A COMPUTERIZED IMAGE ANALYSIS FRAMEWORK FOR DYNAMIC CONTRAST ENHANCED MAGNETIC RESONANCE IMAGING (DCE-MRI) WITH APPLICATIONS TO BREAST CANCER

BY SHANNON CHRISTINE AGNER

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements for the

Degree of Doctor of Philosophy

Graduate Program in Biomedical Engineering

Written under the direction of

Anant Madabhushi

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2011

ABSTRACT OF THE DISSERTATION

A Computerized Image Analysis Framework for Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) with Applications to Breast Cancer

by Shannon Christine Agner

Dissertation Director: Anant Madabhushi

Dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) provides a wealth of information about the anatomy of the breast, particularly in the setting of breast cancer diagnosis. In addition to the images it provides regarding the architecture of breast tissue, it also provides functional information about blood flow by means of the DCE study. The sensitivity of DCE-MRI has been reported at close to 100%, so the difficult tasks for the radiologist in reviewing breast DCE-MRI are: (1) discerning between which lesions are benign and which are malignant; and (2) doing so for a patient study that involves hundreds of images and is 4-dimensional. Because of the great detail and volume of information DCE-MRI provides, computational methods for both extracting and analyzing information derived from the images are useful in distilling the entire patient study down to the most salient images and features for the radiologist to examine. In this dissertation, computer-based methods developed for analyzing the data acquired in a breast DCE-MRI patient study are described.

In the first part, pre-processing methods used for aligning the images of the time-dependent DCE study are explained. Because segmentation is important for describing

the morphology of the lesion as well as the region of interest for any subsequent quantitative analysis of a lesion, as a second step to pre-processing, a spectral embedding based active contour (SEAC) method for segmentation of lesions is developed and tested. A feature developed for extracting the spatiotemporal characteristics of breast lesions, termed textural kinetics, is then described, and its utility is demonstrated for distinguishing benign from malignant lesions as well as in identifying triple negative breast lesions, a lesion type that is extremely aggressive and has no targeted therapies. Finally, these quantitative methods are summarized in a computer aided diagnosis framework that provides insight into the biologic nature of breast lesion subtypes as well as for directing treatment and determining prognosis.

Preface

This dissertation represents the collective published and unpublished works of the author. Chapters 2-5 are primarily composed from the content of conference papers and peer-reviewed journal articles [1–4] as well as unpublished work written by the author of this dissertation over the course of her thesis work.

Acknowledgements

It takes a village to raise an MD/PhD student...

I would first like to thank Anant Madabhushi for his continued support, oversight, and expectation of nothing but my best effort. I also want to thank Dr. Kinzy and Dr. Millonig for their support throughout their tenures as MD/PhD Program Director, and Dr. Terregino, without whose foresight I would not even be at RWJMS and Rutgers.

I want to thank the members of my dissertation committee who have been instrumental in helping me to frame my dissertation in the context of important clinical problems. The members of the Laboratory for Computational Imaging and Bioinformatics also deserve my gratitude for many great discussions about various aspects of my projects. I especially want to thank Jun Xu for helping me to fundamentally understand active contour models and Jonathan Chappelow for countless conversations on image registration.

I want to express my deepest gratitude to my family for continually supporting me throughout these past few years. I also want to thank my friends who always provided respite from the lab in the form of laughter, relaxation, ski days, and hikes.

Dedication

I dedicate this dissertation to my family (my parents, my siblings, my grandmother, and my great aunt), without whom I would not have been able to undertake this long road of schooling.

The Road Not Taken

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;

Then took the other, as just as fair,
And having perhaps the better claim
Because it was grassy and wanted wear,
Though as for that the passing there
Had worn them really about the same,

And both that morning equally lay
In leaves no step had trodden black.
Oh, I marked the first for another day!
Yet knowing how way leads on to way
I doubted if I should ever come back.

I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I,
I took the one less traveled by,
And that has made all the difference.

-Robert Frost

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Despite the increased use of dynamic contrast enhancement magnetic resonance imaging (DCE-MRI) for breast cancer diagnosis, questions regarding the interpretation of breast MRI and its appropriate uses persist. The objective of this work is to demonstrate how the novel computerized image registration, segmentation, feature extraction, and classification tools we are developing in the context of breast DCE-MRI can contribute to personalized medicine for breast cancer detection, diagnosis, prognosis, and treatment. This study comprises four specific aims. Aims 1 and 2 are focused on development and evaluation of image analysis methods for DCE-MRI while Aims 3 and 4 will leverage the methods developed in Aims 1 and 2 for distinguishing malignant from benign lesions and for distinguishing more aggressive breast cancer subtypes from those that traditionally have a positive prognosis. The overarching objective of this study is to be able to develop a tool set that will allow for early detection of highly aggressive molecular subtypes of breast cancer based on radiological imaging features alone.

In Aim 1, we develop registration and segmentation methods for DCE-MRI. Our novel registration method, called spectral embedding based registration (SERg), integrates texture features, spectral embedding, and higher order mutual information to register post-contrast images to pre-contrast DCE-MR images. This method is particularly useful in the setting of DCE-MRI because of imaging characteristics inherent to DCE-MRI. We also derive an automated model-free active contour-based lesion segmentation method, called spectral embedding based active contour (SEAC), which integrates temporal contrast enhancement information and spectral data embedding. In Aim 2, we develop novel quantitative lesions features, called *textural kinetics*, which measure the changes in texture of the lesion of interest as a function of time, analogous

to the kinetic contrast versus time curves that radiologists currently use for diagnosis. However, different from contrast kinetics, textural kinetics provide a measure of the spatial pattern of contrast uptake as a function of time. Our hypothesis is that our computational image analysis methods developed in Aims 1 and 2 will provide insight into breast lesion detection and characterization on DCE-MRI to reveal subtleties of breast lesions traditionally occult to the radiologist. In Aim 3, we leverage the methods developed in Aims 1 and 2 for distinguishing benign from malignant breast lesions while in Aim 4 we focus on distinguishing different molecular subtypes of malignant lesions such as triple negative, estrogen receptor positive, and human epidermal growth factor receptor 2 (HER2) positive lesions.

1.1 Significance of proposed work

While there are available targeted treatments for two common types of breast cancer, there is a third, called triple negative breast cancer that constitutes 10-13% of the population, yet has no targeted treatment and is extremely aggressive [7]. Since the American Cancer Society [8] also recently recommended all women with a family history of breast cancer also have yearly breast MRIs in addition to annual mammograms, the number of breast MRIs performed will continue to increase. Of note, there is significant overlap between BRCA1 mutations [7], which are sometimes found in familial breast cancers, and triple negative breast cancers. In this study, we aim to create a computerized decision support system to assist physicians in assessing suspicious breast lesions on MRI to help guide time to treatment and noninvasive monitoring of the biological effects of therapies that target the aggressive triple negative subtype of breast cancer. The hypothesis of this project is that by finding ways of identifying the aggressive triple negative subtype of breast cancer on MRI, the time to confirmatory biopsy can be decreased, thus increasing likelihood of treatment effectiveness. In addition, by identifying the biological characteristics on the MRI, treatments that are in the research pipeline for triple negative breast cancers can be monitored noninvasively to understand their biological effects. Furthermore, we hope to use the radiologic imaging features developed in this thesis to correlate appearance on DCE-MRI with diagnosis and prognosis

derived from histopathology evaluation of breast tissue biopsy, immunohistochemistry results and molecular assay scores.

The goal of this project is a computerized decision support system that will assist physicians in assessing breast lesions in a noninvasive manner during all aspects of a woman's care: from screening for breast cancer to planning treatment, should she have breast cancer, to monitoring tumor response to therapy. From a research perspective, this study will provide a better understanding of the biology of familial types of breast cancer. In addition, the system could be used as a training tool for medical students and residents. From a healthcare cost perspective, this system will allow for overall reduction in cost by earlier diagnosis and subsequent time-to-treatment of aggressive types of breast cancer.

1.2 Breast Cancer

According to the National Cancer Institute, approximately 207,090 women will be diagnosed with breast cancer in the United States in 2010 [1], and 1 in 8 women will be diagnosed with breast cancer during the course of their lifetime. However, it has been shown that early detection and treatment can reduce the chances of death due to breast cancer by 15 to 30 percent. In 2005, patients with the BRCA1 gene mutation, which has significant overlap with the triple negative breast cancers, were recommended to obtain annual breast MRIs in addition to mammography. As a result, the use of DCE-MRI for breast cancer diagnosis more than doubled between 2003 and 2008 to a frequency of over 645,000 breast MRI exams annually in the United States, growing much faster than the number of MRI exams overall. This demonstrates the potential key role in the personalized medicine of treating breast cancer may play in the future.

The current guidelines recommend that a woman over the age of 40 begin annual mammographic screening for breast cancer. If a suspicious lesion is found on the mammogram, follow up ultrasound or MRI is typically done. If the lesion is still deemed suspicious after secondary imaging, a biopsy is performed, and the tissue sample is sent

[1] http://www.cancer.gov/cancertopics/types/breast

to pathology for definitive diagnosis and prognosis. If cancer is found, immunohistochemistry (IHC) or immunofluorescence (IF) staining is done to look for overexpression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor (HER2) since targeted therapies for both ER and HER2 positive tumors exist. The question we aim to answer is whether or not associated radiologic imaging phenotypes exist that correlate with the cellular protein expression phenotypes.

1.3 Role of DCE-MRI in Breast Cancer Diagnosis

When a typical DCE-MRI is performed, T1-weighted images are acquired before contrast injection. A gadolinium-based contrast agent is then injected into the patient's bloodstream, and subsequent images are acquired at regular intervals after contrast injection. When a radiologist currently evaluates a breast DCE-MRI, she looks for shape, heterogeneity of signal intensity, and intensity kinetics within the lesion of interest. All of these characteristics have been formalized in the BIRADS classification system for DCE-MRI [9]. A malignant lesion tends to have more irregular borders and be more heterogeneous in signal intensity than benign lesions. It also has a rapid uptake and washout of contrast, which differs from benign lesions, which tend to take up contrast agent more slowly and wash contrast out more slowly.

Although contrast allows lesions to be seen more readily on MRI, making DCE-MRI the most sensitive of the 3 imaging modalities currently used in breast radiologic imaging, it also generates many more false positives than ultrasound or mammography. As a result, many more biopsies are done on benign lesions than were previously performed with ultrasound or mammography. This obviates the need for generating computer-aided image analysis systems that more specifically discern benign lesions from malignant ones.

Although the increased sensitivity of DCE-MRI can be perceived as a detriment, it also has its benefits. Radiologists are currently able to see more detail of the breast than ever before, especially in young women who tend to have more dense breasts [10]. Additionally, in recent years, DCE-MRI has been explored at a radiologic imaging

modality capable of discerning histologic cancer subtypes such as lobular cancers from ductal cancers and carcinoma in situ from invasive cancers. It has long been thought that tumor angiogenesis was correlated with cancer invasiveness based on histologic analysis [11], and this is the feature that DCE-MRI intends to capture. As we begin to understand more about breast cancer and its different biological subtypes, there have also been some qualitative observational studies looking at imaging patterns of molecular subtypes such as ER positive, HER2 positive, and triple negative breast cancers [12–14].

The goal of this project is to combine both anatomical and contrast dye uptake information about highly aggressive breast cancer lesions to develop specific and quantitative MRI profiles for these breast cancers. The hope is that identifying these aggressive breast cancers lesions at the time of screening or follow up imaging might allow for faster time to treatment as well as allow clinicians to noninvasively monitor biological effects of chemotherapies that are specifically tailored to a particular type of breast cancer.

1.4 Overview of Dissertation

In this dissertation, new methods for computerized analysis of breast DCE-MRI are developed. In the first part, pre-processing methods used for aligning the images of the time-dependent DCE study and a lesion segmentation method that utilizes the DCE information are described. In the second part, a new feature for quantitatively describing spatiotemporal patterns of contrast enhancement is developed and explained. In the final part, a computer-aided diagnosis framework for first distinguishing benign from malignant and then distinguishing molecular subtypes of breast cancer is constructed.

Image artifacts due to patient movement and anatomical deformation cause inaccuracies when comparing one time point to another in a DCE-MRI time series, and these artifacts are not always rectified by a conventional intensity-based image registration method. We introduce an alternative image representation called spectral embedding (SE)- based registration (SERg) that facilitates registration of each post-contrast image

to the corresponding pre-contrast image. SERg is employed via higher order mutual information (MI) (i.e., alpha-MI) and a B-spline based elastic registration method in order to use multiple eigenvectors that result from SE. We demonstrate on 12 synthetic combinations of noise and image deformations on synthetic brain magnetic resonance imaging (MRI) that SERg performs better in terms of normalized MI and correlation ratio than an intensity based registration scheme. We also demonstrate qualitative improvements using SERg on breast DCE-MRI over intensity-based registration.

Since segmentation is important for describing the morphology of the lesion as well as the region of interest for any subsequent quantitative analysis of a lesion, as a second step to pre-processing, a spectral embedding based active contour (SEAC) method for segmentation of lesions is developed and tested. On a cohort of 50 breast lesions, the SEAC outperformed fuzzy c-means in terms of both boundary (mean absolute difference, MAD) and area-based measures (Dice similarity coefficient, DSC); SEAC had a MAD of 3.2 2.1 pixels and a DSC of 0.74 0.13 compared to FCM which had a MAD of 7.2 7.4 pixels and a DSC of 0.58 0.32. When using these segmentations for driving a morphology-based support vector machine classifier, SEAC demonstrated a classification accuracy of 0.73 whereas FCM demonstrated a classification accuracy of 0.65 in a benign versus malignant lesion classification task.

A feature developed for extracting the spatiotemporal characteristics of breast lesions, termed textural kinetics, is then described, and its utility is demonstrated for distinguishing benign from malignant lesions on a cohort of 41 patients (Area under the curve (AUC) = 0.91) as well as in identifying triple negative (TN) breast cancer (CA), a breast cancer subtype that is extremely aggressive and has no targeted therapies, amongst a cohort of 76 breast lesions in 65 patients (AUC = 0.73).

Finally, these quantitative methods are summarized in a computer aided diagnosis framework that provides insight into the biologic nature of breast lesion subtypes. The attention is focused on distinguishing TN lesions from other molecular subtypes. By determining a quantitative MR imaging signature for TN CAs that is distinct from other CA subtypes (e.g., ER+ and HER2+ CAs), as well as from benign TN mimics (e.g., FAs), CAD algorithms tuned to specific high-risk patient groups (e.g., BRCA1

mutation carriers predisposed to forming TN CAs) can be developed. TN CAs possess certain characteristic features on DCE-MRI that can be detected and quantified via CAD, allowing for highly accurate discrimination of TN CAs from non-TN CAs, as well as between TN CAs and benign FAs. Such CAD algorithms may provide added diagnostic benefit in identifying the highly aggressive TN CA phenotype in high-risk women with DCE-MRI.

Chapter 2

Spectral Embedding- based Registration (SERg) for the Alignment of DCE-MR Images

2.1 Introduction

Alternative data representations for improved registration accuracy and efficiency have been explored with the idea that two images in the transformed space may be more similar than the corresponding images in the raw intensity space. Nyul et al. [15] used the ball-scale concept, where a ball of a certain radius defined a region of homogeneity in an image, to drive registration. Similarly, Saha [16] demonstrated the effectiveness of tensor scale (t-scale), whereby an ellipsoid defines the region of homogeneity, to improve image registration.

Although alternative data representations have been developed previously primarily for motion correction, time-series data involving dynamic contrast enhancement (DCE) introduces another difference between target and template images: changes in contrast. Since the post-contrast images are typically registered to pre-contrast images, comparing images whose dynamic intensity ranges change from time point to time point on the basis of signal intensity may not be an appropriate method for DCE data. This is illustrated in Figure 2.1, which shows the subtraction images of a breast DCE-MRI before (Figures 2.1 (a, b, c)) and after (Figures 2.1 (d, e, f)) intensity based registration at 3 time points.

In this paper, we present a spectral embedding (SE)-based registration (SERg) method for time series data. SE, a nonlinear dimensionality reduction (DR) method, is used to transform the images into an alternative data representation that is amenable to image registration. SE has the advantage over the linear DR method, principle component analysis (PCA), which was previously used by Staring et al. [17] for registration

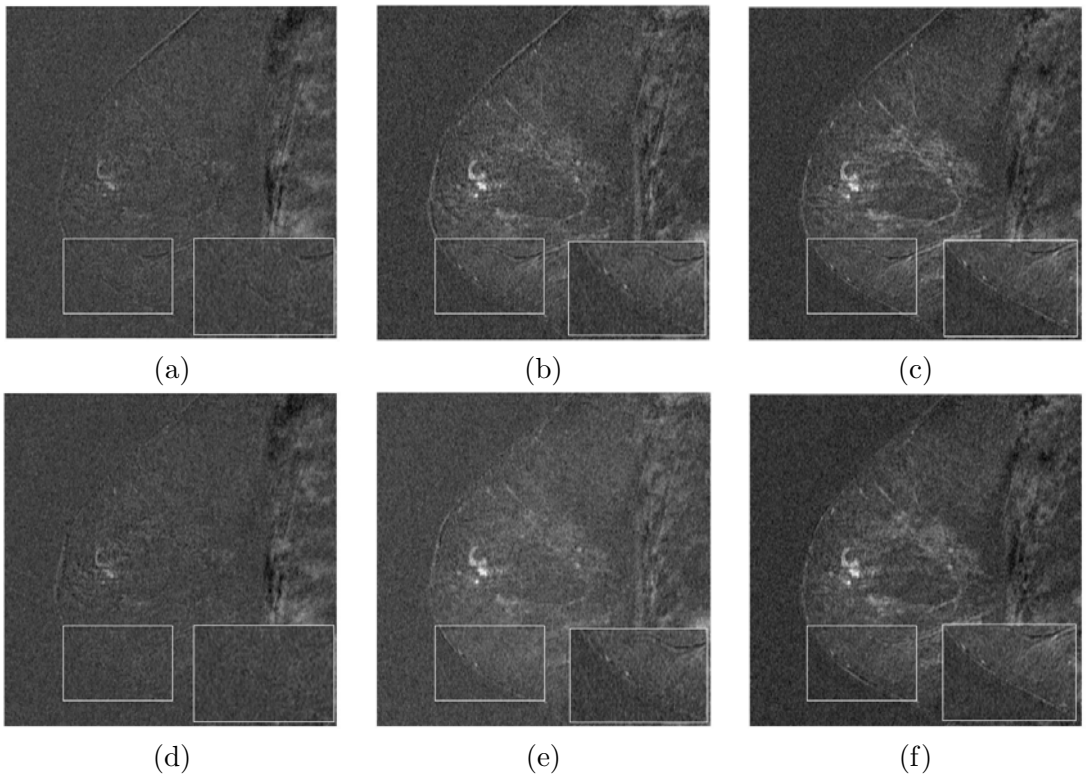


Figure 2.1: Comparison of unregistered post-contrast time point 1 (a), 2 (b), and 3 (c)) and corresponding intensity-based registered ((d), (e) and (f)) subtraction images. Inset in bottom-right corner of each image highlights misalignment demonstrated by high signal intensity at breast edge. Note that misalignment is recovered in the first post-contrast time point and progressively worsens as contrast agent is taken up by the breast.

of time-dependent data, of having the ability to capture the nonlinearities inherent in biological data, especially functional data such as DCE-MRI, where noise in the form of contrast agent is added in a nonlinear fashion to the image. A linear DR method such as PCA would not be able to take such nonlinearities into account when embedding a pre-contrast and a post-contrast image in the same space. However, by embedding the pre- and post-contrast images into the same embedding space, SE allows the embedding eigenvectors to not only separate salient from non-salient regions in the image, but also to identify corresponding regions in both images. The registration is then driven by these areas of correspondence and salience.

Since the SE based transformation of the data allows for a vectorial representation of the scalar MRI scene in terms of the principle eigenvectors at every voxel, higher order entropy measures such as α-MI [18] can be leveraged as the registration objective

function, thereby enabling more accurate capture of the statistics of multivariate data.

We quantify the accuracy of SERg on 12 different combinations of noise and deformations of a synthetic brain MRI data set. We also demonstrate the qualitative accuracy of SERg on breast DCE-MR images.

2.2 Theory and Motivation behind SERg

2.2.1 Formal problem statement

With issues of both motion and changes in image intensity in contrast enhanced MRI in mind, we pose the following problem statement. Let there exist: (1) a pre-contrast template image $\mathcal{C}^{pre} = (C, f^{pre})$, where C is a 2D grid of spatial locations and $f^{pre}(c)$ is the intensity value at location c; and (2) a similarly defined post-contrast target image $\mathcal{C}^{post} = (C, f^{post})$. The question then is whether an image transformation can be determined, $\Lambda(\mathcal{C}^{pre}, \mathcal{C}^{post})$ that yields an alternative image representation, $\hat{\mathcal{C}}^{pre}, \hat{\mathcal{C}}^{post}$ such that, given an image similarity measure, ϕ, $\phi(\mathcal{C}^{pre}, \mathcal{C}^{post}) < \phi(\hat{\mathcal{C}}^{pre}, \hat{\mathcal{C}}^{post})$?

2.2.2 Spectral Embedding

Let $\mathcal{F} = [\mathbf{F}(\mathbf{x}_1), \mathbf{F}(\mathbf{x}_2), \ldots, \mathbf{F}(\mathbf{x}_N) \in \mathbb{R}^{N \times D}]$ be the data matrix of N feature vectors with dimensionality D. The aim of SE is to reduce $\mathcal{F} \in \mathbb{R}^{N \times D}$ to a low d-dimensional space where $d << D$. Let $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N \in \mathbb{R}^{N \times d}]$ be the optimal low dimensional projections [19] and the associated eigenvectors of a given object $\mathbf{x}_i$, where $i \in \{1, \ldots, N\}$, are $\mathbf{v}_i \in \mathbb{R}^{1 \times d}$ where $\mathbf{v}_i = [v_1, v_2, \ldots, v_N]$ and $v_j \in \mathbb{R}^{1 \times 1}, j \in \{1, \ldots, d\}$ is an individual eigenvector for a given $\mathbf{v}_i$. The optimal $\mathbf{V}$ can be obtained by solving,

$$\mathbf{V} = \arg\min_{v}\left(\frac{\sum_r \sum_s \|\mathbf{v}(r) - \mathbf{v}(s)\|^2 w_{rs}}{\sum_r \mathbf{v}(r)^2 \mathbf{d}(r)}\right), \tag{2.1}$$

where w_{rs} is the (r, s) element of the weight matrix $W = [w_{rs}] \in \mathbb{R}^{N \times N}, r, s \in \{1, \ldots, N\}$, which assigns edge weights to characterize similarities between the pairwise observations $\mathbf{x}_r$ and $\mathbf{x}_s$ and $\mathbf{d}(r) = \sum_s w_{rs}, r, s \in \{1, \ldots, N\}$. The graph edge weight of two nodes, r and s, can be formulated as $w_{rs} = e^{\frac{-\|\mathbf{F}(\mathbf{x}_r) - \mathbf{F}(\mathbf{x}_s)\|_2^2}{\sigma_I^2}}$ where σ_I^2 is a scaling parameter. The minimization of Equation 2.1 reduces to a minimum eigenvalue

decomposition problem,

$$(\mathbf{D} - W)\mathbf{v} = \lambda \mathbf{D}\mathbf{v}, \tag{2.2}$$

where $\mathbf{D}$ is a diagonal matrix, $\mathbf{D}_{rr} = \sum_s W_{rs}$. The 3 eigenvectors of $\mathbf{v}_i = [v_1, v_2, v_3]$ where $i \in \{1, \ldots, N\}$ associated with the smallest eigenvalues, $\lambda_1, \lambda_2, \lambda_3$, are used in conjunction with SERg.

2.2.3 SERg Implementation with α-MI

In the conventional manner, we formulate the registration method as an optimization problem such that a function [17],

$$\hat{\mu} = \arg\max_{\mu} \phi(\mathbf{T}_{\mu}(c); \mathcal{C}^A, \mathcal{C}^B), \tag{2.3}$$

describes the optimization, where ϕ is the objective function, measuring similarity between $\mathcal{C}^A$ and $\mathcal{C}^B$ which is maximized when $\mathcal{C}^A$ and $\mathcal{C}^B$ are aligned optimally, and $\hat{\mu}$ is the vector of coordinates that transforms the target image, $\mathcal{C}^B$, to be in alignment with template image, $\mathcal{C}^A$. In the case of 1-dimensional images such as intensity images, MI, ϕ_{MI} is used to optimize image alignment. MI is typically defined as

$$\phi_{MI}(\mathcal{C}^{pre}, \mathcal{C}^{post}) = H(\mathcal{C}^{pre}) + H(\mathcal{C}^{post}) - H(\mathcal{C}^{pre}, \mathcal{C}^{post}), \tag{2.4}$$

where $H(\mathcal{C}) = -\sum p_f \log p_f$ is the Shannon entropy of the intensities of image $\mathcal{C}$, $H(\mathcal{C}^{pre}, \mathcal{C}^{post})$ is the joint entropy, and p_f is the intensity probability at intensity f. However, in the case of multidimensional images, a multivariate approach must be taken to the objective function. This can be done through two methods: (1) histogram-based mutual information (MI), or (2) α-MI via entropic graphs [18]. The joint entropy is typically calculated using a joint histogram; however, issues of sparseness in the joint histogram can lead to inaccuracies in the computation of joint entropy (JE). The use of entropic graphs was recently introduced by Neeumuchwala et al. [18] to overcome issues with using a joint histogram in the calculation of JE, and in this paper, we use α-MI to estimate image similarity in a multivariate fashion. α-MI is calculated as,

$$\alpha - MI = \frac{1}{\alpha - 1} \log \frac{1}{K^{\alpha}} \sum_{c=1}^{K} \left(\frac{\Gamma_c^{pre,post}(\mu)}{\sqrt{\Gamma_c^{pre} \Gamma_c^{post}(\mu)}} \right)^{2\gamma} \tag{2.5}$$

where $\Gamma^{pre,post}$ is neighborhood graph for all points in both the pre- and post-contrast images, Γ^{pre} is the neighborhood graph for the pre-contrast, template image, Γ^{post} is the neighborhood graph for the post-contrast, target image, $K = |c|$ where $|\cdot|$ is the cardinality of any set, $\gamma = D(1-\alpha)$ where D is the length of the feature vector describing c, and $0 < \alpha < 1$ ($\alpha = 0.99$ for all experiments herein as was previously done by Staring et al. [17]). Registration is performed using a free-form deformation previously described [20].

2.2.4 Algorithm for SERg

Algorithm *SERg*

Input: Image scenes $\mathcal{C}^{pre}, \mathcal{C}^{post}$.

Output: Image transformation, $\mathcal{C}^{post}(\mathbf{T}_\mu(C))$.

begin

0. Initialize $\mathcal{C}^{pre}, \mathcal{C}^{post}$;
1. Extract feature scenes, $\mathcal{F} = F(C)$;
2. Apply spectral embedding [19] to $\mathcal{F} = [\mathbf{F}(\mathbf{x}_1), \mathbf{F}(\mathbf{x}_2), \ldots, \mathbf{F}(\mathbf{x}_n) \in \mathbb{R}^{N \times d}]$ to obtain $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^{N \times k}]$;
3. Initialize $\phi(t) = \phi_0$ where $t = 0$;
4. *while* $|\phi(t) - \phi(t-1)| > \epsilon$ *do*
5. Maximize $\phi(\mathbf{T}_\mu(c); \hat{\mathcal{C}}^A, \hat{\mathcal{C}}^B)$.
6. $t = t + 1$;
7. *endwhile*;
8. $\mathcal{C}^{post}(\mathbf{T}_\mu(C))$;

end

2.3 Experimental Design

2.3.1 Data description

Table 2.3.1 shows a summary of the 3 datasets used which sum to 28 image pairs for evaluation of SERg. Synthetic brain data was generated from the MNI BrainWeb database [21]. Contrast dye introduced to the subject can be interpreted as noise in the images, and as such, the image without noise is considered to be the pre-contrast image. The 1st post-contrast time point is simulated by the image with 1% noise, the 2nd post-contrast time point is simulated by the image with 5% noise, and the 3rd post-contrast time point is simulated by the image with 9% noise. Four nonlinear deformations were imposed on the simulated post-contrast images using control points and thin-plate splines [22].

Table 2.3.1: Performance measures to evaluate SEAC

Data type	*Image pairs*	*Data description*
Synthetic brain	12	4 different image deformations, 3 different noise levels
Fat-suppressed breast DCE-MRI	10	Peak post-, pre-contrast enhancement fat-suppression in acquisition
Fat-saturated breast DCE-MRI	6	Peak post-, pre-contrast enhancement; no fat-suppression in acquisition

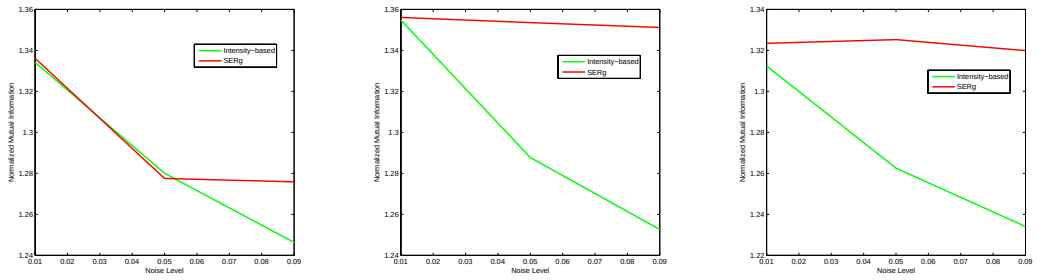
Eight 1st order statistical features along with 7 edge detection filters at a window size of 5×5 voxels are calculated for each pre- and post-contrast image. This feature set plus the original raw intensity image results in a feature vector, $\mathcal{F}$, of length 16 to be used in conjunction with SE.

2.3.2 Measures of registration accuracy

We compared the registration accuracy obtained by SERg to an intensity based registration first proposed by Rueckert et al. [20]. Data was evaluated qualitatively by visual inspection of the difference image resulting from $C^{post} - C^{pre}$. Quantitative measures of registration accuracy were correlation ratio (CR) [23] and normalized MI (NMI).

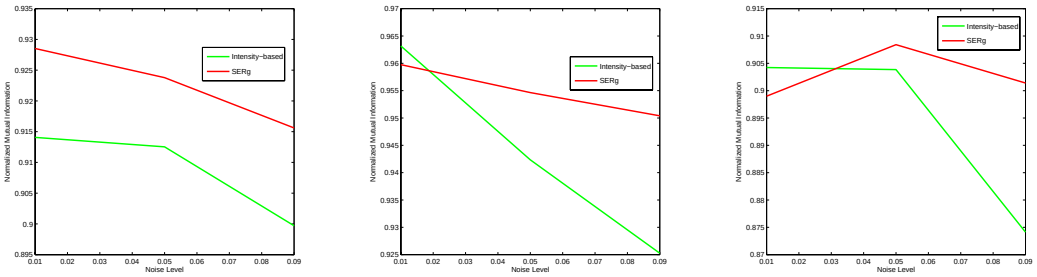
Quantitative Results

Figures 2.2 and 2.3 show the results of the synthetic experiments of varied noise and deformation field. Because we have the images before deformation, a ground truth similarity can be calculated for each image pair. The SERg method was closer to the ground truth in terms of both NMI and CR.



(a) (b) (c)

Figure 2.2: Plots of NMI as a function of noise level for 3 different deformations. Note that the intensity-based registration (green line) results in lesser NMI as more noise is introduced into the deformed image, whereas SERg NMI (red line) remains relatively robust to noise.

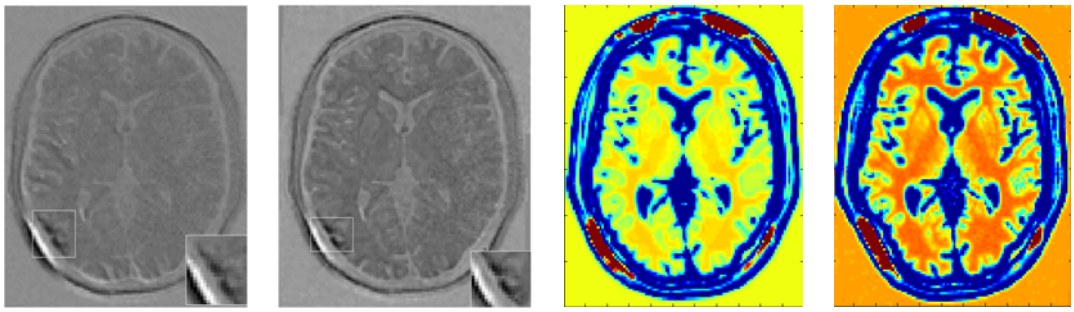


(a) (b) (c)

Figure 2.3: Plots of correlation ratio (CR) as a function of noise level for 3 different deformations. Note that although CR is comparable for the intensity-based (green line) registration and SERg (red line) at 1% noise, as more noise is introduced to the image, the difference in CR between intensity-based registration and SERg grows.

Figure 2.4 shows an example of the recovered deformation using intensity based registration (Figure 2.4(a)) and SERg (Figure 2.4(b)) for the most severe noise level. The deformation spans a longer segment of the interface between the skull and the brain matter in Figure 2.4(a) than in Figure 2.4(b). The eigen image scene that results from the simultaneous spectral embedding of the feature scences are shown for both

the simulated pre-contrast (Figure 2.4(c)) and simulated post-contrast (Figure 2.4(d)) images.



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

Figure 2.4: An example of the synthetic deformation (bottom, left of image in (a, b, and d)) imposed on the brain data. Undeformed image subtracted from 9% noise image containing deformation (a). Undeformed image subtracted from 9% noise image after SERg registration. Eigen image scenes, generated by representing each of the first 3 eigenvectors as color channels of color image, of undeformed noise-free image (c) and 9% deformed image (d) after spectrally embedding in the same space. Note that cerebrospinal fluid is blue in both images, white matter is yellow and orange hues, and grey matter is cyan.

Qualitative Results

Figure 2.5 shows difference images for 3 breast DCE-MRI datasets which require registration. The difference images show that the breast border strongly enhances when images are not aligned. Although intensity-based registration seems unable to align the images (Figure 2.5 (b, e, h)), SERg is able to do so (Figure 2.5 (c, f, i)).

2.4 Concluding Remarks

In this paper we demonstrate that an alternative data representation called spectral embedding (SE) based registration (SERg) can more accurately register images in the setting of noise like that found during contrast uptake of a dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) time series. The intuition behind using SERg is that by maximizing intracluster similarity and minimizing intercluster similarity through SE, areas of correspondence and salience, respectively are best preserved. We have demonstrated via synthetic data with 3 noise and 4 deformations levels that SERg

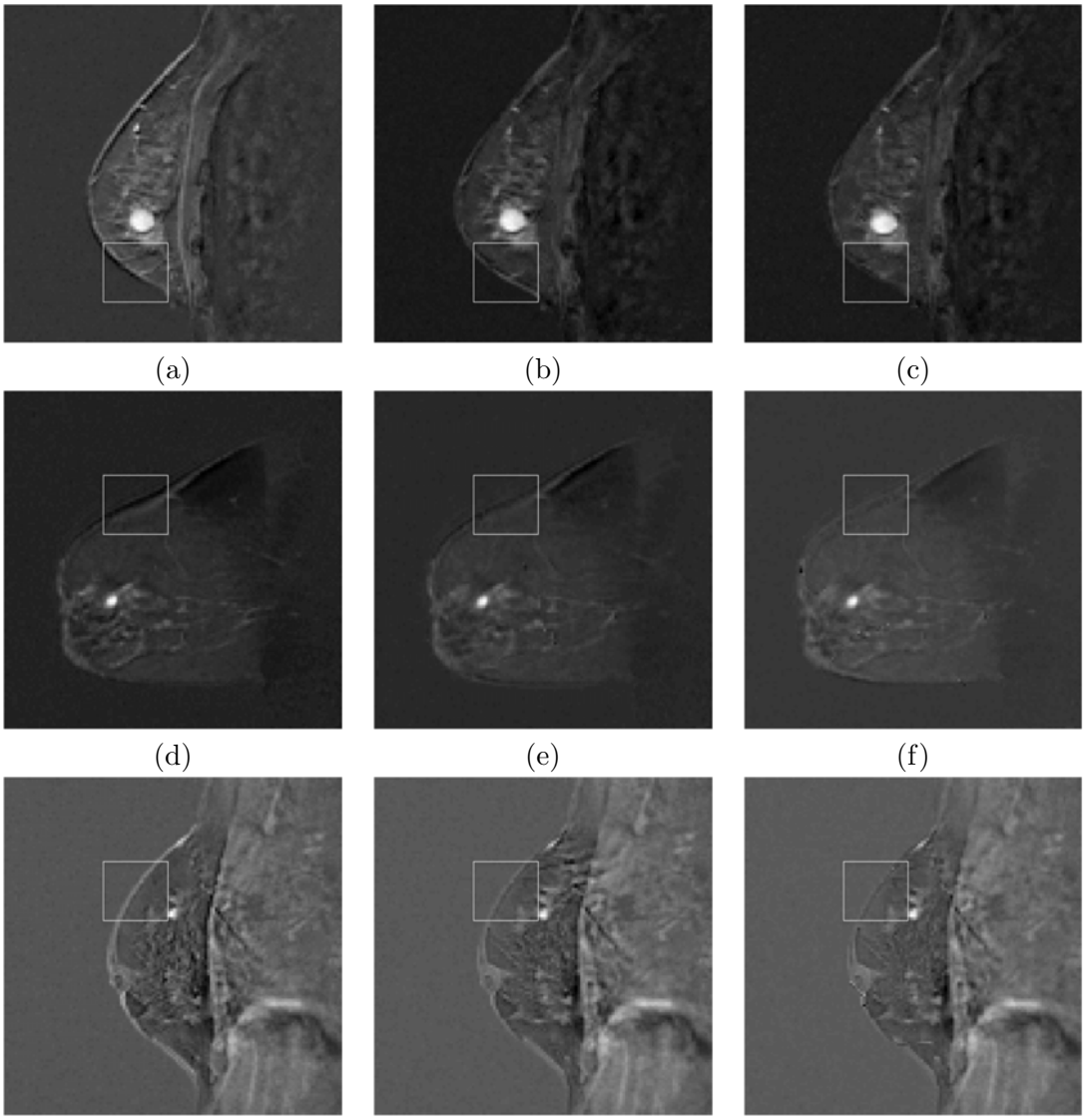


Figure 2.5: Comparison of unregistered (a, d, g), signal intensity based registration (b, e, h), and SERg (c, f, i) and difference images for 3 difference breast DCE-MRI datasets. Inset in bottom-right corner of each image highlights misalignment demonstrated by high signal intensity at breast edge.

is more resilient to these image deficiencies than is an intensity based registration. We also show that in the presence of different levels of noise, SERg outperforms an intensity based registration scheme.

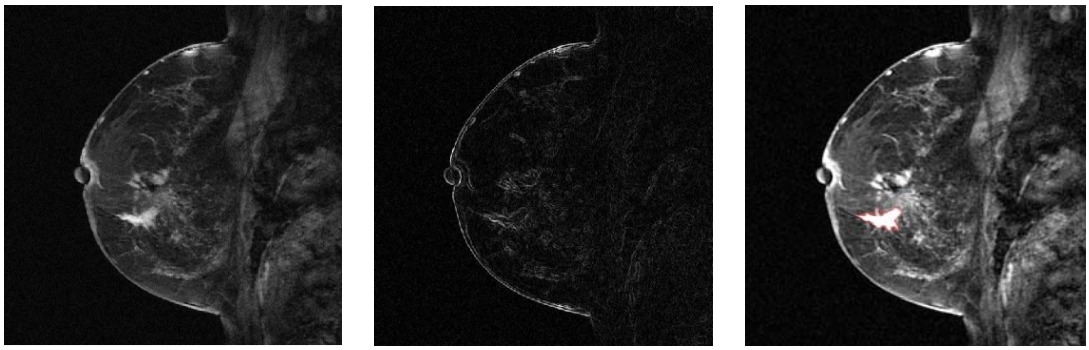
Chapter 3

Spectral Embedding-based Active Contour (SEAC): Applications to Breast Lesion Segmentation

3.1 Introduction

Active contour (AC) schemes are commonly used for medical image segmentation [24]. AC models can be divided into boundary- based [25], region- based [26], and hybrid [27] models, which combine boundary- and region- based methods. The success of boundary-based approaches hinges upon the strength of the image gradient at the boundary between the region of interest (ROI) and the background. This is often a limitation of the AC model when applied to grayscale images such as those found in radiology. Strong image gradients at the lesion border may not always be available, particularly when the ROI has diffuse boundaries as is demonstrated in Figure 3.1 and which often occurs in the case of malignant lesions [28]. For region- based methods, region statistics are computed over the entire image. These statistics may not be very valuable in the case of grayscale radiologic images, especially because grayscale images only have a single channel of information from which to derive the image statistics.

The relatively weak image gradients at the lesion boundary and multiple objects which have similar intensity statistics to the lesion of interest (see Figure 3.1) in the grayscale image beg the question: Can an alternative image representation be determined where stronger image gradients and region statistics for driving an AC model be found? Alternative image representations have been used previously in the context of noise filtering [29], image registration [15], and fuzzy connectedness based segmentation [30]. Nyul et al. [15] employed ball-scale for multi-protocol image registration, where ball-scale [15] is a locally adaptive scale definition such that every image pixel location is parametrized by the radius of the largest ball that satisfies some pre-defined



(a) (b) (c)

Figure 3.1: A DCE-MR image of a malignant breast lesion (a) and its corresponding gradient image (b). Note that the gradient is not uniform along the boundary of the lesion and is very weak in some regions of the lesion boundary. The lesion boundary delineated by an expert radiologist is illustrated in (c).

local homogeneity criterion. Saha [31] defined tensor scale (t-scale) at every spatial location as the largest ellipse that satisfies some pre-defined homogeneity criterion. The t-scale based representation has been employed in the context of image segmentation and image filtering [29, 30]. In each of these methods, transforming the data into another image space allowed for an improvement in the corresponding image processing task. However, to the best of our knowledge, no attempts have been made to explore the utility of manifold learning (ML) schemes to seek improved image representations that would be amenable for use in conjunction with an active contour segmentation scheme.

ML based schemes like spectral embedding (SE) allow for parametrically representing high dimensional data in a reduced dimensional space [19]. Several researchers have explored ML for data classification [32–34]. However, while others have employed SE in the context of image partitioning algorithms such as normalized cuts [19], few studies have addressed alternative data representations for facilitating segmentation of time dependent magnetic resonance imaging (MRI) such as dynamic contrast enhanced (DCE)-MRI. Eyal et al. [32] used the principal eigenvectors derived from principal component analysis (PCA) to determine a parametric representation of breast DCE-MRI data. Since the feature matrix in PCA is a covariance matrix, the eigenvectors associated with the largest eigenvalues rotate the data along axes of maximum variance,

and the gradient functional derived from the PCA eigenvectors would be based on a gradient of deviation from the mean. Image gradients derived in this fashion would most likely not be strong enough to serve as the stopping criterion for an active contour formulation. In contrast, SE uses the eigenvectors corresponding to the minimum eigenvalues based on the eigenvalue decomposition of a weighted affinity matrix, where the affinity matrix represents the pairwise dissimilarity between all pixels in the image obtained via a Gaussian, exponential, or polynomial kernel in the original feature space.

Since SE will partition the images in order to maximize intracluster similarity and minimize intercluster similarity [19], the eigenvectors are oriented along the directions of fundamental patterns of the data. In the context of DCE-MRI, these fundamental patterns are related to the time-intensity curves at each pixel in the image. In the context of DCE-MRI, the time-intensity curves from lesion and non-lesion areas will have different characteristics as previously shown in multiple different studies [35, 36]. In this work, we present a new spectral embedding based active contour (SEAC) scheme for segmentation of lesions on breast DCE-MRI. SEAC results in strong gradients at object boundaries because the AC is evolved on the image scene determined by the embedding of the time series data. The set of orthonormal eigenvectors obtained via SE preserve both local and global image similarities [19], so an SE approach might also yield improved region- based statistics, which in conjunction with stronger boundary gradients results in an improved hybrid active contour scheme.

3.2 Previous Related Work and Novel Contributions

Several studies have shown that quantitative morphological features extracted from breast lesions are helpful for distinguishing between benign and malignant breast lesions [28,37]. Typically, a radiologist's expert delineation of the lesion boundary is considered the gold standard for lesion segmentation. However, manual segmentation is notoriously susceptible to inter-rater variability in breast MRI interpretation [28, 38].

Because accurate lesion segmentation is time consuming, yet necessary for quantitative lesion analysis, many groups have explored various automated segmentation

methods for breast DCE-MRI [35, 36, 39–41]. Szabo et al. [36] used a pixel-wise classifier that used dynamic contrast signal intensities in conjunction with an artificial neural network to identify lesion areas of interest. Other methods that have also used pixel-wise classifiers for segmentation include Twellmann et al. [39], who used the dynamic contrast signal intensities in conjunction with a support vector machine (SVM) classifier and Chen et al. [35], who used a fuzzy c-means (FCM) clustering scheme. Additionally, Wu et al. [40] clustered the time-series data of breast DCE-MRI using Markov random fields. The drawback of such pixel-wise classification methods is that a continuous lesion boundary is only implicitly defined by a cluster of similarly classified pixels, and most of these schemes require post-processing morphological operations such as hole-filling and dilation to provide a smooth, continuous lesion boundary.

An alternative to pixel based approaches is to use an AC model, which guarantees a closed contour. In Shi et al. [41], the image data is first transformed from the intensity space into an alternative data representation by fuzzy c-means. The AC is then initialized and evolved on the resulting FCM clustered space. This allows for incorporation of the time-intensity curve information, spatial clustering of the data by FCM, and the guarantee of a continuous lesion boundary by the introduction of the AC. Shi et al. [41] found that by combining the AC with the FCM method, the lesion segmentation was closer to the manual radiologist segmentation than the method using FCM alone in a cohort of breast cancer patients imaged pre- and post- neoadjuvant therapy.

All of the aforementioned methods operate on the scalar grayscale image intensities. Time-series data such as DCE-MRI contains multiple time points over which the image of the lesion of interest is captured, and typically, only a single time point (usually the time point at which the lesion maximally enhances) is used for segmentation. However, implementations of the AC model have been developed for multi-dimensional images [42–45]. Chan et al. [43] demonstrated an extension of the original scalar image-based CV model to use with vector-valued images. Rousson and Deriche [44] also demonstrated a vector-valued active contour. In a recent application to a medical imaging problem, Xu et al. [45] developed a tensor gradient based AC for use with histopathological images by computing the gradient from vectorial images, and hence

representing the image gradient as tensors. Xu et al. [45] showed that the tensor gradient more completely captured the gradient information in a multi-channel image than using a single channel of the image, thus yielding a more discriminating AC scheme.

In this paper, we present a new hybrid AC model, called SE based AC (SEAC), that employs an alternative data representation derived from SE. In the current study, the tensor-based gradients are derived from the SE eigenvectors, which are fed into the energy functional of a hybrid AC model. We leverage the findings in Xu et al. [45] that the tensor gradient derived from the vectorial image provides stronger gradients for driving the AC model than the corresponding gradient derived from a scalar image. We also believe that the SE data representation provides more descriptive region statistics than could be derived from a scalar grayscale image. Our method for lesion segmentation on DCE-MRI allows for the use of dynamic time series data in the differentiation between normal and lesion tissue via SE. This method is not limited to DCE-MRI data alone and could potentially be used for lesion detection and segmentation on other types of longitudinal or time-series data as well as other types of multi-parametric imaging (e.g., T_1-, T_2-, and diffusion-weighted imaging).

The remainder of the paper is organized as follows: In Section 3, we describe the theory and intuition behind computing gradients in the spectral embedding space, In Section 4 we describe the algorithm and implementation of SEAC. In Section 5, we discuss our experimental design and our boundary-, area-, and classifier-based methods for evaluating segmentation accuracy. In Section 6, we report our results and discussion, and in Section 7, some concluding remarks are presented.

3.3 Theory and Intuition behind Computing Gradients in the Spectral Embedding Space

In this section, we review the calculation of the pixel-wise SE as well as the subsequent tensor gradient calculation.

Table 1: Description of Notation	
Symbol	**Description**
$\mathcal{C}$	2D image scene
C	2D Cartesian grid of pixels $c = (x, y)$
$\mathbf{v}(c)$	Eigenvectors associated with pixel c
ρ_{mq}	Directional component of local structure tensor
$\mathbb{C}$	The zero level set $\mathbb{C} = \{c \in \Theta : \phi(c) = 0\}$
$\gamma(\mathbf{v}(c))$	Tensor gradient computed on $\mathbf{v}$
$H(\phi)$	Heaviside function $H(\phi) = \begin{cases} 1, & \phi(c) \geq 0; \\ 0, & \phi(c) < 0. \end{cases}$
$\delta(\phi)$	Delta function $\delta(\phi) = \begin{cases} +\infty, & \phi(c) = 0; \\ 0, & \phi(c) \neq 0. \end{cases}$
$\Theta_h, h \in \{1, 2\}$	Region membership $\Theta_h = \begin{cases} \Theta_1, & \phi(c) \geq 0; \\ \Theta_2, & \phi(c) < 0. \end{cases}$
$\phi(t; c)$	The level set function

3.3.1 Manifold Learning via Spectral Embedding

Let $\mathcal{F} = [\mathbf{F}(\mathbf{x}_1), \mathbf{F}(\mathbf{x}_2), \ldots, \mathbf{F}(\mathbf{x}_N) \in \mathbb{R}^{N \times D}]$ be the data matrix of N feature vectors with dimensionality D. The aim of SE is to reduce $\mathcal{F} \in \mathbb{R}^{N \times D}$ to a low d-dimensional space where $d << D$. Let $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N \in \mathbb{R}^{N \times d}]$ be the optimal low dimensional projections [19] and the associated eigenvectors of a given object $\mathbf{x}_i$, where $i \in \{1, \ldots, N\}$, are $\mathbf{v}_i \in \mathbb{R}^{1 \times d}$ where $\mathbf{v}_i = [v_1, v_2, \ldots, v_N]$ and $v_j \in \mathbb{R}^{1 \times 1}, j \in \{1, \ldots, d\}$ is an individual eigenvector for a given $\mathbf{v}_i$. The optimal $\mathbf{V}$ can be obtained by solving,

$$\mathbf{V} = \arg\min_{v}\left(\frac{\sum_r \sum_s \|\mathbf{v}(r) - \mathbf{v}(s)\|^2 w_{rs}}{\sum_r \mathbf{v}(r)^2 \mathbf{d}(r)}\right), \tag{3.1}$$

where w_{rs} is the (r, s) element of the weight matrix $W = [w_{rs}] \in \mathbb{R}^{N \times N}, r, s \in \{1, \ldots, N\}$, which assigns edge weights to characterize similarities between the pairwise observations $\mathbf{x}_r$ and $\mathbf{x}_s$ and $\mathbf{d}(r) = \sum_s w_{rs}, r, s \in \{1, \ldots, N\}$. The graph edge weight of two nodes, r and s, can be formulated as $w_{rs} = e^{\frac{-\|\mathbf{F}(\mathbf{x}_r) - \mathbf{F}(\mathbf{x}_s)\|_2^2}{\sigma_I^2}}$ where σ_I^2 is a scaling parameter. The minimization of Equation (1) reduces to a minimum eigenvalue

decomposition problem,

$$(\mathbf{D} - W)\mathbf{v} = \lambda \mathbf{D}\mathbf{v}, \tag{3.2}$$

where $\mathbf{D}$ is a diagonal matrix, $\mathbf{D}_{rr} = \sum_s W_{rs}$. The 3 eigenvectors of $\mathbf{v}_i = [v_1, v_2, v_3]$ where $i \in \{1, \ldots, N\}$ associated with the smallest eigenvalues, $\lambda_1, \lambda_2, \lambda_3$, are used in conjunction with SEAC.

3.3.2 Tensor gradients in the Spectrally Embedded Space

Following the calculation of the eigenvectors by solving the minimization of Equation (1), the gradients of the embedding vectors can be subsequently calculated along the spatial coordinates axes, which results in a tensor gradient function,

$$\nabla \mathbf{V} = \nabla[\arg\min_{v}(\frac{\sum_r \sum_s \|\mathbf{v}(r) - \mathbf{v}(s)\|^2 w_{rs}}{\sum_r \mathbf{v}(r)^2 \mathbf{d}(r)})]. \tag{3.3}$$

The tensor gradient function employed is inspired by the Cumani operator [46], a second-order differential operator for for vectorial images, based on the Di Zenzo multi-valued geometry [47]. For an eigen image $\hat{\mathcal{C}} = (C, \mathbf{v})$, where $\mathbf{v}(c)$ is the associated set of eigenvectors at pixel $c \in C$, the L_2 norm of $\mathbf{v}$ at each $c \in C$ can be written in matrix form as

$$d\mathbf{v}^2 = \begin{bmatrix} dX \\ dY \end{bmatrix}^T \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} dX \\ dY \end{bmatrix}, \tag{3.4}$$

where

$$\begin{aligned} \rho_{11} &= (\frac{\partial \mathbf{v}}{\partial X})^T (\frac{\partial \mathbf{v}}{\partial X}) = (\frac{\partial v_1}{\partial X})^2 + (\frac{\partial v_2}{\partial X})^2 + (\frac{\partial v_3}{\partial X})^2 \\ \rho_{12} &= \rho_{21} = (\frac{\partial \mathbf{v}}{\partial X})^T (\frac{\partial \mathbf{v}}{\partial Y}) = \frac{\partial v_1}{\partial X} \cdot \frac{\partial v_1}{\partial Y} + \frac{\partial v_2}{\partial X} \cdot \frac{\partial v_2}{\partial Y} + \frac{\partial v_3}{\partial X} \cdot \frac{\partial v_3}{\partial Y} \end{aligned} \tag{3.5}$$

ρ_{22} is defined similarly to ρ_{11} along the Y-axis. It is important to note that the gradients, ρ_{11}, ρ_{12}, and ρ_{22} are computed on the eigenvectors in the embedding space. The matrix $[\rho_{mq}] = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$ is the first fundamental form in vector eigenspace and is also referred to as the local structure tensor. For the matrix $[\rho_{mq}]$, the maximum and minimum eigenvalues of the matrix ($\tilde{\lambda}_+$ and $\tilde{\lambda}_-$) representing the extreme rates of

change in the direction of their corresponding eigenvectors. $\tilde{\lambda}_+$ and $\tilde{\lambda}_-$ may be formally expressed by $\tilde{\lambda}_\pm = (\rho_{11} + \rho_{22} \pm \sqrt{\triangle})/2$, where $\triangle = (\rho_{11} - \rho_{22})^2 + 4\rho_{12}^2$. The tensor gradient is defined as [42],

$$\gamma(\mathbf{v}(c)) = \sqrt{\tilde{\lambda}_+ - \tilde{\lambda}_-}. \tag{3.6}$$

From Equations (3.4)-(3.6), it is easy to show that the gray scale gradient $\sqrt{\frac{\partial^2 v_j}{\partial X^2} + \frac{\partial^2 v_j}{\partial Y^2}}$, where $j \in \{1, 2, 3\}$, (widely employed for edge detection) is a special case of the tensor gradient $\gamma(\cdot)$.

Figure 3.2 illustrates the comparison between use of fuzzy c-means (FCM) clustering [35, 41]to drive the active contour Figure (3.2(b)) with the SE-based image representation driven AC (3.2(d)). In comparing Figure 3.2(a) and Figure 3.2(c), one can observe that the tensor gradient at the lesion boundary (shown in the inset in Figure 3.2(b), 3.2(d)) is not strong along the entire lesion boundary in the FCM-derived gradient (Figure 3.2(a)) compared to the gradient computed in the SE space (3.2(c)). However, the tensor gradient is strong along the entire perimeter of the lesion in the SE-based gradient image (Figure 3.2(c)). This leads to a more accurate segmentation of the lesion using SEAC (Figure 3.2(d)) than using the AC in conjunction with FCM (Figure 3.2(b)).

3.4 Spectral Embedding-based Active Contour (SEAC)

In this section, we briefly describe the hybrid active contour model obtained by combining the boundary and region driven energy terms from the spectrally embedded space.

3.4.1 Hybrid Active Contour Energy Function

A hybrid AC model can be employed to incorporate both gradient- and region-based information into the AC model. From the previous section, we can formulate the gradient as $g(c) = \frac{1}{1+\gamma^2}$. The region based component relies on modeling the background and foreground regions as parametric distributions, $p(\mathbf{v}(c)|\Theta_h), h \in \{1, 2\}$. Here the

parameters $\Theta_h = \{\mu_h, \Sigma_h\}$ are approximated as multivariate Gaussian distributions. The energy functional of hybrid AC model can be expressed as,

$$E(\partial\mathcal{C}, \{\Theta_1, \Theta_2\}, \phi) = \int_{\mathcal{C}} \{-\alpha \left[H(\phi) \log p(\mathbf{v}(c)|\Theta_1) + (1 - H(\phi)) \log p(\mathbf{v}(c)|\Theta_2)\right] + \beta g \delta(\phi) |\nabla\phi|\} \, d\mathcal{C}. \tag{3.7}$$

Using calculus of variations, the curve evolution function can be derived by minimizing the energy function (3.7):

$$\frac{\partial\phi}{\partial t} = |\nabla\phi| \{\alpha \left[\log p(\mathbf{v}(c)|\Theta_2) - \log p(\mathbf{v}(c)|\Theta_1)\right] + \beta div(g \frac{\nabla\phi}{|\nabla\phi|})\}, \tag{3.8}$$

where $H(\phi)$ is the Heaviside function, Θ_1 and Θ_2 are the image foreground and background, respectively, $\phi(t; c)$ is the level set function, α and β are positive constant parameters, and $\delta(\phi)$ is the Delta function. From an initial contour ϕ_0, the curve evolution function in Equation (3.8) is evolved until model convergence.

3.4.2 Algorithm for SEAC

A summary of the tensor gradient and active contour algorithms are outlined here.

Algorithm: *SETensorGradient*()

Input: Image scene $\mathcal{C} = (C, \mathcal{F})$.

Output: Eigen image scene with associated tensor gradient: $\hat{\mathcal{C}} = (C, \gamma)$.

begin

0. Initialize $\mathcal{C} = (C, \mathcal{F})$;
1. Apply SE [19] to $\mathcal{F} = [\mathbf{F}(\mathbf{x}_1), \mathbf{F}(\mathbf{x}_2), \ldots, \mathbf{F}(\mathbf{x}_n) \in \mathbb{R}^{N \times D}]$ to obtain $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d \in \mathbb{R}^{N \times d}]$;
2. Compute gradients of the embedding vectors in XY plane, (v_1, v_2, v_3): $\sqrt{\frac{\partial^2 v_j}{\partial X^2} + \frac{\partial^2 v_j}{\partial Y^2}}$, where $j \in \{1, 2, 3\}$;
3. Calculate tensor gradient: $\gamma(\mathbf{v}(c)) = \sqrt{\tilde{\lambda}_+ - \tilde{\lambda}_-}, \forall c \in C$;

end

Algorithm: *SEAC*

Input: $\hat{\mathcal{C}} = (C, \gamma)$.

Output: Final AC: ϕ_T.

begin

0. SE Tensor Gradient($\hat{\mathcal{C}}$);
1. Calculate $g(c) = \frac{1}{1+\gamma^2}$;
2. Model $p(\mathbf{v}(c)|\Theta_h), h \in \{1, 2\}$ as multivariate Gaussians.;
3. Formulate energy functional using $p(\mathbf{v}(c)|\Theta_h), h \in \{1, 2\}$ and $g(c)$.;
4. Initialize $\phi(t; c) = \phi_0$ where $t = 0$;
5. *while* $|\phi(t; c) - \phi(t-1; c)| > \epsilon$ *do*
6. Minimize curve evolution function, $\frac{\partial \phi}{\partial t}$
7. $t = t + 1$;
8. *endwhile*;
9. $\phi_T = \phi(t; c)$;

end

3.5 Experimental Design

3.5.1 Data Description

A total of 50 (40 malignant, 10 benign) breast DCE-MRI studies were obtained from the Hospital at the University of Pennsylvania. All of these were clinical cases where a screening mammogram revealed a lesion suspicious for malignancy. All studies were collected under Institutional Review Board approval, and lesion diagnosis was confirmed by biopsy and histologic examination. Sagittal T1-weighted, spoiled gradient echo sequences with fat suppression consisting of one series before contrast injection of Gd-DTPA (precontrast) and 3 to 8 series after contrast injection (postcontrast) were acquired at 1.5 Tesla (Siemens Magnetom). Single slice dimensions were 384×384 or 512×512 with a slice thickness of 3 cm. Temporal resolution between postcontrast acquisitions was a minimum of 90 seconds. The region of interest (ROI) associated with the lesion was then manually segmented via MRIcro imaging software [48] by an

attending radiologist (M.A.R.) with expertise in MR mammography. The radiologist selected a 2D slice of the MRI volume that was most representative of each lesion, and the analyses were performed only for that 2D slice.

For each pixel, c, in each image, a dynamic signal intensity vector was created consisting of the signal intensity values of the pixel and each time point in the time series. $\mathcal{F}^t$ is the function that assigns a signal intensity value at every pixel c at each time point $t \in \{0, 1, 2, \ldots, T-1\}$ where T is the number of time points the DCE-MRI time series. $t = 0$ refers to the time at which the precontrast image is acquired and $t \in \{1, \ldots, T-1\}$ refer to the times at which the subsequent postcontrast images are acquired.

3.5.2 Comparative Segmentation Strategies

Fuzzy c-means (FCM) clustering as described in Chen et al. [35] is widely used for automated segmentation of breast lesions on DCE-MRI [49, 50]. FCM is a data clustering scheme similar to k-means in that data are clustered around a prescribed number of centroids. However, unlike k-means, the resulting class membership is a fuzzy membership to each cluster. We implement the method in Shi et al. [41] (referred to as FCM+AC) to compare the AC driven by FCM to that driven by SE in SEAC.

3.5.3 Measures for Evaluating Segmentation Performance

In order to evaluate the accuracy of the automated lesion segmentation resulting from SEAC, the final SEAC lesion boundaries were compared to the manual segmentation provided by a radiologist. The manual segmentation was considered here to be the ground truth (GT) for the lesion boundary. The lesion boundary resulting from the automated segmentation is denoted as $G_1^a(\hat{\mathcal{C}})$ and the set of pixels inside $G_1^a(\hat{\mathcal{C}})$ is $G_2^a(\hat{\mathcal{C}})$. The lesion boundary resulting from the manual segmentation is denoted as $G_1^b(\mathcal{C})$, and the set of pixels inside $G_1^b(\mathcal{C})$ is denoted as $G_2^b(\mathcal{C})$. Lesion boundary accuracy was evaluated by boundary, area, and classifier based performance measures, described below.

3.5.4 Boundary-based

Mean absolute difference (MAD) is calculated by evaluating the mean difference between each point, c_z, on $G_1^a(\hat{\mathcal{C}}) = \{c_z | z \in \{1, \ldots, Z\}\}$ (SEAC or FCM+AC) where $Z = |G_1^a(\hat{\mathcal{C}})|$ and $|\cdot|$ is the cardinality of any set, and the corresponding closest point on $G_1^b(\mathcal{C}) = \{c_\psi | \psi \in \{1, \ldots, |G_1^b(\mathcal{C})|\}$ such that,

$$MAD = \frac{1}{Z} \sum_{z=1}^{Z} [\min_{\psi} \|c_z - c_\psi\|_2]. \tag{3.9}$$

Lower values of MAD reflect a more similar segmentation to the ground truth manual segmentation.

3.5.5 Area-based

Dice similarity coefficient (DSC) is calculated as follows:

$$DSC = \frac{2|G_2^a(\hat{\mathcal{C}}) \cap G_2^b(\mathcal{C})|}{|G_2^a(\hat{\mathcal{C}})| + |G_2^b(\mathcal{C})|}. \tag{3.10}$$

The closer the DSC value is to 1, the more similar the automated lesion segmentation is to the GT segmentation.

3.5.6 Classifier-based

Because an accurate lesion segmentation is necessary for accurate morphological feature extraction [28, 37], improved classification accuracy would be the ultimate test to demonstrate SEAC's improvement over previous methods. Morphological features [2] based on the lesion contour are extracted and used in conjunction with a support vector machine (SVM) classifier to determine if morphological features based on SEAC contours can yield similar classifier accuracy to morphological features extracted from (a) manually segmented boundaries; and (b) determine if morphological features based on SEAC segmentations will result in higher classifier accuracy compared to morphological features based on FCM+AC segmentations of breast lesions. We calculated 6 morphological features [2] on 40 datasets (10 benign fibroadenomas; 13 ER+; 7 HER2+; 10 triple negative, ER-/PR-/HER2-) for which molecular phenotyping on histopathology had been performed. The features extracted [2] include:

1. Area overlap ratio;
2. Normalized average radial distance ratio;
3. Standard deviation of normalized distance ratio;
4. Variance of distance ratio;
5. Compactness; and
6. Smoothness.

The morphological feature calculation requires a pre-defined lesion boundary, so the boundaries resulting from SEAC, FCM+AC, and the manual segmentation were used for morphological feature extraction. The features for each lesion were then used in conjunction with a SVM classifier [51] with leave-one-out cross validation to determine the lesion diagnosis.

Although the SVM classifier typically produces a hard classification [51], a pseudo-threshold can be generated for the SVM by converting the distance of each object classified to the SVM decision hyperplane into a soft likelihood of belonging to the object class [2]. Thus the greater the distance of the object from the hyperplane, the higher the likelihood that it belongs to a particular class whereas the proximity of an object to the hyperplane reflects higher ambiguity in the class assignment. The receiver-operator characteristic curves (ROC) curves for SVMs can be generated by varying the location of the decision hyperplane. As the distance of the objects from the decision hyperplane changes, the corresponding object-class probabilities also change. At each location of the decision hyperplane, classification sensitivity and specificity estimates are obtained. The trade-off between true positive rate (Y-axis of ROC curve) and false positive rate (X-axis of ROC curve) obtained at each of the different locations of the hyperplane is used to generate a ROC curve. The area under the curve (AUC) for the ROC curves for the extracted morphological features sets derived from each of the segmentation methods is reported in Table 3.

3.6 Results and Discussion

3.6.1 Comparing SEAC and FCM + AC via Boundary- and Area-based Metrics

Figure 3.3 shows the comparison between the ground truth boundary for the lesion obtained via a radiologist and the two automated segmentation methods. The SEAC segmentations in Figures 3.3(c), (f), and (i) are more similar to the manual segmentations in Figures 3.3 (a), (d), and (g) compared to the FCM initialized AC (FCM+AC) segmentations in Figures 3.3 (b), (e), and (h). In particular, Figure 3.3(f) shows that the SEAC model was able to capture the spicularity of the lesion whereas the FCM+AC segmentation (Figure 3.3(e)) was not. The mean and variance of both MAD and DSC for FCM+AC and SEAC in segmenting lesions over 50 studies are shown in Table 2 and demonstrate that SEAC is closer to the GT segmentation in terms of both MAD, a boundary-based metric, and DSC, a region-based metric.

Table 2: Evaluation of SEAC vs. FCM+AC

Segmentation Method	*MAD* ($\mu \pm \sigma$)	*DSC* ($\mu \pm \sigma$)
FCM + AC	7.2±7.4	0.58±0.32
SEAC	$\mathbf{3.2 \pm 2.1}$	$\mathbf{0.74 \pm 0.13}$

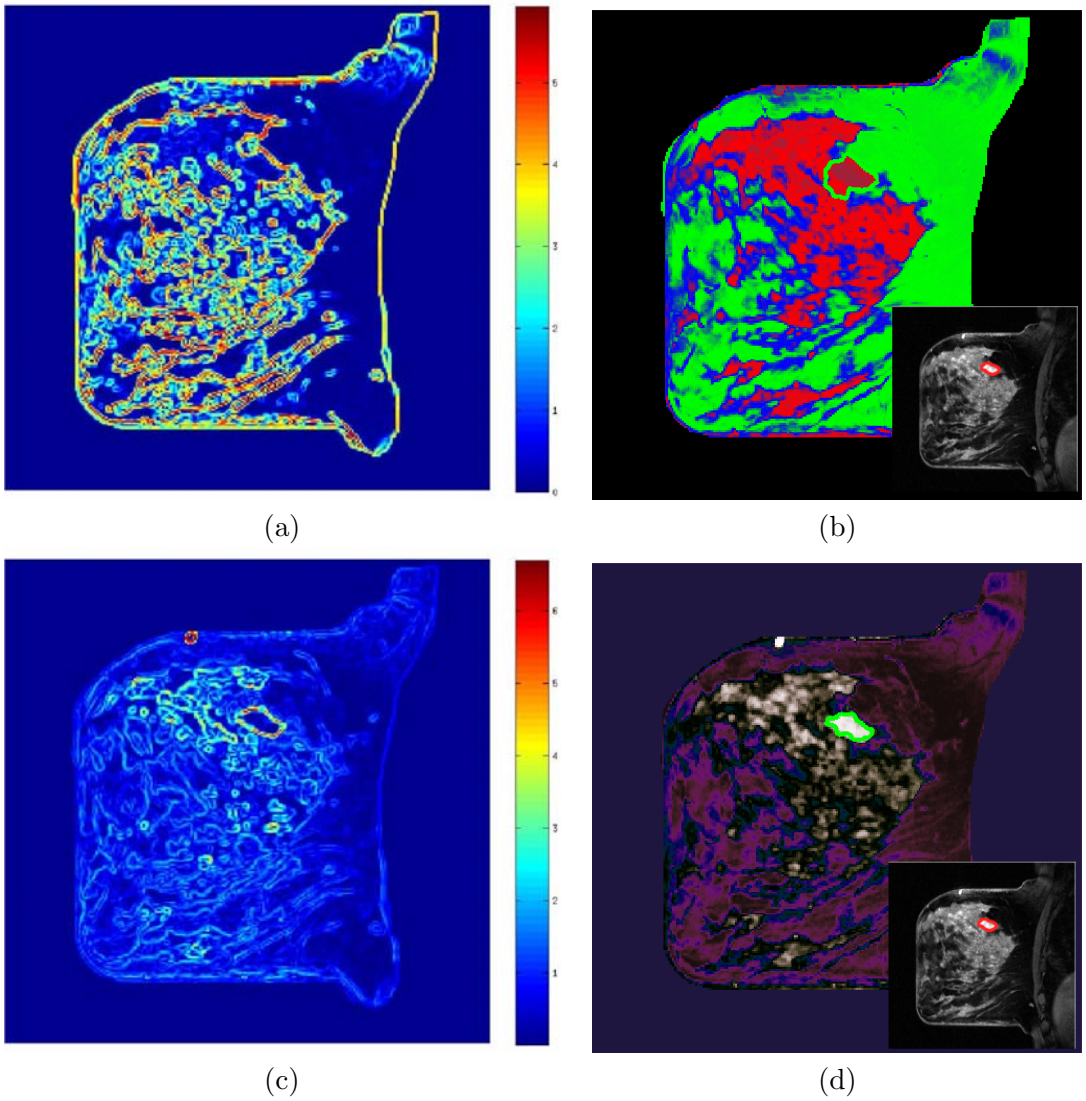
3.6.2 Classifier Performance based Comparison of SEAC and FCM + AC

The AUC values calculated based on the SVM classifier performance in discriminating between benign and malignant lesions via morphological features extracted (a) manually, (b) via FCM+AC, and (c) via SEAC are shown in Table 3. The AUC for the manual segmentation was best, whereas the AUC for the SEAC morphological features was second, and the morphological features based on FCM+AC had the lowest AUC. In conjunction with the results in Table 2, it is clear that accurate lesion segmentation leads to better classifier performance when the lesion classifier is based on morphology features.

Table 3: Benign vs. Malignant Lesion Classification	
Segmentation Method	*AUC*
Manual Segmentation	0.83
FCM + AC	0.65
SEAC	0.73

3.7 Concluding Remarks

In this paper we presented a new active contour (AC) model (spectral embedding based AC, (SEAC)) involving use of an alternative image representation obtained via a manifold learning scheme that results in stronger boundary gradients and improved region statistics, in turn providing improved stopping criteria for the AC. SE transforms the high dimensional DCE-MRI time series data to a reduced dimensional space that is comprised of an orthogonal basis set of eigenvectors. This transformed space provides strong tensor gradients and improved region statistics compared to those that might be obtained from the original grayscale image alone. On a cohort of 50 breast DCE-MRI studies, we showed that SEAC outperformed segmentations based on a fuzzy c-means based AC. We also demonstrated that on a cohort of 40 lesions, the morphological features derived from SEAC yielded better lesion classification compared to morphological features derived from FCM-based segmentation. While in this work we demonstrate the use of SEAC with breast DCE-MRI data, SEAC could be easily applied to segmenting structures on other high dimensional, time-series imaging data as well. In future work, we plan to automatically optimize the parameters of the energy function in the AC and expand the data cohort to test the robustness of SEAC with a diverse group of lesion types.



(a) (b)

(c) (d)

Figure 3.2: Comparison of segmentation using (b) FCM+AC, and (d) SEAC. The tensor gradient associated with the FCM image is shown in (a) and the tensor gradient associated with SEAC is shown in (b). Note that in regions of the automated segmentation driven by the FCM image, the areas where the contour fails to stop are the locations where the gradient is weak (a),(b). Conversely, the tensor gradient derived from SE is strong at all locations along the lesion boundary, resulting in a final automated segmentation that is very similar compared to the manual segmentation (inset image in (b) and (d)).

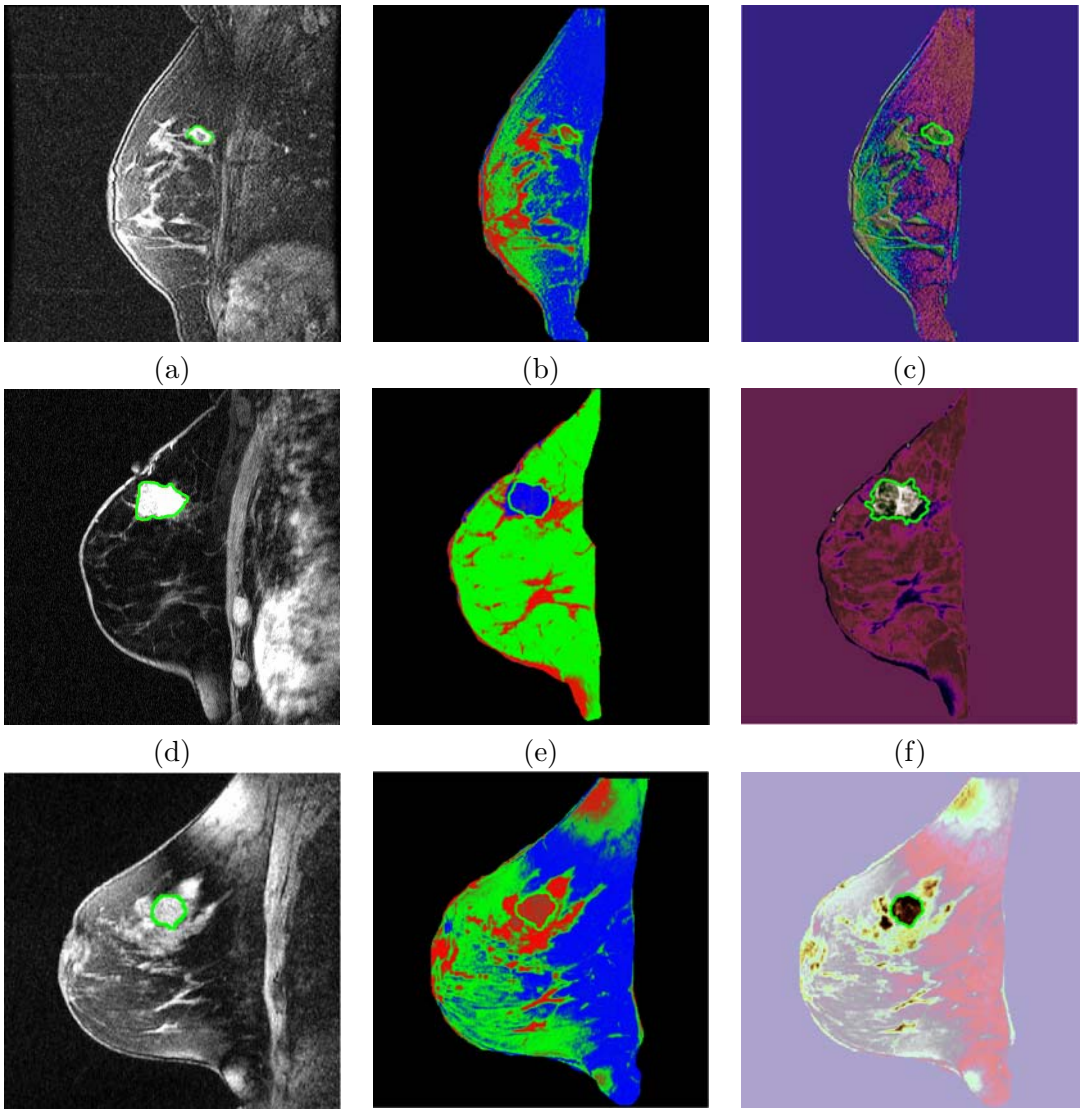


Figure 3.3: Comparison of segmentation of two different invasive ductal carcinomas and one benign lesion via (a), (d), (g) manual delineation, (b), (e), (h) FCM+AC, and (c), (f), (i) SEAC. Note that in (b), (e), and (h) the FCM image does not provide a strong enough gradient at the lesion boundary and hence is unable to discriminate between the lesion and the other breast tissue. However, SEAC (c), (f), (i) yields a lesion segmentation that is very close to the manual delineation of the tumor in (a), (d), (g). Note that the colormaps displayed for both the FCM and SEAC methods ((b), (e), (h) and (c), (f), (i), respectively) only reflect the pixel similarities as determined by the 2 schemes, pixels with similar time-intensity curves being assigned similar colors.

Chapter 4

Distinguishing Benign from Malignant Breast Lesions

4.1 Introduction

Magnetic Resonance Imaging (MRI) is currently used as a complement to conventional x-ray mammography in diagnosis of breast lesions [8]. X-ray mammography remains the gold standard for breast cancer screening and offers high 2-dimensional resolution, which is advantageous for detecting small variations in tissue composition, such as microcalcifications [52]. However, due to the constraints of imaging a 3-dimensional structure in a single plane, ultrasound or breast dynamic contrast-enhanced (DCE)-MRI is often used as a secondary imaging technique when a suspicious lesion is found on mammography [52]. Ultrasound is very good at detecting tissue composition and hence is able to provide additional information to the mammogram in situations where the breast tissue is dense or a cystic mass needs to be ruled out [53]. DCE-MRI is also very good at imaging dense breasts, but its major advantages over mammography and ultrasound are the ability to: (a) image the entire breast as thin slices that comprise the entire breast volume and (b) measure variations in contrast uptake that provide information about the vascularity of the breast tissue. Since malignant tumors often have a high density of blood vessels that are poorly formed and thus leaky, they take up contrast dye at a different rate from benign lesions, allowing radiologists to distinguish malignant from benign lesions based on corresponding differences in contrast kinetics [8, 54].

On account of breast DCE-MRI's high 3-dimensional resolution and its ability to acquire kinetic contrast information, its lesion detection sensitivity is close to 100% [10], much higher than that of either mammography or ultrasound [8]. However, specificity of breast DCE-MRI is low, with rates of between 30% and 70% [10, 49] having been

reported. High false positive detection rates on MRI often lead not only to anxiety for the patient, but may also result in an unnecessary invasive biopsy [8, 10]. In a review of the literature, Saslow et al. [8] found biopsy rates were between 3 and 15% when MRI was used for breast cancer screening. In addition to the problem of low specificity, another shortcoming of breast MRI is that only experienced radiologists are able to accurately distinguish benign from malignant tumors [8, 28]. This often leads to high rates of interobserver variability [28]. Ikeda et al. [28] reported this variability with kappa statistics between 0.21 and 0.40, where 1.0 represents complete agreement and 0.0 represents no agreement above the level expected by random chance. Therefore, one of the challenges in facilitating increased acceptance of breast DCE-MRI as a screening modality is reducing false positive detection errors, thereby boosting detection specificity. Additionally, the interobserver variability for breast DCE-MRI must be minimized.

To address the issues of low specificity and high interobserver variability in breast DCE-MRI, the American College of Radiology proposed the Breast Imaging Reporting and Data System (BIRADS) [9], a semi-quantitative classification protocol for evaluating breast lesions. Lesions are evaluated on the basis of shape, margin morphology, internal enhancement, and kinetic or time-intensity curve characteristics [37, 38]. Assuming that the imaging is complete, the radiologist gives each lesion seen on DCE-MRI a score between 1 and 6, where 1 is *Negative* and 6 is *Known cancer* [9] based on the combination of lesion characteristics. Although the BIRADS system has helped to standardize the diagnosis of breast lesions, studies [8, 28, 38] continue to report significant variability in lesion interpretation among radiologists.

The remainder of this paper is organized as follows. In Section 2, we discuss the previous work in the analysis of breast DCE-MRI and computer-aided diagnosis (CAD) for breast DCE-MRI as well as the motivation for the methods proposed in this paper. In Section 3, we provide a description of the data and notational convention employed and also describe our feature extraction schemes. We also provide details on the classifier methods used to quantify feature performance in Section 3. In Section 4, the experiments performed and the metrics used to evaluate the features are described.

Quantitative and qualitative results showing the performance of the individual descriptors are presented in Section 5. Concluding remarks and future directions are presented in Section 6.

4.2 Previous Work and Motivation

In clinical decision-making, changes in signal intensity kinetics are an important descriptor for breast lesion characterization in DCE-MRI [28,37,38,52,54,55]. DCE-MRI involves first injecting a contrast agent such as gadolinium diethylenetriamine-pentaacid (Gd-DTPA) into the patient's bloodstream and concurrently acquiring a time series of MR images of the breast. Since malignant lesions tend to grow leaky blood vessels in abundance, the contrast agent is taken up by tumors preferentially [56]. Kuhl et al. [54] found that data in the time series MR images could be plotted as single data points on a temporal curve, where the shape of the curve is reflective of the lesion class. It was shown [54] that malignant lesions had a characteristic curve with a steep positive initial slope indicating rapid uptake of contrast agent and a subsequent negative slope indicating rapid washout. Benign lesions had slow contrast uptake (small positive initial slope) and then plateaued or did not reach peak intensity during the image acquisition period. Although signal intensity kinetics offer a great advantage to DCE-MRI for studying the functional attributes of breast lesions compared to other modalities, features derived from contrast enhancement data contribute to the high false positive rates reported for breast DCE-MRI [10]. For instance, while both benign and malignant neoplastic tissue frequently have contrast enhancement patterns that differ from normal breast tissue, it is often difficult for radiologists to differentiate between benign and malignant lesions simply by visually inspecting the contrast-enhanced lesion on the postcontrast MRI. Consequently, several quantitative and semi-quantitative models have been proposed to measure the manner in which a lesion takes up the contrast dye [6,35,36,39,55,57–68].

Several computer based image analysis systems have been proposed [35, 36, 39, 57–67, 69, 70] with the aim of reducing interobserver variability on breast DCE-MRI. CAD approaches for breast MRI are typically either for automatically (a) detecting (Computer aided detection, CADe) [35, 36, 39, 57–59, 69, 70] or (b) classifying a lesion as

benign or malignant (Computer aided diagnosis, CADx) [35, 36, 60–67]. Automated CADe approaches usually exploit the fact that malignant lesions typically have different signal intensity kinetic profiles on DCE-MRI compared to normal parenchyma [35, 36, 39, 57–59, 69, 70]. Some of these methods have been shown to have a detection accuracy comparable to manual detection [35, 36, 39, 57–59, 69, 70]. However, a CADx system assumes that the lesion detection has been solved either manually or via CADe, and it is typically comprised of two modules: (1) a quantitative feature extractor and (2) a classifier that employs the attributes extracted from the lesion to discriminate lesion classes. Several different CADx classifiers for DCE-MRI have been proposed including linear discriminant analysis [58], artificial neural networks [36, 49, 60, 61], and support vector machine (SVM) classifiers [62]. Feature descriptors employed by CADx systems have typically included morphological [61], lesion texture [58, 64], contrast enhancement [35, 36, 62], or a combination of morphological and contrast enhancement descriptors [49, 60, 63]. In [61], Meinel et al. found mean volume, area, radial length, spiculation, perimeter length, and compactness to be among the best morphological features, and their results with a back-propagation neural network classifier yielded an AUC of 0.9748 on a dataset of 80 lesions using the leave-one-out method. Zheng et al. [58] reported an AUC of 0.97 using temporal enhancement texture features on a cohort of 36 lesions. Gibbs et al. [64] reported an AUC of 0.80 using co-occurrence texture features. However, by including patient data in their regression model, they achieved an accuracy of 92% and an AUC of 0.92. Using contrast enhancement alone, Chen et al. [35] achieved an AUC of 0.85 over 121 studies, and Levman et al. [62] obtained an AUC of 0.74 using empirical enhancement features such as signal enhancement ratio and time to peak enhancement over a cohort of 94 studies.

The three timepoint (3TP) model [65] and the pharmacokinetic model [6] are two common classifier-based approaches that focus on the kinetic contrast-enhancement data and have been proposed for automated lesion diagnosis on DCE-MRI. The 3TP model developed by Degani et al. [65] assigns a color (red, green, or blue) to the slope of the contrast uptake portion of the kinetic curve and a color intensity between 0.0 and 1.0 to the contrast washout portion of the kinetic curve. The colormaps allow for

a parametric visualization of the contrast enhancement profile by displaying each pixel as red if malignant, blue if benign, and green if suspicious for malignancy. Recently on a dataset of 127 lesions [67], the 3TP model yielded a sensitivity of 75.0% and a specificity of 83.1%, improving somewhat upon previously reported specificity results [10, 49], but sacrificing sensitivity. The pharmacokinetic (PK) model differs from the 3TP method in that it attempts to provide a physiologic interpretation of the data by determining parameters such as K^{trans} (the transfer constant between the plasma and tissue compartments), v_e (the extracellular extravascular volume fraction), and k_{ep} (the ratio of K^{trans}/v_e) [6]. Szabo et al. [36] reported 71% sensitivity and 100% specificity using features derived from the Hayton PK model. Veltman et al. [71] reported an AUC of 0.83 with the Tofts PK model. While K^{trans},v_e, and k_{ep} have been shown to discriminate between lesion classes [6], the computed values are highly sensitive to the choice of initial conditions [68]. Moreover, it has been noted in some recent studies [10, 35, 52] that the heterogeneity of lesion enhancement poses problems for the correct selection of pixels for the calculation of signal enhancement features.

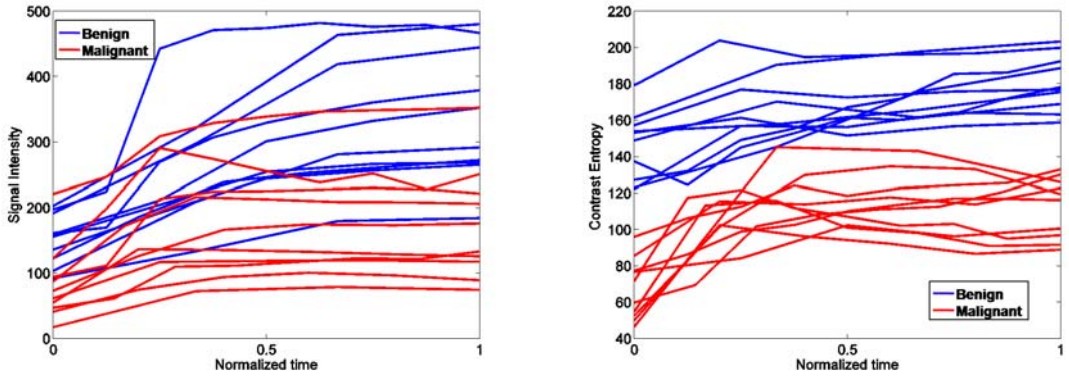
The use of signal intensity kinetic profiles for lesion classification is also limited by other technical hurdles including MR artifacts such as bias field inhomogeneity [72] and intensity nonlinearity [73]. An alternative to using temporal signal intensity profiles to characterize the lesion is by quantifying the lesion texture, a somewhat nebulous term broadly used to refer to localized spatial variations in signal intensity.

Lesion texture has also been acknowledged as an important lesion descriptor as evidenced the incorporation of *internal enhancement* as a BIRADS descriptor for breast MRI lesion classification [9]. While internal enhancement is intended for the assessment of the MR image at a single timepoint after contrast injection, it may also be useful to capture a measure of the change in this feature as a function of contrast enhancement. The concept of studying spatiotemporal textural patterns has been previously explored in [58] and [59]. Zheng et al. [58] computed the discrete Fourier transform coefficients of kinetic changes in Gabor filter features to create parametric maps of the lesions. The authors reported an AUC of 0.97 using a leave-one-out LDA classifier on a cohort of 36 lesions. Woods et al. [59] computed a 4-dimensional co-occurrence matrix to

calculate texture features in a pixel-wise fashion to differentiate between normal and malignant tissue. They were able to distinguish malignant and benign tissue areas with a sensitivity of 96.22% and a specificity of 99.85% on 4 invasive ductal carcinomas.

In this paper, a scheme is presented to analyze *textural kinetic* curves which quantify the spatiotemporal patterns of lesion texture during contrast uptake and diffusion. Instead of reducing the data into a single 2-dimensional image representation as in [58] and [59], the data here is presented in a manner familiar to radiologists and analogous to the signal intensity kinetic curves. Hence, the texture measures associated with the lesion at each pre- and postcontrast enhancement timepoint are plotted on a time series curve. Unlike [58, 59], the textural kinetic curves can be defined in multiple different parametric spaces (including Gabor, first order statistical, and Haralick). Parameters obtained from model fitting these textural kinetic curves are employed in conjunction with a classifier to distinguish lesion classes. This allows for powerful meta-classifiers involving parameters from multiple textural and morphological representations to be easily constructed, unlike in [58, 59], which only consider spatiotemporal changes of certain attributes (Gabor [58] and co-occurrence [59], respectively). To illustrate the discriminability associated with the textural kinetic features, the signal intensity and corresponding textural kinetic curves for a second order textural kinetic feature (Contrast Entropy) for 10 benign (blue curves) and 10 malignant lesions (red curves) were plotted in Figures 1(a) and 1(b), respectively. Each lesion was first manually segmented by an expert radiologist. The mean signal intensity as well as the texture at each pre- and postcontrast timepoint were calculated and then plotted. Figures 1(a) and 1(b) reveal that the textural kinetic feature was able to separate the benign from the malignant lesions better than signal intensity kinetics. The improved separation of lesion classes using the textural kinetic features appears to reinforce the fact that textural representations of image intensity are robust to bias field and intensity non-standardness [73].

The main components of our methodology for assessing the performance of textural kinetic features as a lesion classifier on breast DCE-MRI is illustrated in the flowchart shown in Figure 2. Textural kinetic features are compared to additional signal intensity kinetic, morphologic, and precontrast static texture descriptors of the lesion



(a) (b)

Figure 4.1: (a) Signal intensity and (b) a second order textural kinetic curves (Contrast Entropy) for ten malignant (red) and ten benign (blue) tumors over the course of contrast administration. Time=0 is precontrast; progression along the Time axis denotes postcontrast time points. Time is normalized due to variability in the number of postcontrast timepoints among the datasets.

in distinguishing between benign and malignant breast lesions. The texture operators used include Gabor filters, first order textural features, and 2nd order textural features. Gabor filters have been modeled on the patterning of the human visual cortex [74] and have found widespread application in image analysis [58, 74, 75]. First order textural features (mean, median) give a global picture of lesion enhancement, whereas standard deviation and range yield insight into lesion heterogeneity. Second order textural features, calculated via co-occurrence matrices [76], reflect regional heterogeneity in the lesion. This may be particularly important, for example, in deciding if the malignant-type signal enhancement in a single pixel location in a lesion is an artifact or if there are neighboring pixels with similar signal intensities that corroborate the enhancement characteristics of a malignant lesion. The classification performance of each feature is assessed with both an SVM classifier, which yields a hard benign or malignant classification, and a probabilistic boosting tree (PBT) classifier, which assigns a probability of malignancy to each lesion. The SVM classifier is a kernel-based method that has the ability to capture the nonlinear relationships that occur between biological data [34]. Both the SVM and PBT classifiers have been successfully employed for a variety of biomedical applications [77]. In addition to the classifiers, graph embedding [19], a

nonlinear dimensionality reduction method, is used to observe the clustering of data in a reduced-dimensional embedding space. The separability between the lesion in the reduced dimensional embedding space is directly related to the original high dimensional feature space to which graph embedding is applied.

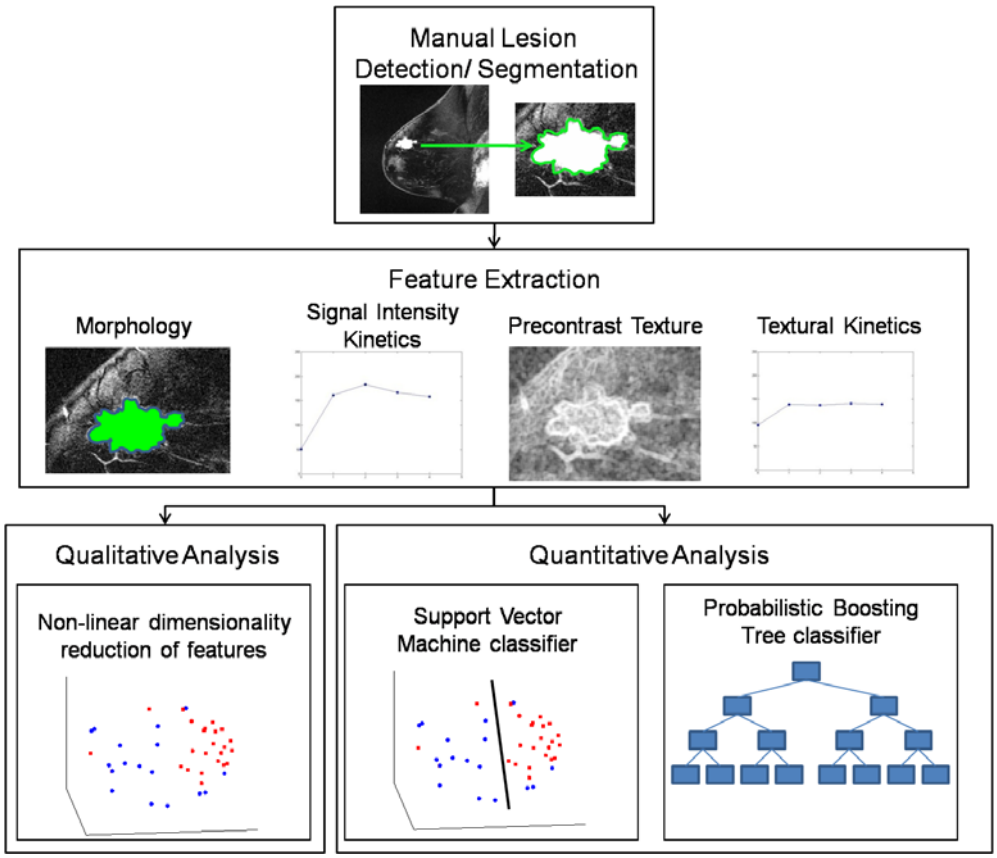


Figure 4.2: Flowchart illustrating the steps comprising the methodology presented in this paper. Following manual lesion detection and segmentation, 3 different feature classes are extracted (morphological, signal intensity kinetics, precontrast texture) to compare against textural kinetics. Quantitative evaluation of the 4 feature classes is done via SVM and PBT classifier accuracy while graph embedding is used for qualitative evaluation.

4.3 Materials and Methods

4.3.1 Data Description

A total of 41 (24 malignant, 17 benign) breast DCE-MRI studies were obtained from the Hospital at the University of Pennsylvania. All of these were clinical cases where a screening mammogram revealed a lesion suspicious for malignancy. All studies were collected under Institutional Review Board approval, and lesion diagnosis was confirmed by biopsy and histologic examination. Sagittal T1-weighted, spoiled gradient echo sequences with fat suppression consisting of one series before contrast injection of Gd-DTPA (precontrast) and 3-8 series after contrast injection (postcontrast) were acquired at either 1.5 Tesla or 3 Tesla (Siemens Magnetom or Trio, respectively). Single slice dimensions were 384×384, 512×512, or 896×896 pixels with a slice thickness of 3cm. Temporal resolution between postcontrast acquisitions was a minimum of 90 seconds. The region of interest (ROI) associated with the lesion was then manually segmented via MRIcro imaging software [48] by an attending radiologist with expertise in MR mammography. The radiologist selected a 2D slice of the MRI volume that was most representative of each lesion, and the analyses were performed only for that 2D slice.

4.3.2 General Notation Used

We define a 2D section of a 3D MRI volume as $\mathcal{C} = (C, f^t)$, where C is a spatial grid of pixels $c \in C$, and f^t is the function that assigns a signal intensity value at every pixel $c \in C$ at each time point $t \in \{0, 1, 2, \ldots, T-1\}$ in the DCE-MRI time series. $t = 0$ refers to the time at which the precontrast image is acquired and $t \in \{1, \ldots, T-1\}$ refer to the times at which the subsequent postcontrast images are acquired. The segmentation performed by the radiologist defines the boundary of the lesion, where the set of boundary points, $\mathcal{R} = \{d^{(0)}, d^{(1)}, \ldots, d^{(n-1)}, d^{(n)}\}$, is a subset of the pixels contained in the lesion, $\mathcal{L}$, where $\mathcal{L} \subset C$. The set of pixels in $\mathcal{R} = \{d^{(0)}, d^{(1)}, \ldots, d^{(n-1)}, d^{(n)}\}$ constitute a closed path such that for $i \in \{0, \ldots, n-1\}, d^{(0)} = d^{(n+1)}$, and $\|d^{(i)} - d^{(i+1)}\|_2 \leq \sqrt{2}$, assuming unit spacing between the pixels in C. The coordinates of the centroid of $\mathcal{L}$ are defined as $\bar{c} = (\bar{x}, \bar{y})$, where $\bar{c} = \frac{1}{|\mathcal{L}|} \sum_{c \in \mathcal{L}} c$ and $|\mathcal{L}|$ is the cardinality of set $\mathcal{L}$. A list

Symbol	*Description*
$\mathcal{C}$	2D MR image scene.
$h_q(c)$	Texture feature for $c \in \mathcal{L}, q \in \{1, \ldots, Q\}$.
C	2D grid of pixels, $c \in C$.
$\mathcal{N}_s(c)$	Square neighborhood of length s associated with each $c \in \mathcal{L}$.
c	Spatial location of a pixel in C, where $c = (x, y)$.
Φ^γ	Feature operator, where $1 \leq \gamma \leq \mathcal{K}$.
$\mathcal{L}$	Set pixels corresponding to a lesion.
$\mathcal{F}^\gamma$	Average feature value for operator Φ^γ over all $c \in \mathcal{L}$.
$\mathcal{R}$	A set of pixels defining the boundary of a lesion, $\mathcal{L}$.
$\phi \in \{SVM, 3TP, PBT\}$	Classifier type.
$\bar{c}$	Centroid of a lesion, defined by the 2D center of mass.
$\mathcal{V}_\phi$	Classifier output where $\mathcal{V}_\phi \in \{-1, +1\}$.
r	Maximum radial distance of the lesion.
$\{N_{TP,\phi}, N_{TN,\phi}, N_{FP,\phi}, N_{FN,\phi}\}$	Number of lesions identified as True Positive, True Negative, False Positive, and False Negative, respectively, using classifier ϕ.
t	A time point in the MRI time series, $t \in \{0, 1, 2, \ldots, T-1\}$.
$Y \in \{-1, +1\}$	Ground truth label of lesion, $\mathcal{L}$.
$f^t(c)$	Signal intensity value associated with a pixel, c, at timepoint, $t \in \{0, 1, 2, \ldots, T-1\}$.
$\mathbf{E} = \{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_\mathcal{M}\}$	Dataset comprised of $\mathcal{M}$ lesions.

Table 4.1: List of notation and symbols commonly used in this paper.

of notation and symbols commonly used in this paper is shown in Table 1.

4.3.3 Feature Extraction

Texture Features

A combination of 92 precontrast texture features and 92 textural kinetic features are calculated to describe the texture of each lesion in the dataset. We explore multiple different texture operators, $\Phi^\gamma, 7 \leq \gamma \leq \mathcal{K}$, where $\mathcal{K} = 191$ represents the total number of features. The application of $\Phi^\gamma, \gamma \in \{7, \ldots, \mathcal{K}\}$ to each lesion $\mathcal{L}$ yields a unique feature, $\mathcal{F}^\gamma$. Both the texture of the lesion before contrast agent injection (precontrast

texture) and the textural kinetic feature classes are obtained via application of several steerable, non-steerable, and statistical filters. Table 2 summarizes all texture features employed in our work.

A. Gradient Features

Eleven non-steerable gradient features were obtained using Sobel and Kirsch edge filters and first order spatial derivative operations. The Gabor filters, comprising the steerable class of gradient features, are defined by the convolution of a 2D Gaussian function with a cosine [78]. Hence, for every $c \in \mathcal{L}$, where $c = (x, y)$ [79, 80], the Gabor filter bank response can be expressed as,

$$l_{\Omega,\Lambda,\varphi}(c) = e^{-\frac{x'^2+y'^2}{2\Omega^2}} cos(2\frac{\pi}{\Lambda}x'), \tag{4.1}$$

where Λ is the wavelength of the sinusoid which controls the spatial frequency (scale) of the oscillations. The width of the Gaussian envelope Ω is used to define filters as a function of Λ such that $\Omega = 0.56\Lambda$ as derived in [74]. Filter orientation, φ, dictates the coordinate transformations: $x' = x\cos\varphi + y\sin\varphi$ and $y' = -x\sin\varphi + y\cos\varphi$. Six different scales ($\Lambda \in \{\frac{\pi}{2\sqrt{2}}, \frac{\pi}{4}, \ldots, \frac{\pi}{16}\}$) and 8 orientations ($\varphi \in \{0, \frac{\pi}{8}, \ldots, \frac{7\pi}{8}\}$) were considered in constructing the Gabor filter bank.

B. First Order Statistical Features

Four first order statistical features (mean, median, standard deviation, and range) for 3 different square window sizes, $s \in \{3, 5, 7\}$, were calculated for the gray values of pixels within the sliding window $\mathcal{N}_s$. At every $c \in C, \mathcal{N}_s(c) = \{d \in C | d \neq c, \|d-c\|_2 \leq s\}$, and $\|\cdot\|_2$ is the L_2 norm. Hence, average intensity, $\bar{f}(c)$ within window $\mathcal{N}_s(c)$ is calculated as,

$$\bar{f}(c) = \frac{1}{|\mathcal{N}_s(c)|} \sum_{d \in \mathcal{N}_s(c)} f(d), \tag{4.2}$$

where $c \in C$ is the center pixel of the square window $\mathcal{N}_s(c)$. Median, standard deviation, and range of image intensities within each $\mathcal{N}_s(c)$ for each $c \in C$ are also calculated.

Texture Feature Class	*Individual Attributes*	*Parameters*
Gabor Filters	6 scales 8 orientations	$\Lambda \in \{\frac{\pi}{2\sqrt{2}}, \frac{\pi}{4}, \ldots, \frac{\pi}{16}\}$, $\varphi \in \{0, \frac{\pi}{8}, \ldots, \frac{7\pi}{8}\}$
Kirsch Filters	X-direction Y-direction XY-diagonal	$s = 3$
Sobel Filters	X-direction Y-direction XY-diagonal YX-diagonal	$s = 3$
Grey Level	Mean Median Standard Deviation Range x-Gradient y-Gradient Magnitude of Gradient Diagonal Gradient	$s \in \{3, 5, 7\}$
Haralick	Contrast Energy Contrast Inverse Moment Contrast Average Contrast Variance Contrast Entropy Intensity Average Intensity Variance Intensity Entropy Entropy Energy Correlation Info. Measure of Correlation 1 Info. Measure of Correlation 2	$s = 3$, $g = \max_{d \in C}[f(d)]$

Table 4.2: Summary of all textural features considered in this paper with associated parameter values. These features were used in the calculation of precontrast, postcontrast, and kinetic textural features.

C. Second Order Statistical Features

To calculate the second order statistical (Haralick) feature scenes [76], a pixel window, $\mathcal{N}_s$, $s = 3$ is defined. The parameter $s = 3$ was chosen to capture spatial variations at a high resolution because some of the smaller lesions in the dataset have an area just above 100 pixels. We then compute from each $\mathcal{N}_s(c)$, $c \in C$, a $g \times g$ spatial gray level co-occurrence matrix $\mathbf{G}_c$, where g is the maximum grayscale intensity of the image, C ($g = \max_{d \in C}[f(d)]$). The value $\mathbf{G}_c[u, w]$ at any location, $u, w \in \{1, \ldots, g\}$, represents the frequency with which two distinct pixels, $d, k \in \mathcal{N}_s(c)$ with associated image intensities $f(d) = u, f(k) = w$ are adjacent (i.e., within the same 8-pixel neighborhood of $\mathcal{N}_s(c)$). A total of 13 second order statistical [76] features (see Table 2) were extracted within each $\mathcal{N}_s(c)$ for every pixel $c \in C$ for $s = 3$.

Textural Kinetic Features

For each $d \in \mathcal{L}$, $h_q^t(d)$ represents each of the Q different pixel-based pre- and post-contrast texture feature values, where $q \in \{1, \ldots, Q\}$ and $t \in \{0, \ldots, T-1\}$. The mean feature value, $\bar{h}_q^t$, within each lesion $\mathcal{L}$ and at each timepoint t is then expressed as $\bar{h}_q^t = \frac{1}{|\mathcal{L}|} \sum_{d \in \mathcal{L}} h_q^t(d), q \in \{1, \ldots, Q\}$, and a corresponding textural kinetic vector, $\hat{h}_q = [\bar{h}_q^0, \bar{h}_q^1, \ldots, \bar{h}_q^{T-1}]$ is created. A third order polynomial is fitted to $\hat{h}_q$ to characterize its shape as,

$$\tilde{h}_q^t = \rho_{q,3} t^3 + \rho_{q,2} t^2 + \rho_{q,1} t + \rho_{q,0}, \tag{4.3}$$

where $[\rho_{q,3}, \rho_{q,2}, \rho_{q,1}, \rho_{q,0}]$ are the model coefficients obtained by minimizing the root mean squared difference error between $\hat{h}_q$ and $\tilde{h}_q^t$, where $t \in \{0, \ldots, T-1\}$ and $q \in \{1, \ldots, Q\}$.

Signal Intensity Kinetic Feature

Signal intensity kinetic curves are calculated from the mean signal intensity within the lesion ROI in a manner similar to the curve generated and described in Section 2.3.2. The average lesion intensity is obtained as $[\bar{f}^0, \bar{f}^1, \bar{f}^2, \ldots, \bar{f}^{T-1}], t \in \{0, \ldots, T-1\}$.

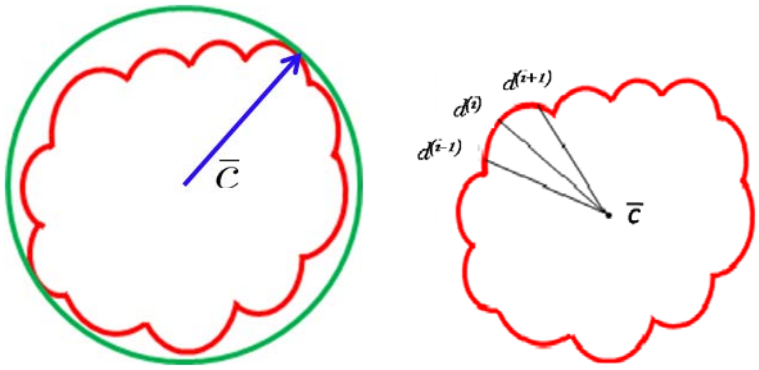
The model coefficients $[\eta_3, \eta_2, \eta_1, \eta_0]$ of a third order polynomial are obtained by the minimization procedure described in Section 2.3.2.

Morphological Features

For mass-like lesions, two important lesions descriptors in the BIRADS lexicon are lesion shape (e.g., round, oval, lobular, irregular) and lesion margin (e.g., smooth, irregular, spiculated) [9]. In this study, we consider 6 different quantitative descriptors modeled on the BIRADS attributes (Table 3) [81,82]. The *Area Overlap Ratio* is a measure of lesion roundness, and the *Normalized Average Radial Distance Ratio, Standard Deviation of Normalized Distance Ratio, Variance of Distance Ratio, Compactness, and Smoothness* are all descriptors for quantifying irregularity of the lesion boundary.

Morphological feature	*Description*
Area Overlap Ratio	$\frac{\lvert\mathcal{L}\rvert}{\pi r^2}$ where $r = \max_{d\in\mathcal{R}}[\lVert d-\bar{c}\rVert]$ (See Figure 3(a)).
Normalized Average Radial Distance Ratio	$\frac{\frac{1}{\lvert\mathcal{R}\rvert}\sum_{d\in\mathcal{R}}\lVert d-\bar{c}\rVert}{\max_{d\in\mathcal{R}}[\lVert d-\bar{c}\rVert]}$.
Standard Deviation of Normalized Distance Ratio	$\sigma_\Gamma = \sqrt{\frac{1}{\lvert\mathcal{R}\rvert}\sum_{d\in\mathcal{R}}(\Gamma(d)-\mu_\Gamma)^2}$, where $\Gamma(d) = \frac{\lVert d-\bar{c}\rVert}{\max_{d\in\mathcal{R}}[\lVert d-\bar{c}\rVert]}$ and $\mu_\Gamma = \frac{1}{\lvert\mathcal{R}\rvert}\sum_{d\in\mathcal{R}}\Gamma(d)$.
Variance of distance ratio	Square of standard deviation, σ_Γ^2
Compactness	$\frac{[\mathcal{D}(\mathcal{R})]^2}{\lvert\mathcal{L}\rvert}$, where $\mathcal{D}(\mathcal{R})$ = the perimeter of $\mathcal{R}$, $\mathcal{D}(\mathcal{R}) = \sum_{d\in\mathcal{R}_{i=0}}^{n-1}\lVert d^{(i+1)}-d^{(i)}\rVert$.
Smoothness	$\sum_{d\in\mathcal{R},i\in\{0,\ldots,n-1\}}\mathcal{B}(d^{(i)})$, where: $\mathcal{B}(d^{(i)}) = \lvert\lVert d^{(i)}-\bar{c}\rVert - \frac{\lVert d^{(i-1)}-\bar{c}\rVert+\lVert d^{(i+1)}-\bar{c}\rVert}{2}\rvert$ (see Figure 3(b)).

Table 4.3: List of morphological features and their mathematical descriptions.



(a) (b)

Figure 4.3: Schematics illustrating the calculation of morphological features: (a) Lesion boundary (red) and circle enclosing it (green) used to calculate Area Overlap Ratio; (b) vectors used for calculation of lesion smoothness. $\bar{c}$ is the lesion centroid, and $d^{(i-1)}, d^{(i)}, d^{(i+1)}, i \in \{1, \ldots, n-1\}$, are consecutive points on the lesion boundary.

4.3.4 Classification

Support Vector Machine (SVM):

The SVM classifier, $\mathcal{V}_{SVM}$, is employed to evaluate the ability of the lesion descriptors to discriminate between benign and malignant breast lesions on DCE-MRI. We construct the $\mathcal{V}_{SVM}$ by using a kernel function (Π) to project training data $E_{tra} \subset \mathbf{E}$, where $\mathbf{E} = \{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_{\mathcal{M}}\}$ is the set of all lesions, onto a higher-dimensional space. This higher-dimensional space allows the SVM to construct a hyperplane to separate the 2 data classes (benign and malignant in our case). The $\mathcal{V}_{SVM}$ is then evaluated by projecting testing data, $E_{tes} \subset \mathbf{E}$, where $E_{tes} \cap E_{tra} = \emptyset$ into the same space and recording the location of the newly embedded datapoint with respect to the hyperplane. In our implementation, the radial basis function (RBF) kernel was employed to project the attributes with $\mathcal{F}^{\gamma}(\mathcal{L}_i)$ and $\mathcal{F}^{\gamma}(\mathcal{L}_j)$, $\gamma \in \{1, \ldots, \mathcal{K}\}$, where $i, j \in \{1, \ldots, \mathcal{M}\}$ and $i \neq j$, into a higher dimensional space. The functional form of the RBF is given by [83],

$$\Pi(\mathcal{F}^{\gamma}(\mathcal{L}_i), \mathcal{F}^{\gamma}(\mathcal{L}_j)) = e^{-\delta \|\mathcal{F}^{\gamma}(\mathcal{L}_i) - \mathcal{F}^{\gamma}(\mathcal{L}_j)\|_2^2}, \tag{4.4}$$

where δ is a scaling parameter. The general form of the SVM classifier is given as

$$\Theta(\mathcal{L}_i) = \sum_{\zeta=1}^{\tau} \xi_{\zeta} Y(\mathcal{L}_{\zeta}) \Pi(\mathcal{F}^{\gamma}(\mathcal{L}_i), \mathcal{F}^{\gamma}(\mathcal{L}_{\zeta})) + \mathbf{b}, \tag{4.5}$$

where $\zeta \in \{1, 2, \ldots, \tau\}$ represents the τ marginal training samples (i.e., support vectors), $\mathbf{b}$ is the hyperplane bias estimated for E_{tra}, and ξ_ζ is the model parameter determined by maximizing an objective function subject to constraints which control the trade-off between empirical risk and model complexity [34, 51]. $Y(\mathcal{L}_j) \in \{+1, -1\}$ represents the class labels, malignant and benign, respectively. $\Theta(\mathcal{L}_i)$ represents the displacement from image $\mathcal{L}_i$ to the hyperplane, and the output of the SVM classifier, $\mathcal{V}_{SVM}(\mathcal{L}_i)$, is equal to $sign[\Theta(\mathcal{L}_i)] \times 1$.

Probabilistic Boosting Trees (PBT):

AdaBoost [84] is one of the most commonly used ensemble machine learning algorithms which yields a class label prediction by combining the outputs from several weak classifiers. However, in AdaBoost the weighting scheme sometimes penalizes samples that are misclassified by a weak classifier even if they were previously correctly classified by a different weak classifier. Additionally, the order of features considered during classification is not preserved via Adaboost. The probabilistic boosting tree (PBT) algorithm [85] addresses these issues by iteratively generating a tree structure of length, B, in the training stage where each node of the tree is boosted with H weak classifiers. The hierarchical tree is obtained by dividing training samples, E_{tra} into two subsets of $E_{tra,Right}$ and $E_{tra,Left}$ based on the learned strong classifier at each node using the standard AdaBoost algorithm [84] and recursively training the left and right sub-trees. To avoid overfitting, the error parameter ϵ is introduced such that samples falling in the range $[\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon]$ are assigned to both subtrees with probabilities $p(\mathcal{F}^\gamma(\mathcal{L})|+1) \longrightarrow E_{tra,Right}$ and $p(\mathcal{F}^\gamma(\mathcal{L})|-1) \longrightarrow E_{tra,Left}$, where the function $p(\mathcal{F}^\gamma(\mathcal{L})|+1)$ represents the posterior class conditional probability of $\mathcal{L}$ belonging to class $+1$ (malignant lesion). The algorithm stops when misclassification error hits a predefined threshold, θ. ϵ was set to 0.1 and $\theta = 0.45$ as suggested in [85]. During testing, the posterior class conditional probability of the sample being malignant is calculated at each node based on the learned hierarchical tree. The discriminative model is obtained at the top of the tree by combining the probabilities associated with probability propagation of the sample at various nodes. The output of the PBT classifier, $\mathcal{V}_{PBT}$, is defined such that if

$p(\mathcal{L}|+1) > \alpha$, then $\mathcal{V}^{\alpha}_{PBT} = +1$, else $\mathcal{V}^{\alpha}_{PBT} = -1$, where $\alpha \in [0, 1]$.

For both the $\mathcal{V}_{SVM}$ and $\mathcal{V}_{PBT}$, we use leave-one-out strategy for classifier training and evaluation. The training dataset, E_{tra}, is related to the test dataset, E_{tes}, by $E_{tes} = \mathbf{E} - E_{tra}, E_{tes} \cap E_{tra} = \emptyset$. During each iteration, E_{tes} contains only one lesion, one that was not considered as E_{tes} during previous iterations.

Three Timepoint (3TP) Modeling

We compared our kinetic texture classifier to the popular 3TP classifier ($\mathcal{V}_{3TP}$). The methods described in [65–67] were used to create a parametric map of signal intensity kinetics in the Hue, Saturation, Value (HSV) color space for each pixel d in each lesion $\mathcal{L}$. Thus the contrast washout rate (F_{wout}) is assigned to the Hue channel, and the contrast uptake rate (F_{win}) is assigned twice to both the Saturation and Value channels. Each pixel, $d \in \mathcal{L}$ is assigned a red hue when representing highest likelihood of malignancy, green when representing moderate likelihood of malignancy, or blue when representing low likelihood of malignancy. At each pixel, $d \in \mathcal{L}$, $F_{win}(d) = \frac{f^1(d) - f^0(d)}{t_1 - t_0}$ and $F_{wout}(d) = \frac{f^2(d) - f^1(d)}{t_2 - t_1}$, where t_0 is the precontrast timepoint, t_1 is the first postcontrast timepoint, and t_2 is the second postcontrast timepoint. $F_{win}(d)$ and $F_{wout}(d)$ were rescaled between 0 and 1 for all $d \in \mathcal{L}$. The empirical thresholds for the hue channel were as follows: if $F_{wout}(d) < 0.4$, then d is assigned red (0 radian); if $F_{wout}(d) > 0.5$, then d is assigned blue ($\pi/3$); else if $0.4 \leq F_{wout}(d) \leq 0.5$ then d is assigned green ($2\pi/3$). Both the saturation and value channels are set equal to the normalized F_{win} value. A lesion was classified as malignant if it contained any red pixels ($\mathcal{V}_{3TP} = +1$) and benign if it contained no red pixels ($\mathcal{V}_{3TP} = -1$).

4.4 Experiments and Performance Measures

4.4.1 Experiments

3.1(a) Discriminating benign vs. malignant lesions based on individual attributes from morphological, signal intensity kinetics, precontrast texture, and textural kinetics feature classes.

A total of 41 images (17 benign, 24 malignant) were analyzed. The separability of lesion classes (17 benign, 24 malignant) using individual descriptors from the feature classes (textural kinetic, precontrast texture, morphological, and signal intensity kinetics) was first qualitatively evaluated using graph embedding, a nonlinear dimensionality reduction technique. We then compared the 191 individual descriptors to discriminate between benign and malignant lesions using 2 different quantitative classifiers, SVMs and PBTs. We also compare the SVM and PBT classification results to the 3TP classifier.

3.1(b) Discriminating benign vs. malignant based on combination of best-performing features.

Following identification of the top-performing features (Experiment 3.1(a)), these attributes are used to construct both a combined SVM classifier and a combined PBT classifier. The performance of each combined meta-classifier (SVM, PBT) is then evaluated against individual attributes.

4.4.2 Performance Measures

Qualitative Evaluation via Graph Embedding

Graph embedding (GE) is a nonlinear dimensionality reduction scheme that is used to transform the high-dimensional set of image features into a low-dimensional embedding while preserving relative distances between images in the original feature space [19, 34]. Given lesions $\mathcal{L}_i$ and $\mathcal{L}_j$ with corresponding feature vectors $\mathcal{F}^\gamma(\mathcal{L}_i)$ and $\mathcal{F}^\gamma(\mathcal{L}_j)$, where $i, j \in \{1, \ldots, \mathcal{M}\}$ and $\gamma \in \{1, \ldots, \mathcal{K}\}$, an $\mathcal{M} \times \mathcal{M}$ confusion matrix $\mathcal{W}[i,j] = e^{-\|\mathcal{F}^\gamma(\mathcal{L}_i) - \mathcal{F}^\gamma(\mathcal{L}_j)\|_2} \in \mathbb{R}^{\mathcal{M}\times\mathcal{M}}$ is constructed. The optimal embedding vector, $\mathbf{X}$, is obtained from the maximization of the following function:

$$\mathcal{E}(\mathbf{X}) = 2(\mathcal{M}-1) \cdot trace\left[\frac{\mathbf{X}^\mathsf{T}(\mathcal{A}-\mathcal{W})\mathbf{X}}{\mathbf{X}^\mathsf{T}\mathcal{A}\mathbf{X}}\right], \tag{4.6}$$

where $\mathcal{A}$ is the diagonal matrix where each diagonal element is defined as $\mathcal{A}[i,i] = \sum_j \mathcal{W}[i,j], \forall a \in \{1, 2, \ldots, \mathcal{M}\}$. The lower-dimensional embedding space is defined by the Eigenvectors corresponding to the β smallest Eigenvalues of $(\mathcal{A}-\mathcal{W})\mathbf{X} = \lambda\mathcal{A}\mathbf{X}$.

The matrix $\mathbf{X}(\mathbf{E}) \in \mathbb{R}^{\mathcal{M} \times \beta}$ of the first β Eigenvectors is constructed such that $\mathbf{X}(\mathbf{E}) = \{\mathbf{X}(\mathcal{L}_1), \mathbf{X}(\mathcal{L}_2), \ldots, \mathbf{X}(\mathcal{L}_{\mathcal{M}})\}$. In our case, $\beta = 3$ so that the embedding basis vectors can be denoted, e_1, e_2, e_3 for any $\mathbf{X}(\mathcal{L}_i)$, where $i \in \{1, \ldots, \mathcal{M}\}$. Embedding plots of the data reduced to three dimensions were used to visualize each feature's ability to cluster the lesions into their appropriate diagnoses.

Quantitative Evaluation

For all three classifiers, $\mathcal{V}_\phi$, where $\phi = \{SVM, PBT, 3TP\}$, each lesion is identified as either a true positive (TP), false positive (FP), false negative (FN), or a true negative (TN) by comparing the classifier output, $\mathcal{V}_\phi(\mathcal{L})$, to the true label, $Y(\mathcal{L})$. If $\mathcal{V}_\phi(\mathcal{L}) = Y(\mathcal{L}) = -1$, lesion $\mathcal{L}$ is identified as a TN; if $\mathcal{V}_\phi(\mathcal{L}) = Y(\mathcal{L}) = +1$, lesion $\mathcal{L}$ is identified as a TP; if $\mathcal{V}_\phi(\mathcal{L}) = +1$ and $Y(\mathcal{L}) = -1$, lesion $\mathcal{L}$ is identified as a FP error; and if $\mathcal{V}_\phi(\mathcal{L}) = -1$ and $Y(\mathcal{L}) = +1$, lesion $\mathcal{L}$ is identified as a FN error. For each classifier, $\phi \in \{SVM, PBT, 3TP\}$, the number of TP ($N_{TP,\phi}$), TN ($N_{TN,\phi}$), FP ($N_{FP,\phi}$), and FN ($N_{FN,\phi}$) lesions over the entire set $\mathbf{E}$ are calculated. Sensitivity (SN_ϕ), specificity (SP_ϕ), and accuracy (AC_ϕ) for each classifier are then calculated as,

$$SN_\phi = \frac{N_{TP,\phi}}{N_{TP,\phi} + N_{FN,\phi}}, SP_\phi = \frac{N_{TN,\phi}}{N_{TN,\phi} + N_{FP,\phi}}, \text{ and}$$

$$AC_\phi = \frac{N_{TP,\phi} + N_{TN,\phi}}{|\mathbf{E}|},$$

where $|\mathbf{E}|$ is the cardinality of set, $\mathbf{E}$.

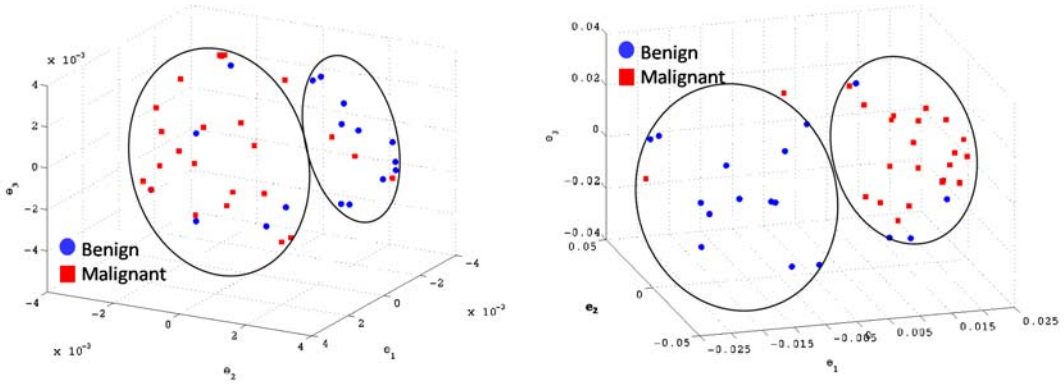
Receiver Operator Characteristic Curves

Receiver Operating Characteristic (ROC) curves representing the trade-off between sensitivity and specificity for breast cancer diagnosis can be generated for $\mathcal{V}_{PBT}$ since α can be varied. Each point on the ROC curve corresponds to the sensitivity (SN^α_{PBT}) and 1-specificity ($1 - SP^\alpha_{PBT}$) over $\mathbf{E}$ for some probability threshold $\alpha \in [0, 1]$, where the interval between α values, $\Delta\alpha$, is 0.05.

4.5 Results

4.5.1 Qualitative Results

Figure 4 shows the embedding plots for the 2 features that separated the benign from malignant lesions best. The morphologic features and precontrast texture features did not cluster the lesions appropriately, whereas the signal intensity kinetic feature and various textural kinetic features did separate the data reasonably well into benign and malignant lesion categories. Although signal intensity kinetics produces a clustering of data classes that is similar to textural kinetic features, the clusters appear better separated in the textural kinetic embedding space (Figure 4(b)).



(a) (b)

Figure 4.4: Embedding plots obtained by plotting the three graph embedding vectors (e_1, e_2, e_3) for (a) Signal intensity kinetics and (b) Gradient in the X-direction kinetics. Note the increased separation of benign (blue) and malignant (red) lesions for the textural kinetic feature compared to signal intensity kinetics.

Parametric Maps

Figure 5 shows representative images for two types of benign lesions and one type of malignant lesion. Each row shows (left to right) the precontrast image, the postcontrast image corresponding to the peak lesion enhancement (maximum signal intensity across the time series), the 3TP parametric map, and a postcontrast texture image for Intensity Entropy, a second order statistical (Haralick) feature. Note the differences in internal

intensity in the postcontrast texture maps (see Fig. 5(d), 5(h), 5(l)), especially between the malignant lesion and the two benign lesions. The malignant lesion in Figure 5(l) is overall brighter than the two benign lesions, indicating a higher heterogeneity in internal lesion enhancement. This is not as easily discernible from the parametric representations shown in Figure 5(c), 5(g), and 5(k).

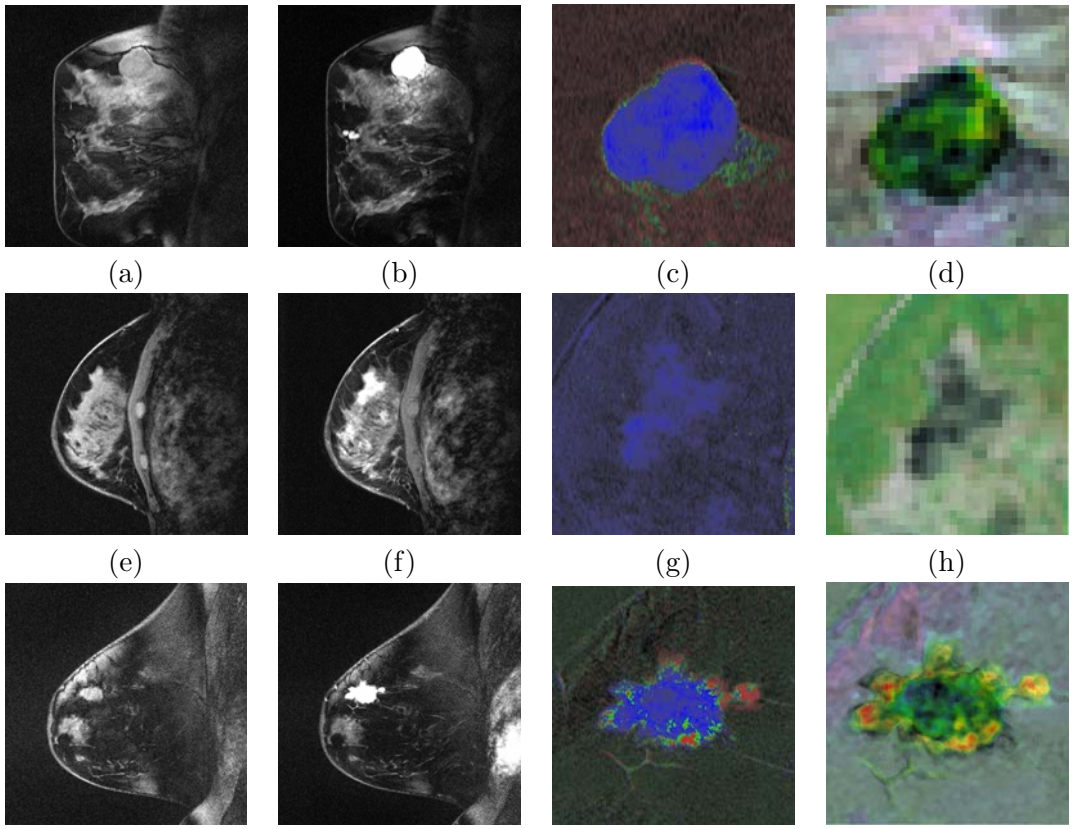


Figure 4.5: Examples of the contrast enhancement patterns associated with (a)-(d) a benign fibroadenoma; (e)-(h) a benign sclerosing adenosis; and (i)-(l) a malignant lesion. Figures 5 (a), (e), and (i) show the full precontrast image. Figures 5 (b), (f), and (j) show the postcontrast timepoint showing peak enhancement. Figures 5 (c), (g), and (k) show the feature 3TP maps corresponding to the studies in (a), (e), and (i). Figures 5 (d), (h), and (l) show the postcontrast texture maps for the Intensity Entropy feature for the peak enhancement timepoint. The 3rd and 4th columns are magnified to accentuate the lesion. Note that Figures 5(d), (h), and (l) are able to illustrate lesion heterogeneity (highest in 5(l) as indicated by the brightness of the lesion compared to 5(d) and (h)) by a simple grayscale value, whereas human visual interpretation of the red, green, and blue values in 5(c), (g), and (k) is required to reach a similar conclusion.

Feature class	*Feature*	AC_{SVM}	SN_{SVM}	SP_{SVM}	AUC_{SVM}
Morphological	Smoothness	0.73	0.88	0.53	0.77
1st Order Textural Kinetics	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{\pi}{8}$	0.71	0.67	0.76	0.73
	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$	0.76	0.75	0.76	0.78
	Median gray level	0.76	0.75	0.76	0.78
2nd Order Textural Kinetics	Contrast inverse moment	0.73	0.88	0.52	0.70

Table 4.4: Results of SVM classifier for top 5 performing individual features in distinguishing benign from malignant lesions using leave-one-out validation.

4.5.2 Quantitative Results

Classification of Lesions Using Individual Features with a Support Vector Machine Classifier

Table 4 shows that in conjunction with the SVM classifier the best textural kinetic feature (X-gradient) had greater values of accuracy than the best-performing morphology feature, smoothness, and signal intensity kinetics performed worse than smoothness. For the morphological feature class, smoothness showed higher accuracy than all other morphological features. We also observed that variants of the X-direction gradient feature, including the Sobel and Kirsch edge filters and 1st order derivative operations, all performed comparably. Gradients in the Y- direction and diagonal gradients (results not shown here) did not perform as well. Note that when the $\mathcal{V}_{SVM}$ was employed, the best-performing 1st and 2nd order (Contrast inverse moment) textural kinetic features also outperformed signal intensity kinetics in terms of accuracy, sensitivity, and specificity (Table 4).

Classification of Lesions Using Individual Features with Probabilistic Boosting Trees

For $\mathcal{V}_{PBT}$, we found similar results to those obtained using $\mathcal{V}_{SVM}$. Using results from the operating point on the ROC curve (defined by the point on the curve that minimizes

Feature class	*Feature*	AC_{SVM}	SN_{SVM}	SP_{SVM}	AUC_{SVM}
Morphological	Smoothness	0.73	0.88	0.53	0.77
Precontrast Texture	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$	0.63	0.90	0.25	0.65
Postcontrast Texture	Intensity Variance	0.68	0.83	0.47	0.70
Signal Intensity	Signal Intensity Kinetics	0.63	0.67	0.59	0.75
1st Order Textural Kinetics	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$	0.76	0.75	0.76	0.78
2nd Order Textural Kinetics	Contrast inverse moment	0.73	0.88	0.52	0.70

Table 4.5: Results of SVM classifier for top-performing individual attributes in distinguishing benign from malignant lesions using leave-one-out validation.

the Euclidean distance from the feature's ROC curve to the ideal 100% sensitivity, 100% specificity point on the graph), Table 5 shows that Contrast inverse moment performed the best from among a Gabor-filtered precontrast image, X-gradient and smoothness. The ROC curves in Figure 6 show that contrast inverse moment had the highest accuracy, sensitivity, and specificity at the operating point among the different feature classes. This was also reflected in the average area under the curve (AUC_{PBT}) in Table 5.

Classification of Lesions Using 3TP parametric maps

The classification of all lesions in the dataset using the 3TP parametric maps ($\mathcal{V}_{3TP}$) produced an accuracy of 78%, sensitivity of 92%, and specificity of 59%.

4.5.3 Classification Performance of Combinations of Features

We selected the best performing features from across all feature classes that had an accuracy rate above 70% to create a combined metaclassifier using the SVM and PBT classifiers. By combining smoothness and the textural kinetic feature for X-gradient in conjunction with an SVM classifier ($\mathcal{V}_{SVM}$), a classification accuracy of 85%, sensitivity of 88%, and specificity of 82% using the leave-one-out strategy was obtained.

Feature class	*Feature*	AC_{PBT}	SN_{PBT}	SP_{PBT}	AUC_{PBT}
Morphological	Smoothness	0.85	0.91	0.76	0.91
1st Order Textural Kinetics	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{\pi}{8}$	0.89	0.98	0.76	0.78
	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$	0.89	0.99	0.74	0.86
	Median gray level	0.90	0.97	0.81	0.83
2nd Order Textural Kinetics	Contrast inverse moment	0.90	0.95	0.82	0.92

Table 4.6: Results of PBT classifier for top-performing individual attributes in distinguishing benign from malignant lesions using leave-one-out validation. Note that the AC_{PBT}, SN_{PBT}, and SP_{PBT} values reported here are for the operating point on each feature's respective ROC curve.

Feature class	*Feature*	AC_{PBT}	SN_{PBT}	SP_{PBT}	AUC_{PBT}
Morphological	Smoothness	0.85	0.91	0.76	0.91
Precontrast Texture	Gabor filter, $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$	0.84	0.94	0.71	0.86
Postcontrast Texture	Intensity Variance	0.70	0.92	0.41	0.58
Signal Intensity	Signal Intensity Kinetics	0.79	0.94	0.59	0.78
1st Order Textural Kinetics	X-gradient	0.83	0.88	0.76	0.85
2nd Order Textural Kinetics	Contrast inverse moment	0.90	0.95	0.82	0.92

Table 4.7: Results of PBT classifier for top-performing individual attributes in distinguishing benign from malignant lesions using leave-one-out validation.

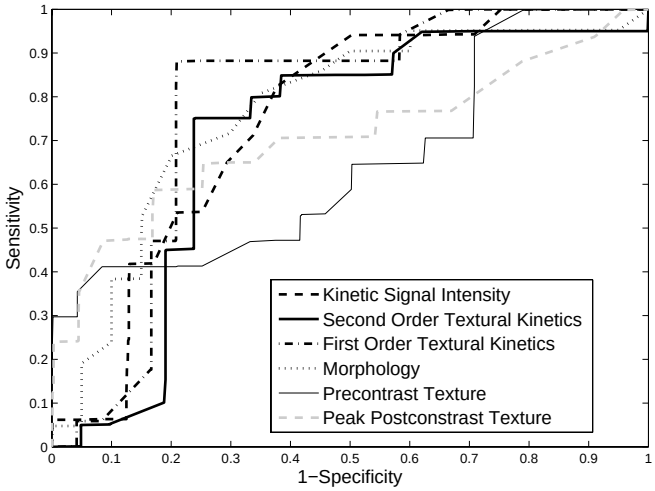


Figure 4.6: ROC curves generated for $\mathcal{V}_{SVM}$ by the decision plane for the top-performing feature in each feature class. The individual features for which ROC curves have been plotted are: second order textural kinetic feature, Contrast inverse moment; first order textural kinetic feature, X-gradient; morphology feature, Smoothness; and the Gabor filter channel corresponding to $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$ for the precontrast image.

Classifier combination	*AC*	*SN*	*SP*	*AUC*
Smoothness + X-direction Sobel filter $+\mathcal{V}_{SVM}$	0.82	0.92	0.71	0.78
Smoothness + Contrast inverse moment + $\mathcal{V}_{PBT}$	0.89	0.99	0.76	0.91

Table 4.8: Results of classifiers obtained by combination of multiple attributes in distinguishing benign from malignant lesions using leave-one-out validation (N/A: not applicable).

By combining smoothness and the textural kinetic feature, Contrast inverse moment kinetics in conjunction with a PBT classifier ($\mathcal{V}_{PBT}$), a classification accuracy of 89%, sensitivity of 99%, specificity of 76%, and an $AUC = 0.91$ using the leave-one-out strategy was obtained. In Figure 7, the embedding plot of the reduced feature space of the metaclassifier using graph embedding further corroborates the $\mathcal{V}_{SVM}$ and $\mathcal{V}_{PBT}$ results, revealing good separation between the lesion classes.

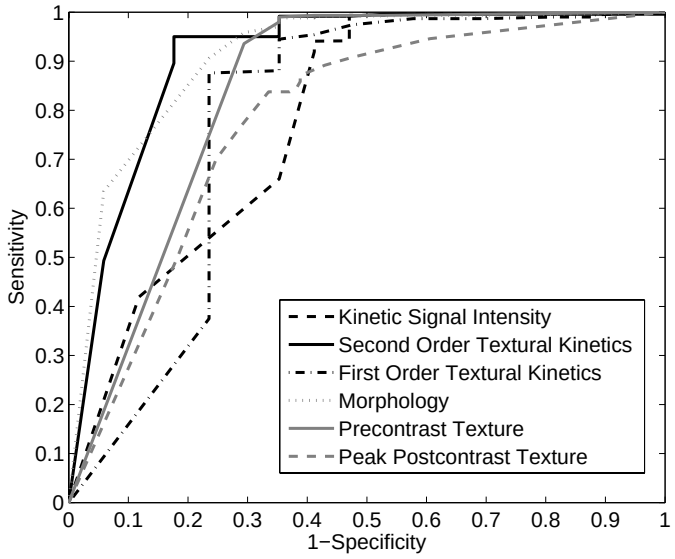


Figure 4.7: ROC curves generated for $\mathcal{V}_{PBT}$ by varying probability threshold, $\alpha \in [0, 1]$, for the top-performing feature in each feature class. The individual features for which ROC curves have been plotted are: second order textural kinetic feature, Contrast inverse moment; first order textural kinetic feature, X-gradient; morphology feature, Smoothness; and the Gabor filter channel corresponding to $\Lambda = \frac{\pi}{8\sqrt{2}}, \varphi = \frac{7\pi}{8}$ for the precontrast image.

4.6 Concluding Remarks

In this paper we presented a new attribute, textural kinetics, for discriminating between benign and malignant lesions by quantifying the spatiotemporal patterns of lesion texture during the contrast enhancement time series. We showed that textural kinetic features outperformed the signal intensity kinetics feature on a dataset of 41 (17 benign, 24 malignant) breast lesions in terms of accuracy, sensitivity, and specificity. An SVM classifier in conjunction with the textural kinetic descriptors yielded an accuracy of 83%, sensitivity of 79%, and specificity of 88%, and the PBT classifier yielded accuracy of 90%, sensitivity of 95%, specificity of 82%, and area under the curve (AUC) of 0.92. Compared to related texture-based approaches for lesion classification [58, 86], textural kinetics perform better than enhancement variance dynamics introduced by Chen et al. [86], which reported an AUC of 0.85. While the textural kinetic features performed marginally worse compared to spatio-temporal enhancement profiles introduced in Zheng et al. [58], the study in [58] employed a smaller dataset (36 lesions), and

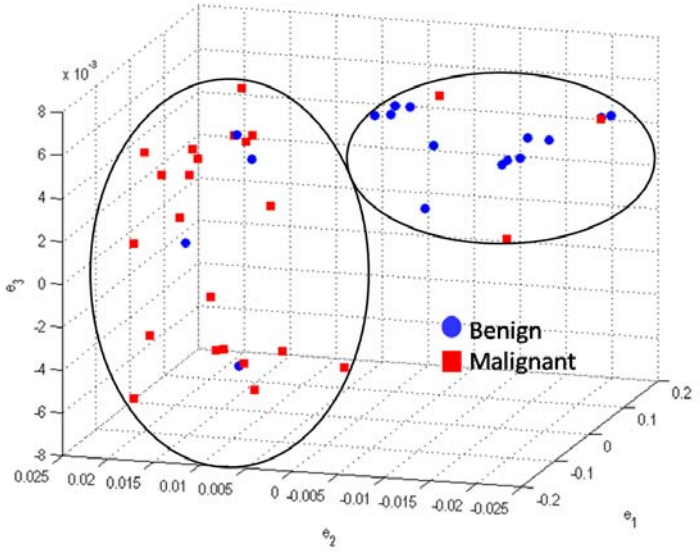


Figure 4.8: Embedding plot obtained by plotting the three graph embedding vectors (e_1, e_2, e_3) for the combination of the morphology feature, Smoothness, with the first order textural kinetic feature, X-gradient. Note the increased separation between the lesions compared to Figure 4.

the accuracy of their classifier decreased to less than 90% when the dataset decreased by 5 lesions, indicating some sensitivity of their approach to the composition of the dataset. The textural kinetic features yielded consistently good classification performance for both the SVM and PBT classifiers. When the textural kinetic attributes were combined with morphologic descriptors, the resulting SVM classifier yielded an 85% accuracy, sensitivity of 88%, and specificity of 82%, and the resulting PBT classifier yielded an 89% accuracy, sensitivity of 99%, specificity of 76%, and an AUC of 0.91, demonstrating that pairing of morphology and signal intensity kinetic features with orthogonal lesion attributes such as textural kinetics could result in improved diagnosis of breast cancer on breast DCE-MRI. Based on previous commentary in the literature [72, 87], textural attributes are likely more robust to MRI artifacts such as bias field and intensity non-standardness. In addition, textural kinetics provide information that is not visually intuitive to the radiologist, yet perform better than any of the intuitive BIRADS descriptors, suggesting the importance of CAD as a complement to radiological diagnosis. It is important to note that what is presented here is not a

full-fledged CAD system, but rather a study in the utility of textural kinetics in distinguishing benign from malignant lesions. In future work, we plan to more rigorously test the robustness of the features and combinations of features on a larger cohort. We also plan to incorporate automated lesion detection and segmentation into the current workflow.

Chapter 5

Distinguishing Molecular Subtypes of Breast Cancer

5.1 Introduction

Triple negative (TN) breast cancer (CA) has recently been identified as an important CA subtype due to its prevalence in the population (10-20% of all diagnosed CA [7]), lack of options for targeted molecular therapies [88], and poor prognosis [89]. Dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) has been recently shown to be sensitive for detecting TN CA [14, 90] and screening BRCA mutation carriers [8, 14, 71, 91, 92] who often develop TN CA [7], particularly BRCA1 mutation carriers.

Studies comparing phenotypic differences between TN and non-TN CA on DCE-MRI show that many TN CAs have smooth, round margins compared to the spiculated borders of the more common ER positive (ER+) CAs [93]. TN CAs also show rim enhancement on post contrast T1-weighted (T1-w) imaging and central high signal intensity on T2-weighted (T2-w) imaging [12, 93–95]. Interestingly, the characteristic features of TN CAs, as identified by Uematsu et al. [93] and Wang et al. [12] do not conform to the attributes commonly ascribed to malignancies in the BIRADS MRI lexicon [9]. In fact, Uematsu et al. [93] reported that TN lesions may have morphologic and kinetic features on DCE-MRI that are similar to those of benign fibroadenomas (FAs).

Computer-aided diagnosis (CAD) methods, which provide diagnostic information based on quantitative descriptors of the lesion under inspection, have been proposed to (a) reduce inter-rater variability [61, 96] and (b) to increase diagnostic specificity [13, 97–100] on DCE-MRI. Recently, breast CAD solutions have started to move beyond solely distinguishing benign from malignant lesions [33, 35, 50, 61] to addressing more complex diagnostic questions. For instance, Bhooshan et al. [50] proposed the use

of textural, morphological, and kinetic descriptors for distinguishing ductal carcinoma in situ (DCIS) from invasive ductal carcinoma (IDC). Makkat et al. [13] similarly proposed the use of imaging descriptors such as the quantification of tumor blood flow for differentiating between HER2 positive (HER2+) and HER2 negative tumors [13].

In [2], the term textural kinetics was introduced to describe the dynamic variation of textural features of breast lesions during contrast uptake. Textural kinetics outperformed standard morphologic, static texture, and kinetic intensity features in distinguishing benign from malignant lesions [2]. Textural kinetics features are calculated by computing a series of texture features at each time point in the DCE-MRI series. For each static texture feature, a statistic such as the mean or median feature value over all pixels within the lesion at each time point is calculated and then plotted as a function of time. The shape of this curve, akin to the shape of the signal intensity kinetic curves first described by Kuhl [54] and Kinkel et al. [38], provide a description of textural variations as a function of contrast uptake within the lesion. In this study, we introduce a new computerized image analysis, feature extraction, and classifier framework for describing the imaging behavior of breast lesions on DCE-MRI. By determining a quantitative MR imaging signature for TN CAs that is distinct from other CA subtypes (e.g., ER+ and HER2+ CAs), as well as from benign TN mimics (e.g., FAs), CAD algorithms tuned to specific high-risk patient classes (e.g., BRCA1 mutation carriers predisposed to forming TN CAs) can be developed. In this study, the success of this framework for evaluating the TN imaging signature using a CAD system is tested on a cohort 76 breast lesions from 65 patients.

5.2 Materials and Methods

Breast MRI data was prospectively collected in an Institutional Review Board-approved study at the University of Pennsylvania between 2002 and 2007. In this study, women underwent MRI either to further investigate a suspicious lesion identified on screening mammography or for local staging prior to surgery. This study examined MRI characteristics in 76 solid lesions from 65 patients for whom pathology results and, where applicable, ER, PR, and HER2 results were available. Confirmatory diagnosis was

made by histopathologic examination of tissue obtained by either core biopsy sampling or lumpectomy. Of the 76 lesions, 12 were benign (i.e., 12 FAs) and 64 were invasive carcinomas. All of the carcinomas were immunohistochemically stained for hormone receptors and HER-2/neu. In cases in which staining for HER-2/neu was inconclusive, amplification was confirmed with fluorescence in situ hybridization, Of the 64 carcinomas, 21 were TN (ER-/PR-/HER2-) CA, 18 were HER2+ (14 ER-/HER2+, 4 ER+/HER2+) CA, and 25 were ER+ (ER+/HER2-) CA. Patient MRIs were acquired at either 1.5 or 3 Tesla (Siemens Sonata or Trio, respectively). Sagittal imaging included fat-saturated three-dimensional T1-w DCE imaging before and after 20 mL/kg Gd-DTPA (Omniscan) contrast injection. Imaging parameters for DCE-MRI varied over time and magnet type (matrix size: 256x256-896x896, with in-plane resolution 0.20-0.70 mm/pixel, slice thickness 2-5 mm, TR=7-26 ms, TE=1.8-6.5 ms, flip angle 25-30 degrees). DCE-MRI data sets were acquired once before contrast injection and at 90 second time intervals upon bolus contrast injection, for a total scan duration of between 5-8 minutes.

5.2.1 Lesion Segmentation and Feature Extraction

A representative slice of the DCE-MRI volume was chosen and confirmed by a radiologist with 12 years experience in breast MRI interpretation (M.A.R.) who was blinded to pathologic diagnosis. The lesion boundary was manually delineated based on the early post-contrast image demonstrating greatest lesion conspicuity from neighboring tissues. Morphologic features (i.e. shape features, margins) were calculated based on this boundary. All other features (e.g. static textural, intensity kinetics, and textural kinetics) were computed based on the pixels enclosed by the lesion segmentation. Table 1 lists and describes the computer extracted features used in this study. Lesion sizes were generally evenly distributed among the lesion classes (FA (17.2 12.1 mm), TN (26.0 17.6 mm), HER2+ (29.0 24.4 mm), and ER+ (29.8 19.4 mm)). Figure 5.1 shows the overall framework.

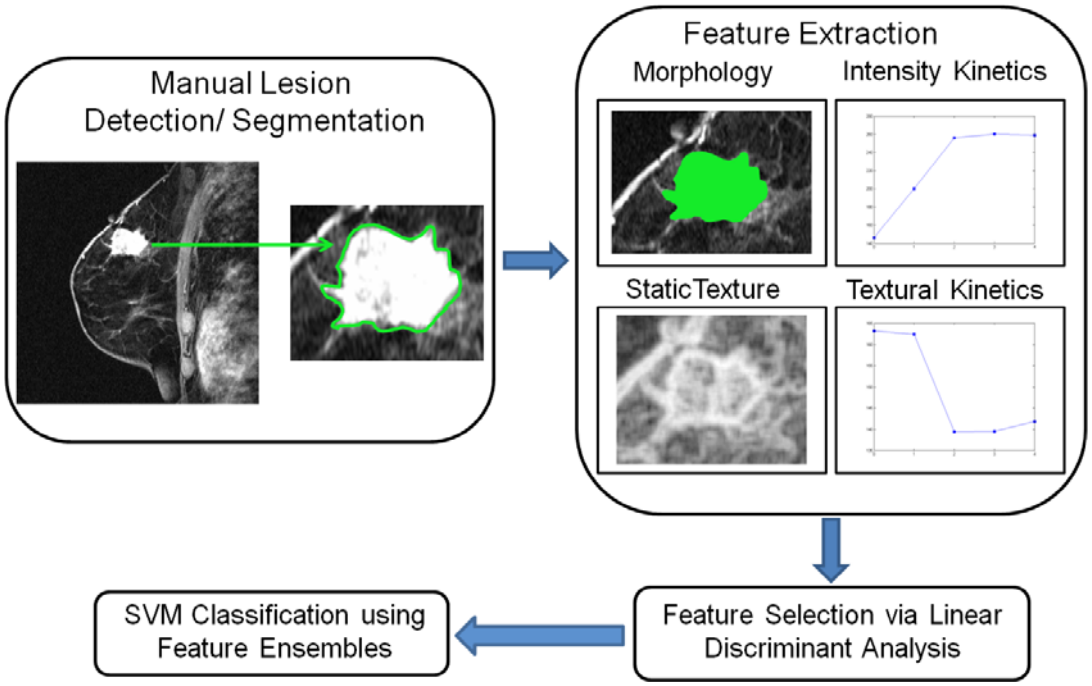


Figure 5.1: Flowchart showing workflow of lesion classification for the 5 classification tasks (TN vs. all non-TN, TN vs. non-TN CA, TN vs. ER+, TN vs. HER2+, TN vs. benign-FA) addressed in this study.

5.2.2 Morphologic features

Six morphologic features [2, 3] relating to the boundary between the lesion and the surrounding tissue were determined: 1) area overlap ratio, 2) normalized average radial distance ratio, 3) standard deviation of normalized distance ratio, 4) variance of distance ratio, 5) compactness, and 6) smoothness. These morphological features were employed to analyze the roundness, spiculation, regularity, and smoothness of the lesion boundary along with lesion shape. In addition, 2 features used previously for breast lesion analysis on DCE-MRI (margin sharpness and variance in margin sharpness [50, 86]) were also extracted.

5.2.3 Intensity kinetics

A total of 4 intensity kinetic (maximal uptake, time to peak, uptake rate, and washout rate) features [50, 58, 86, 101] were calculated in order to measure the amount and rate of contrast uptake [54].

5.2.4 Static texture

Static texture features were computed from the peak enhancing image for each lesion as was done in previous studies [2, 50, 95]. A set of 22 textural features were calculated for each lesion, reflecting heterogeneity of tissue types in a lesion based on first order statistics [2], derivative operations, and grey level co-occurrence matrices (GLCMs) [76].

5.2.5 Textural Kinetics

Textural kinetics, features related to the dynamically changing textural appearance of the enhancing breast lesion as it passes from pre-, to peak-, to equilibrium-phase enhancement. The average value of each of the 22 texture features was plotted as a function of time, and a 3rd order polynomial was fit to the curve, yielding a vector of four coefficients per kinetic textural feature (see Figure 5.2) [2]. These four coefficients represent the corresponding textural kinetic feature for each static texture feature. Textural kinetic features were calculated for each of the 22 textural features described in the static texture section.

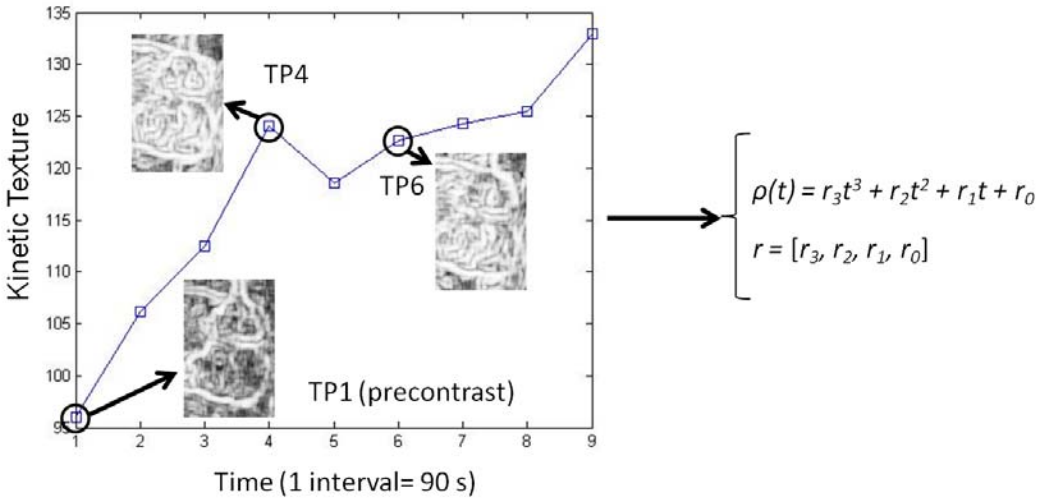


Figure 5.2: Calculation of textural kinetics feature, contrast entropy, for a TN CA. The mean feature value is calculated and plotted as a function of time. A third order polynomial is then fit to the curve, and the 4 coefficients resulting from curve fitting are used to represent the textural kinetics feature for that lesion. $\rho(t)$ represents the vector of mean texture values at each time point, and r is the vector of 4 coefficients that results from fitting a cubic polynomial to the texture vs. time plot; TP: time point.

5.2.6 Support Vector Machine Classifier

In order to determine quantitative imaging features on DCE-MRI (see Table 1) that best discriminated TN CA from other molecular subtypes of CA and from benign FAs, feature selection was performed in a feed-forward manner using linear discriminant analysis (LDA) [102, 103]. The important features identified during the LDA feature selection process were combined with equal weighting and used in conjunction with an SVM classifier [51]. The SVM assigned a likelihood value to each lesion of belonging to a specific class by exploiting the distance of each lesion to the SVM decision hyperplane [51], defined as a multidimensional decision boundary used to discriminate between lesion classes: the greater the distance of a lesion from the hyperplane, the higher the likelihood that the lesion belongs to a particular class. As the distance of the objects from the decision hyperplane changes, the corresponding object-class probabilities also change. By varying the position of the decision hyperplane, classification sensitivity and specificity estimates at each location are obtained, in turn allowing for the calculation of a receiver operator characteristic (ROC) curve for the classifier. The SVM classifier mechanism employs a leave-one-out strategy, and area under the ROC curve (AUC) [2] for each of the 5 classification tasks (TN vs. all non-TN, TN vs. non-TN CA, TN vs. ER+, TN vs. HER2+, and TN vs. FA) was evaluated.

5.3 RESULTS

LDA for Feature Selection Different textural and morphological image features were deemed important for each of the different lesion classification tasks (TN vs. all non-TN, TN vs. non-TN CA, TN vs. ER+, TN vs. HER2+, TN vs. FA). For each classification task, the most discriminating features with the associated LDA classification accuracies as each feature is added in the feed-forward feature selection process are shown in Figure 5.3.

5.3.1 TN vs. all non-TN

For TN vs. all non-TN lesions (FA, HER2+, ER+), the morphology feature compactness (ratio of lesion perimeter to area) and 2 textural features, namely a combination of static intensity average and kinetic X-direction Sobel filter yielded the best discriminability in the feature selection process. Overall, TN lesions had lower values of the GLCM feature intensity average indicating more lesion heterogeneity at peak contrast relative to that of non-TN lesions. Conversely, the kinetic X-direction Sobel filter feature set indicated that the TN lesions steadily increased in homogeneity of lesion appearance over time, whereas non-TN lesions varied more across time. In addition, TN lesions were more compact than non-TN lesions.

5.3.2 TN vs. non-TN CA

For the TN vs. non-TN CA (HER2+ and ER+) task, the static texture intensity average, and kinetic textures X-direction Sobel filter and GLCM feature, energy, were identified as the most important features. Consistent with the analysis for the TN vs. non-TN lesions, non-TN CAs still had higher values of the GLCM feature, intensity average texture, and the textural kinetic feature, X-direction Sobel filter, had a set of coefficient values that reflected more variation in contrast uptake across the non-TN CAs than that of the TN CAs.

5.3.3 TN vs. FA

For the TN CA vs. FA classification task, static texture features, intensity average and intensity variance appeared to have greatest relevance. FAs had higher localized texture values for both intensity average and intensity variance than TN CAs, indicating TNs were far more heterogeneous at peak contrast than were FAs.

5.3.4 TN vs. ER+

For the task of distinguishing TN CAs vs. ER+ CAs, the morphological feature smoothness, the static GLCM texture intensity average and the textural kinetics of X-direction

Sobel filter were important descriptors. Figure 5.4 illustrates the difference in lesion shape smoothness between a TN CA and an ER+ CA. The mean feature values for this comparison indicate that ER+ CAs tend to have more irregular borders and are more homogeneous at peak contrast based on the higher values of smoothness and higher intensity average values, respectively. However, the coefficient values for the kinetic X-direction Sobel filter feature indicate that homogeneity varies more as a function of contrast uptake in ER+ CAs than in the case of TN CAs.

5.3.5 TN vs. HER2+

In distinguishing TN CAs from HER2+ CAs, a combination of the morphology feature, compactness, the static texture feature, energy, and the textural kinetics feature contrast energy appeared to be most relevant. The higher value of compactness for the TN CAs than HER2+ CAs suggests that TN CAs have smoother borders than HER2+ CAs. HER2+ CAs are more heterogeneous at peak contrast uptake as demonstrated by the energy values, but the heterogeneity of contrast uptake in TN CAs varies more as a function of time than the heterogeneity of contrast uptake in HER2+ CAs, as reflected by the textural kinetics feature, contrast energy. In Figure 5.5, examples of contrast energy versus time plots and associated pre-, early post-, and late post-contrast enhancement are shown for both a HER2+ and a TN CA.

Performance of the feature set for discriminating lesion classes ROC curves corresponding to each of the 5 classification tasks (TN vs. all non-TN, TN vs. non-TN CA, TN vs. ER+, TN vs. HER2+, TN vs. FA) are shown in Figure 5.6, and the corresponding AUC values are listed in Table 2. The selected feature set was best able to distinguish between TN CAs and FAs, with an AUC of 0.97. However, the selected features were also was able to segregate TN CAs from other CAs (TN vs. non-TN CA: AUC = 0.74) and distinctly from ER+ (TN vs. ER: AUC = 0.79) and HER2+ (TN vs. HER2+: AUC = 0.74) cancers.

5.4 DISCUSSION

A CAD system, based on the aforementioned quantitative DCE-MRI features, was developed to distinguish TN CAs (a) from different molecular subtypes of CA, and (b) from benign TN mimics (i.e., FAs), on DCE-MRI. CAD systems, like the one presented, have the potential to provide insight into the underlying tumor biology along with providing information about how a patients CA may respond to targeted therapies [104]. Although differences in lesion appearance on DCE-MRI have been previously suggested in qualitative comparisons of molecular subtypes of CA [13,93,95], to our knowledge, this is the first attempt to quantify imaging differences between TN and ER+ CAs and TN and HER2+ CAs. Furthermore, this is the first systematic comparison of TN CA and FA, a common benign lesion with morphologic appearance and intensity kinetics features that resemble those of TN CA, confounding accurate diagnosis on DCE-MRI [14, 93].

The particular lesion classification tasks (TN vs. all non-TN, TN vs. non-TN CA, TN vs. ER+, TN vs. HER2+, TN vs. FA) were chosen based on (i) distinguishing TN CAs from FAs and for (ii) relevance to individualized therapy (e.g., TN CA compared to HER2+ CA or ER+ CA) . As expected, there was significant overlap in the morphologic features of TN CA and FA. Thus, morphology features were not useful for distinguishing these lesions. Instead, static texture features were found to differ sufficiently between these two lesion classes. This result echoes the findings in Schrading and Kuhl [14], who showed radiologist- determined morphology features did not contribute to the discernment between FAs and TN CAs on DCE-MRI. In contrast, lesion morphology was useful for distinguishing TN CAs from both the ER+ and HER2+ CA classes. This is also consistent with the findings of Wang et al. [12], who showed that TN CAs have smooth, circumscribed borders, while HER2+ CAs are spiculated and the findings of Uematsu et al. [93], who showed that TN CAs tend to have smooth borders, while ER+ CAs tend to have irregular borders.

In all classification tasks except for distinguishing FAs from TN CAs, textural kinetics features in addition to static texture features were identified as important attributes.

The peak contrast (i.e., static) texture feature intensity average was typically lower for TN CAs than for ER+ CA. This is similar to the finding in Uematsu et al. [93] that TN CAs were more likely to demonstrate regional enhancement TN CAs, which can often be described as rim enhancement. Additionally, the feature selection results via LDA suggest that kinetic heterogeneity of lesion enhancement described by the textural kinetics [2] of the X-direction Sobel filter feature also contribute to distinguishing between TN and ER+ CAs. This finding appears to suggest that in addition to static texture, which helps to describe lesion heterogeneity, the evolution over time of the geographic pattern of contrast uptake by the lesion also provides insight into the lesions molecular identity.

Although HER2+ CAs were found to be more heterogeneous at peak enhancement (i.e., in static texture), this textural characteristic of HER2+ CAs did not change over the course of contrast uptake (i.e., in textural kinetics). Conversely, the degree of textural heterogeneity of TN CAs was more variable over time as demonstrated by textural kinetics. This finding likely reflects differences in neovascular characteristics between HER2+ and TN CAs, as suggested by Makkat et al. [13] with TN CAs having a less uniformly distributed network of blood vessels. Thus, despite the global similarity in the kinetic contrast behavior of TN and HER2+ CAs (relative to those of ER+ CAs), differences between these classes are revealed by analysis of textural kinetics. Such a finding may become a useful biomarker of tumor response as vascular- and other molecularly-targeted therapies are introduced in the neoaduvant setting. Textural kinetics, a relatively new feature, was previously shown to be useful in distinguishing benign from malignant lesions [2]. In this paper, it was also useful for discriminating TN CAs from other breast lesions. It is possible that textural kinetics, which measures spatio-temporal changes in breast lesion appearance during contrast uptake, capture unique aspects of the biological heterogeneity of breast cancers. Whereas intensity (i.e. overall contrast) kinetics features such as contrast uptake, washout, and time to peakdid not distinguish TN CAs from other molecular subtypes of CA, textural kinetics provided such discrimination in all comparisons of TN CAs with other CA subtypes. As modern MR imaging techniques evolve to allow for greater degrees of image acceleration

without sacrificing image resolution, it is expected that textural kinetic analyses will provide an even richer source of data to explore the biologic heterogeneity of breast lesions.

Our study had its limitations. The number of patients studied was small, although in the same range as analogous studies by Schrading and Kuhl [14] (68 patients) , Wang et al. [12] (56 patients), or Makkat et al. [13] (57 patients). Both 1.5 T and 3 T examinations, with varying image protocols, were evaluated in this study. However, images were treated in the same manner by the CAD system, and neither image acquisition method nor magnet strength appeared to affect the image analysis. While there were no obvious trends in classification errors, some of the misclassified lesions were those with diameters less than 10 mm, suggesting that further work is required to determine the success in classifying small lesions on MRI. However, not all, small lesions were misclassified in our study. In future work, we plan to test this lesion classification framework on an independent dataset obtained at a different clinical site. In future work, we plan to quantify rim enhancement, a feature previously shown to be an important descriptor of TN CAs in the recent paper by Uematsu et al. [93], in order to further refine the CAD classification of TN CAs.

CAD systems have been previously proposed for distinguishing benign from malignant lesions on DCE-MRI. This is, to our knowledge, the first CAD study focused on differentiating molecular subtypes of breast cancer on DCE-MRI. Using these techniques, we have identified morphologic as well as both static and kinetic textural imaging phenotypes via DCE-MRI that may be able to segregate breast cancer subtypes. The correlation between imaging and histopathologic signatures of TN CA has previously been suggested in Tchou et al. [105], who correlated fluoro-deoxyglucose uptake in positron emission tomography scans with proliferation markers in the histopathological sections of TN CA. We intend to similarly explore quantitative correlation between DCE-MRI and analytic histopathologic signatures of breast cancer in future work.

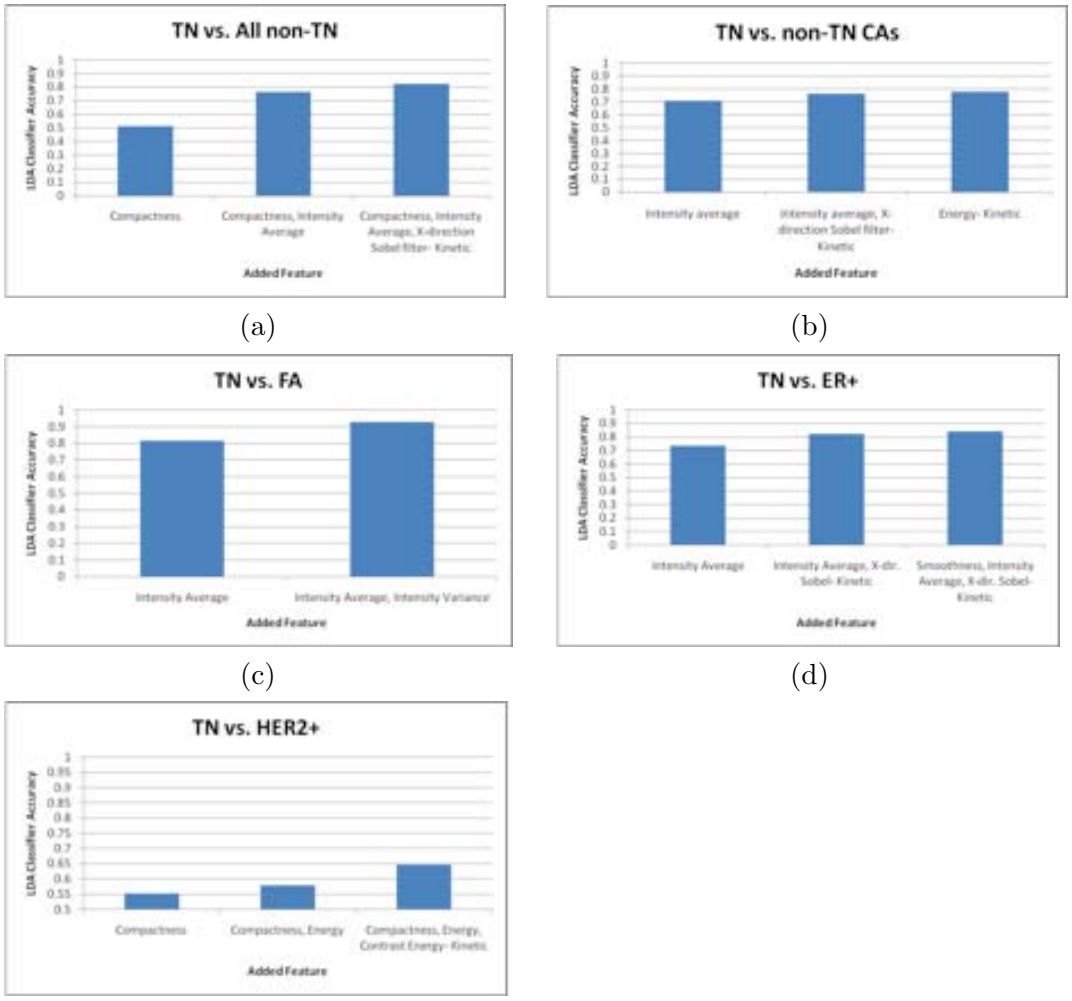
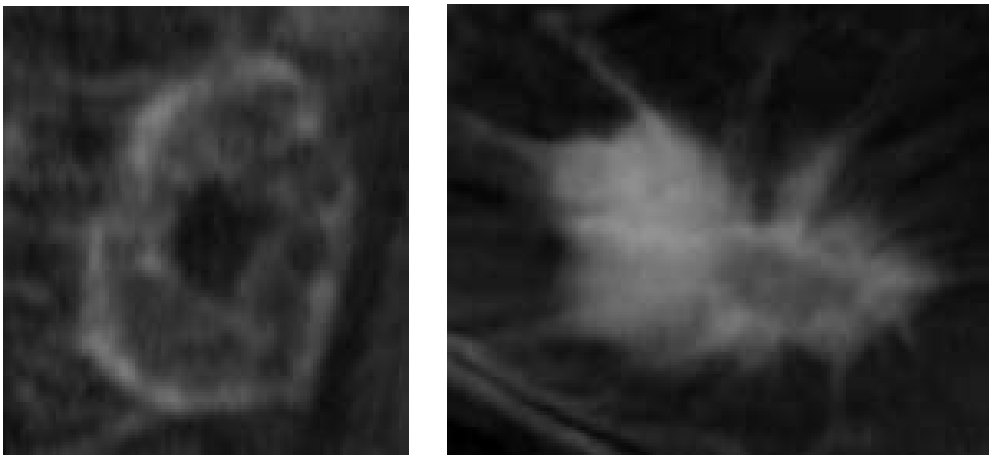


Figure 5.3: Increase in linear discriminant analysis (LDA) classifier accuracy as a function of feature inclusion for (a) TN vs. all non-TN lesions, (b) TN vs. non-TN CAs, (c) TN vs. FA, (d) TN vs. ER+, and (e) TN vs. HER2+ lesions. The best performing feature (in terms of classifier accuracy) is first chosen and subsequent features are included based on their ability to improve existing classifier accuracy.



(a) (b)

Figure 5.4: Morphology of a lobulated TN CA (a) compared to an irregularly bordered ER positive (b) CA.

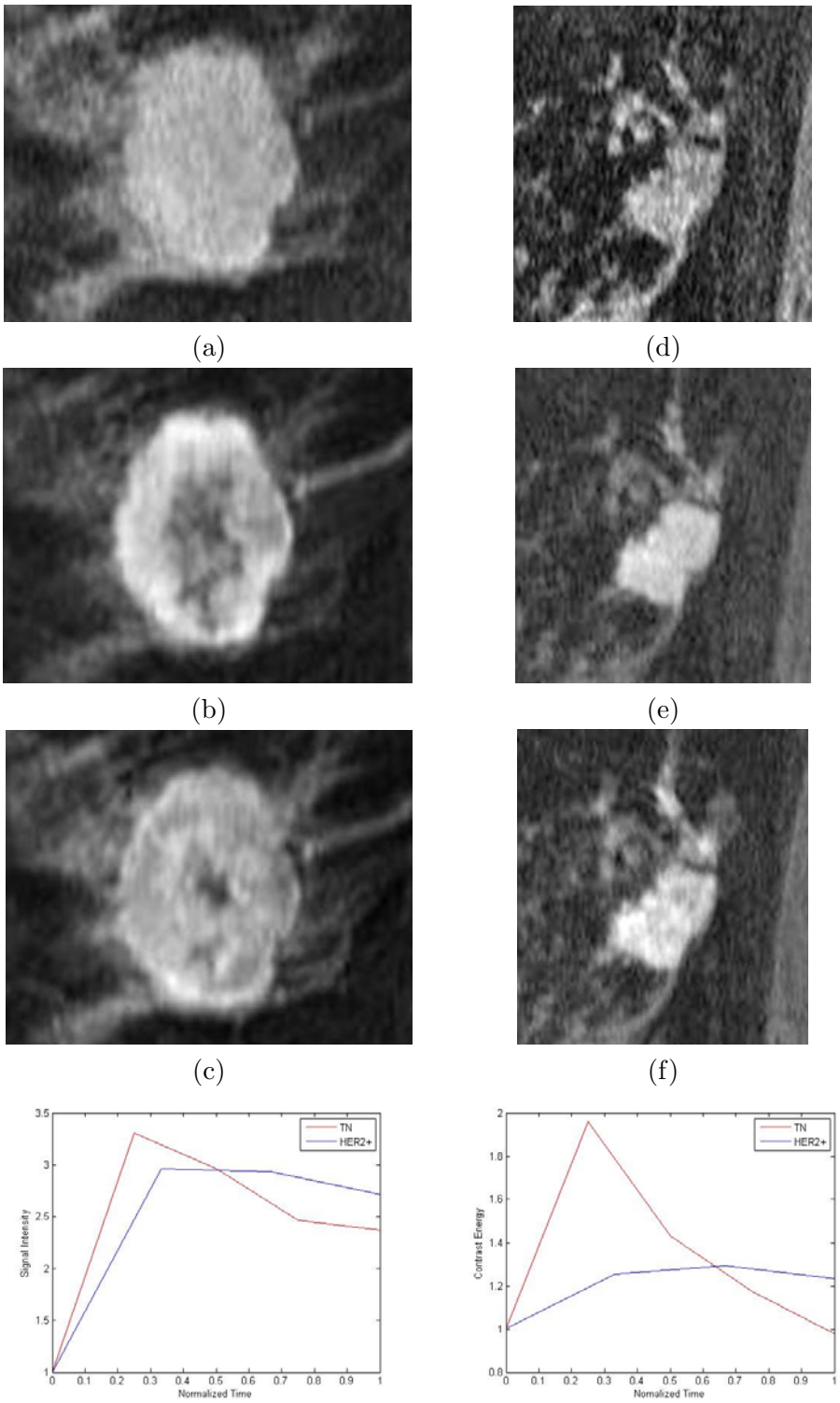
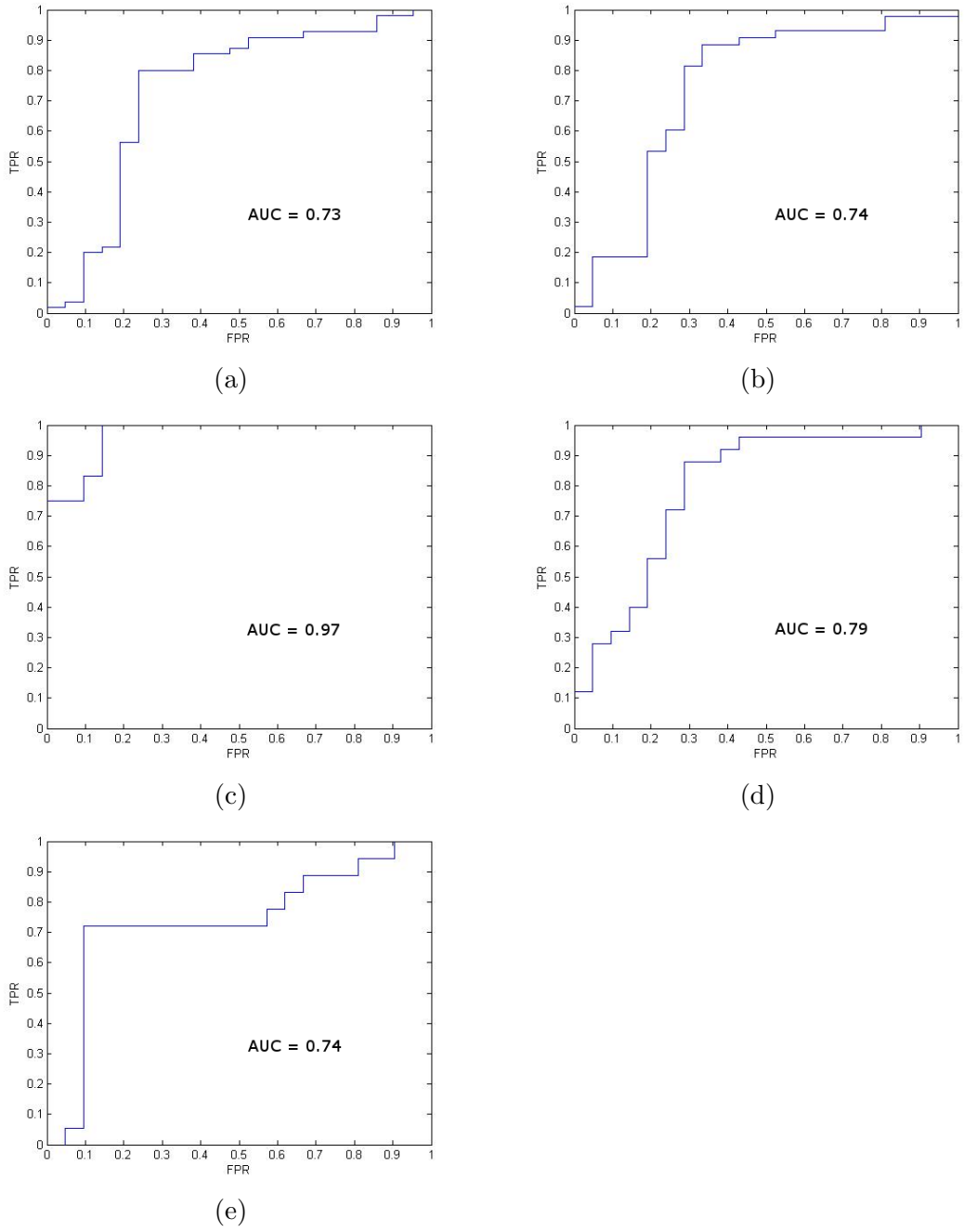


Figure 5.5: Figure 5. Pre- (a, d), early- (b, e) and late- (c, f) contrast enhancement of (a-c) TN CA and (d-f) HER2+ CA and their associated (g) intensity vs. time and (h) textural kinetics (contrast energy) vs. time curves. Note the varying uniformity in the TN CA compared to the HER2+ CA, which steadily increases in homogeneity of contrast enhancement. While the intensity vs. time curves for the two lesions are similar, the contrast energy vs. time curves are drastically different.



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

Figure 5.6: Increase in linear discriminant analysis (LDA) classifier accuracy as a function of feature inclusion for (a) TN vs. all non-TN lesions, (b) TN vs. non-TN CAs, (c) TN vs. FA, (d) TN vs. ER+, and (e) TN vs. HER2+ lesions. The best performing feature (in terms of classifier accuracy) is first chosen and subsequent features are included based on their ability to improve existing classifier accuracy.

Chapter 6

Future work

Over the course of this dissertation we have presented methods for pre-processing DCE-MRI data as well as methods for distinguishing benign from malignant lesions and identifying molecular subtypes of breast cancer on DCE-MRI.

For the spectral embedding based registration (SERg) method, we showed that such a method is appropriate for settings of increased noise such as that presented by DCE-MRI. Although simulated brain MRI data demonstrated the resilience and robustness of SERg to noise compared to using a intensity-based registration [20], demonstrated improvements by SERg in actual breast MRI data were quite small. The current working model for SERg has many moving parts such as the α-MI formulation, the method for image transformation after the registration optimization has been performed, and the features used as input to the spectral embedding. In addition, a hierarchical or local (rather than global) registration scheme may be more advantageous in the context of breast MRI since the ultimate goal is to align time point images most accurately for kinetic feature analysis.

The spectral embedding based active contour (SEAC) segmentation has produced excellent results compared to the fuzzy c-means "state of the art." A further improvement, however, would be to extract optimal weights for the region and boundary terms of the hybrid active contour by using image-specific information such as the initial region and edge information for a particular image. In Chapter 3, we learn the optimal weights for a given initialization using a brute force method and based on the best segmentations compared to ground truth for a testing set of data. Due to the computational cost of such a brute force method, we feel this could be done in a more intelligent fashion.

For both SERg and SEAC, the novel aspects are derived from the use of spectral embedding to transform the DCE-MRI to an alternative image space. The beauty of this method is that it is model-free and undersupervised (in other words, no domain-specific information is required such as is needed for a pharmacokinetic model). As has been demonstrated for many biological applications outside of this thesis [34], spectral embedding also seems to capture the nonlinearities of biological data better than an linear dimensionality reduction method like PCA. The disadvantage, however, of spectral embedding is the computational memory required for the large matrices generated by MR images. For example, for a standard 512×512 image, the similarity matrix alone requires $(512)^2 \times (512)^2 \times 16 = 109.9GB$ of memory. For most of the images analyzed in Chapter 3, we actually had to downsample the images just to calculate the similarity matrices. Future work should include optimization of the spectral embedding and the possibilities of using a sparse matrix with some constraints imposed for spatial or similarity neighborhoods or perhaps even using an approximation to the similarity matrix such as the Nystrom approximation presented in Chen et al. [106] and Fowlkes et al. [107]. Importantly, because the ultimate use is in a medical diagnostic test, the trade-offs in accuracy for using such approximations must also be thoroughly explored.

In Chapter 4, a benign versus malignant lesion classification is presented in which *textural kinetics* are first presented. The utility of such features in the context of benign versus malignant lesion classification is presented. What we see with textural kinetics is that patterns of internal enhancement have just as much impact on lesion diagnosis as the more traditional morhpological and kinetic signal intensity features. Future work in textural kinetics would be in the manner of measuring texture over time. We empirically chose a cubic fit to the texture versus time curve, but more sophisticated models to fit the curve could be explored.

In Chapter 5, we take on the most challenging of our proposal specific aims in identifying molecular subtypes of breast cancer on DCE-MRI. Similar to the benign versus malignant lesion study, we found that morphology and textural kinetics play a large role in distinguishing TN CAs from other subtpyes of breast cancer as well as distinguishing TN CAs from benign mimic lesions (i.e., fibroadenomas). One feature

type that we were hoping to quantify in this study was *rim enhancement*. Uematsu et al. [93] observed this as a highly correlative lesion feature for TN CAs. This feature would also be interesting from the perspective of linking the MRI signatures to the histopathologic signatures. Yamaguchi et al. [108] anecdotally demonstrated that such a "ring-like appearance," or rim enhancement is correlated with an acellular center of the tumor, often thought to be a central necrosis. This feature is also interesting because it presents the opportunity to also use information from the T2-weighted MRI, which is typically only used for confirmation of benign lesions such as cysts or fibroadenomas. This also provides an interesting segue to the bridge between MRI features and histopathologic features.

As genotypic and phenotypic information as well as prognosis become more important factors in breast cancer diagnosis and treatment, there is increasing interest in bridging the gap between diagnostic tools and confirmatory diagnosis, or ground truth. Baltzer et al. [109] recently performed a study in which the kinetic features of invasive breast carcinomas on DCE-MRI were compared with histologic features. They found that volume enhancement (defined by fraction of tumor that enhances) was connected to lymph node status and ER and HER2 status. They also found that total lesion volume and plateau voxel volume (defined by fraction of tumor that shows plateaued enhancement) were predictors of ER and HER2 status. Strongest initial enhancement was most predictive of ER- status, and TTP was most indicative of positive lymph node status. This connection between diagnosis and prognosis is the next step to fortifying DCE-MRI's position in the breast cancer diagnosis and prognosis spectrum, and a computer-aided analysis solution including features from DCE-MRI in conjunction with features from histopathology would lend itself well to this problem. This will likely come in the form of kinetic texture on MRI combined with cellular architecture on pathology to reflect a better biological understanding of the phenomena observed on DCE-MRI. Such analyses will undoubtedly provide insight not only to the biological architecture of a tumor, but also the prognosis for the tumor as observed on DCE-MRI.

Appendix A

Additional Experiments on Pharmacokinetic Modeling

A.1 Implementing the Brix model

A.1.1 Introduction

Previous Work

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is currently used as a secondary tool to conventional x-mammography in breast lesion evaluation. X-ray mammography is still the current gold standard for screening and offers high 2-dimensional resolution, which is advantageous for detecting small variations in tissue composition, such as microcalcifications. However, due to the constraints of imaging a 3 dimensional structure in a single plane, ultrasound or breast DCE-MRI is often used as a secondary imaging technique when a suspicious lesion is found on mammography. Ultrasound is very good at detecting tissue composition, so it is able to provide additional information to the mammogram in situations where the breast tissue is dense or a cystic mass needs to be ruled out. DCE-MRI is also very good at imaging dense breasts, but its major advantages over mammography and ultrasound are the ability to image the entire breast as thin slices that comprise the entire breast volume and the ability to measure variations in contrast uptake that provide information about the vascularity of the breast tissue; because malignant tumors often have a high density of blood vessels that are poorly formed, and thus, leaky, they take up contrast dye at a different rate from benign ones, and radiologists are able to observe this difference in the DCE-MR images [54].

It has been shown in clinical decision-making that signal intensity kinetics are important for breast lesions classification, and a large body of the research literature

is devoted to this. Contrast enhancement is performed by first injecting gadolinium diethylenetriamine-pentaacid (Gd-DTPA) into the patient's bloodstream and concurrently acquiring MRI images of the breast. Since malignant lesions tend to grow leaky blood vessels in abundance, contrast agent is taken up by tumors perferentially [56]. Kuhl et al. [54] found that data in the temporal MRIs could be plotted as single data points on a time series curve that was reflective of the lesion type. It was shown that malignant lesions had a characteristic curve, showing rapid uptake of contrast (steep positive initial slope) and rapid washout (subsequent negative slope). Benign lesions had slow contrast uptake (small positive initial slope) and then plateaued or did not reach peak intensity during the time series. These phenomena are illustrated in Figure A.1. This description of the DCE-MRI data is now considered convention in radiologic interpretation of breast DCE-MRI. Although this is a great advantage of DCE-MRI over the other modalities and contributes to the high sensitivity rates reported for breast DCE-MRI, both benign and malignant neoplastic tissue frequently have contrast enhancement patterns that differ from normal breast tissue, and these abnormalities are highlighted in the time-dependent MRIs. As such, it is often difficult for radiologists to differentiate between benign and malignant lesions simply by visually inspecting the contrast-enhanced lesion on the postcontrast MRI, which can be appreciated from Figure A.1 in comparing the malignant lesion curve with that of the curve for fibrocystic changes, a type of benign lesion. Nevertheless, DCE-MRI is a main component of breast lesion characterization,

One of the ways in which the DCE-MRI data is quantified is by using pharmacokinetic (PK) models. PK models assume that the contrast agent used in DCE-MRI migrates through the bloodstream and the body's tissues in a manner similar to that of a drug. In the Hayton pharmacokinetic model [5], a two compartment model based on diffusion is implemented to determine tissue properties related to the rate of contrast uptake and washout of the lesion in question. Figure A.2 shows a diagram of the model. The central compartment represents the blood plasma, or the blood vessels where the blood and injected contrast agent freely flow. When the contrast agent reaches the

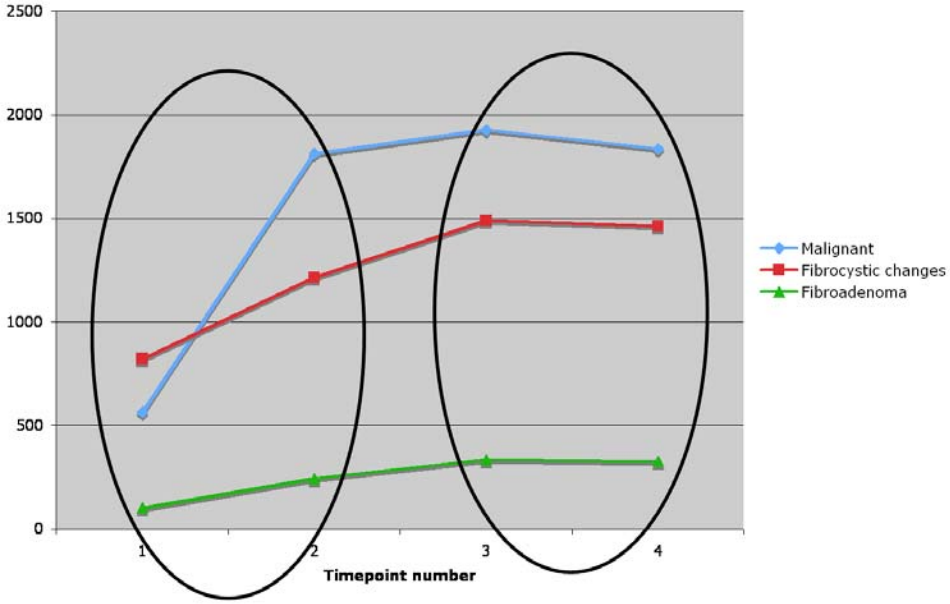


Figure A.1: Sample signal intensity versus time curves for one malignant and two benign lesion types.

capillaries, it moves across the capillary wall and into the peripheral compartment, representing the body's tissues. Material also moves from the peripheral compartment into the central compartment. This exchange of materials is crucial for delivery of oxygen and nutrients to the tissues and for the elimination of carbon dioxide and waste materials from the tissues. This is also the exchange of which DCE-MRI takes advantage in determining tissue composition of suspicious lesions.

Although pharmacokinetic models are popular methods of analysis for DCE-MRI in tumor analysis, many current softwares used by clinicians are proprietary, making it difficult to replicate a PK analysis performed by another research group. For this study, implementing a PK model was undertaken in order to have a commonly used DCE-MRI model to compare to experimental methods currently being developed in our laboratory. In this study, we implement the Hayton 2-compartment pharmacokinetic model and apply it to a dataset of actual clinical breast DCE-MR images. Since PK models are typically used to differentiate between benign and malignant lesions based

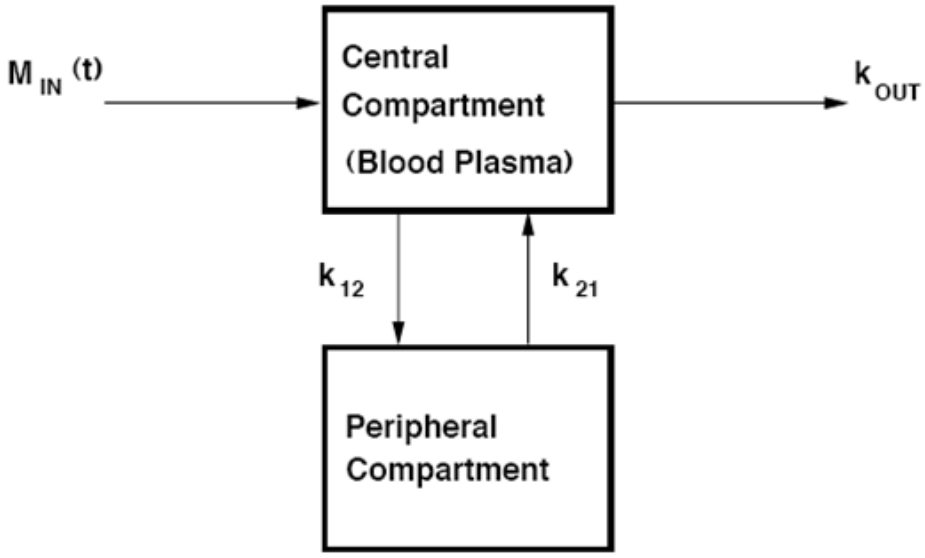


Figure A.2: Physiologically-based two compartment model used by Hayton [5] to represent material exchange across capillary.

on parameter values derived from the PK model, these data are then plotted in a 3-dimensional space to examine differences between derived parameters for different lesion types.

A.1.2 Materials and Methods

Data Description

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 and histologic examination. 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) at either 1.5 Tesla or 3 Tesla (Siemens Magnetom or Trio, respectively).

Temporal resolution between post-contrast acquisitions was 45-90 seconds. The lesion region of interest (ROI) was then manually segmented in MRIcro [48] 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.

Derivation of the Model

From the two-compartment model, two differential equations describing the diffusion of materials across the capillary membrane can be written as the following set of equations.

$$V_c \frac{dC_c}{dt} = k_{21} V_p C_p - (k_{12} + k_{out}) V_c C_c + M_{in} \tag{A.1}$$

$$V_p \frac{dC_p}{dt} = k_{12} V_c C_c - k_{21} V_p C_p \tag{A.2}$$

where C_c, C_p, V_c, and V_p are the concentrations and volumes of the central and peripheral compartments, respectively. By substituting $k'_{out} = k_{out} + k_{12}$ and $X(s)$ as the Laplace transform of M_{in}/V_c, we find the transfer function relating the concentration of contrast agent, $C_p(s)$ to the input function of contrast agent, $X(s)$ to be:

$$\frac{C_p(s)}{X(s)} = \frac{A}{(s+a)(s+b)} \tag{A.3}$$

We assume a bolus injection of contrast agent, which allows us to set $X(s) = \mathcal{L}(\delta(t)) = 1$. By substituting for k_1 and k_2 parameters and and taking the inverse Laplace transform, the final time-dependent model equation is

$$C(t) = \frac{A}{a-b}(e^{-bt} - e^{-at}) \tag{A.4}$$

This is the equation that is used to fit the data. See Hayton et al. [5] for full details of derivation.

Parameter Estimation

To perform a comparison between lesions in the dataset, a single curve is used to represent each lesion. The mean of the signal intensity within each lesion's ROI was calculated at each timepoint before and after contrast injection. The mean signal

intensity versus time is then plotted for each lesion. Once the exponential equation needed for curve-fitting is derived, the fit() function in MATLAB is used to obtain A, a, and b for each lesion.

A.1.3 Results

Figure A.3 shows an example of the curve-fitting perfomeed for each datapoint, using MATLAB. It was suprising that the model fit the data so well, given the simplicity of the Hayton model. From these curve-fits, the parameters A, a, and b were determined.

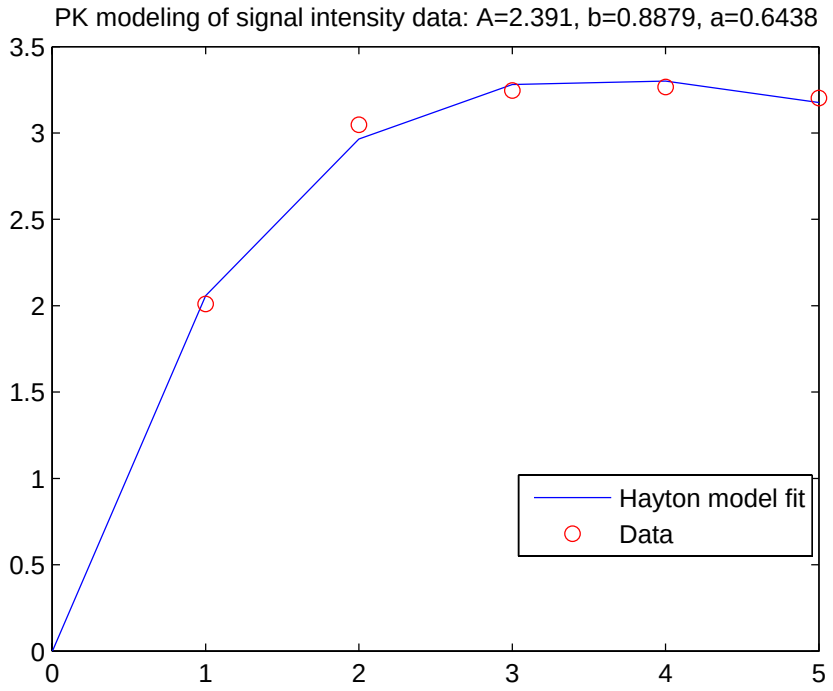


Figure A.3: Sample curve-fitting of datapoints using derived exponential model.

Figure A.4 shows a three-dimensional plot of A vs. a vs. b for all lesions in the dataset. Although the A parameter seemed reasonably different between the benign and malignant lesions, there was a lot of similarity within a values and b values for both the benign and malignant lesion classes.

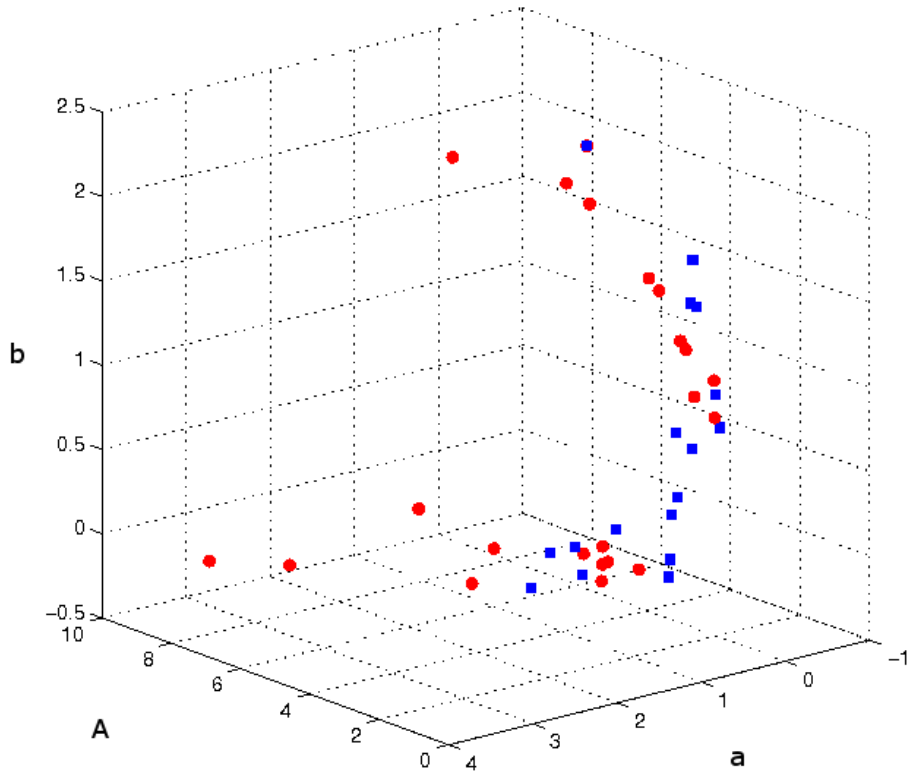


Figure A.4: Three dimensional plot showing parameters A,a, and b for all lesions in dataset.

A.1.4 Concluding Remarks

In this study, the Hayton [5] pharmacokinetic model was implemented for quantifying breast DCE-MRI data in a physiologically-based manner. The model coefficient estimation did not produce any obvious data classifiers. However, since we used a mean signal intensity value for each timepoint, it might be more useful to calculate these curves at each pixel to create a parametric map of the lesion rather than using this model for global lesion characterization. As a preliminary case, we calculated the parametric maps for one benign and one malignant lesion (see Figure A.5).

It was found in this specific case, that the a and b values for the majority of the pixels for the benign lesion were an order of magnitude smaller than those of the malignant

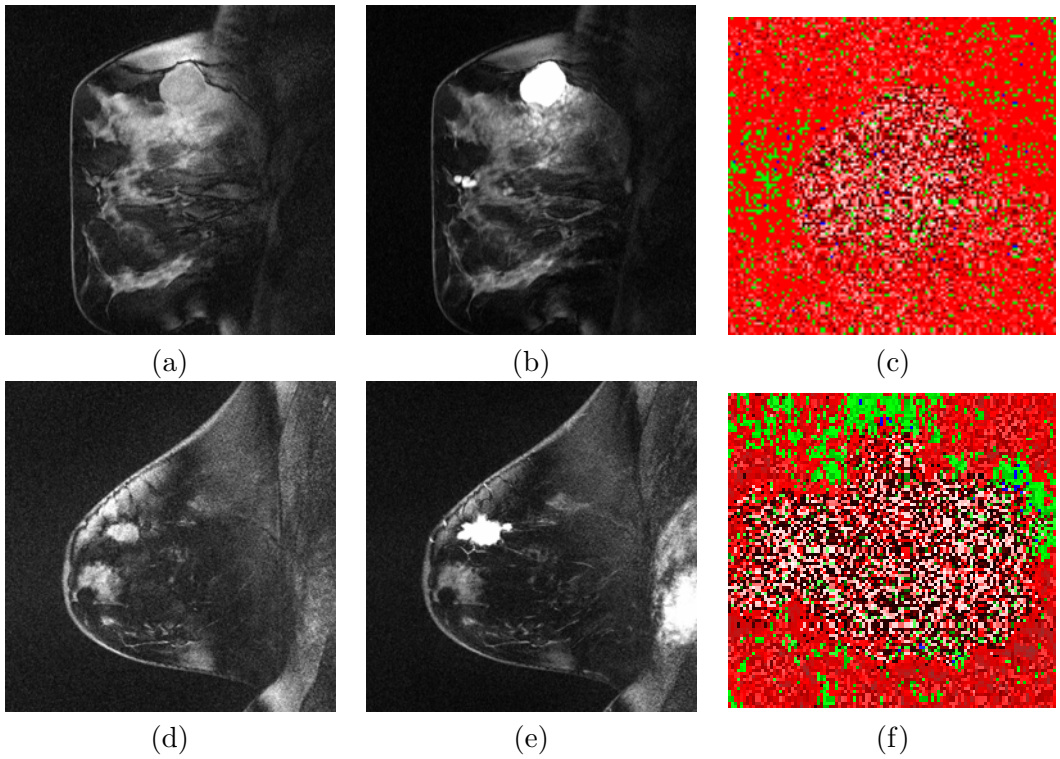


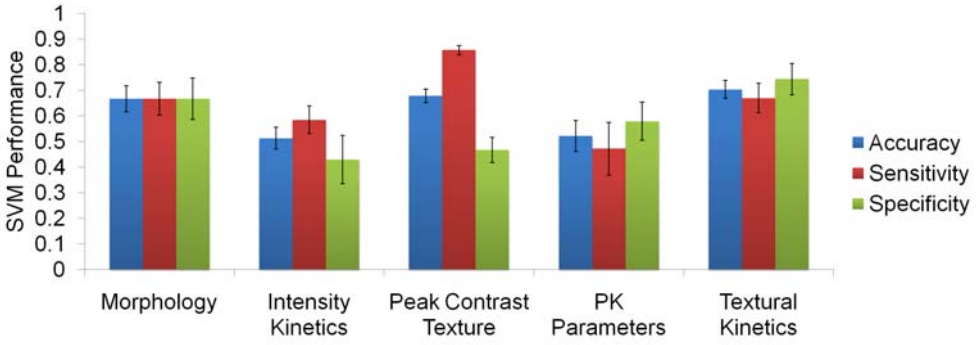
Figure A.5: An example of the contrast enhancement of (a-c) a benign fibroadenoma and (d-f)a malignant lesion are shown. The lesions show (from left to right) the full precontrast image (a,d), the postcontrast timepoint showing peak enhancement, (b,e), simulated pharmacokinetic maps,(c,f).

lesion. In addition, Figure A.5(c) and (f) shows an observable difference in the lesion brightness from the surrounding tissue brightness. This was observed for both the benign and the malignant lesion. If the trend follows for other lesions in the dataset, this may prove to be helpful in lesion segmentation in the future rather than targeting the PK models as a classification tool. Other future directions for this project include finding ways to increase the speed of computation, since pixel-wise lesion analysis takes approximately 1-2 hours on a server with a 2.33GHz processor. In addition, it may also be useful to now explore other PK models such as the Tofts model [6] or the Kety [6] model, which are more complicated models that take additional physiological parameters such as vessel leakage into account, to find the best model to fit my dataset.

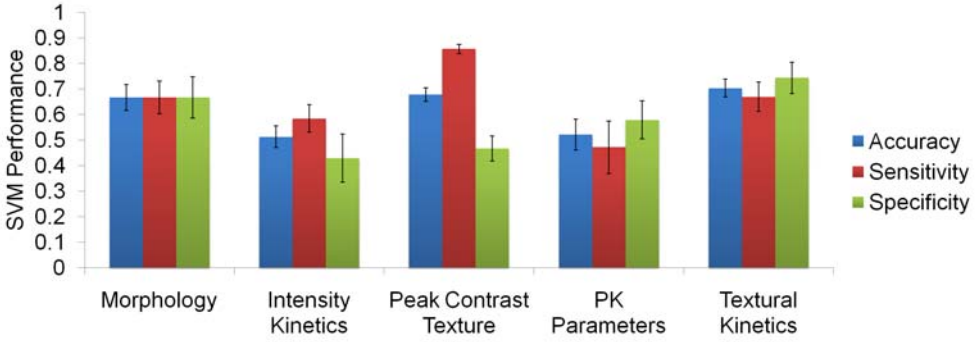
A.2 Implementing the Tofts Model

The Tofts model is the most widely used pharmacokinetic model. A biexponential decay is often used as a standard arterial input function to the system such that [110]: $C_p(t) = D\sum_{i=1}^{2} a_i^T e^{-m_i t}$ where D is the dose (mmole/kg), and $a_1^T = 3.99$ kg/L, $a_2^T = 4.78$ kg/L, $m_1 = 0.144 min^{-1}$, and $m_2 = 0.011 min^{-1}$. These constants are based on experiments performed by Weinmann [111] for normal control subjects and have been used in many studies [6, 110, 112]. The pharmacokinetic model is then modeled as [112]: $C_t(t) = \frac{K^{trans} C_{p0}}{K_{ep}} = v_e C_{p0} e^{-k_{ep} t}$ The model is then fitted to the data using the Levenberg-Marquardt algorithm for curve fitting. From the curve fitting, K_{trans} and v_e are extracted. However, pharmacokinetic models tend to have some drawbacks. First, the most accurate way to approximate the physiologic model is to actually measure an arterial input function from the patient [113]. In addition, studies have shown that temporal resolution of the images as well as image artifacts [114] such as bias field can greatly affect the accuracy of measurements. Di Giovanni et al. [114] showed that on a 3 Tesla scanner, incorrect T1 estimation and radiofrequency (RF) field inhomogeneity can cause errors ranging from 66% up to 531% for K_{trans} and 74-233% for v_e. Temporal resolution seemed to have the least effect, demonstrating errors up to 48% in K_{trans} and negligible errors for v_e when decreasing the temporal resolution from 10 to 70 seconds between time points.

Figure A.7 shows 9 examples of datasets for which the Tofts parameters were calculated in a pixel-wise fashion. The resulting colormaps showed that K_{ep} and K_{trans} were not informative in this setting. However, v_e did demonstrate reasonable results. The parenchyma can be reasonably differentiated from the rest of the breast. In a 2010 *ISMRM* abstract [1], we also demonstrated the utility of these parameters in a lesion-wise fashion by comparing the classification accuracy of distinguishing TN lesions from FAs and then also from ER+ CAs by means of PK parameters and textural kinetics. It was found that textural kinetics outperform both morphological features and PK parameters in distinguishing TN lesions from other lesion subtypes see Figure A.6.



(a)



(b)

Figure A.6: Results of classification of (a) TN CAs versus FAs and (b) TN CAs versus ER+ CAs using 5 different feature types in conjunction with a support vector machine classifier.

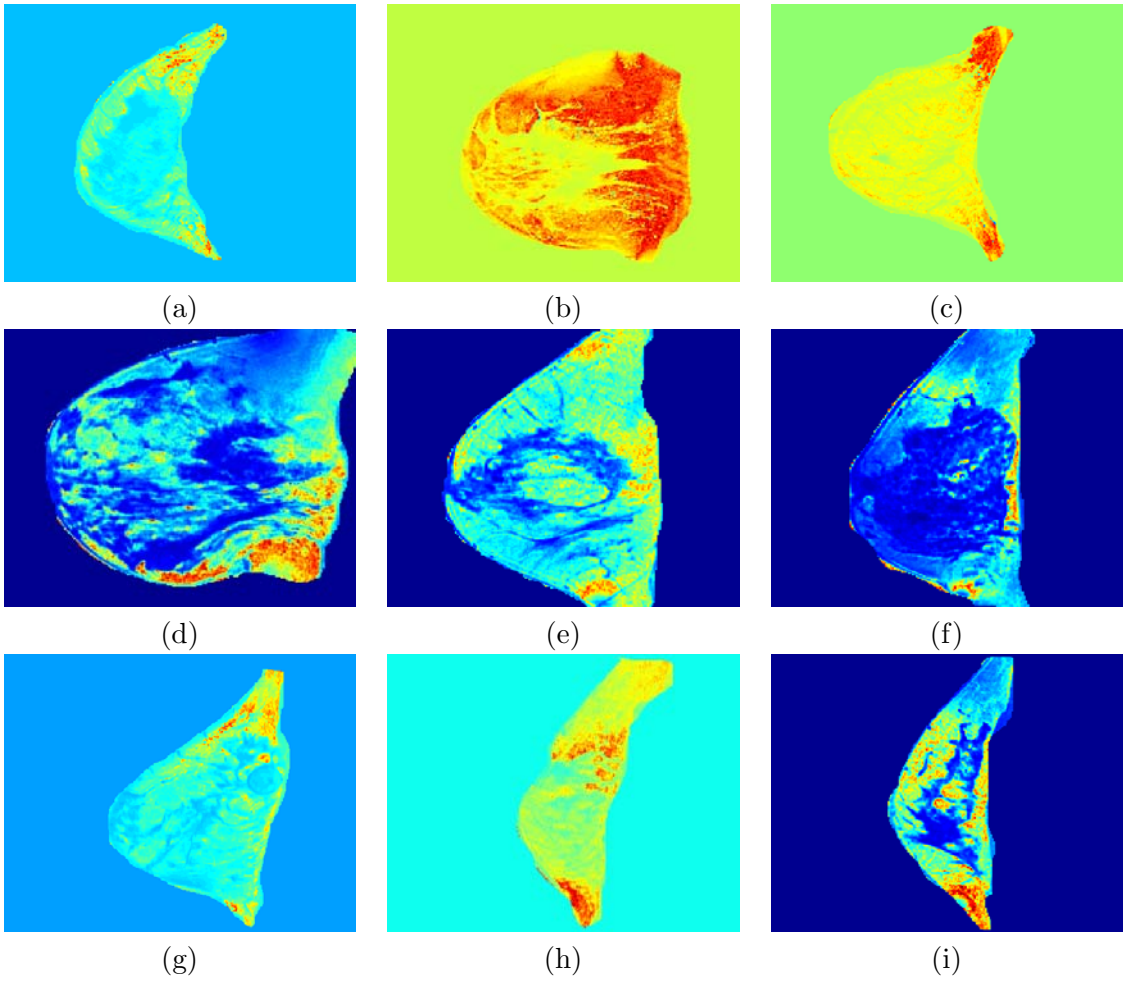


Figure A.7: Nine different examples of the v parameter for the Tofts [6] pharmacokinetic model.

Appendix B

Machine Learning Summary

Throughout the course of this dissertation, various methods for computing data relationships are used. In this appendix, spectral embedding (SE), a manifold learning technique for understanding data relationships, is reviewed. Support vector machines, which provide the final measures of lesion classification in Chapters 4 and 5, are also reviewed.

B.1 Dimensionality reduction

In this thesis, a nonlinear dimensionality reduction method called spectral embedding, which is calculated via an eigenvalue decomposition of a similarity matrix, is frequently used. Over the course of this dissertation, we have been asked numerous times why we do not use principal component analysis (PCA) or how PCA performs compared to SE. For clarity, we provide a brief review of principal component analysis (PCA) and then provide a brief comparison.

B.1.1 Pricipal component analysis

PCA [115] attempts to reduce the dimensionality of the data while retaining maximum variance of the dataset. PCA is most popularly implemented by performing an eigenvalue decomposition of a covariance matrix generated from the original data. The resulting eigenvectors are then considered to be the principle components, and the first few retain the maximum variance in the original dataset. In addition, if the eigenvectors are chosen to be orthonormal, then the variance captured by a given eigenvector is reflected by the corresponding eigenvalue.

Using the same notation found in Chapter 3 for SE, let $\mathcal{F} = [\mathbf{F}(\mathbf{x}_1), \mathbf{F}(\mathbf{x}_2), \ldots, \mathbf{F}(\mathbf{x}_N) \in$

$\mathbb{R}^{N\times D}]$ be the data matrix of N feature vectors with dimensionality D. The aim of PCA is to reduce $\mathcal{F} \in \mathbb{R}^{N\times D}$ to a low d-dimensional space where $d << D$. Let $\mathbf{V_{PCA}} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N \in \mathbb{R}^{N\times d}]$ be the optimal low dimensional projections [115] and the associated eigenvectors of a given object $\mathbf{x}_{i,PCA}$, where $i \in \{1, \ldots, N\}$, are $\mathbf{v}_i \in \mathbb{R}^{1\times d}$ where $\mathbf{v}_{i,PCA} = [v_{1,PCA}, v_{2,PCA}, \ldots, v_{N,PCA}]$ and $v_j \in \mathbb{R}^{1\times 1}, j \in \{1, \ldots, d\}$ is an individual eigenvector for a given $\mathbf{v}_i$. The optimal $\mathbf{V_{PCA}}$ can be obtained by solving the eigenvalue decomposition of,

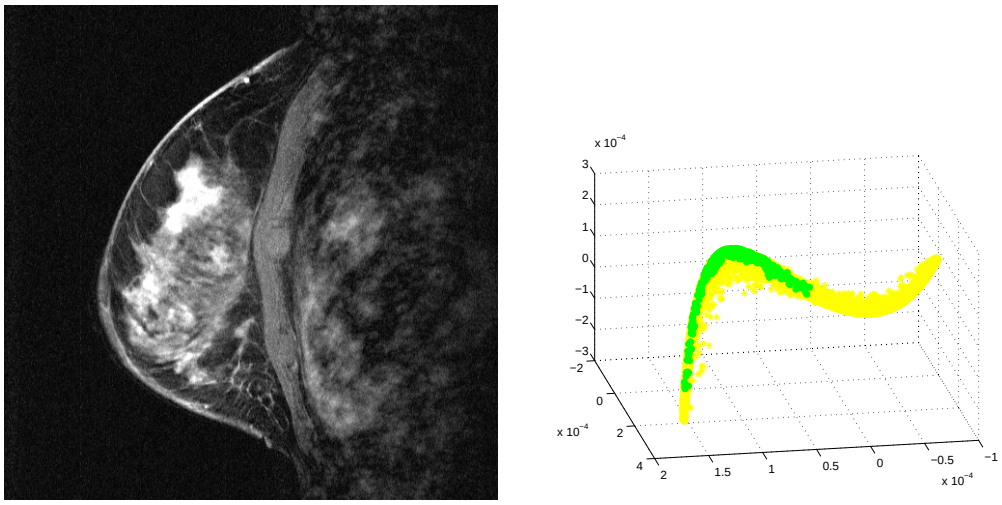
$$\Sigma_{PCA}\mathbf{v_{PCA}} = \lambda_{PCA}\mathbf{v_{PCA}} \tag{B.1}$$

where Σ_{PCA} is the covariance matrix of $\mathcal{F}$. This can be compared to the eigenvalue decomposition for SE in Equation 3.2: $(\mathbf{D} - W)\mathbf{v} = \lambda\mathbf{D}\mathbf{v}$.

B.1.2 Spectral embedding

SE is a nonlinear dimensionality reduction method that allows one to transform a high-dimensional data space to a low-dimensional data representation. The implementation that we chose to use is in the form of normalized cuts first proposed by Shi and Malik [19]. Normalized cuts was originally designed as a graph partitioning problem, where the connectivity, or similarity between every point and every other point in the image space is represented by the similarity between associated feature vectors. The optimal partition between data is one which minimizes intercluster similarity and maximizes intracluster similarity.

Although both PCA and SE are derived from an eigenvalue decomposition, PCA is a linear DR method while SE is a nonlinear DR method. This means that when performing PCA, any data point in the original data space can be recapitulated as a linear combination of the principle components. This is an advantage of PCA. However, PCA is often unable to capture nonlinear relationships between data in the original data space, and we have found this to be the case, both in performed pixel-wise and lesion-wise DR. Figure B.2 demonstrates the comparison in a lesion-wise DR using PCA (FigureB.2(a)) and SE (FigureB.2(b)) where the aim was to distinguish between TN and non-TN lesions. One can see the improved separation between lesion classes in



(a) (b)

Figure B.1: Post- contrast time point (a) of a breast lesion and (b) associated spectral embedding manifold (yellow, non-lesion pixels; green, lesion pixels). When associated with their appropriate spatial locations, a separation between lesion and non-lesion area can be resolved.

FigureB.2(b).

B.2 Support Vector Machines

The computer-based classifier most frequently used in this dissertation is a support vector machine (SVM) classifier. SVMs are a kernel-based method that projects the data from the original space into a kernel space where the data can be separated via a hyperplane (see Figure B.3). The optimal hyperplane is found by maximizing the distance between the decision hyperplane and the so-called support vectors [102], which are determined by the data points in the kernel space which are closest to the decision hyperplane. At a very simplistic level, SVMs can be thought of as the opposite of dimensionality reduction in implementation because the mapping function actually maps the data into a higher dimensional space in order to separate the data classes optimally. This is demonstrated in Figure B.3 where the data is projected from a 2-dimensional to a 3-dimensional space so that a 2-D plane can separate the data.

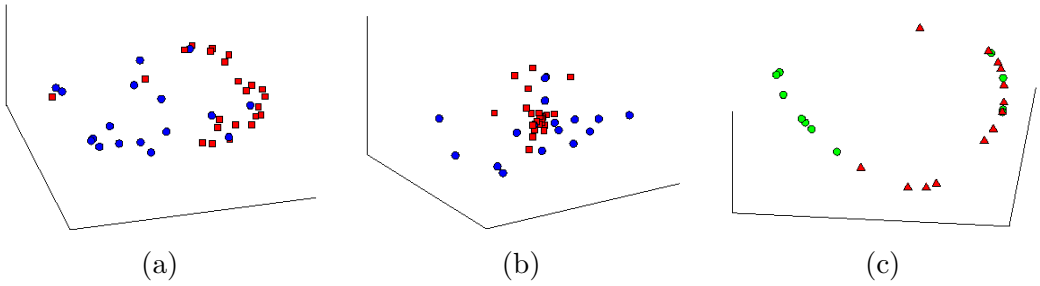


Figure B.2: Embedding plots showing: (a,b) all lesions from first order kinetic texture feature space (malignant: red squares, benign: blue circles) in (a) SE- reduced dimensional space and (b) PCA- reduced dimensional space; (c) malignant lesions only (TN: red triangles, non-TN: green circles) using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

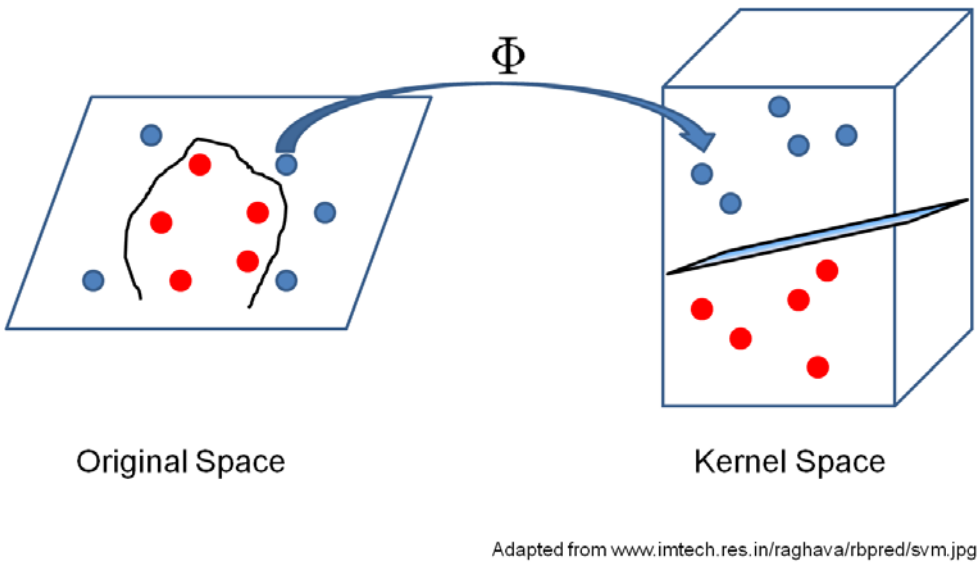


Figure B.3: Post- contrast time point (a) of a breast lesion and (b) associated spectral embedding manifold (yellow, non-lesion pixels; green, lesion pixels). When associated with their appropriate spatial locations, a separation between lesion and non-lesion area can be resolved.

Vita

Shannon Christine Agner

2011 Ph.D. in Biomedical Engineering from Rutgers, The State University of New Jersey

2004 M.S. in Engineering Sciences from Dartmouth College

2002 A.B. in Engineering Sciences from Dartmouth College

1998 Graduated from Cherry Hill High School East, Cherry Hill, New Jersey

2008-present Graduate Fellow, Department of Biomedical Engineering, Rutgers, The State University of New Jersey

Patents System and Method for Automated Segmentation, Characterization, and Classification of Possibly Malignant Lesions and Stratification of Malignant tumors, Anant Madabhushi, Shannon Agner, Mark Rosen, PCT/US09/34505, Feb. 2009.

Peer-reviewed Journal Papers Computerized Image Analysis for Identifying Triple Negative Breast Cancers and Distinguishing Triple Negative Breast Cancers from Other Molecular Subtypes of Breast Cancer on DCE-MRI: A Feasibility Study. Under Review.

Madabhushi A, Agner S, Basavanhally A, Doyle S, Lee G. Computer-aided prognosis: Predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data. Computerized Medical Imaging and Graphics. In Press.

Agner, S, Soman, S, Libfeld, E, McDonald, M, Thomas, K, Englander, S, Rosen, M, Chin, D, Nosher, J, Madabhushi, A, "Textural Kinetics: A Novel Dynamic Contrast Enhanced (DCE)-MRI Feature for Breast Lesion Classification", Journal of Digital Imaging, 2010.

Basavanhally, A. N., Ganesan, S., Agner, S., Monaco, J. P., Feldman, M. D., Tomaszewski, J. E., Bhanot, G., Madabhushi, A., "Computerized Image-Based Detection and Grading of Lymphocytic Infiltration in HER2+

Breast Cancer Histopathology," Biomedical Engineering, IEEE Transactions on, vol.57 (3), pp. 642-653, 2010.

Peer-reviewed Conference Papers Agner SC, Xu J, Karthigeyan S, Rosen, MA, Englander S, Madabhushi, A. Spectral Embedding based Active Contour (SEAC): Application to Breast Lesion Segmentation on DCE-MRI. SPIE Medical Imaging 2011. Best Student Paper Award Finalist.

Madabhushi A, Basavanhally A, Doyle S, Agner S, Lee G., "Computer-aided prognosis: Predicting patient and disease outcome via multi-modal image analysis", Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on (ISBI) pp. 1415-1418, 2010.

Agner, S, Xu, J, Fatakdawala, H, Ganesan, S, Madabhushi, A, Englander, S, Rosen, M, Thomas, K, Schnall, M, Feldman, M, Tomaszewski, J, "Segmentation and classification of triple negative breast cancers using DCE-MRI," Biomedical Imaging: From Nano to Macro, pp.1227-1230, 2009.

Shannon C. Agner, Salil Soman, Edward Libfeld, Margie McDonald, Mark A. Rosen, Mitchell D. Schnall, Deanna Chin, John Nosher, and Anant Madabhushi. Novel kinetic texture features for breast lesion classification on dynamic contrast enhanced (DCE) MRI. Proc. SPIE 6915, 69152C (2008).

Doyle, S.; Agner, S.; Madabhushi, A.; Feldman, M.; Tomaszewski, J., Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features, Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on , vol., no., pp.496- 499, 14-17 May 2008.

Naik, S.; Doyle, S.; Agner, S.; Madabhushi, A.; Feldman, M.; Tomaszewski, J., "Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology," Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on , vol., no., pp.284-287, 14-17 May 2008.

References

[1] Shannon Agner, Mark Rosen, Sarah Englander, Diana Sobers, Kathleen Thomas, John Tomaszewski, Michael Feldman, Shridar Ganesan, Mitchell Schnall, and Anant Madabhushi. Distinguishing molecular subtypes of breast cancer based on computer-aided diagnosis of dce-mri. *International Society for Magnetic Resonance in Medicine Annual Meeting*, 2010.

[2] Shannon Agner, Salil Soman, Edward Libfeld, Margie McDonald, Kathleen Thomas, Sarah Englander, Mark Rosen, Deanna Chin, John Nosher, and Anant Madabhushi. Textural kinetics: A novel dynamic contrast-enhanced (dce)-mri feature for breast lesion classification. *Journal of Digital Imaging*, pages 1–18, 2010. 10.1007/s10278-010-9298-1.

[3] Shannon C. Agner, Salil Soman, Edward Libfeld, Margie McDonald, Mark A. Rosen, Mitchell D. Schnall, Deanna Chin, John Nosher, and Anant Madabhushi. Novel kinetic texture features for breast lesion classification on dynamic contrast enhanced (dce) mri. volume 6915, page 69152C. SPIE, 2008.

[4] Shannon C. Agner, Jun Xu, Mark Rosen, Sudha Karthigeyan, Sarah Englander, and Anant Madabhushi. Spectral embedding based active contour (seac): application to breast lesion segmentation on dce-mri. volume 7963, page 796305. SPIE, 2011.

[5] Paul Hayton, Michael Brady, Lionel Tarassenko, and Niall Moore. Analysis of dynamic mr breast images using a model of contrast enhancement. *Medical Image Analysis*, 1(3):207–224, 1997.

[6] Paul S. Tofts, Bruce Berowitz, and Mitchell D. Schnall. Quantitative analysis of dynamic gd-dtpa enhancement in breast tumors using a permeability model. *MRM*, 33:564–568, 1995.

[7] J. S. Reis-Filho and A. N J Tutt. Triple negative tumours: a critical review. *Histopathology*, 52(1):108–118, 2008.

[8] Debbie Saslow, Carla Boetes, Wylie Burke, Steven Harms, Martin O. Leach, Constance D. Lehman, Elizabeth Morris, Etta Pisano, Mitchell Schnall, Stephen Sener, Robert A. Smith, Ellen Warner, Martin Yaffe, Kimberly S. Andrews, and Christy A. Russell. American cancer society guidelines for breast screening with mri as an adjunct to mammography. *CA: A Cancer Journal for Clinicians*, 57:75–89, 2007.

[9] *American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BIRADS Atlas).* American College of Radiology, Reston, VA, 2003.

[10] C.W. Piccoli. Contrast-enhanced breast mri: factors affecting sensitivity and specificity. *European Radiology*, 7 (Suppl. 5):S281–S288, 1997.

[11] Noel Weidner, Joseph P. Semple, William R. Welch, and Judah Folkman. Tumor angiogenesis and metastasis- correlation in invasive breast carcinoma. *New England Journal of Medicine*, 324(1):1–8, 1991.

[12] Y. Wang, D. M. Ikeda, B. Narasimhan, T. A. Longacre, R. J. Bleicher, S. Pal, R. J. Jackman, and S. S. Jeffrey. Estrogen receptor-negative invasive breast cancer: imaging features of tumors with and without human epidermal growth factor receptor type 2 overexpression. *Radiology*, 246(2):367–75, 2008.

[13] S. Makkat, R. Luypaert, T. Stadnik, C. Bourgain, S. Sourbron, M. Dujardin, J. De Greve, and J. De Mey. Deconvolution-based dynamic contrast-enhanced mr imaging of breast tumors: correlation of tumor blood flow with human epidermal growth factor receptor 2 status and clinicopathologic findings–preliminary results. *Radiology*, 249(2):471–82, 2008.

[14] S. Schrading and C. K. Kuhl. Mammographic, us, and mr imaging phenotypes of familial breast cancer. *Radiology*, 246(1):58–70, 2008. Schrading, Simone Kuhl, Christiane K Comparative Study United States Radiology Radiology. 2008 Jan;246(1):58-70.

[15] Laszlo G Nyul, Jayaram K Udupa, and Punam K Saha. Incorporating a measure of local scale in voxel-based 3-d image registration. *IEEE Trans Med Imaging*, 22(2):228–237, Feb 2003.

[16] Punam K. Saha, Hui Zhang, Jayaram K. Udupa, and James C. Gee. Tensor scale-based image registration. volume 5032, pages 314–324. SPIE, 2003.

[17] Marius Staring, Uulke A van der Heide, and et al. Registration of cervical mri using multifeature mutual information. *IEEE TMI*, 28(9):1412–1421, Sep 2009.

[18] A. Hero H. Neeumuchwala and P. Carson. Image matching using alpha-entropy measures and entropic graphs. volume 85, pages 277–296, 2005.

[19] Jianbo Shi and J. Malik. Normalized cuts and image segmentation. *IEEE PAMI*, 22(8):888–905, 2000.

[20] Daniel Rueckert, L. I. Sonoda, and et al. Non-rigid registration using free-form deformations: Application to breast mr images. *IEEE TMI*, 18(8):712–721, 1999.

[21] A.C. Evans R.K.-S. Kwan and G.B. Pike. Mri simulation-based evaluation of image-processing and classification methods. *IEEE Trans. on Med. Imaging*, 18(11):1085–97, 1999.

[22] F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE PAMI*, 11(6):567–585, 1989.

[23] Alexis Roche, Grgoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *MICCAI'98,LNCS1496*, pages 1115–1124. Springer Verlag, 1998.

[24] Jr. Yezzi, A., S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum. A geometric snake model for segmentation of medical imagery. *Medical Imaging, IEEE Transactions on*, 16(2):199 –209, 1997.

[25] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22:61–79, 1997. 10.1023/A:1007979827043.

[26] T.F. Chan and L.A. Vese. Active contours without edges. *Image Processing, IEEE Transactions on*, 10(2):266 –277, February 2001.

[27] Nikos Paragios and Rachid Deriche. Coupled geodesic active regions for image segmentation: A level set approach. In David Vernon, editor, *Computer Vision ECCV 2000*, volume 1843 of *Lecture Notes in Computer Science*, pages 224–240. Springer Berlin / Heidelberg, 2000.

[28] Debra M. Ikeda, Nola Hylton, Karen Kinkel, Mary G. Hochman, Christiane Kuhl, Werner A. Kaiser, Jeffrey C. Weinreb, Stanley F. Smazal, Hadassah Degani, Petra Viehweg, John Barclay, and Mitchell D. Schnall. Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging. *Journal of Magnetic Resonance Imaging*, 13:889–895, 2001.

[29] P.K. Saha and J.K. Udupa. Scale-based diffusive image filtering preserving boundary sharpness and fine structures. *Medical Imaging, IEEE Transactions on*, 20(11):1140 –1155, 2001.

[30] Punam K. Saha, Jayaram K. Udupa, and Dewey Odhner. Scale-based fuzzy connected image segmentation: Theory, algorithms, and validation. *Computer Vision and Image Understanding*, 77(2):145 – 174, 2000.

[31] Punam Kumar Saha. Tensor scale: A local morphometric parameter with applications to computer vision and image processing. *Computer Vision and Image Understanding*, 99(3):384 – 413, 2005.

[32] Erez Eyal, Daria Badikhi, Edna Furman-Haran, Fredrick Kelcz, Kevin J Kirshenbaum, and Hadassa Degani. Principal component analysis of breast dce-mri adjusted with a model-based method. *J Magn Reson Imaging*, 30(5):989–998, Nov 2009.

[33] Andrew R Jamieson, Maryellen L Giger, Karen Drukker, Hui Li, Yading Yuan, and Neha Bhooshan. Exploring nonlinear feature space dimension reduction and data representation in breast cadx with laplacian eigenmaps and t-sne. *Med Phys*, 37(1):339–351, Jan 2010.

[34] George Lee, Carlos Rodriguez, and Anant Madabhushi. Investigating the efficacy on nonlinear dimensionality reduction schemes in classifying gene and protein expression studies. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 5:368–384, 2008.

[35] Weijie Chen, Maryellen L. Giger, Ulrich Bick, and Gillian M. Newstead. Automatic identification and classification of characteristic kinetic curves of breast lesion on dce-mri. *Medical Physics*, 33(8):2878–2887, 2006.

[36] Botond K. Szabo, Peter Aspelin, and Maria Kristoffersen Wiberg. Neural network approach to the segmentation and classification of dynamic magnetic resonance images of the breast: Comparison with empiric and quantitative kinetic parameters. *Academic Radiology*, 11:1344–1354, 2004.

[37] Mitchell D. Schnall, Jeffrey Blume, David A. Bleumke, Gia A. DeAngelis, Nanette DeBruhl, Steven Harms, Sylvia H. Heywang-Kobrunner, Nola Hylton, Christiane Kuhl, Ella D. Pisano, Petrina Causer, Stuart J. Schnitt, David Thickman, Carol B. Stelling, Paul T. Weatherall, Constance Lehman, and Constantine A. Gastonis. Diagnositc architectural and dynamic features at breast mr imaging: Multicenter study. *Radiology*, 238(1):42–53, 2006.

[38] Karen Kinkel, Thomas H. Helbich, Laura J. Esserman, John Barclay, Ellen Schwerin, Edward A. Sickles, and Nola M. Hylton. Dynamic high-spatial-resolution mr imaging of suspicious breast lesion: Diagnostic criteria and interobserver variability. *Am. J. Roentgenol.*, 175(1):35–43, 2000.

[39] Thorsten Twellmann, Anke Meyer-Baese, Oliver Lange, Simon Foo, and Tim W. Nattkemper. Model-free visualization of suspicious lesions in breast mri based on supervised and unsupervised learning. *Eng. Appl. Artif. Intell.*, 21:129–140, 2008.

[40] Qiu Wu, Marcos Salganicoff, Arun Krishnan, Donald S. Fussell, and Mia K. Markey. Interactive lesion segmentation on dynamic contrast enhanced breast mri using a markov model. volume 6144, page 61444M. SPIE, 2006.

[41] Jiazheng Shi, Berkman Sahiner, Heang-Ping Chan, Chintana Paramagul, Lubomir M. Hadjiiski, Mark Helvie, and Thomas Chenevert. Treatment response assessment of breast masses on dynamic contrast-enhanced magnetic resonance scans using fuzzy c-means clustering and level set segmentation. *Medical Physics*, 36(11):5052–5063, 2009.

[42] Guillermo Sapiro. Color snakes. *CVIU*, 68(2):247–253, 1997.

[43] Tony F. Chan, B. Yezrielev Sandberg, and Luminita A. Vese. Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11(2):130 – 141, 2000.

[44] M. Rousson and R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. In *Motion and Video Computing, 2002. Proceedings. Workshop on*, pages 56 – 61, 2002.

[45] Jun Xu, Andrew Janowczyk, Sharat Chandran, and Anant Madabhushi. A weighted mean shift, normalized cuts initialized color gradient based geodesic active contour model: applications to histopathology image segmentation. volume 7623, page 76230Y. SPIE, 2010.

[46] Cumani Aldo. Edge detection in multispectral images. pages 40–51. Academic Press, Inc., 1991.

[47] Silvano Di Zenzo. A note on the gradient of a multi-image. volume 33, pages 116–125. Academic Press Professional, Inc., San Diego, CA, USA, 1986.

[48] Chris Rorden and Matthew Brett. Stereotaxic display of brain lesions. *Behav Neurol*, 12(4):191–200, 2000.

[49] Ke Nie, Jeon-Hor Chen, Hon J. Yu, Yong Chu, Orhan Nalcioglu, and Min-Ying Su. Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast mri. *Academic Radiology*, 15(12):1513 – 1525, 2008.

[50] N. Bhooshan, M. L. Giger, S. A. Jansen, H. Li, L. Lan, and G. M. Newstead. Cancerous breast lesions on dynamic contrast-enhanced mr images: computerized characterization for image-based prognostic markers. *Radiology*, 254(3):680–90, 2010.

[51] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[52] S.H. Heywang-Kobrunner, P. Viehweg, A. Hienig, and Ch. Kuchler. Contrast-enhanced mri of the breast: accuracy, value, controversies, solutions. *European Journal of Radiology*, 24:94–108, 1997.

[53] Edward A. Sickles, Roy A. Filly, and Peter A. Callen. Benign breast lesions: Ultrasound detection and diagnosis. *Radiology*, 151:467–470, 1984.

[54] Christiane Katharina Kuhl, Peter Mielcareck, Sven Klaschik, Claudia Leutner, Eva Wardelmann, Jurgen Gieske, and Hans H. Schild. Dynamic breast mr imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions. *Radiology*, 211:101–110, 1999.

[55] Ka-Loh Li, Roland Henry, Lisa J. Wilmes, Jessica Gibbs, Xiaoping Zhu, Ying Lu, and Nola M. Hylton. Kinetic assessment of breast tumors using high spatial resolution signal enhancement ratio (ser) imaging. *Magnetic Resonance in Medicine*, 58(3):572–581, 2007.

[56] CK Kuhl. Mri of breast tumors. *European Journal of Radiology*, 10:46–58, 2000.

[57] Mark J. Stoutjesdijk, Jeroen Veltman, Henkjan Huisman, Nico Karssemeijer, Jelle O. Barentsz, Johan G. Blickman, and Carla Boetes. Automated analysis of contrast enhancement in breast mri lesions using mean shift clustering for roi selection. *Journal of Magnetic Resonance Imaging*, 26(3):606–614, 2007.

[58] Yuanjie Zheng, Sarah Englander, Sajjad Baloch, Evangelia I. Zacharaki, Yong Fan, Mitchell D Schnall, and Dinggang Shen. Step: Spatiotemporal enhancement pattern for mr-based breast tumor diagnosis. *Medical Physics*, 37(7):3192–3204, 2009.

[59] Brent J. Woods, Bradley D. Clymer, Tahsin Kurc, Johannes T. Heverhagen, Robert Stevens, Adem Orsdemir, Orhan Bulan, and Michael V. Knopp. Malignant-lesion segmentation using 4d co-occurrence texture analysis applied to dynamic contast-enhanced magnetic resonance breast image data. *Journal of Magnetic Resonance Imaging*, 25:495–501, 2007.

[60] Christine E. McLaren, Wen-Pin Chen, Ke Nie, and Min-Ying Su. Prediction of malignant breast lesions from mri features: A comparison of artificial neural network and logistic regression techniques. *Academic Radiology*, 16:842–851, 2009.

[61] Lina Arbash Meinel, Alan H. Stolpen, Kevin S. Berbaum, Laurie L. Fajardo, and Joseph M. Reinhardt. Breast mri lesion classification: Improved performance of human readers with a backpropagation neural network computer-aided (cad) system. *Journal of Magnetic Resonance Imaging*, 25(1):89–95, 2007.

[62] Jacob Levman, Tony Leung, Petrina Causer, Don Plewes, and Anne L. Martel. Classification of dynamic contast-enhanced magnetic resonance breast lesions by support vector mechines. *IEEE Transactions on Medical Imaging*, 27:688–696, 2008.

[63] Alan Penn, Scott Thompson, Rachel Brem, Constance Lehman, Paul Weatherall, Mitchell Schnall, Gillian Newstead, Emily Conant, Susan Ascher, Elizabeth Morris, and Etta Pisano. Morphologic blooming in breast mri as a characterization of margin for discriminating benign from malignant lesions. *Academic Radiology*, 13:1344–1354, 2006.

[64] Peter Gibbs and Lindsay W. Turnbull. Texture analysis of contast-enhanced mr images of the breast. *Magnetic Resonance in Medicine*, 50:92–98, 2003.

[65] H. Degani, V. Gusis, D. Weinstein, S. Fields, and S. Strano. Mapping pathophysiological features of breast tumors by mri at high spatial resolution. *Nature Medicine*, 3(7):780–782, 1997.

[66] D. Weinstein, S. Strano, P. Cohen, S. Fields, J. M. Gomori, and H. Degani. Breast fibroadenoma: mapping of pathophysiologic features with three-time-point, contrast-enhanced mr imaging–pilot study. *Radiology*, 210(1):233–240, 1999.

[67] E. A M Hauth, H. Jaeger, S. Maderwald, A. Muhler, R. Kimmig, and M. Forsting. Quantitative 2- and 3-dimensional analysis of pharmacokinetic model-derived variables for breast lesions in dynamic, contrast-enhanced mr mammography. *European Journal of Radiology*, 66(6):300–308, 2008.

[68] Anne L Martel. A fast method of generating pharmacokinetic maps from dynamic contrast-enhanced images of the breast. *Int Conf Med Image Comput Comput Assist Interv*, 9(Pt 2):101–108, 2006.

[69] Anna Vignati, Valentina Giannini, Alberto Bert, Massimo Deluca, Lia Morra, Diego Persano, Laura Martincich, and Daniele Regge. A fully automatic lesion detection method for dce-mri fat-suppressed breast images. In *SPIE Medical Imaging*, 2009.

[70] Ernesto Coto, Soren Grimm, Stefan Bruckner, Meister Eduard Groller, Armin Kanitsar, and Omaira Rodriguez. Mammoexplorer: An advanced cad application for breast dce-mri. In *Proceedings of Vision, Modeling and Visualization*, 2005.

[71] J. Veltman, R. Mann, T. Kok, I. M. Obdeijn, N. Hoogerbrugge, J. G. Blickman, and C. Boetes. Breast tumor characteristics of brca1 and brca2 gene mutation carriers on mri. *Eur Radiol*, 18(5):931–8, 2008.

[72] Anant Madabhushi, Jayaram K. Udupa, and Andre Souza. Generalized scale: Theory, algorithms, and application to image inhomogeneity. *Computer Vision and Image Understanding*, 101(2):100–121, 2006.

[73] Anant Madabhushi, Michael Feldman, Dimitris N. Metaxas, John Tomaszewski, and Deborah Chute. Automated detection of prostate adenocarcinoma from high-resolution ex vivo mri. *IEEE Transactions on Medical Imaging*, 24(12):1611–1625, 2005.

[74] Peter Kruizinga and Nikolay Petkov. Nonlinear operator for oriented texture. *IEEE Transactions on Image Processing*, 8(10):1395–1407, 1999.

[75] Anant Madabhushi, Jianbo Shi, Mark Rosen, John E Tomaszeweski, and Michael D Feldman. Graph embedding to improve supervised classification and novel class detection: application to prostate cancer. *Int Conf Med Image Comput Comput Assist Interv*, 8(Pt 1):729–737, 2005.

[76] Robert M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, 1973.

[77] Pallavi Tiwari, Mark Rosen, and Anant Madabhushi. Dimensionality reduction scheme for detection of prostate cancer from magnetic resonance spectroscopy (mrs). *Medical Physics*, 2009, *In press*.

[78] Dennis Gabor. Theory of communication. *Journal of the Institute of Electrical Engineers*, 93:429–457, 1946.

[79] Simona Grigorescu, Nicolai Petkov, and Peter Kruizinga. Comparison of texture features based on gabor filters. *IEEE Transactions on Image Processing*, 11:1160–1167, 2002.

[80] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. In *Proc. IEEE International Conference on Systems, Man and Cybernetics*, pages 14–19, 4–7 Nov. 1990.

[81] W. Nick Street, William H. Wolberg, and O. L. Mangasarian. Nuclear feature extraction for breast tumor diagnosis. *IS&T/SPIE 1999 International Symposium on Electrical Imaging: Science and Technology*, 1905:861–870, 1993.

[82] Shivang Naik, Scott Doyle, Michael Feldman, John Tomaszewski, and Anant Madabhushi. Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information. *The second international workshop on Microscopic Image Analysis with Applications in Biology*, 2007.

[83] Olivier Chapelle, Patrick Haffner, and Vladimir Vapnik. Support vector machines for histogram-based image segmentation. *IEEE Transactions on Neural Networks*, 10:1055–1064, 1999.

[84] Yoav Freund and Robert E. Schapire. A decision- theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139, 1997.

[85] Zhuowen Tu. Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV '05)*, 2:1589–1596, 2005.

[86] Weijie Chen, Maryellen Giger, Li Lan, and Ulrich Bick. Computerized interpretation of breast mri: Investigation of enhancement-variance dynamics. *Medical Physics*, 31:1076–1082, 2004.

[87] Devrim Unay, Ahmet Ekin, Mujdat Cetin, Radu Jasinschi, and Aytul Ercil. Robustness of local binary patterns in brain mr image analysis. *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, pages 2098–2101, 2007.

[88] Gert Stockmans, Karen Deraedt, Hans Wildiers, Philippe Moerman, and Robert Paridaens. Triple-negative breast cancer. *Curr Opin Oncol*, 20(6):614–620, 2008.

[89] B. G. Haffty, Q. Yang, M. Reiss, T. Kearney, S. A. Higgins, J. Weidhaas, L. Harris, W. Hait, and D. Toppmeyer. Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J Clin Oncol*, 24(36):5652–7, 2006.

[90] J. Tchou, L.-P. Wang, M. Sargen, S. Sonnad, J. Tomaszewski, and M. Schnall. Do triple-negative breast cancers have a distinct imaging phenotype? In *29th Annual San Antonio Breast Cancer Symposium*, page Abstract no. 551108, San Antonio, TX, 2006.

[91] Constance D. Lehman, Constantine Gastonis, Christiane K. Kuhl, R. Edward Hendrick, Etta D. Pisano, Lucy Hanna, Sue Peacock, Stanley F. Smazal, Daniel D. Maki, Thomas B. Julian, Elizabeth R. DePeri, David A. Bleumke, and Mitchell D. Schnall. Mri evaluation of the contralateral breast in women with recently diagnosed breast cancer. *New England Journal of Medicine*, 356(13):1295–1303, 2007.

[92] P. Shah, M. Rosen, J. Stopfer, J. Siegfried, R. Kaltman, B. Mason, K. Armstrong, K. L. Nathanson, M. Schnall, and S. M. Domchek. Prospective study of breast mri in brca1 and brca2 mutation carriers: effect of mutation status on cancer incidence. *Breast Cancer Res Treat*, 118(3):539–46, 2009.

[93] T. Uematsu, M. Kasami, and S. Yuen. Triple-negative breast cancer: correlation between mr imaging and pathologic findings. *Radiology*, 250(3):638–47, 2009.

[94] B. E. Dogan, A. M. Gonzalez-Angulo, M. Gilcrease, M. J. Dryden, and W. T. Yang. Multimodality imaging of triple receptor-negative tumors with mammography, ultrasound, and mri. *AJR Am J Roentgenol*, 194(4):1160–6, 2010.

[95] J-H. Chen, G. Agrawal, B. Feig, H-M. Baek, P. M. Carpenter, R. S. Mehta, O. Nalcioglu, and M-Y. Su. Triple-negative breast cancer: Mri features in 29 patients. *Ann Oncol*, 18(12):2042–2043, 2007.

[96] E. E. Deurloo, S. H. Muller, J. L. Peterse, A. P. Besnard, and K. G. Gilhuijs. Clinically and mammographically occult breast lesions on mr images: potential effect of computerized assessment on clinical reading. *Radiology*, 234(3):693–701, 2005.

[97] C. Meeuwis, S. M. van de Ven, G. Stapper, A. M. Fernandez Gallardo, M. A. van den Bosch, W. P. Mali, and W. B. Veldhuis. Computer-aided detection (cad) for breast mri: evaluation of efficacy at 3.0 t. *Eur Radiol*, 20(3):522–8, 2010.

[98] T. Arazi-Kleinman, P. A. Causer, R. A. Jong, K. Hill, and E. Warner. Can breast mri computer-aided detection (cad) improve radiologist accuracy for lesions detected at mri screening and recommended for biopsy in a high-risk population? *Clin Radiol*, 64(12):1166–74, 2009.

[99] T. C. Williams, W. B. DeMartini, S. C. Partridge, S. Peacock, and C. D. Lehman. Breast mr imaging: computer-aided evaluation program for discriminating benign from malignant lesions. *Radiology*, 244(1):94–103, 2007.

[100] C. D. Lehman, S. Peacock, W. B. DeMartini, and X. Chen. A new automated software system to evaluate breast mr examinations: improved specificity without decreased sensitivity. *AJR Am J Roentgenol*, 187(1):51–6, 2006.

[101] S. Basu, W. Chen, J. Tchou, A. Mavi, T. Cermik, B. Czerniecki, M. Schnall, and A. Alavi. Comparison of triple-negative and estrogen receptor-positive/progesterone receptor-positive/her2-negative breast carcinoma using quantitative fluorine-18 fluorodeoxyglucose/positron emission tomography imaging parameters: a potentially useful method for disease characterization. *Cancer*, 112(5):995–1000, 2008.

[102] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. John Wiley and Sons, New York, 2nd edition, 2001.

[103] R. Hupse and N. Karssemeijer. The effect of feature selection methods on computer-aided detection of masses in mammograms. *Phys Med Biol*, 55(10):2893–904, 2010.

[104] S. M. Bentzen. Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol*, 6(2):112–7, 2005.

[105] Julia Tchou, Seema Sonnad, Meredith Bergey, Sandip Basu, John Tomaszewski, Abass Alavi, and Mitchell Schnall. Degree of tumor fdg uptake correlates with proliferation index in triple negative breast cancer. *Molecular Imaging and Biology*, 12(6):657–662, 2010.

[106] Wen-Yen Chen, Yangqiu Song, Hongjie Bai, Chih-Jen Lin, and Edward Y. Chang. Parallel spectral clustering in distributed systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:568–586, 2011.

[107] Charless Fowlkes, Serge Belongie, Fan Chung, and Jitendra Malik. Spectral grouping using the nystrom method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:214–225, 2004.

[108] Rin Yamaguchi, Hidemi Furusawa, Hiroshi Nakahara, Masuko Inomata, Kiyoshi Namba, Maki Tanaka, Kazuaki Ohkuma, Kohsuke Tayama, Teruhiko Fujii, Hirohisa Yano, Masayoshi Kage, and Masamichi Kojiro. Clinicopathological study of invasive ductal carcinoma with large central acellular zone: Special reference to magnetic resonance imaging findings. *Pathology International*, 58(1):26–30, 2008.

[109] Pascal Baltzer, Tibor Vag, Matthias Dietzel, Sebastian Beger, Christian Freiberg, Mieczyslaw Gajda, Oumar Camara, and Werner Kaiser. Computer-aided interpretation of dynamic magnetic resonance imaging reflects histopathology of invasive breast cancer. *European Radiology*, 20:1563–1571, 2010. 10.1007/s00330-010-1722-x.

[110] Paul S. Tofts. Modeling tracer kinetics in dynamic gd-dtpa mr imaging. *Journal of Magnetic Resonance Imaging*, 7(1):91–101, 1997.

[111] HJ Weinmann, M Laniado, and W Mützel. Pharmacokinetics of gddtpa/dimeglumine after intravenous injection into healthy volunteers. *Physiol Chem Phys Med NMR.*, 16(2):167–72, 1984.

[112] Paul S. Tofts, Gunnar Brix, David L. Buckley, Jeffrey L. Evelhoch, Elizabeth Henderson, Michael V. Knopp, Henrik B.W. Larsson, Ting-Yim Lee, Nina A. Mayr, Geoffrey J.M. Parker, Ruediger E. Port, June Taylor, and Robert M. Weisskoff. Estimating kinetic parameters from dynamic contrast-enhanced t1-weighted mri of a diffusable tracer: Standardized quantities and symbols. *Journal of Magnetic Resonance Imaging*, 10(3):223–232, 1999.

[113] Ruediger E. Port, Michael V. Knopp, and Gunnar Brix. Dynamic contrast-enhanced mri using gd-dtpa: Interindividual variability of the arterial input function and consequences for the assessment of kinetics in tumors. *Magnetic Resonance in Medicine*, 45(6):1030–1038, 2001.

[114] P Di Giovanni, C A Azlan, T S Ahearn, S I Semple, F J Gilbert, and T W Redpath. The accuracy of pharmacokinetic parameter measurement in dce-mri of the breast at 3 t. *Physics in Medicine and Biology*, 55(1):121, 2010.

[115] I.T. Jolliffe. *Principal component analysis*. Springer series in statistics. Springer-Verlag New York, 2nd ed. edition, 2002.