $\mathcal{N}_\kappa(c)$ associated with every voxel location $c \in C$ and centered on c; $c \notin \mathcal{N}_\kappa(c)$. Gabor gradient features were calculated at 13 scales ($r \in \{-\frac{\pi}{16}, -\frac{\pi}{8\sqrt{2}}, \ldots, \frac{\pi}{16}\}$), 8 orientations ($\theta \in \{0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}\}$) and 3 window sizes ($\kappa \in \{3, 5, 8\}$).

B. First Order Statistical Features Four first order statistical features for 3 different window sizes were calculated. They included the mean, median, standard deviation, and range for the gray values of pixels within the sliding window neighborhood $\mathcal{N}_\kappa$, $\kappa \in \{3, 5, 8\}$.

C. Second Order Statistical Features To calculate the second order statistical (Haralick) feature scenes, we compute a $G \times G$ co-occurrence matrix $P_{d,c,\kappa}$ associated with $\mathcal{N}_\kappa(c_i)$, where G is the maximum gray scale intensity in $\mathcal{C}$. The value at any location $[g_1, g_2]$ in $P_{d,c,\kappa}$, where $g_1, g_2 \in \{1, \ldots, M\}$, represents the frequency with which two distinct voxels $c_i, c_j \in \mathcal{N}_\kappa(c), i, j \in \{1, \ldots, |C|\}$ 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 energy, entropy, inertia, contrast, correlation, sum average, sum variance, sum entropy, difference average, difference variance, difference entropy, local homogeneity and average deviation were extracted at every voxel $c \in C$, based on $P_{d,c,\kappa}$, for $\kappa \in \{3, 5, 7\}$, $d = 1$ and $G \in \{64, 128, 256\}$.

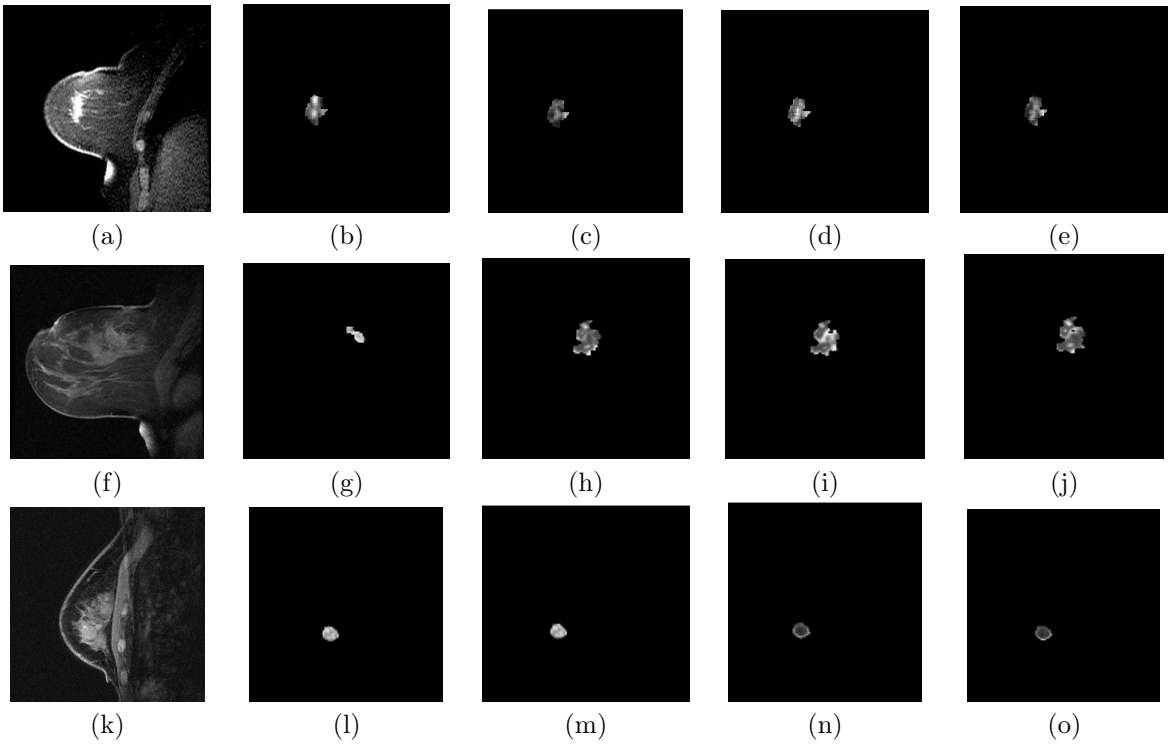


Figure 2. Sample images showing various morphologies as defined by the segmentation boundaries. An example of the texture kinetics of (a-e) a malignant lesion; (f-j) a lesion showing fibrocystic changes, and (k-o) a fibroadenoma are shown. The lesions show (from left to right) the full precontrast image (a,f,k) and the second order energy texture maps for the precontrast (b,g,l) and subsequent 3 postcontrast time points (c-e,h-j,m-o). Note the changes in texture at each time point within a single lesion.

3.2 Morphological Features

The segmentation performed by the radiologist defines the boundary of the lesion, where the boundary points, $\mathcal{B}$, are a subset of the voxels contained in the lesion $L \subset C$. 2D morphological features that calculate the following based on Green's theorem where the centroid coordinates $\bar{c} = (\bar{x}, \bar{y})$ are calculated using the boundary, $\mathcal{B}$, of the lesion as defined by the radiologist. Figure 3 shows examples of 3 different lesion boundaries that are used in this analysis. From the centroid and area values, the following features are computed:[17]

1. *Area overlap ratio*: The area enclosed by $\mathcal{B}$ divided by the area of the smallest circle enclosing $\mathcal{B}$.

2. *Average distance ratio*: Ratio of average distance to maximum distance from the centroid of $\mathcal{B}$:

$$\frac{\frac{1}{|\mathcal{B}|}\sum_{d\in\mathcal{B}}\|\bar{c}-d\|}{\max_{d\in\mathcal{B}}[\|\bar{c}-d\|]} \tag{4}$$

where $\|\cdot\|$ is the L_2 norm.

3. *Standard deviation of distance ratio*: For each $d \in \mathcal{B}$, compute the distance ratio:

$$f_{d,\gamma}(d) = \frac{\|d-\bar{c}\|}{\max_{d\in\mathcal{B}}[\|d-\bar{c}\|]}. \tag{5}$$

Standard deviation of distance ratio is then defined as $\underset{d\in\mathcal{B}}{\mathrm{STD}}\,[f_{d,\gamma}(d)]$, where STD refers to the standard deviation function.

4. *Variance of distance ratio*: $[\underset{d\in\mathcal{B}}{\mathrm{STD}}\, f_{d,\gamma}(d)]^2$,

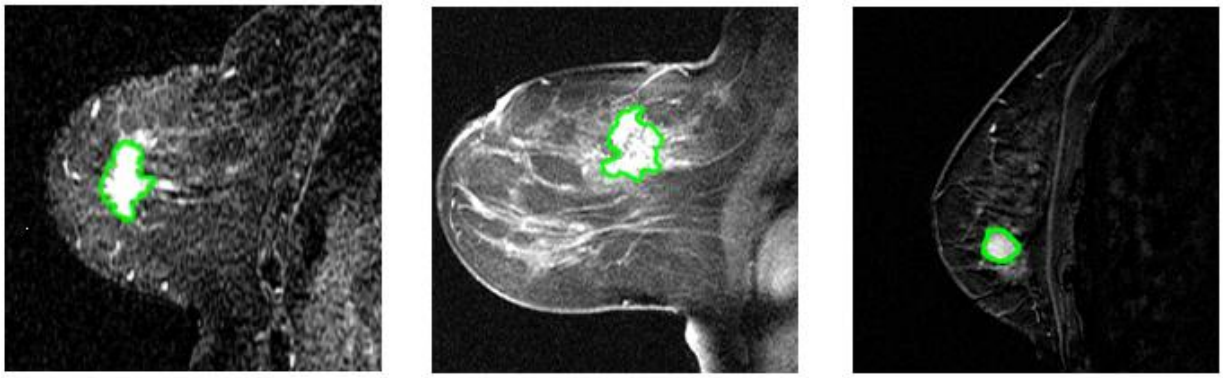
5. *Perimeter ratio*: The ratio of estimated length of $\mathcal{B}$ to the true length $\mathcal{B}$. Estimated length is computed using linear interpolation between 5-10 points, depending on the number of points on $\mathcal{B}$, that are sampled at equal intervals from $\mathcal{B}$. True length is computed using all points lying on $\mathcal{B}$.

6. *Compactness*: The true length of $\mathcal{B}$ squared divided by the area enclosed by $\mathcal{B}$.

7. *Smoothness*: For points d_{v-1}, d_v, and d_{v+1} on $\mathcal{B}$, where d_{v-1} is immediately adjacent and counter-clockwise to point d_v and point d_{v+1} is immediately adjacent and clockwise from point d_v,

$$S_{d_v} = |\mathbf{d}(d_v,\bar{c}) - \frac{[\mathbf{d}(d_{v-1},\bar{c}) + \mathbf{d}(d_{v+1},\bar{c})]}{2}|, \tag{6}$$

where $\mathbf{d}(d_v,\bar{c})$ is the Euclidean distance between d_v and $\bar{c}$. Smoothness is defined as $\sum_{d_v\in\mathcal{B}} S_{d_v}$.



(a) (b) (c)

Figure 3. Sample images showing various morphologies as defined by the segmentation boundaries. Figure 3(a) shows a malignant lesion; figure 3(b) shows a lesion with fibrocystic changes, and figure 3(c) shows a fibroadenoma. Note the variation in the contour between the three lesions.

4. FEATURE RANKING VIA SUPPORT VECTOR MACHINE CLASSIFIER

4.1 Dimensionality Reduction

For each of the 18 lesions in the dataset, there is an associated 507-dimensional feature vector. Since the number of features is much greater than the number of samples in the dataset, the dimensionality of the feature space must be reduced in order to perform classification. The dimensionality reduction methods used were Principal Component Analysis (PCA) and Locally Linear Embedding (LLE).

A. Principal Components Analysis (PCA) PCA[18] is widely used to visualize high-dimensional data and discern relationships by finding orthogonal axes that contain the greatest amount of variance in the data. These orthogonal eigenvectors corresponding to the largest eigenvalues are called 'principal components' and are obtained in the following manner. Each data point $c \in C$ is first centered by subtracting the mean of all the features for each observation c from its original feature value $f_u(c)$ as shown in Equation 7.

$$f_u(c) = f_u(c) - \frac{1}{n}\sum_{i=1}^{n} f_u(c), \tag{7}$$

for $u \in \{1, 2, ..., M\}$. From feature values $\tilde{f}_u(c)$ for each $c \in C$, a new n x M matrix $\mathcal{Y}$ is constructed. The matrix $\mathcal{Y}$ is then decomposed into corresponding singular values as shown in Equation 8.

$$\mathcal{Y} = U\lambda V^T, \tag{8}$$

where via singular value decomposition, an n x n diagonal matrix λ containing the eigenvalues of the principal components and an m x n left singular matrix U and M x n matrix V are obtained. The eigenvalues in λ represent the amount of variance for each eigenvector g_w^{PCA}, $w \in \{1, 2, ..., m\}$ in matrix V^T and are used to rank the corresponding eigenvectors in the order of greatest variance. Thus, the first m eigenvectors are obtained, as they contain the most variance in the data while the remaining eigenvectors are discarded so each data sample $c \in C$ is now described by an m-dimensional embedding vector $G^{PCA}(c)$.

B. Locally Linear Embedding (LLE) LLE[19] operates by assuming that objects in a neighborhood of a feature space are locally linear. Consider the set of feature vectors $\mathcal{F} = \{\mathbf{F}(c_1), \mathbf{F}(c_2), \ldots, \mathbf{F}(c_n)\}$, $n = |C|$, where $\mathbf{F}(\mathbf{c_i}) = [\mathbf{f_u}(\mathbf{c_i})|\mathbf{u} \in \{\mathbf{1}, \ldots, |\mathbf{C}|\}$. We wish to map the set $\mathcal{F}$ to the set $\mathcal{X} = \{X_{LLE}(c_1), X_{LLE}(c_2), \ldots, X_{LLE}(c_n)\}$ of embedding co-ordinates. For all objects $c \in C$, LLE maps the feature vector $\mathbf{F}(c)$ to the embedding vector $X_{LLE}(c)$. Let $\{c_{\eta_i(1)}, \ldots, c_{\eta_i(k)}\}$ be the k nearest neighbors of c_i where $\eta_i(k)$ is the index of the location of the k^{th} neighbor of c_i in C, where $i \in \{1, \ldots, |C|\}$. The feature vector $\mathbf{F}(c_i)$ and its k nearest neighbors (kNN), $\{\mathbf{F}(c_{\eta_i(1)}), \mathbf{F}(c_{\eta_i(2)}), \ldots, \mathbf{F}(c_{\eta_i(k)})\}$ are assumed to lie on a patch of the manifold that is local linearly, allowing us to use Euclidean distances between the neighbors. Each $\mathbf{F}(c_i)$ can then be approximated by a weighted sum of its kNN. The optimal reconstruction weights are given by the sparse matrix W_{LLE} (subject to the constraint $\sum_j W_{LLE}(i,j) = 1$) that minimizes

$$S_1(W_{LLE}) = \sum_{i=1}^{|C|} \left\| \mathbf{F}(c_i) - \sum_{j=1}^{k} W_{LLE}(i, \eta_i(j)) \mathbf{F}(c_{\eta_i(j)}) \right\|_2 . \tag{9}$$

Having determined the weighting matrix W_{LLE}, the next step is to find a low-dimensional representation of the points in $\mathcal{F}$ that preserves this weighting. Thus, for each $\mathbf{F}(c_i)$ approximated as the weighted combination of its kNN, its projection $X_{LLE}(c_i)$ will be the weighted combination of the projections of these same kNN. The optimal $\mathcal{X}_{LLE}$ in the least squares sense minimizes

$$S_2(\mathcal{X}_{LLE}) = \sum_{i=1}^{|C|} \left\| X_{LLE}(c_i) - \sum_{j=1}^{|C|} W_{LLE}(i,j) X_{LLE}(c_j) \right\|_2 = \mathrm{tr}\left(\mathcal{X}_{LLE} L \mathcal{X}_{LLE}^{\mathsf{T}}\right), \tag{10}$$

where tr is the trace operator, $\mathcal{X}_{LLE} = \{X_{LLE}(c_1), X_{LLE}(c_2), \ldots, X_{LLE}(c_n)\}$, $E = (I - \mathcal{X}_{LLE})\left(I - \mathcal{X}_{LLE}^{\mathsf{T}}\right)$ and I is the identity matrix. The minimization of (10) subject to the constraint $\mathcal{X}_{LLE}\mathcal{X}_{LLE}^{\mathsf{T}} = I$ (a normalization constraint that prevents the solution $\mathcal{X}_{LLE} \equiv \mathbf{0}$) is an eigenvalue problem whose solutions are the eigenvectors of the Laplacian matrix E. Since the rank of E is n–1, the first eigenvector is ignored and the second smallest eigenvector represents the best one-dimensional projection of all the samples. The best two-dimensional projection is given by the eigenvectors with the second and third smallest eigenvalues, and so forth.

4.2 Support Vector Machine (SVM)

SVM methods were 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 classifier contained two stages: (a) training and (b) testing. The low dimensional eigenvalues obtained by individually projecting features corresponding to each of the 4 classes were used as inputs to the classifier. From the training data, a hyperplane is created in the eigenspace obtained via PCA or LLE that optimally separates the data into benign and malignant lesion classes.

Given a set of labeled training data from two distinct classes, an SVM will project the data into a high dimensional space constructed by a kernel function, Ψ, operating on the training data. Testing data are then classified according to where they fall in relation to the hyperplane when operated on by the same kernel function Ψ. The objects of each class that lie closest to this hyperplane are the "support vectors." The general form of the SVM classifier is:[20]

$$\mathcal{V}(\mathbf{x}) = \sum_{\tau=1}^{N_s} \xi_\tau y_\tau \Psi(\mathbf{x}, \mathbf{x}_\tau) + \mathbf{b}, \tag{11}$$

where $\mathbf{x}$ is the input training data, $\mathbf{x}_\tau$, $\tau \in \{1, 2, \cdots, N_s\}$ denotes the support vectors, $y \in \{-1, 1\}$ as the training labels, $\Psi(\cdot, \cdot)$ is a positive, definite, symmetric kernel function, $\mathbf{b}$ is a bias obtained from the training set to maximize the distance between the support vectors, and ξ is a model parameter chosen to maximize the objective function:

$$\Delta(\xi) = \sum_{\tau=1}^{N_s} \xi_\tau - \frac{1}{2} \sum_{\tau,\rho=1}^{N_s} \xi_\tau \xi_\rho y_\tau y_\rho \Psi. \tag{12}$$

The kernel function $\Psi(\cdot, \cdot)$ defines the nature of the decision hyperplane. A common kernel called the radial basis function (RBF) was used in this study. The parameters ρ and $\mathbf{b}$ are found through empirical training and testing of the classifier.

The efficacy of the feature set is then evaluated based on the ability of the feature set to correctly classify each lesion, using the remaining lesions in the dataset as the training set.[16,21] Both accuracy of classification and positive predictive value (PPV) are calculated for each feature set using both PCA and LLE. Accuracy is defined as (tp + tn)/(tp+tn+fp+fn) and positive predictive value (PPV) is defined as tp/(tp+fp), where tp is the number of true positives, tn is the number of true negatives, fp is the number of false positives, and fn is the number of false negatives.

5. RESULTS

5.1 Qualitative Results

5.1.1 Comparison of Kinetic Curves

For comparison purposes, the DCE Intensity and the Gradient Texture feature from the First Order Statistical Texture Feature set were plotted as a function of time for five of the nine benign lesions and five of the nine malignant lesions (see Figures 4(a), 4(d) and Figures 4(b), 4(e), respectively). The same datasets were plotted in both instances. The benign datasets showed greater similarity within the benign lesion class on the Gradient Texture plot compared to the DCE Intensity plot. This suggests that the texture feature kinetics may reveal intraclass similarity better than DCE signal intensity kinetics. In addition, when comparing benign lesions and malignant lesions on the same plot (Figure 4(c), 4(f)), Gradient Texture appears to reveal a greater difference between benign and malignant lesions while simultaneously revealing the same increase in intra-class similarity.

This result has promising implications in lesions classification, particularly in clustering schema used for classification. The clustering trend was also present when the dimensionality reduction methods, PCA and LLE, were applied to the various feature sets and then in combination.

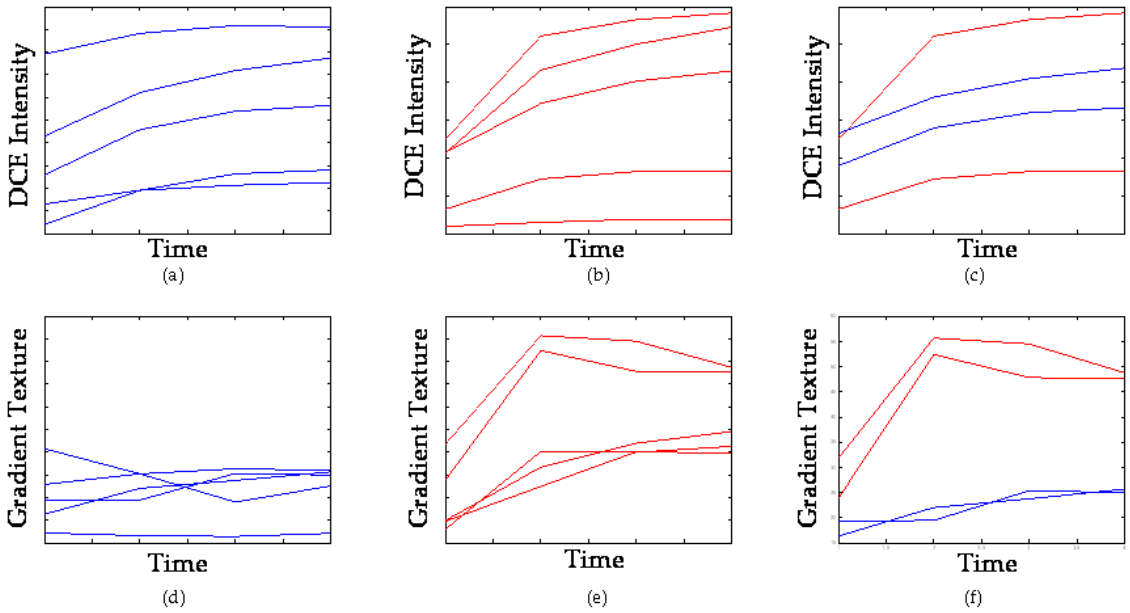


Figure 4. Signal intensity (a-c) and kinetic texture feature (d-f) plotted over the course of contrast administration. Time=0 is precontrast; progression along the Time axis denotes postcontrast time points.

5.1.2 Comparison of Low-dimensional Embedding Plots

Figure 5 shows projections of the reduced feature space for many of the feature types studied. The X, Y, and Z axes of the three-dimensional plots each represent the eigenvalues of the reduced feature space. Because the dimensionality reduction methods are used to preserve relationships between data in the 507-dimensional feature space, proximity of two objects in the reduced feature space represents similarity in the unreduced feature space, so two lesions that are close to one another on the three dimensional plot have similar features in the unreduced feature space. As we observed in the kinetic data plots, greater separation between malignant and benign lesions and greater clustering of within-class lesions is observed in the morphology plot (Figure 5(a)) and the two kinetic feature types shown (Figures 5(c) and 5(d)). This process was also repeated using LLE (not shown).

5.2 Quantitative Results

Table 1 shows the accuracy and PPV results obtained with a support vector machine classifier for morphological, kinetic (signal intensity), kinetic textural, steerable filters (Gabor), first order statistical texture (Grey), second order statistical texture(Haralick), and all features (as a concatenated feature vector). By using PCA, morphological and kinetic textural features performed best, both with an accuracy of 78%, and positive predicitve values (PPV) of 73% and 86%, respectively. When LLE was used, baseline second order statistical (Haralick) features and kinetic textural features performed best, both with accuracy values of 72% and PPV values of 83%.

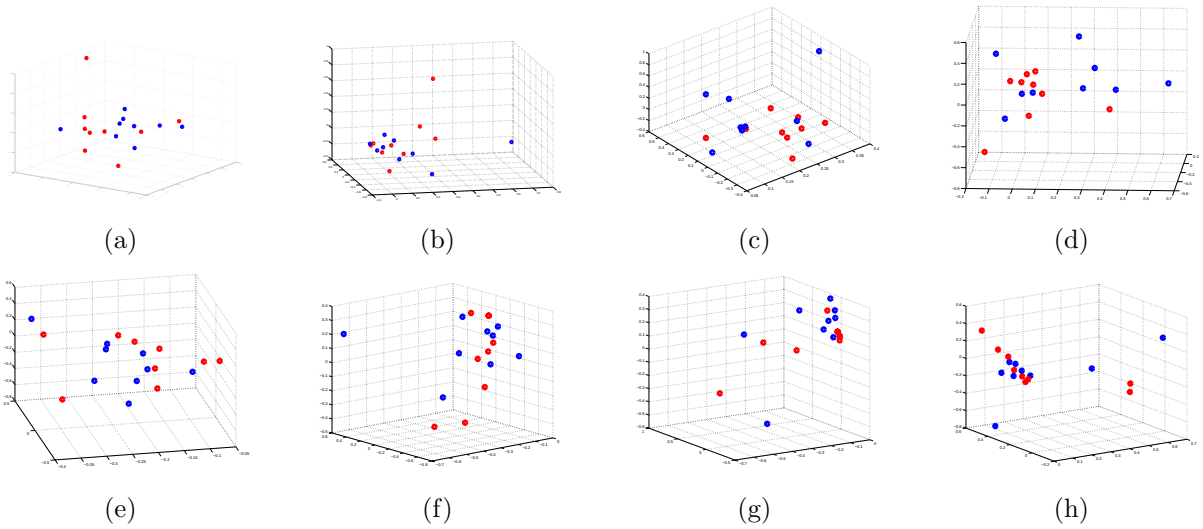


Figure 5. Reduced feature space plots, showing dimensionality reduction results using PCA for various feature types: (a) Morphological Features, (b) Signal Intensity, (c) Kinetic Texture using First Order Statistical Features, (d) Kinetic Texture using Gradient Features, (e) Gabor Features, (f) First Order Statistical Features, (g) Second Order Statistical Features, (h)All Features. Blue dots each represent a benign lesion in the dataset; red dots each represent a malignant lesion in the dataset.

Table 1: SVM Evaluation of the Feature Subsets

DR Method	*Feature Set*	*Accuracy (%)*	*PPV (%)*	*DR Method*	*Feature Set*	*Accuracy (%)*	*PPV (%)*
PCA	Morphological	78	73	LLE	Morphological	61	55
	Kinetic	67	71		Kinetic	61	56
	Kinetic Textural	78	86		Kinetic Textural	72	83
	Gabor	72	75		Gabor	78	50
	1st Order Statistical	57	57		1st Order Statistical	56	55
	2nd Order Statistical	72	83		2nd Order Statistical	72	83

6. CONCLUDING REMARKS

In this paper we presented a novel image feature called the *kinetic texture feature* to distinguish between malignant and benign lesions on breast DCE-MRI. We evaluated this feature against kinetic signal intensity features which are currently employed by most breast MRI radiologists and against several hundred other textural and morphological attributes in distinguishing between benign and malignant lesions on 18 breast lesions as imaged using DCE-MRI. Our primary findings were:

- Kinetic texture appears to outperform kinetic signal intensity, suggesting it may be more robust to MRI artifacts such as bias field and non-standardness.
- Kinetic texture appears to perform better than morphological features, suggesting its potential clinical use and relevance in building an integrated breast DCE-MRI CAD technique.

It is recognized by the authors that these findings come from a limited dataset, and future work must include a larger sample size. In addition, a feature pruning analysis should be performed to parse out any suspected poor classifiers that might be overrepresented in some of the feature sets, decreasing those features sets' performance. We believe this work is an important component of building a comprehensive CAD system, comprising features from morphological, textural, and kinetic classes.

ACKNOWLEDGMENTS

This work was possible due to grants from the Coulter Foundation (WHCF4-29349,WHCF 4-29368), the Busch Biomedical Award, the Cancer Institute of New Jersey, the New Jersey Commission on Cancer Research and the National Cancer Institute (R21CA127186-01, R03CA128081-01), and the Society for Imaging Informatics in Medicine (SIIM).

REFERENCES

1. D. Saslow, "American cancer society guidelines for breast screening with mri as an adjunct to mammography," *CA: A Cancer Journal for Clinicians* **57**, pp. 75–89, 2007.
2. D. M. Ikeda, N. Hylton, K. Kinkel, M. G. Hochman, C. Kuhl, W. A. Kaiser, J. C. Weinreb, S. F. Smazal, H. Degani, P. Viehweg, J. Barclay, and M. D. Schnall, "Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging," *Journal of Magnetic Resonance Imaging* **13**, pp. 889–895, 2001.
3. K. Kinkel, T. H. Helbich, L. J. Esserman, J. Barclay, E. Schwerin, E. A. Sickles, and N. M. Hylton, "Dynamic high-spatial-resolution mr imaging of suspicious breast lesion: Diagnostic criteria and interobserver variability," *Am. J. Roentgenol.* **175**(1), pp. 35–43, 2000.
4. M. D. Schnall, J. Blume, D. A. Bleumke, G. A. DeAngelis, N. DeBruhl, S. Harms, S. H. Heywang-Kobrunner, N. Hylton, C. Kuhl, E. D. Pisano, P. Causer, S. J. Schnitt, D. Thickman, C. B. Stelling, P. T. Weatherall, C. Lehman, and C. A. Gastonis, "Diagnositc architectural and dynamic features at breast mr imaging: Multicenter study," *Radiology* **238**, pp. 42–53, January 2006.
5. L. A. Meinel, A. H. Stolpen, K. S. Berbaum, L. L. Fajardo, and J. M. Reinhardt, "Breast mri lesion classification: Improving performance of human readers with a backpropagation neural network computer-aided (cad) system," *Journal of Magnetic Resonance Imaging* **25**(1), pp. 89–95, 2007.
6. T. W. Nattkemper, B. Arnrich, O. Lichte, W. Timm, A. Degenhard, L. Pointon, C. Hayes, and M. O. Leach, "Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods," *Artificial Intelligence in Medicine* **34**, pp. 129–139, June 2005.
7. J. Wei, L. M. Hadjiiski, B. Sahiner, H.-P. Chan, J. Ge, M. A. Roubidoux, M. A. Helvie, C. Zhou, Y.-T. Wu, C. Paramagul, and Y. Zhang, "Computer-aided detection systems for breast masses: Comparison of performances on full-field digital mammograms and digitized screen-film mammograms," *Academic Radiology* **14**, pp. 659–669, June 2007.
8. W. Chen, M. L. Giger, H. Li, U. Bick, and G. M. Newstead, "Volumetric texture analysis of breast lesions on contrest-enhanced magnetic resonance images," *Magnetic Resonance in Medicine* **58**(3), pp. 562–571, 2007.
9. M. L. Giger, H. Al-Hallaq, Z. Huo, C. Moran, D. E. Wolverton, C. W. Chan, and W. Zhong, "Computerized analysis of lesions in us images of the breast," *Academic Radiology* **6**, pp. 665–674, November 1999.
10. P. S. Tofts, B. Berowitz, and M. D. Schnall, "Quantitative analysis of dynamic gd-dtpa enhancement in breast tumors using a permeability model," *Magnetic Resonance in Medicine* **17**, pp. 357–267, 1991.
11. K.-L. Li, R. Henry, L. J. Wilmes, J. Gibbs, X. Zhu, Y. Lu, and N. M. Hylton, "Kinetic assessment of breast tumors using high spatial resolution signal enhancement ratio (ser) imaging," *Magnetic Resonance in Medicine* **58**(3), pp. 572–581, 2007.
12. L. G. Nyul and J. K. Udupa, "On standardizing the mr image intensity scale," *Magnetic Resonance in Medicine* **42**, pp. 1072–1081, 1999.
13. A. Madabhushi, J. K. Udupa, and A. Souza, "Generalized scale: Theory, algorithms, and application to image inhomogeneity," *Computer Vision and Image Understanding* **101**, pp. 100–121, February 2006.
14. A. Madabhushi, M. Feldman, D. N. Metaxas, J. Tomaszewski, and D. Chute, "Automated detection of prostate adenocarcinoma from high-resolution ex vivo mri," *IEEE Transactions on Medical Imaging* **24**(12), pp. 1611–1625, 2005.
15. J. Chappelow, A. Madabhushi, M. Rosen, J. Tomaszewski, and M. Feldman, "Multimodal image registration of ex vivo 4 tesla mri with whole mount histology for prostate cancer detection," in *Proceedings of SPIE*, J. P. Pluim and J. Reinhardt, eds., *Proceedings of SPIE: Medical Imaging* **6512**, SPIE, March 2007.

16. A. Madabhushi and D. N. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation for ultrasonic breast lesions," *IEEE Transactions on Medical Imaging* **22**, pp. 155–169, February 2003.
17. MIAAB, *Gland Segmentation and Computerized Gleason Grading of Prostate Histology by Integrating Low-, High-Level and Domain Specific Information*, (Piscataway, NJ), September 2007.
18. G. Lee, "An empirical comparison of dimensionality reduction methods for classifying gene and protein expression datasets," ISBRA, 2007.
19. S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding.," *Science* **290**, pp. 2323–2326, 2000.
20. C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning* **20**(3), pp. 273–297, 1995.
21. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery* **2**, pp. 121–167, 1998.