Integrated Graph Theoretic, Radiomics, and Deep Learning Framework for Personalized Clinical Diagnosis, Prognosis, and Treatment Response Assessment of Body Tumors

by

Vishwa S. Parekh

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

October, 2018

Abstract

Purpose: A new paradigm is beginning to emerge in radiology with the advent of increased computational capabilities and algorithms. The future of radiological reading rooms is heading towards a unique collaboration between computer scientists and radiologists. The goal of computational radiology is to probe the underlying tissue using advanced algorithms and imaging parameters and produce a personalized diagnosis that can be correlated to pathology. This thesis presents a complete computational radiology framework (I-GRAD) for personalized clinical diagnosis, prognosis and treatment planning using an integration of graph theory, radiomics, and deep learning.

Methods: There are three major components of the I-GRAD framework–image segmentation, feature extraction, and clinical decision support.

Image Segmentation: I developed the multiparametric deep learning (MPDL) tissue signature model for segmentation of normal and abnormal tissue from multiparametric (mp) radiological images. The segmentation MPDL network was constructed from stacked sparse autoencoders (SSAE) with five hidden layers. The MPDL network parameters were optimized using k-fold cross-validation. In addition, the MPDL segmentation network was tested on an independent dataset.

Feature Extraction: I developed the radiomic feature mapping (RFM) and contribution scattergram (CSg) methods for characterization of spatial and inter-parametric relationships in multiparametric imaging datasets. The radiomic feature maps were created by filtering radiological images with first and second order statistical texture filters followed by the development of standardized features for radiological correlation to biology and clinical decision support. The contribution scattergram was constructed to

visualize and understand the inter-parametric relationships of the breast MRI as a complex network. This multiparametric imaging complex network was modeled using manifold learning and evaluated using graph theoretic analysis.

Feature Integration: The different clinical and radiological features extracted from multiparametric radiological images and clinical records were integrated using a hybrid multiview manifold learning technique termed the Informatics Radiomics Integration System (IRIS). IRIS uses hierarchical clustering in combination with manifold learning to visualize the high-dimensional patient space on a two-dimensional heatmap. The heatmap highlights the similarity and dissimilarity between different patients and variables.

Results: All the algorithms and techniques presented in this dissertation were developed and validated using breast cancer as a model for diagnosis and prognosis using multiparametric breast magnetic resonance imaging (MRI). The deep learning MPDL method demonstrated excellent dice similarity of 0.87±0.05 and 0.84±0.07 for segmentation of lesions on malignant and benign breast patients, respectively. Furthermore, each of the methods, MPDL, RFM, and CSg demonstrated excellent results for breast cancer diagnosis with area under the receiver (AUC) operating characteristic (ROC) curve of 0.85, 0.91, and 0.87, respectively. Furthermore, IRIS classified patients with low risk of breast cancer recurrence from patients with medium and high risk with an AUC of 0.93 compared to OncotypeDX, a 21 gene assay for breast cancer recurrence.

Conclusion: By integrating advanced computer science methods into the radiological setting, the I-GRAD framework presented in this thesis can be used to model radiological imaging data in combination with clinical and histopathological data and produce new tools for personalized diagnosis, prognosis or treatment planning by physicians.

Primary Reader: Dr. Michael A. Jacobs

Secondary Readers: Dr. Russell H. Taylor and Dr. Jerry L. Prince

Acknowledgements

The credit for guiding me through my dissertation goes to Dr. Michael Jacobs. Thank you, Dr. Jacobs, for being an amazing mentor. The best thing about Dr. Jacobs was his approachability. He would always have time to generally talk, brain storm ideas, or resolve any problems I am facing. He gave me freedom to independently explore problems and come up with my own solutions. Thank you very much, Dr. Jacobs, for providing me with a rare opportunity to work on problems I felt strongly about.

Dr. Jacobs took a keen interest not only in my research abilities but also in my overall development. He taught me everything from the research process to networking and establishing collaborations. More importantly, Dr. Jacobs always treated me like family. He made sure I had no problems both within and outside my research life. He always told me that there is nothing more important than family. This is one teaching of his that I will always remember.

I would like to thank Dr. Russell Taylor and Dr. Jerry Prince for co-advising me during my Ph. D. and providing me with valuable feedback through the years and on my thesis.

I am thankful to Drs. Vladimir Braverman and Paul Nagy for taking the time to serve on my GBO committee. I would also like to thank Drs. Nathan Crone and Anna Korzeniewska from the department of Neurology for my initial training in research.

I would like to thank Prof. J P Misra, who has been my mentor right from the day I started my academic life. Thank you, Prof. J P Misra, for guiding me through my academic life.

There have been a lot of people who have made this journey both enjoyable and a memorable one. Pratik Kothari, the first person I met when I landed in Baltimore. He was my only family here in Baltimore and over the past seven years, Pratik has been a constant source of support and motivation. Jakir Hossain was one of the first people I met here at Hopkins. He introduced me to Dr. Jacobs and also helped me understand MR concepts. He has been a great friend to talk about research or life in general. I am very lucky to have an amazing group of friends here in Baltimore. I would like to thank them all for always being around. They never made me feel I was away from home and in a different country.

I would like to thank my wife, Kesha for walking the final leg of my Ph.D. with me. You made it much easier. Thank you for being patient with me, especially during the tough times.

Finally, I would like to thank my parents, Sanjay and Rita Parekh. Thank you for your love and support over all these years. Thank you for standing with me and guiding me during the most difficult of times. I could not have done this without you. This dissertation is, most definitely, part of your work.

Dedication

Dedicated to my parents, Sanjay and Rita Parekh.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Symbol = Definition

Radiological Imaging

ADC = Apparent Diffusion Coefficient

CT = Computed Tomography

DWI = Diffusion weighted Imaging

DCE = Dynamic Contrast Enhancement

MRI = Magnetic Resonance Imaging

PET = Positron Emission Tomography

SE = Spin Echo

ST = Slice Thickness

T = Tesla

T1 = Spin-lattice relaxation time

T2 = Spin-spin relaxation time

TI = Inversion Time

TE = Echo Time

TR = Repetition Time

Computer Science methods

CNN = convolutional neural network

CSg = Contribution Scattergram

DfM = Diffusion Mapping

EI = Eigenimage Filter

Isomap = Isometric Feature Mapping

LLE = Locally Linear Embedding

MLP = multilayer perceptron

MPDL = Multiparametric Deep Learning

NLDR = Nonlinear dimensionality reduction

SSAE = Stacked Sparse Autoencoder

Radiomics

FOS = first order statistics

GLCM = Gray level co-occurrence matrix

GLRLM = Gray level run length matrix

NGTDM = neighborhood gray tone difference matrix

RFM = Radiomic Feature Mapping

MPRAD = Multiparametric Radiomics

Computer aided clinical decision support systems

I-GRAD = Integrated graph theoretic, radiomics, deep learning

IRIS = Informatics Radiomics Integration System

LIST OF ALGORITHMS

Chapter 1. Introduction

Given the rapid changes in computational algorithms and the data rich environment in clinical medicine, the key question is "How to extract, evaluate and visualize the high-dimensional clinical data in a meaningful way with quantitative metrics for tumor characterization?" in particular, in radiology, which is rich in digital imaging methods. Moreover, the radiological reading room is undergoing a paradigm shift to a symbiosis of computer science and radiology. This thesis presents an outline for computer science to enrich radiology and other parts of the clinical environment using advanced computational algorithms.

1.1 Computational Radiology

Radiological imaging methods are used to non-invasively probe tissue of interest in different regions of the body to detect and characterize potential deadly pathology, such as malignant cancer. These radiological methods produce large volumes of complex digital imaging data corresponding to an area or region of interest making "reading and interpreting" the image data challenging. However, recent advances in the field of computer vision and machine learning have led to a new paradigm shift to computational radiology in the field of medicine, where, the future radiological reading rooms will have a unique collaboration between computer scientists and radiologists (human experts) [22]. In the near future, advanced computational methods will be assisting radiologists in various aspects of radiological decision making, such as identification and segmentation of different tissue types, tumor characterization, and diagnosis. For example, large data sets, such as breast and brain can be quickly triaged into different groups by using advance computer science methods, where the potential "worst" cases are looked at first. This could lead to increased

efficacy and confidence of the radiologist making the right diagnosis. These methods can be further developed to integrate advanced radiological parameters with multivariate information from different sources such as the electronic heath record, pathology, and clinical history to assist treating clinicians in arriving at the most accurate diagnosis. This integration of multivariate information would give the clinician an improved understanding of the complex nature of a disease and develop improved personalized predictive biomarkers or group similar patients for personalized treatment planning and disease prognosis.

Radiological imaging techniques are powerful noninvasive tools used for the detection, differentiation, and diagnosis of different tissue characteristics in patients. These imaging methods include X-Ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Nuclear Medicine (NM) Positron Emission Tomography (PET), and Ultrasound (US). Each of these modalities creates a different tissue contrast based on the mechanism of acquisition and the inherent properties of the tissue, enabling the method to determine whether the tissue is normal or abnormal. For example, multiparametric MRI (mpMRI) of breast tissue using advanced quantitative parameters of dynamic contrast enhanced (DCE) MRI and diffusion weighted imaging (DWI) provide non-invasive information about the underlying tissue biology. The DCE-MRI provides information about vascularity and permeability of the tissue, which is characterized by rapid uptake of contrast agent, followed by fast washout. On the other hand, the movement of intra- and inter-cellular water in normal and tumor environments of breast tissue are characterized by DWI and the apparent diffusion coefficient (ADC) of water map. The ADC map provides a quantitative biophysical parameter that measures the cellularity of tissue.

My goal is to extract all the information present in the radiological images and develop radiological biomarkers for characterization of a patient's diagnosis, prognosis, or treatment response. A biomarker is defined as "A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" [23]. Correspondingly, radiological biomarkers are derived from images to form a feature or a set of features modeled to characterize a biological property of the underlying tissue of interest. The complete set of imaging features that extracted from multiparametric radiological images can be broadly categorized into two categories: spatial features (texture and shape) and inter-parametric relationship features. The spatial appearance of the tissue of interest is evaluated using radiomics while the inter-parametric relationships are evaluated using a contribution scattergram. Both these methods are discussed in detail in Part II of the thesis. Briefly, radiomics characterizes and visualizes the information about the gray-scale patterns, inter-pixel relationships, and shape-based properties of the region of interest (ROI) which may be correlated to tissue biology [24-30]. I am the first to relate radiomic characteristics of both normal and lesion tissue to the underlying tissue biology.

A contribution scattergram models the inter-parametric relationships between different imaging parameters as a complex network model. Different radiological imaging parameters are represented as nodes on the contribution scattergram and their relationships, as contribution scattergram edges. I am the first to model the multiparametric imaging data as a complex network and evaluate its underlying organizational structure which could potentially provide a unique network signature for every patient or a group of patients with similar pathology.

Integrating and visualizing radiological biomarkers obtained from each of these different imaging methods would be beneficial in defining imaging phenotypes for personalized clinical decision support. The integration of seemingly disparate data and modeling the complex interactions between different features within the data can be accomplished using advanced multiview learning algorithms [31]. Furthermore, integrating radiological biomarkers with histopathology, genomics, and clinical history will create a complete picture of a patient's diagnosis and prognosis.

1.2 Computational Radiology Framework

This thesis presents a novel integration of computer science within radiology using advanced graph theoretic, radiomic, and deep learning (I-GRAD) techniques to develop a personalized radiological diagnosis, prognosis, and treatment assessment framework. **Figure 1.1** illustrates an outline of the I-GRAD framework. As shown in **Figure 1.1**, there are three major components of the I-GRAD framework. The first step involves segmenting the tissue of interest in the acquired multiparametric radiological imaging data. Accurate identification and segmentation of the tissue of interest presents the first challenge as all the subsequent steps depend on the accuracy of the segmentation algorithm.

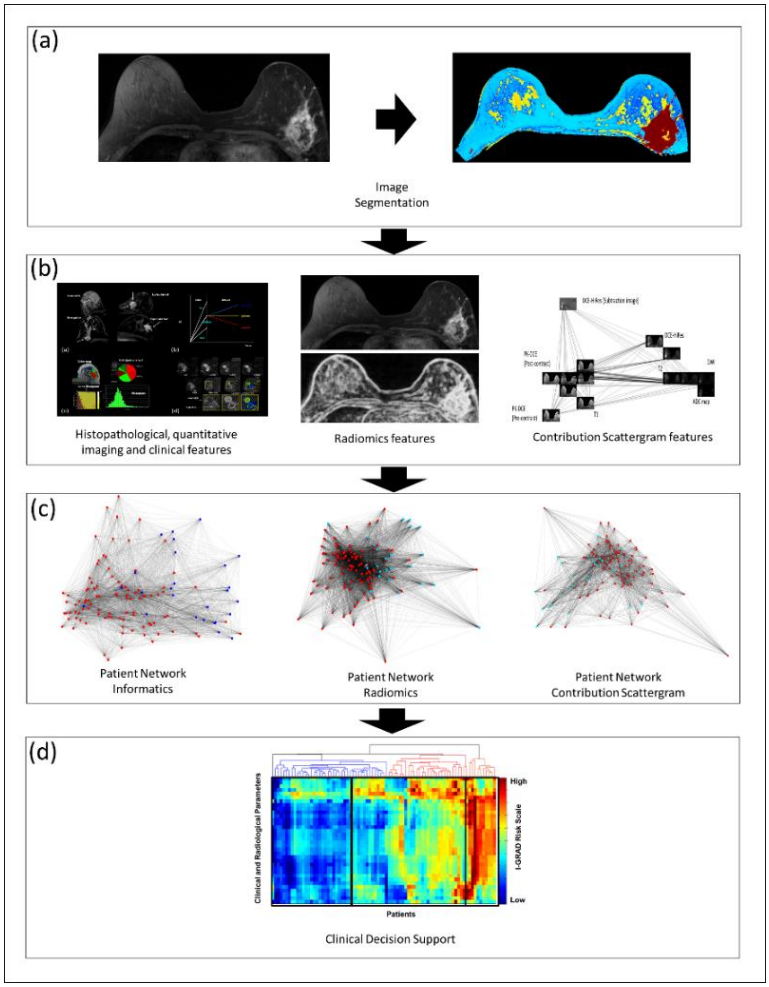


Figure 1.1 Illustration of the I-GRAD Framework for personalized radiological diagnosis and prognosis. There are four major components of the I-GRAD framework. (a) The first component is image segmentation which is developed using multiparametric deep learning (MPDL). (b) The second step involves extraction of intrinsic information present in radiological images using radiomics (shape and texture), manifold learning, and graph theory (complex interactions between imaging parameters). (c) In the third step, a patient space corresponding to each feature is constructed for visualization and analysis of similar and dissimilar patients. (d) The final step involves transforming each patient space into a heatmap visualization tool for patient diagnosis and analysis of the contribution from each feature space.

The second step involves extraction of all the relevant information from the region of interest (ROI) segmented in the previous step. The complete set of features that can be extracted from radiological imaging using advanced computational methods can be categorized into two categories:

1. Morphological properties (texture and shape) of the tissue of interest.
2. Complex interactions between the different imaging parameters acquired in the study.

Integration of these two features into a single model comprises the final step of this framework. These features are extracted from different sources and hence require a multiview learning algorithm that creates a personalized radiological signature correlated to pathology for the application defined by the user, for example, distinguishing malignant from benign tumors.

The objectives of this dissertation are as follows:

1. Develop algorithms for segmentation and classification of different tissue types using radiological imaging methods.
2. Extract biological and structural information from the segmented tissue from the radiological images.
3. Integrate clinical data with all the extracted information into a clinical decision support model using multiview learning for transforming the radiology reading room.

1.3 List of Contributions

1.3.1 Image Segmentation

My first and second contributions address the problem of segmentation of different tissues from multiparametric radiological imaging datasets. For my first contribution, I developed the multiparametric deep learning (MPDL) tissue signature model for segmentation of different tissue types from multiparametric radiological imaging datasets. I further evaluated the potential of MPDL as a generalizable organ invariant tissue segmentation architecture for my second contribution.

Multiparametric deep learning tissue signature model

Supervised segmentation algorithms require pre-defined labels for all the tissue types found in the organ of interest which poses significant challenges. Pre-labeling the data with all possible tissue types is not practical and very time consuming, which could lead to under-sampling of the correct objects or tissue types in a dataset. However, if we use all the data (labeled and unlabeled), we can develop a classifier that has "seen" all the tissue types.

The MPDL model is based on stacked sparse autoencoders (SSAE). Sparse autoencoders are unsupervised neural networks that learn an intrinsic representation of the input by attempting to reconstruct it. Furthermore, the activation of each node in the hidden layer of the autoencoder is specialized to activate in response to only a subset of input data by the introduction of sparsity constraint. The unsupervised nature of SAE allows us to train them on the complete set of labeled and unlabeled data. In MPDL, a set of five SAEs were stacked together with each subsequent SAE trained on the output of the previous SAE and a final softmax classifier to classify the data into different tissue types. The MPDL architecture was then fine-tuned on the labeled data to generate segmentation masks. I

trained and tested the MPDL for segmentation and classification of multiparametric breast Magnetic Resonance Imaging (mpMRI). Furthermore, the MPDL tissue signature model was validated on an independent dataset [32-34] that had different imaging parameters (DCE, T1, T2, DWI for training, and only DCE for testing) and imaging plane orientations (Axial plane for training and Sagittal plane for testing) than the training dataset, thereby demonstrating the robustness of MPDL and eliminating the need to retrain the model every time a new dataset is introduced to the model. A journal paper manuscript describing MPDL has been tentatively accepted, pending revision to Scientific Reports. The manuscript can be found on arXiv preprint server [35]. This work is described in **Chapter 2.**

Organ Invariant Tissue Segmentation

The MPDL model resolves the problem of segmentation of different tissues from a multiparametric radiological imaging dataset, however, it is very narrow in that it is limited to segmentation of tissue from a specific organ. On the other hand, if we look at deep learning models developed for computer vision applications, they are more general. For example, object recognition algorithms based on deep learning are not limited to recognition of sub-groups of objects (e.g. vehicles, food, and furniture), but are generalized to detect all kinds of objects, irrespective of their type, scale, or origin [36].

The goal of this contribution was two-fold. First, I wanted to understand how these deep networks encode the intrinsic representations within their network architecture. As an example, when radiologists develop an intrinsic tissue signature representation of "fat" in their brain, they store the tissue signature as "bright on T1 and dark on T2". Similarly, an intrinsic representation of fluid would be "dark on T1 and bright on T2". The question here is "Does

the MPDL also form intrinsic representations similar to radiologists?" If not, can we train it to learn similar tissue signature representations?

Chapter 3 discusses different techniques for visualization and evaluation of intrinsic representations formed by deep neural networks. My second goal was to utilize these intrinsic representations and their consistency across different organs to evaluate the feasibility of a generalized segmentation architecture. I evaluated the efficacy of MPDL model by testing a *breast*-trained network for segmentation of *brain* mpMRI and vice versa. The results from this preliminary work were presented at the International Society for Magnetic Resonance in Medicine conference in Paris, 2018 and detailed in **Chapter 4.**

My third and fourth contributions of this thesis address the problem of feature extraction from segmented radiological images and quantification of these features.

1.3.2 Feature Extraction

Feature extraction from multiparametric radiological images can be broadly categorized into two categories: inter-voxel analysis and inter-parametric analysis. I developed the methods of radiomic feature mapping and contribution scattergram for inter-voxel and inter-parametric analysis, respectively as detailed in the following sub-sections.

Radiomic Feature Mapping

Radiomics refers to extraction of quantitative textural and shape-based features from radiological images [37]. Traditionally, radiomic features provide information about the gray-scale patterns, inter-pixel relationships, shape, and spectral properties within regions of interest on radiological images [24-30]. Most of the studies in radiomics are focused on extraction of a single quantitative texture value corresponding to all the voxels within the tumor [37]. As a result, these methods of texture analysis have an inherent dependence on

the size of the tissue being evaluated. For example, the range of values that the first order entropy feature can take varies between 0 and $\log_2 N$, where N is the number of voxels in the tissue. Furthermore, the current radiomic features are not standardized nor can they be visualized. Finally, there has been no correlation between textural heterogeneity and tissue biology of the tumor and the surrounding normal tissue. My **third contribution** entails the development and implementation of the radiomic feature mapping (RFM) framework to overcome the limitations of current methods in radiomics analysis. RFM transforms radiological images into complete texture images for visualization and interpretation of tumor heterogeneity. The RFMs provide voxel-wise radiomic features which are not dependent on size of the tissue being evaluated. Furthermore, I developed standardized temporal evolution curves from DCE-MRI RFMs for evaluation of vascular heterogeneity and standardized diffusion evolution curves from DWI RFMs for evaluation of cellular heterogeneity. Finally, I analyzed the diagnostic capabilities of RFM features for prediction of breast tumors as benign or malignant using a novel multiview feature embedding and classification model. The radiomic feature mapping framework is described in **Chapter 6.** A journal paper manuscript describing RFMs has been published in NPJ Breast Cancer, 2017 [38].

The radiomic feature mapping framework, however, produces thirty radiomic feature maps for each radiological image. For a multiparametric dataset with twenty to thirty radiological images, the radiomic feature mapping framework produces 600-900 RFMs rendering the analysis and visualization of RFMs impractical. In addition, texture analysis of a single image is not indicative of the "true texture" of the underlying tissue. Instead, single RFMs only provide textural information from a single point of view. In

multiparametric radiological imaging datasets, tissue signatures (TS) encode the characteristics of a tissue instead of a grayscale value. Therefore, texture analysis of a high-dimensional multiparametric dataset based on tissue signatures would provide multiview textural information, indicative of the true underlying texture of the tissue of interest. To that end, I developed a multiparametric imaging radiomics (MPRAD) framework for extracting radiomics features from multiparametric and multimodal imaging data. I tested the MPRAD framework on multiparametric brain and breast MRI datasets for comparison to single parameter radiomics. The MPRAD framework is described in **Chapter 7** [39].

Contribution Scattergram

Multiparametric data acquisition techniques produce a diversity of different parameters with each parameter quantifying a certain aspect of structural and functional property of the underlying data type. Different parameters acquired interact with each other based on complex high-level relationships. These relationships between parameters could provide important insight into the data being acquired.

The current methods for extraction of multiparametric relationships from MRI are pharmacokinetic analysis from DCE-MRI and apparent diffusion coefficient (ADC) mapping from DWI. These methods are based on mathematical models of the underlying tissue biology. These methods, however, do not attempt to uncover the complete underlying network structure of the different relationships. In addition, these methods are involved with analysis of a single imaging parameter and do not extract the inter-parameter relationships among them. For example, none of the available models are capable of encoding inter-parameter relationships such as “Fat is bright on T1 and dark on T2”.

Finally, currently available models are constructed using voxel-wise analysis and are not capable of evaluating the tissue of interest as a whole.

To overcome these limitations of the current methods, I developed a technique termed the contribution scattergram (CSg) for modeling inter-parametric relationships. The CSg represents high-dimensional data (e.g. multiparametric MRI) as a complex network model. The complex network models the vertices on the CSg to represent different radiological images and the edges to represent the relationship between each of these radiological images. The idea behind contribution scattergram is to visualize and evaluate the organizational structure of the multiparametric MRI image space corresponding to different patients and pathologies, bringing a completely new perspective to radiological diagnosis and prognosis.

The contribution scattergram is constructed by transforming the high-dimensional multiparametric space into a meaningful representation of its intrinsic dimensionality using manifold learning [40-43]. The resulting intrinsic dimensionality displays the complex interactions between the different parameters, which are analyzed using advanced graph theoretic techniques. I extended the CSg using sub-graphs from DCE-MRI and DWI for the analysis of vascularity and cellularity of different tissue types. **Chapter 8** lays down the mathematical foundation and the algorithm for the contribution scattergram. This chapter further demonstrates the potential of contribution scattergram in computer aided decision support for precision medicine with the example application of breast cancer diagnosis using mpMRI. A journal manuscript describing this work is currently ready for submission [44].

1.3.3 Multiview Feature Integration

For my final contribution, I developed the informatics radiomics integration system (IRIS) for integration of all the features extracted from different sources (e.g. radiomics, contribution scattergram, and quantitative imaging parameters) into a clinical decision support system.

The high dimensionality of the dataset formed by clinical and radiological features presents significant challenges for statistical data analysis. First, as these features are acquired from different sources, concatenating them into a single vector for analysis would lead to potential overfitting. Furthermore, associating biological meaning to a concatenated vector of features would be very difficult. Second, for application to clinical setting, the feature integration model needs to be transparent and interpretable, as opposed to a black box model which just produces a final outcome probability. Third, the feature integration model should be capable of visualization and analysis of relationships between different feature spaces and their contribution to the final outcome. Finally, the feature integration model should be capable of operating in unsupervised mode when no ground truth information has been provided.

IRIS is a multiview manifold learning algorithm to visualize and learn relationships between different patients in each of the subspaces formed by parameters obtained from different sources [44]. Furthermore, IRIS also analyzes the association between different parameters for improved diagnosis or for treatment response. IRIS was tested on multiparametric breast MRI to predict recurrence in breast cancer and correlate with OncotypeDX gene array test. **Chapter 9** details the IRIS methodology and the application to breast cancer prognosis. A journal manuscript describing this work is ready for

submission [44]. Finally, **Chapter 10** provides a summary of the results and directions for future work.

1.4 List of Publications and Patents

The major contributions of this dissertation in the form of published (or submitted) manuscripts and disclosed (or pending) patents are listed below.

1.4.1 Journal Publications

1. P. Athamanolap*, **V. Parekh*,** S. I. Fraley*, V. Agarwal*, D. J. Shin, M. A. Jacobs, et al., "Trainable high resolution melt curve machine learning classifier for large-scale reliable genotyping of sequence variants," PloS one, vol. 9, p. e109094, 2014. (* - equal contribution) [45].
2. **V. Parekh** and M. A. Jacobs, "Radiomics: a new application from established techniques," Expert review of precision medicine and drug development, vol. 1, pp. 207-226, 2016. [37].
3. **V. S. Parekh** and M. A. Jacobs, "Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI," NPJ breast cancer, vol. 3, p. 43, 2017 [38].
4. **V. S. Parekh,** K. J. Macura, S. Harvey, I. Kamel, R. EI-Khouli, D. A. Bluemke, et al., "Multiparametric Deep Learning Tissue Signatures for a Radiological Biomarker of Breast Cancer: Preliminary Results," arXiv preprint arXiv:1802.08200, 2018 [35].
5. L. Peng, **V. Parekh,** P. Huang, D. D. Lin, K. Sheikh, B. Baker, et al., "Distinguishing True Progression from Radionecrosis after Stereotactic

Radiotherapy for Brain Metastases with Machine Learning and Radiomics," International Journal of Radiation Oncology• Biology• Physics, 2018 [46].

6. **V. Parekh** and M. Jacobs, "MPRAD: A Multiparametric Imaging Radiomics Framework," Submitted, Nature Communications(NCOMMS-18-26956), 2018 [39].
7. **V. Parekh,** A. Akhbardeh, M. Jacobs, "Contribution Scattergram: A complex network model based on graph theory and manifold learning for creating a unique signature for high dimensional multiparametric data," Ready for Submission, 2018 [44].
8. M. A. Jacobs, **V. Parekh**, C. Umbricht, K. J. Macura, R. EI-Khouli, S. Harvey, et al., "Novel advanced machine learning informatics modeling using clinical and radiological imaging metrics for characterizing breast tumor characteristics with the OncotypeDX gene array," Ready for Submission, 2018 [47].

1.4.2 Conference Publications

1. **V. S. Parekh,** J. R. Jacobs, and M. A. Jacobs, "Unsupervised nonlinear dimensionality reduction machine learning methods applied to multiparametric MRI in cerebral ischemia: preliminary results," in International Society for Optics and Photonics, Medical Imaging, 2014, p. 90342O [48].
2. **V. S. Parekh** and M. A. Jacobs, "A multidimensional data visualization and clustering method: Consensus similarity mapping," in IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016, pp. 420-423 [49].

1.4.3 Abstracts

1. A. Akhbardeh, **V. S. Parekh**, and M. A. Jacobs. "A Novel 3D Registration Method for Multiparametric Radiological Images." Medical Physics 42.6 (2015): 3605-3606.
2. M. A. Jacobs, K. J. Macura, C. Umbricht, **V. Parekh,** R. H. El Khouli, A. C. Wolff, "Novel Informatics Modeling using Clinical and Radiological Imaging Metrics for Characterization of Breast Tumors with the OncotypeDX gene array." Presented at the Radiological Society of North America, 2015
3. **V. Parekh** and M. Jacobs, "Multidimensional Imaging Radiomics-Geodesics: A Novel Manifold Learning Based Automatic Feature Extraction Method for Diagnostic Prediction in Multiparametric Imaging," Medical Physics, vol. 43, pp. 3373-3374, 2016.
4. P. Carlo, **V. Parekh,** S. Harvey, C. B. Umbricht, A. C. Wolff, M. A. Jacobs, "Integration of mammographic, ultrasound, and clinical metrics for characterization of breast lesions using novel informatics modeling with comparison to OncotypeDX". Presented at the Radiological Society of North America, 2016
5. K. J. Macura, **V. Parekh,** S. S. Dianat, M. A. Jacobs, "Novel Informatics Modeling of Magnetic Resonance Imaging Metrics for Characterizing Prostate Lesions with Pathology Correlation". Presented at the International Society for Magnetic Resonance in Medicine, 2017
6. Michael A. Jacobs, **Vishwa S. Parekh**, Katarzyna J Macura, Susan Harvey, Ihab Kamel, Rihal EI-Khouli, David A. Bluemke, "Multiparametric deep learning tissue

signatures for a radiological biomarker of breast cancer: preliminary results', Presented at the Radiological Society of North America, 2017

7. Michael A. Jacobs**,** Dariya I. Malyarenko, David C. Newitt, **Vishwa S. Parekh**, Nola M. Hylton, Thomas L. Chenevert, "Multisite reproducibility of radiomics and ADC measurements for temperature-controlled phantom: preliminary results". Presented at the International Society for Magnetic Resonance in Medicine, 2018
8. **V. Parekh,** K. J. Macura, M. A. Jacobs, "General AI for organ invariant multiparametric MRI tissue segmentation and characterization: preliminary results". Presented at the International Society for Magnetic Resonance in Medicine, 2018

1.4.4 Patents

1. Vatsal Agarwal, Pornpat Athamanolap, Stephanie Fraley, Michael A. Jacobs, **Vishwa Parekh,** Jeff Tza-Huei Wang, Samuel Yang. "Melt Curve Classifier for Reliable Large-scale Genotyping of Sequence Variants." (WO/2015/164517)
2. M. A. Jacobs, **V. S. Parekh.** "Informatics Radiomics Integration System (IRIS): A novel combined informatics and radiomics method for integration of many types of data for classification into different groups for improved visualization." (US Patent App. 15/334,407, 2016)
3. **V S. Parekh**. M. A. Jacobs, "MIRAGE: Radiomics-Geodesic Features Extraction Decision Support System." (Patent Disclosure JHU: D14297, 2016)
4. **V S. Parekh.** M. A. Jacobs, "Radiological Imaging Physics Encoder (RIPE) System: General artificial intelligence system to model the biological composition

of pathological tissue by encoding its underlying imaging physics for medical applications." (Patent Disclosure JHU: D13500, 2015)

Part I: Image Segmentation

Automatic segmentation of medical images has been a topic of significant interest in the field of computer science for a number of years. A large number of algorithms ranging from supervised to unsupervised have been developed for medical image segmentation over the years [50-52]. There are distinct differences between supervised and unsupervised approaches to image segmentation. Supervised segmentation algorithms require pre-defined labels for the potential tissue classes present in the radiological imaging dataset. This is not practical as there can be tissue classes that are not present in the training dataset, but present in the test dataset (e.g. breast implants). On the other hand, unsupervised segmentation algorithms do not require a pre-defined set of class labels. However, they may produce spurious clusters. Supervised segmentation algorithms map all the voxels to a set of pre-defined class labels which makes them suitable for an automated clinical decision support system where the segmentation step is followed by feature extraction and modeling steps. On the other hand, unsupervised segmentation requires an addition step of mapping obtained clusters to actual class labels before doing any feature extraction and modeling. In Part I, the problem of image segmentation using supervised learning is addressed. **Chapters 2-4** describe the general problem of image segmentation in multiparametric MRI and my contributions to the field of multiparametric radiological image segmentation.

Chapter 2. Multiparametric Deep Learning Tissue Signature Model

2.1 Introduction

Magnetic Resonance Imaging (MRI) has a unique ability to form multiple images influenced by different types of intrinsic tissue contrast (i.e., proton density, T1, and T2) based on MR parameters (TE, TR, and flip angles). As a result, multiparametric MRI (mpMRI) can be obtained, typically consisting of anatomical images of T1- and T2-weighted images in addition to quantitative images such as dynamic contrast enhanced (DCE) MRI and diffusion weighted imaging (DWI). Therefore, it is unlikely that a single MRI parameter can completely characterize the complexity of the tissue of interest. Radiologists' use these different "tissue contrasts" to develop tissue signatures to quickly recognize normal from abnormal tissue after years of training. Consequently, if all the imaging parameters can be integrated using advanced computational algorithms to construct "tissue signatures" of different tissue types, more accurate characterization is possible [53-61]. Therefore, the idea is to use tissue signatures to train our image segmentation algorithm to better assist radiologist in the future reading rooms.

Supervised segmentation requires pre-defined labels for all the tissue classes present in the dataset of interest, which requires direct interaction with radiologists to draw segmentation masks or labels on the different tissues present in the radiological dataset of each patient. This task of creating complete labeled datasets is very time consuming and expensive and not likely to happen. As a result, the task of pre-labeling radiological datasets has sometimes been modified to that of labeling a few voxels per tissue of interest per patient. However, this leads to potential under-sampling of different tissues within the dataset and

could produce inaccurate segmentation results. The goal of this work was to overcome these issues of labeling and potential under-sampling of different tissues using the newly devised algorithms.

Conventional supervised machine learning algorithms attempt to model the relationship between the input training data and output labels. As a result, they are not capable of handling unlabeled data. Unlabeled data can be handled by adding an unsupervised component (e.g. representation learning) as a pre-processing step to the segmentation model [62]. The objective of the segmentation component of my thesis was developing an unsupervised component to learn an intrinsic representation of all the tissue signatures (labeled and unlabeled) present in the training dataset. To that end, I implemented stacked sparse autoencoders (SSAE) for dimensionality reduction. The reason behind the choice of SSAE instead of other non-linear dimensionality reduction algorithms such as Isomap, diffusion map, locally linear embedding [40-43] was the interpretability of the lower-dimensional representation for SSAE as further detailed in **Chapter 3**. Furthermore, SSAE produces specialized intrinsic representations that activate in response to a subset of input tissue signatures, making it more appropriate for this application.

I developed the multiparametric deep learning (MPDL) tissue signature model based on SSAE for segmentation of multiparametric radiological images. Sparse autoencoders (SAE) are unsupervised neural networks that learn an intrinsic representation of the input by attempting to reconstruct it [63]. Furthermore, the activation of each node in the hidden layer of the autoencoder is specialized to activate in response to only a subset of input data by the introduction of a sparsity constraint (see **Section 2.3.2**). The unsupervised nature of an SAE allows us to train them on the complete set of labeled and unlabeled data. In MPDL,

a set of five SAEs were stacked together with each subsequent SAE trained on the output of the previous SAE and a final classifier to classify the data into different tissue types. The complete MPDL architecture was then fine-tuned on the labeled data to generate segmentation masks.

In this chapter, I establish the use of the MPDL tissue signature model for radiological biomarkers of breast tissue. **Section 2.2** provides a background on machine and deep learning techniques followed by the stacked sparse autoencoders based MPDL segmentation algorithm in **Section 2.3**. **Section 2.4** details the clinical data, multiparametric imaging protocol, image registration, tissue signatures, and the training parameters developed for the algorithms developed in this chapter. The results from the application of MPDL to breast cancer tissue segmentation and diagnosis are detailed in **Section 2.5**. This chapter concludes with the potential of MPDL tissue signature model in the "precision" medicine applications for different pathologies.

Publication from this work and author contributions

Publication

V. S. Parekh, K. J. Macura, S. Harvey, I. Kamel, R. EI-Khouli, D. A. Bluemke, M. A. Jacobs., "Multiparametric Deep Learning Tissue Signatures for a Radiological Biomarker of Breast Cancer: Preliminary Results," arXiv preprint arXiv:1802.08200, 2018 (Positive Reviews in Scientific Reports – have resubmitted).

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

I developed the multiparametric deep learning (MPDL) tissue signature model and did the entire implementation of all the four algorithms discussed in this chapter.

The clinical data was acquired by Dr. Macura, Dr. Harvey, Dr. Kamel, Dr.-EI-Khouli, Dr. Bluemke, and Dr. Jacobs. This included patient recruitment, scanning, and radiological reporting **(Sections 2.4.1 and 2.4.2).**

2.2 Background

Deep learning techniques have gained popularity in recent years with the development of advanced optimization techniques and an increase in computational efficiency. Deep learning has begun to play an integral part in many different aspects of modern society and has produced excellent results in a variety of fields such as object detection and recognition, text generation, music composition, and autonomous driving to name a few [64]. Computer assisted clinical and radiological decision support have a huge potential for impact from these technologies, with some excellent initial results. A detailed review on the application of deep learning to radiological applications can be found in [52].

The major difference between deep learning and conventional machine learning algorithms lies in the fact that deep learning algorithms do not require an intermediate feature extraction or engineering step in order to learn the relationship between the input x (e.g. gray level intensity values on radiological images) and the corresponding labels y (the tissue type corresponding to these intensity values). **Figure 2.1** illustrates the differences between conventional machine learning and deep learning. Conceptually, machine learning algorithms model the relationship between the input x and labels y using a probability distribution, p over x and y. Machine learning algorithms can be broadly classified into generative and discriminative methods depending on the form of p [65].

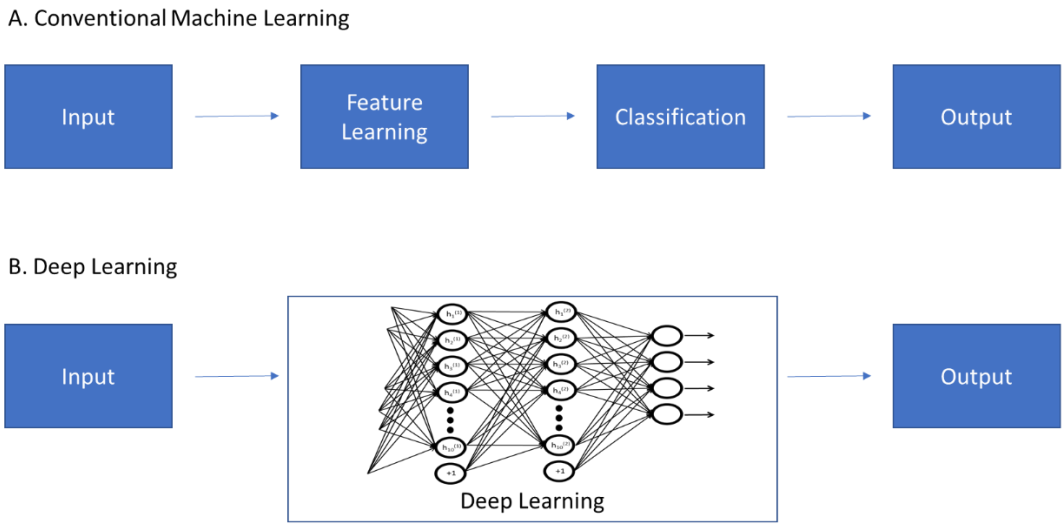


Figure 2.1 Schematic diagram of the differences between conventional machine learning and deep learning methods.

Generative models learn the joint probability distribution $p(x, y)$ in order to estimate the posterior probability $p(y|x)$. Some examples of generative deep learning algorithms include generative adversarial network, variational autoencoder and deep belief networks. In contrast, discriminative models estimate the posterior probability $p(y|x)$ directly without calculating the intermediate joint probability distribution. In other words, discriminative models learn a direct mapping between x and y. Convolutional neural networks, stacked autoencoders and feed forward neural networks comprise the typical examples of discriminative deep learning algorithms [64]. The decision regarding which type of machine learning model to apply for an application depends on the specific problem we are trying to solve. If the problem requires us to only predict the labels y from x, then discriminative models are a better choice as they are not concerned with modeling of $p(x, y)$ and hence may model parameters more effectively to model $p(y|x)$, thereby producing a classifier with higher accuracy. However, discriminative models may not be used if the input, x consists of a large

number of missing values or data points and would require data imputation. In addition, generative models allow the generation of new synthetic data and model different relationships within the input data. For example, if the goal is to classify a lesion as benign or malignant, discriminative deep learning would be a better choice, however, if our goal is to develop a synthetic CT image from multiparametric MRI, a better choice of algorithm would be one of the generative deep learning algorithms.

For my thesis, I am interested in segmentation of different tissue types from multiparametric MRI which makes discriminative deep learning a better choice of algorithm. Furthermore, I am also interested in automatically discovering intrinsic patterns from lesion segmentation that can be useful in classifying benign from malignant lesions. The data in this thesis consists of a large number of unlabeled voxels as compared to labeled voxels, thereby making stacked sparse autoencoders the best choice for my application [66].

2.3 Multiparametric Deep Learning Segmentation Algorithm

2.3.1 Multiparametric Deep Learning Tissue Signature

A voxel tissue signature vector is defined as the vector of gray level intensity values corresponding to that voxel position in each image in the entire data sequence (e.g. breast MRI, $n = 23$ images). Mathematically, the MPDL tissue signature is defined as follows:

$$MPDL\ Tissue\ Signature = [T_1, T_2, DWI, \cdots DCE_n]^T$$

The multiparametric deep learning network was trained on the breast tissue signatures defined using EI identified on all the original breast MRI images as demonstrated in **Figure 2.2**. The MPDL network builds a composite feature representation of the breast tissue signatures of the underlying breast tissue as described in the following subsections.

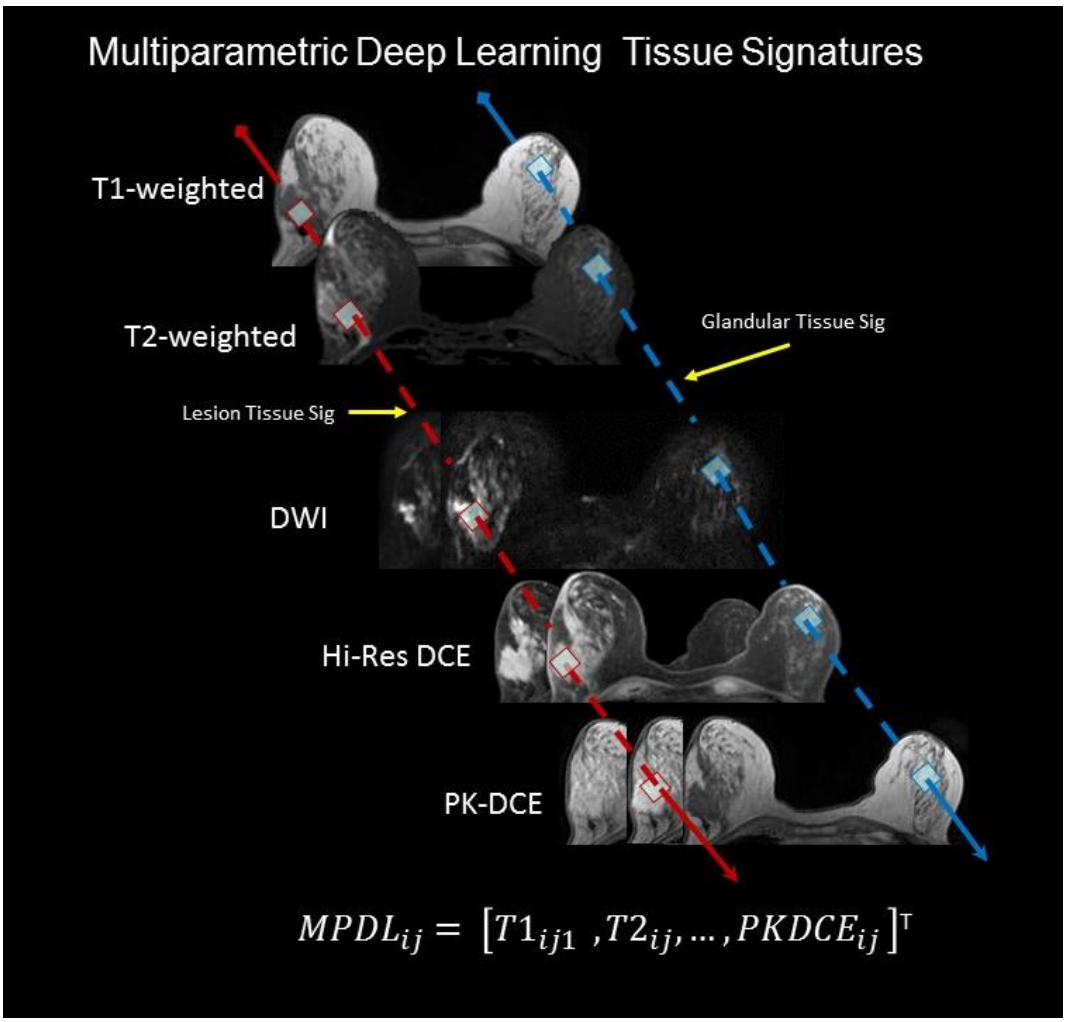


Figure 2.2. Demonstration of the multiparametric deep learning tissue signatures on axial mpMRI of the breast. Representative tissue signatures from the normal and abnormal tissue are obtained on each of the input MRI to create the vector signature $MPDL_{ij}$.

2.3.2 Autoencoder

Autoencoders are unsupervised neural networks that are trained to create a compact or a low-dimensional representation of its input via a hidden layer. For example, if $X = \{x^{(1)}, x^{(2)}, \ldots, x^{(N)}\} \in R^D$ represents an input dataset consisting of N D-dimensional tissue signatures, the goal of an autoencoder would be to create a neural network with d nodes in

the hidden layer such that $Y = \{y^{(1)}, y^{(2)}, \ldots, y^{(N)}\} \in R^d$ characterizes the low-dimensional representation of X.

The network architecture for autoencoders is demonstrated in **Figure 2.3** using an example multiparametric MRI dataset. An autoencoder has two parts: an encoder and a decoder.

The encoder maps the vector X to the vector $h^{(1)}$ representing the hidden layer as follows:

$$h^{(1)} = f\ (WX + b)$$

where f is the transfer function for the encoder, $W \in R^{d \times D}$ is the matrix of weights and b is the bias vector. The function f can be the identity function in case of linear mapping or the sigmoid in case of nonlinear mapping.

Similarly, the decoder maps $h^{(1)}$ back to X using the following equation:

$$\hat{X} = g(W'h^{(1)} + b')$$

The values of W' and b' are equal to the transpose of W and b in case the weights are tied between encoder and decoder.

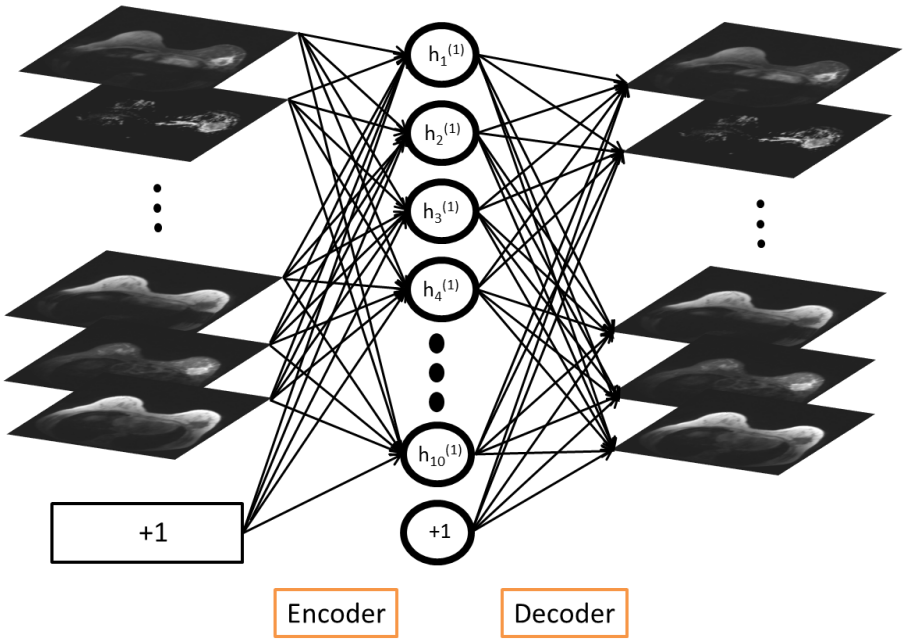


Figure 2.3 Illustration of an autoencoder used to learn a low (ten) dimensional representation of the high-dimensional ($D = 23$) multiparametric MRI data.

2.3.2 Sparse autoencoder

Sparse encoding represents a special type of encoding in which each node in the hidden layer activates in response to only a subset of the total number of training examples. For example, after training a sparse autoencoder on a multiparametric MRI dataset, node 1 may have "specialized" in activating only in response to a fatty tissue signature while node 2 may have "specialized" in activating only in response to a glandular tissue signature. Sparse autoencoders are actualized by adding a sparsity regularizer to the cost function. A sparsity regularizer is defined as follows:

Let $\widehat{\rho_j}$ denote the average activation of a neuron j given by the following equation

$$\widehat{\rho_j} = \frac{1}{N}\sum_{i=1}^{N} h_j^{(1)}(x_i)$$

where N is the total number of training samples.

If ρ denotes the desired average activation or the sparsity proportion of the neuron j across all the training samples, our goal is to impose the constraint $\widehat{\rho_j} = \rho$. Consequently, the sparsity regularization term added to the cost function is given as

$$R_S = \beta \sum_{j=1}^{d} \left(\rho \log \hat{\rho}_j + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_j} \right)$$

where β is the sparsity regularization penalty.

2.3.3 Stacked Sparse autoencoder

The sparse autoencoder discussed in the previous subsection can be stacked to form a deep learning architecture. Stacking the autoencoders provides the deep network with a set of pre-trained weights. The pre-trained weights are especially useful when the application is limited by the number of available pre-labeled training examples.

The first sparse autoencoder in the stack trains each D-dimensional input $x^{(i)}$ into its primary feature representation $h^{(1)}$. The features or the hidden unit activations, $h^{(1)}$ from the first sparse autoencoder are now used as input to the second sparse autoencoder, which is trained to form a composite representation $h^{(2)}$. The composite features $h^{(2)}$ act as raw input to the softmax classifier, training it to classify the initial tissue signature, $x^{(i)}$ as its respective tissue type. A softmax classifier is a generalization of binary logistic regression for classification of input data into multiple classes. The output of the softmax layer is given by $y = (WX + b)$. The output y can be interpreted as a vector of unnormalized log probabilities for each class, with our loss function given by cross-entropy as follows:

$$J = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{4} t_{ij} \ln y_{ij} + \left(1 - t_{ij}\right) \ln(1 - y_{ij})$$

where t_{ij} is the target class and y_{ij} is the output of the deep network at the softmax classification layer. The algorithm for multiparametric deep learning segmentation is represented as a pseudocode in **Algorithm 2.1.** **Figure 2.4** shows the deep architecture formed by stacked sparse autoencoders used for segmentation of multiparametric MRI into different tissue types.

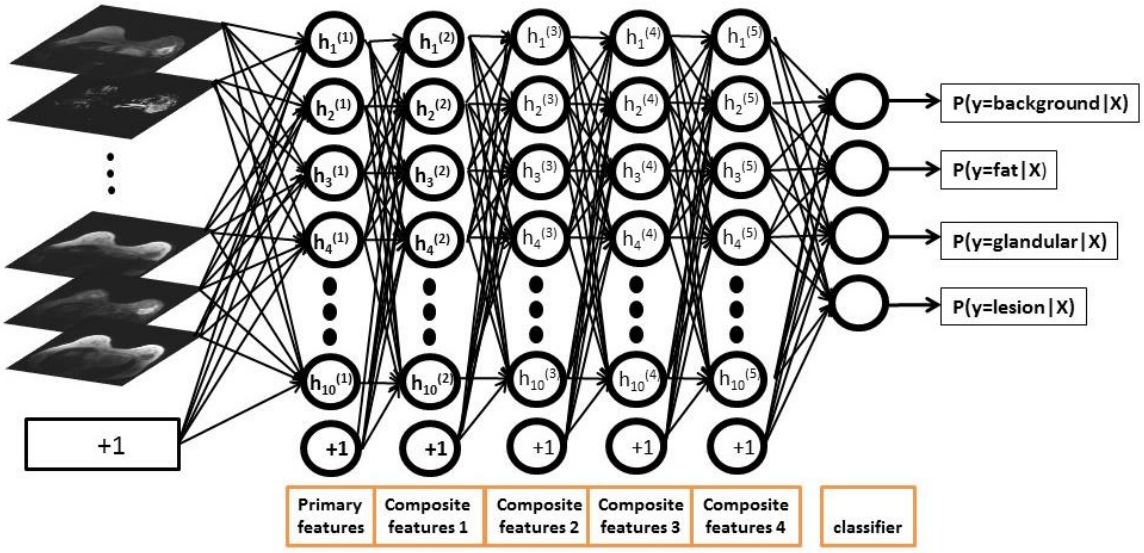


Figure 2.4 Multiparametric MRI segmentation deep network architecture trained to segment multiparametric breast MRI into regions of different tissue types and background. The stacked sparse encoder deep network is constructed of five hidden layers with ten nodes each and a softmax classification layer that outputs the probability of different tissue types for the input tissue signature.

Algorithm 2.1 MPDL segmentation algorithm

Inputs:
X: N multiparametric MRI datasets
$SSAE_{arch}$: Network architecture for $SSAE$. This initialization only contains the network connections.
Output: trained $SSAE$ segmentation network
Initialization: $TS_{train} = \phi, TS_{labels} = \phi$
1: // Generate tissue signatures for training
2: for every dataset, x in X
3: Run Eigenfilter segmentation [67] to generate tissue signature, S_k for every tissue type, k
4: Add S_k to TS_{train}
5: Add k to TS_{labels}
6: Initialize $SSAE$ with $SSAE_{arch}$
7: // Train the $SSAE$ layers in an unsupervised fashion over the complete dataset
8: for every layer in $SSAE$
9: if layer is the first hidden layer
10: train SAE on input X
11: else
12: train SAE on previous layer's output
13: Set layer weights to encoder weights of SAE
14: Set layer output by running SAE encoder on the layer's input
15: $SSAE$ = train $SSAE$ on TS_{train} using TS_{labels} as ground truth
16: return $SSAE$

2.4 Materials and Methods

*The clinical data from our institution was acquired by Dr. Macura, Dr. Harvey, Dr. Kamel, Dr. EI-Khouli, Dr. Bluemke, and Dr. Jacobs. This included patient recruitment, scanning, and radiological reporting (**Sections 2.4.1 and 2.4.2**).*

2.4.1 Clinical subjects

All studies were performed in accordance with the institutional guidelines for clinical research under a protocol approved by the Johns Hopkins University School of Medicine Institutional Review Board (IRB) and all HIPAA agreements were followed for this retrospective study. One hundred and ninety-two women were part of this study. Of the one hundred and ninety-two patients, one hundred and forty-two patients (97 malignant,

41 benign and four normal (no-lesion)) were scanned at our institution. Malignancy was determined by pathology in all cases. The remaining fifty ($n = 50$) cases were obtained from University of California San Francisco (UCSF) for an independent deidentified test data set. These fifty cases were from a Phase 3 clinical trial for women receiving neoadjuvant chemotherapy for locally advanced breast cancer defined by histology [33, 34]. I used the baseline study before initiation of the therapeutic regimen.

2.4.2 Multiparametric MRI imaging protocol

MRI scans were performed on a 3T magnet, using a dedicated phased array breast coil with the patient lying prone with the breast in a holder to reduce motion. MRI sequences consisting of fat suppressed (FS) T2WI spin echo ($TR/TE = 5700/102\ ms$) and fast spoiled gradient echo (FSPGR) T1WI ($TR/TE = 200/4.4\ ms$, $Field\ of\ View\ (FOV) = 35 \times 35\ cm$, $matrix\ size\ (MS) = 256 \times 256$, $slice\ thickness\ (ST) = 4\ mm$, 1 mm gap); diffusion-weighted ($TR/TE = 5000/90ms$, $b = 0 - 800$, $matrix\ size = 192 \times 192$, $ST = 6mm$); and finally, pre- and post-contrast enhanced images FSPGR T1WI ($TR/TE = 20/4\ ms$, $MS = 512 \times 512$, $ST = 3\ mm$) were obtained after intravenous administration of a GdDTPA contrast agent ($0.2mL/kg(0.1\ mmol/kg)$). The contrast agent was injected over 10 seconds, with MRI imaging beginning immediately after completion of the injection and the acquisition of 14 phases. The contrast bolus was followed by a 20cc saline flush. The DCE protocol included two minutes of high temporal resolution (15 sec per acquisition) imaging to capture the wash-in phase of contrast enhancement. A high spatial resolution scan for two minutes then followed, with additional high temporal resolution images (15 sec per acquisition) for an additional two minutes to characterize the wash-out slope of the kinetic curve for

pharmacokinetics(PK) [68]. Total scan time for the entire protocol was less than 45 minutes.

The independent validation breast MRI scans were acquired on a different 1.5T magnets using a dedicated breast RF coils and obtained from the I-SPY clinical trial. The images used for validation were fat suppressed, T1 weighted dynamic contrast enhanced series obtained unilaterally in the sagittal orientation with $TR \leq 20ms$, $TE = 4.5ms$, $flip\ angle \leq 45°$, $FOV = 16 - 18\ cm$, $MS > 256 \times 192$, $ST \leq 2.5\ mm$.

2.4.3 Multiparametric Image Registration

The mpMRI were coregistered using a hybrid registration algorithm that combines 3D wavelet transformation for 3D reslicing and rescaling of the MRI volumes with nonlinear affine transformation to minimize the loss of information in image transformations [61]. The pre-contrast image of the DCE dataset was used as the reference image for all the other MRI images.

2.4.4 Multiparametric MRI tissue signature generation

The Eigenimage filter (EI) segmentation algorithm was used to segment the breast lesions from the post contrast DCE image [67]. The EI is a linear filter that maximizes the projection of the desired tissue (lesion tissue) while it minimizes the projection of undesired tissues (glandular tissue) onto a composite image called an Eigenimage [69].

For this study, four sets of tissue signature vectors were defined. The first set of tissue signatures for normal tissue, $Normal = [N_1, N_2, \cdots . N_n]^T$ was chosen from the glandular tissue, second one for fatty tissue, $Fatty = [F_1, F_2, \cdots . F_n]^T$, a third one for lesion tissue, $Lesion = [L_1, L_2, \cdots . L_n]^T$ and a fourth one for background noise. Each set of MPDL tissue signature vectors created automatically using the following multiparametric region growing

algorithm. The initialization to the region growing algorithm is provided by the operator by identifying a pixel within the tissue of interest. The algorithm attempts to grow the region within a 5x5 neighborhood of the initialization provided by the user. A point in the neighborhood is added to the region of interest if the percentage difference between the signal intensity at that pixel location and the signal intensity of the operator initialization is less than 5%. The final ROI is created by computing a logical AND operation between the ROIs generated from region growing on each of the MR images (**Figure 2.2, Page 26**). By using several images concurrently, the probability of a pixel from another tissue being included in the final ROI (due to noise, partial volume, and nonuniformities) is reduced [70]. The computer time required for producing the final ROI was less than a second for each tissue type.

2.4.5 MPDL segmentation network training and evaluation

The training parameters of the multiparametric MRI segmentation deep network were set as follows:

- Number of hidden layers = 5
- Number of nodes in each hidden layer = 10
- L2 regularization penalty = 0.001
- Sparsity proportion = 0.25
- Sparsity regularization = 4
- The transfer function for the autoencoder nodes was selected as the saturating linear function given as

$$f(x) = \begin{cases} 0, if\ x \leq 0 \\ x, if\ 0 < x < 1 \\ 1, if\ x \geq 1 \end{cases}$$

I tested the multiparametric breast MRI deep network for segmentation of breast into different tissue types using two-fold cross validation. The balance between number of tissue signatures used to train different tissue types was maintained by sampling uniformly at random equal number of tissue signatures corresponding to each tissue type from each patient. In order to evaluate the trained deep networks, the dice similarity between the deep network tumor segmentations and the ground truth obtained using the eigen filter algorithm was computed on the hold out dataset during each cross-validation fold [71]. The dice similarity (DS) between the deep network tumor segmentation, A and ground truth EI segmentation, B is given by the following equation

$$DS = \frac{2(A \cap B)}{n(A) + n(B)}$$

The EI segmentation was obtained by thresholding the EI contrast image. The threshold was obtained by evaluating the EI contrast image histogram and using the mean and a 95% confidence interval. If the images have full overlap, then the $DS = 1.0$ and if there is no overlap, then the $DS = 0$.

Validation of the SSAE Deep Network

I used 50 patients with calculated volumes from the UCSF I-Spy ACRIN study to test the MPDL network [32, 33, 72]. To compare with our dataset, I used the baseline DCE contrast imaging session from the study. The UCSF data was registered to the DCE pre-contrast image. After application of the MPDL, the segmented MPDL and UCSF volumes were compared and analyzed.

2.4.6 Comparative methods

I compared the performance of the SSAE architecture with three other deep learning architectures – deep SSAE without fine tuning ($SSAE_{untuned}$), deep multilayer perceptron (MLP), and deep convolutional neural networks (CNN) [73, 74]. The deep $SSAE_{untuned}$ and deep MLP were implemented with the same architecture as SSAE (five layers with ten nodes each). The deep $SSAE_{untuned}$ was trained without supervised fine tuning and the deep MLP was trained in a completely supervised fashion.

Patch-based 2D-CNN was trained on image patches of size $5 \times 5 \times N$ corresponding to each N-dimensional tissue signature segmented using the Eigenfilter algorithm. The $5x5$ image patch of a tissue signature corresponds to the immediate 5×5 neighborhood of that voxel position. The 2D-CNN consisted of four layers with 128, 64, 32, and 16 filters respectively, followed by a fully connected layer and a softmax layer. Each layer of the 2D-CNN had the following components:

- Convolutional layer with trainable filters of size 3×3
- ReLU activation function given by the following equation

$$f(x) = \begin{cases} 0, if\ x < 0 \\ x, if\ x \geq 0 \end{cases}$$

- Max pooling layer with a 2×2 window

2.4.8 Statistical Methods

The MPDL was tested and validated using UCSF and DCE segmentations. The percent difference and overlap in the lesion areas segmented from the data were computed. Statistical analysis was performed using linear regression to correlate the total lesion areas.

All area values are presented as mean ± standard deviation. A Student's t-test was employed to determine statistical significance between the correlation of lesion areas and the percent difference between the lesion boundaries. Bland-Altman tests were run to insure no bias in the data. Statistical significance was assigned for $p < 0.05$.

2.5 Results

2.5.1 Patient Demographics

The MPDL network was trained on the patient cohort of 142 women with breast lesions for evaluation of their efficacy in segmenting out different tissue types and classifying benign from malignant lesions. The mean age of the patients was 52 years (range: 24-80 years). Ninety-seven patients had biopsy proven malignant breast lesions (68%) while forty-one patients had benign breast lesions (29%) and four had no lesions identified (3%).

2.5.2 Quantitative mpMRI

In the training data set only the DWI and DCE sequences provided quantitative radiological metrics. There were significant differences ($p < 0.001$) between the ADC map values for malignant and benign breast lesions. ADC values for malignant cases were (mean and standard deviation) 1.26 ± 0.13 $(mm^2 \times 10^{-3}/s)$ and benign lesions were 1.74 ± 0.17 $(mm^2 \times 10^{-3}/s)$. Glandular tissue ADC values for malignant and benign lesions were similar, 2.16 ± 0.46 and 2.34 ± 0.33 $(mm^2 \times 10^{-3}/s)$, respectively. The DCE PK values were significantly different ($p < 0.05$) between malignant and benign lesions. The K^{trans} values were 0.55 ± 0.32 $(1/min)$ and EVF were 0.30 ± 0.16 for malignant cases and 0.25 ± 0.19 $(1/min)$ and 0.22 ± 0.13 for benign cases, respectively.

2.5.3 In-house dataset

Multiparametric Deep Learning Tissue Signature Model

The MPDL tissue signatures were defined for different tissue types and applied to the 142 mpMRI breast cases. **Figure 2.5** demonstrates the mpMRI deep network segmentation results on five representative malignant patients. Similarly, **Figure 2.6** illustrates the mpMRI deep network segmentation results on five benign patients. Glandular and lesion tissue was segmented with high accuracy for all the lesions and in all quadrants of the breast.

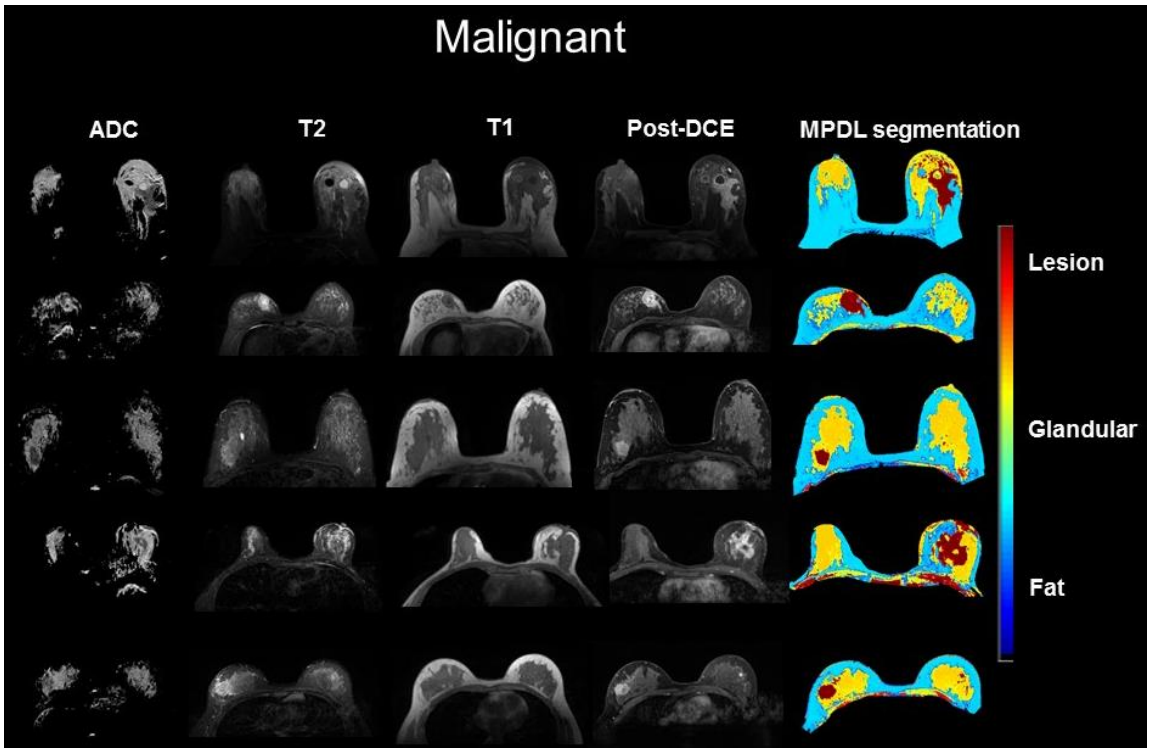


Figure 2.5 Illustrates the use of MPDL network on axial breast mpMRI in five representative malignant patients. The color coding is shown to the right of the images.

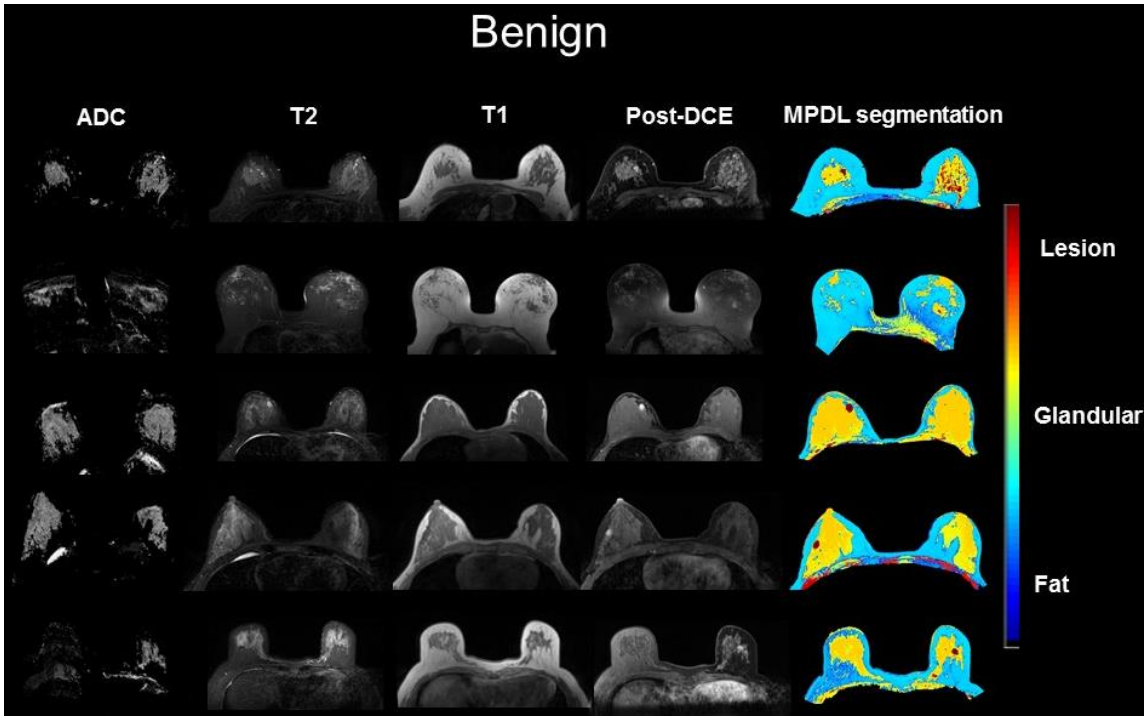


Figure 2.6 Illustrates the use of MPDL network on axial breast mpMRI in five representative benign patients. The color coding is shown to the right of the images.

The dice similarity index between the lesion segmentations defined by Eigenfilter and MPDL demonstrated excellent overlap with mean and standard deviation (SD) 0.87 ± 0.05 for malignant patients and 0.84 ± 0.07 for benign patients and shown in **Figure 2.7.**

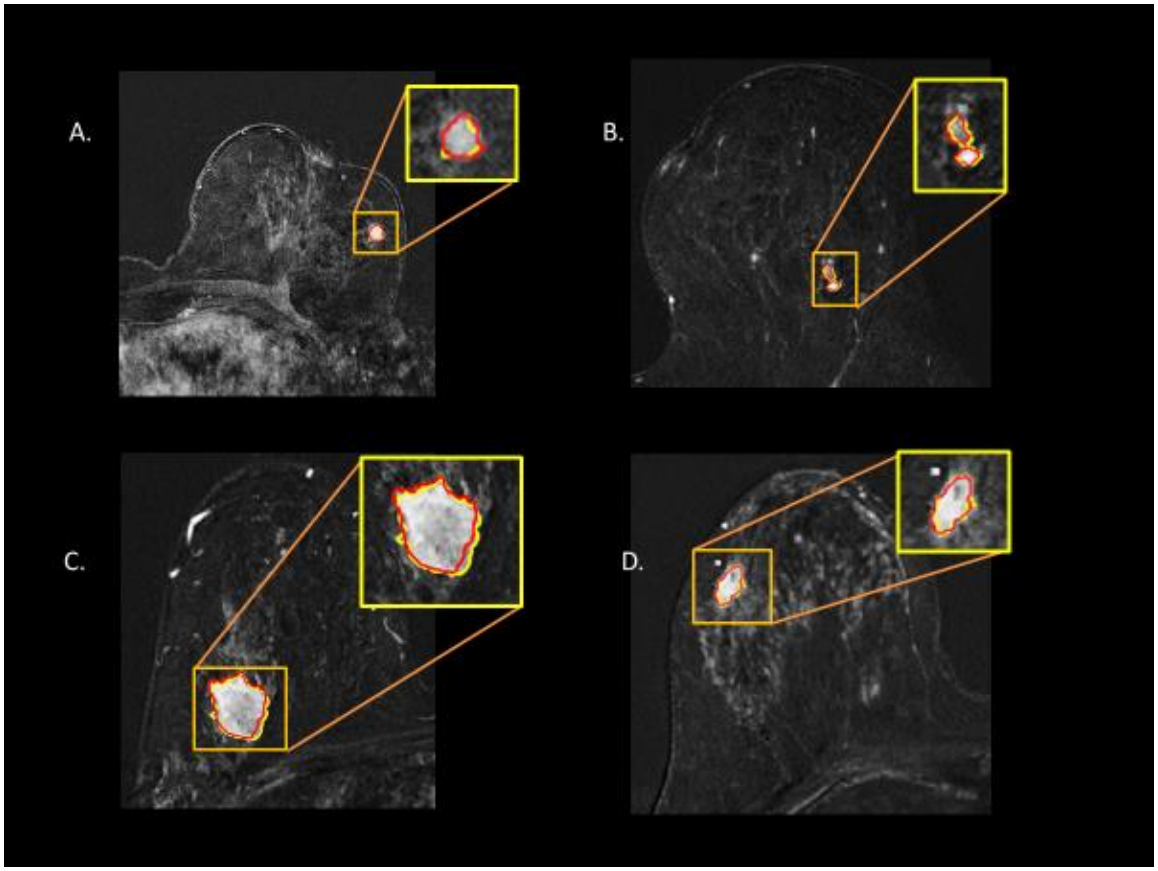


Figure 2.7 Demonstration of overlap between Eigenimage and MPDL segmentation masks of two benign (A and B) and two malignant (C and D) patients overlaid on the dynamic contrast enhanced subtraction image. The Eigenfilter segmentation boundary is displayed in yellow while the MPDL segmentation boundary is displayed in red. On the overlap masks, the blue region corresponds to the overlap between the two methods, yellow represents the area segmented by the MPDL alone while red represents the area segmented by Eigenfilter alone.

Comparison to other deep network architectures

Figure 2.8 demonstrates the performance of the four deep learning architectures on three example patients. As shown in **Figure 2.8**, the segmentations from the SSAE, MLP, and 2D-CNN architectures are very similar, however, the MLP and 2D-CNN architectures appear to be more prone to false positives. Furthermore, the SSAE architecture without fine-tuning missed the major tissue classes (normal and abnormal) in some cases reaffirming the need for supervised fine tuning. The dice similarity for the remaining three architecture on a randomly selected subset of twenty patients (ten benign and ten malignant) were as follows:

1. MLP: 0.84 ± 0.08
2. 2D-CNN: 0.85 ± 0.05
3. SSAE: 0.89 ± 0.04

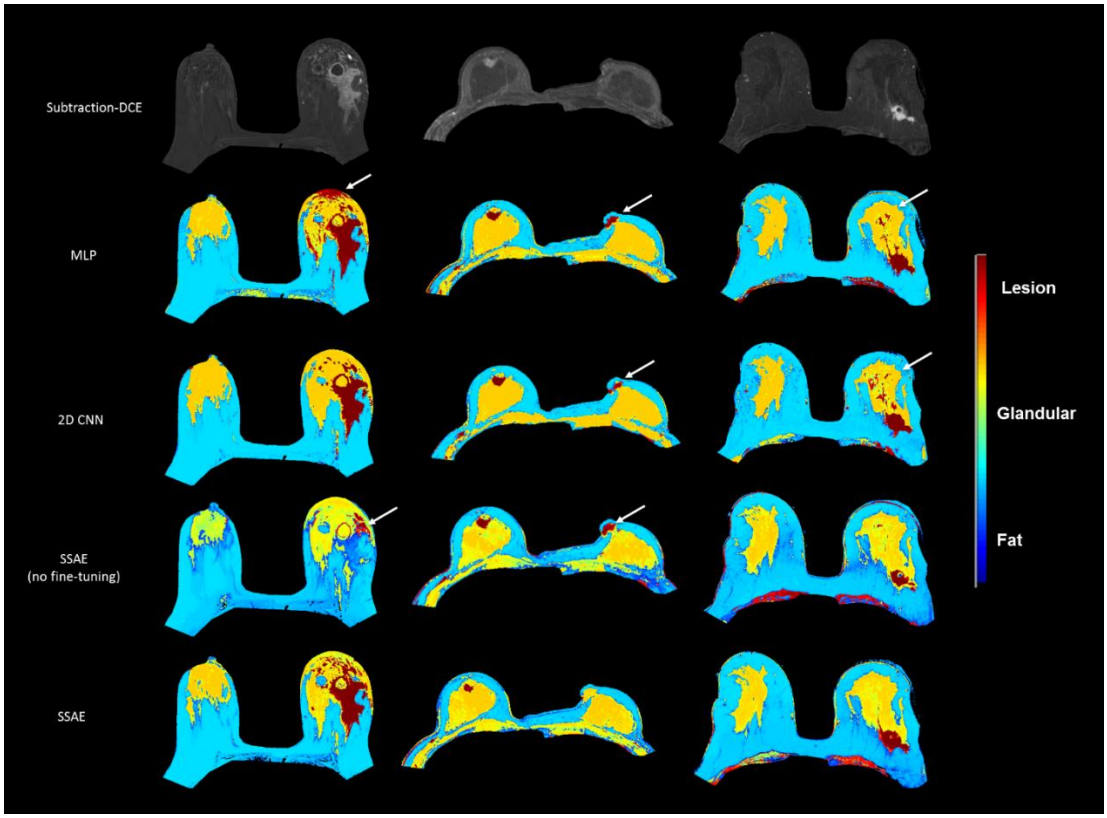


Figure 2.8. Comparison between the breast tissue segmentations obtained from different deep learning architectures viz. multi-layer perceptron (MLP), 2D convolutional neural network (2D CNN), and stacked sparse autoencoders (SSAE) with and without fine-tuning on three example patients. The differences in the lesion segmentations of the three networks are indicated by arrows.

2.5.4 Validation dataset

The trained MPDL network was tested on an independent cohort of 50 patients [32, 34, 72]. The validation of the MPDL segmentation on an independent clinical (UCSF) data was excellent. **Figure 2.9** illustrates representative cases comparing the segmented tissue regions of UCSF validation data set using the MPDL tissue signatures.

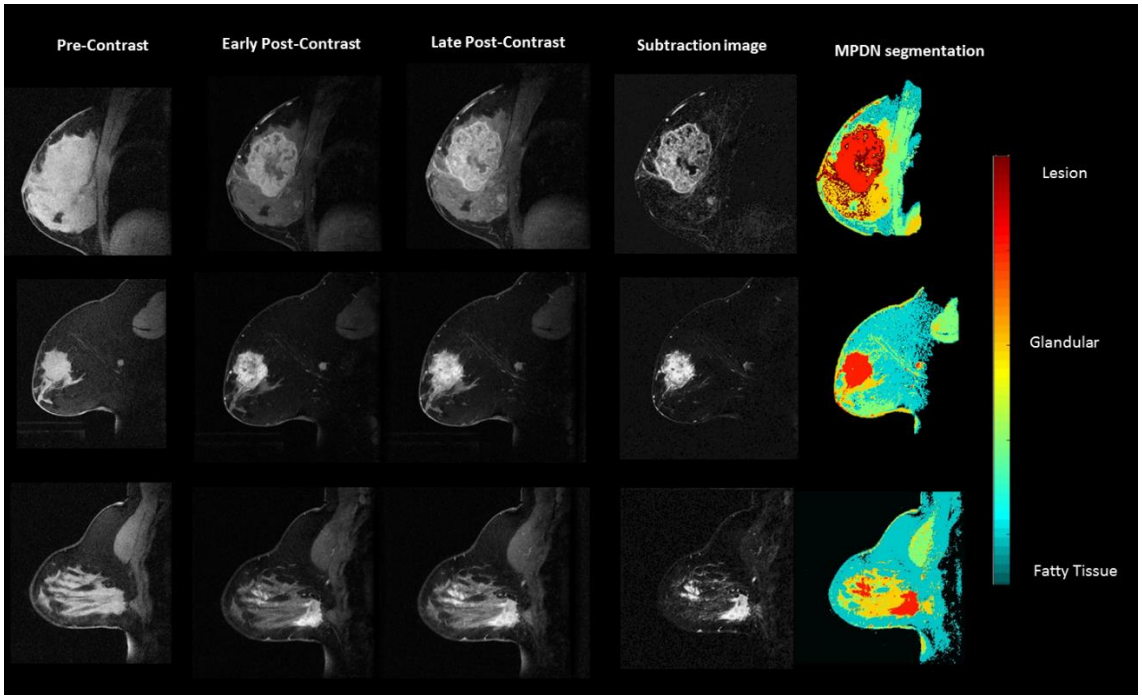


Figure 2.9 Demonstration of three representative cases from the I-SPY2 validation cohort and the resulting MPDL segmentations. In all cases, the segmented regions were highly correlated between each other. The color coding is shown to the right of the images.

The segmented volumes of the UCSF set resulted in a low percent difference of $4.4 \pm 3.9\%$ between lesion volumes defined by MPDL and USCF. Bland-Altman plots are shown in **Figure 2.10**, demonstrating excellent agreement between the data sets. Bland Altman plots are useful in identifying the presence of any fixed bias in our data (e.g. if the difference between the lesion volumes measured by MPDL and UCSF increases as the true volume of the tumor increases). The mean percentage difference in the volumes segmented by MPDL and UCSF was not significantly different from 0 ($p = 0.18$) indicating there was no fixed bias in our data.

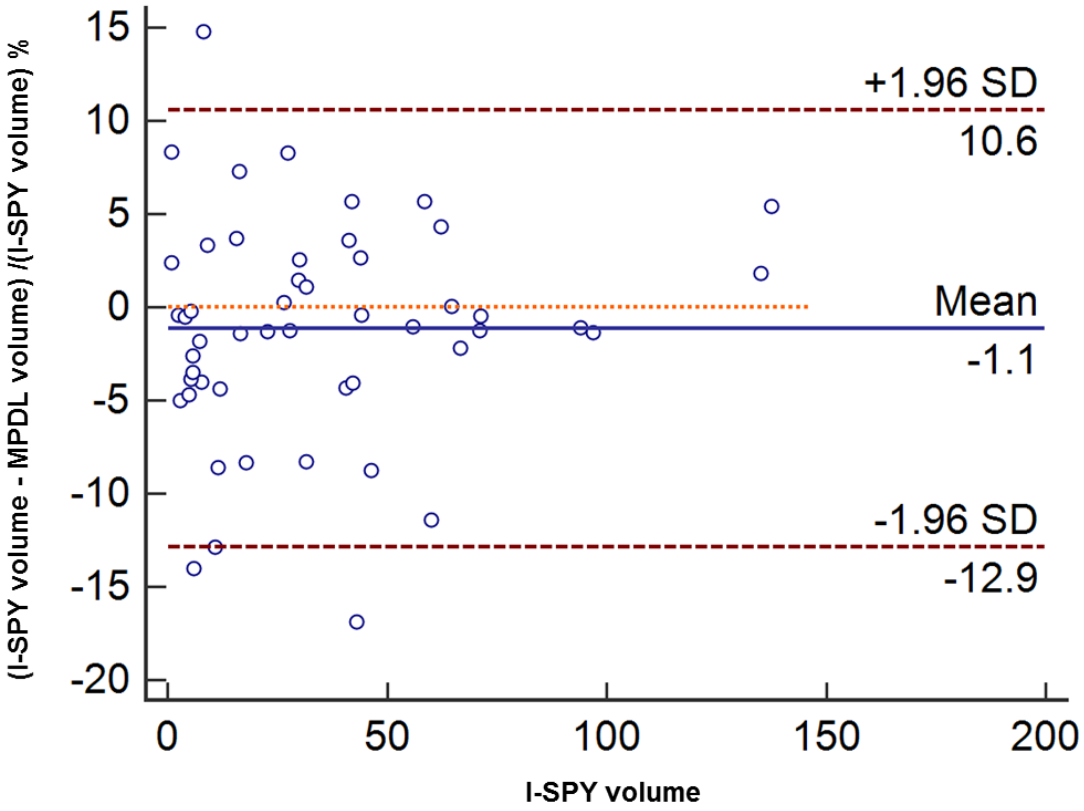


Figure 2.10 Bland-Altman plots demonstrating the limits of agreement of the percent differences on the representative sagittal breast cases from the validation data set and MPDL segmentations. The mean is shown by the center line and the confidential intervals (±2SD) are shown at 10.6% and -12.9%. The plot shows excellent agreement between the two measurements.

2.6 Discussion

I have developed, tested, and validated a cognitive computing platform that organizes, integrates, and interprets imaging information using a MPDL tissue signature model. The application of the MPDL tissue signature model resulted in excellent segmentation and classification of different tissue classes. This study employed an integrated multiparametric breast MRI deep learning model in the clinical setting and demonstrates that MPDL tissue signatures define normal and abnormal tissue with excellent accuracy.

The MPDL model was able to accurately segment breast tissue irrespective of the magnetic field strength (3T for our model and 1.5T for the validation set). Furthermore, the MPDL model was invariant to the imaging orientation as our data was in the axial plane while the validation set was in sagittal plane. This invariance is due to the underlying depiction of the tissue using tissue signature vectors, which captures the tissue underlying characteristics and allows for the "adjustment" to different MRI input. Moreover, the MRI parameters, as well as, the time resolution of the DCE used to train the MPDL model were different for our dataset and the validation dataset reasserting the robust nature of the MPDL model and eliminates the need to "retrain" the MPDL model.

I evaluated the effect of unsupervised pre-training and fine tuning by implementing two additional models. First, I implemented a multi-layer perceptron (MLP) with exactly the same architecture as MPDL to evaluate the efficacy of unsupervised pre-training. Next, I evaluated the results of the SSAE algorithm without fine-tuning the complete architecture. The MPDL algorithm outperformed both MLP and SSAE without fine-tuning demonstrating the effectiveness of the complete MPDL architecture. Furthermore, the MPDL algorithm also compared favorably to the patch based 2D CNN algorithm. The

number of trainable parameters in a 2D CNN architecture was significantly higher than the number of trainable parameters for MPDL. As a result, the size of the training dataset was not sufficient to optimally train the CNN, thereby producing suboptimal results with potentially spurious segmentations.

Currently, it is very unlikely that machine learning and deep learning will replace radiologists as has been suggested by some, yet there may be a role for improved efficiency in the workflow and accuracy of interpretation. Using advanced computational methods allow for this coming change to be better managed within radiology. My results demonstrate that the MPDL method can be used on an independent data set acquired from different institutions. The I-Spy trial is one of the largest MRI trials and incorporates many different MRI field strengths. Moreover, the ability for the MPDL to learn different tissue signatures allows it to adapt to different data sets with highly accurate results. This was shown with the high dice similarity of the validation data using different input MRI data.

There are, however, some technical limitations to the use the MPDL network in practice. First, increased computational power on the graphical processor units ((GPU) > 2500 cores, 12GB) used here may not be widely available. However, the use of advanced GPU computing is rapidly finding applications in many different radiological datasets [52]. More specific to the present study, any assessment of the clinical value of MPDL network will require additional studies in a larger patient population. Moreover, a prospective trial with subsequent follow-up and pathological correlation using MPDL will provide us with new data to explore the exact application and methods to apply to larger studies**.**

In conclusion**,** I have demonstrated that integrated MPDL method accurately segmented and classified different breast tissue from multiparametric breast MRI. The

MPDL images allow for improved visualization of different tissue characteristics based on multiple radiological parameters.

Chapter 3. Multiparametric Deep Learning: Looking under the hood

3.1 Introduction

In **Chapter 2**, I introduced the multiparametric deep learning approach for segmentation of different tissue types in a multiparametric radiological imaging dataset with excellent results on multiparametric breast MRI. However, we have treated the MPDL as a black box until now; i.e., we do not know how the MPDL organizes, integrates, and interprets the imaging information. For example, when radiologists look at radiological imaging datasets, they integrate them in their brain and use their "brain" to form intrinsic knowledge about the data (e.g. fluid is dark on T1WI and bright on T2WI). The goal of this chapter is to decompose the trained MPDL model, open the black box, and answer the following questions:

1. How does the MPDL model encode the intrinsic representations?
2. Can we visualize them?
3. Can we reconstruct the input from the intrinsic representations?
4. Can we predict when the MPDL model would fail?

In this work, I implement different methods for visualization of layer weights and reconstruction of inputs from the activation patterns in different layers. **Section 3.2** describes the methods used to visualize and reconstruct inputs and outputs from different layers. The results from application of these methods are detailed in **Section 3.3**. The final section discusses the results from opening the deep learning black box and the

potential of MPDL tissue signature model in developing a general tissue recognition system.

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

I developed the SAE correlation reduction algorithm and implemented all the techniques discussed in this chapter for visualization of layer weights and input reconstruction.

3.2 Methods

3.2.1 Visualization of layer weights

Consider an example neural network shown in **Figure 3.** This is a simple neural network that directly maps input layer to the output layer with no hidden layer. Mathematically, the input layer is given as $X = \{x^{(1)}, x^{(2)}, \dots, x^{(N)}\} \in R^D$ and the output layer is given as $Y = \{y^{(1)}, y^{(2)}, \dots, y^{(N)}\} \in R^d$, where D is the dimensionality of the input and d is the dimensionality of the output. The activation of any node $y_j^{(i)}$ in the output layer is given by the following equation

$$y_j^{(i)} = f(W_j . x^{(i)} + b)$$

Here W_j is the vector of weights connected to node, y_j and f is the activation function. The weight vector acts as a pattern recognizer and determines the input patterns that would activate the corresponding node. The goal here is to plot the weight vector W_j for all different nodes j in the MPDL tissue signature model and understand what kind of patterns are encoded at different layers of the MPDL neural network.

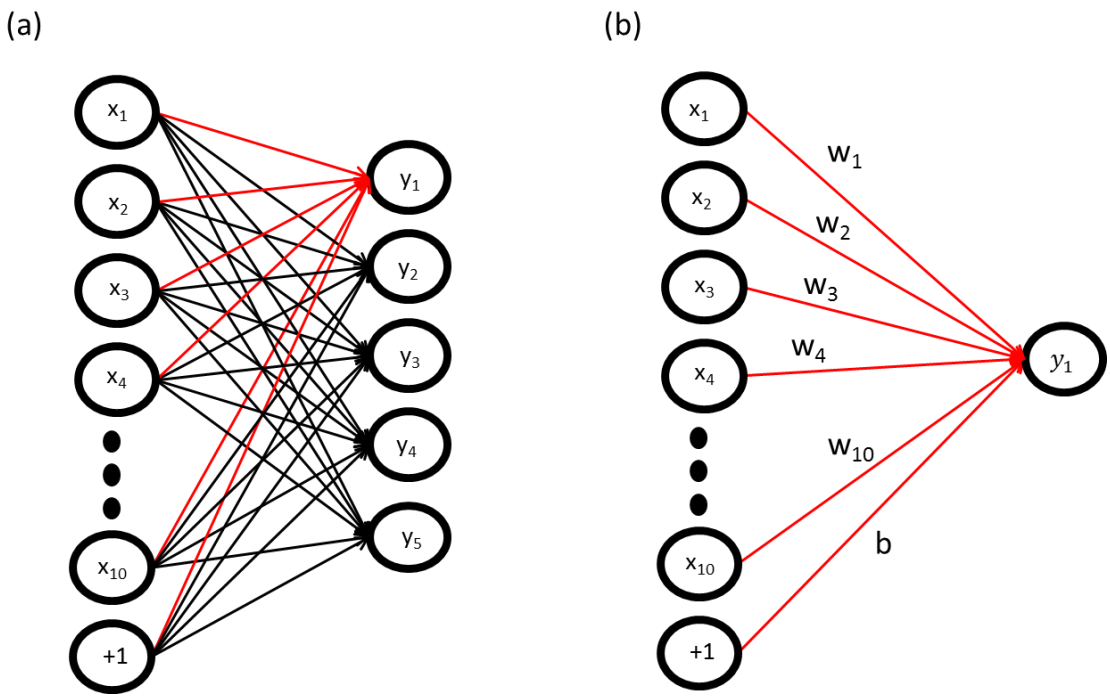


Figure 3.1 (a) Illustration of a simplified neural network model with no hidden layers. The inputs are directly connected to the outputs. (b) Illustration of the weight vector $[w_1, w_2, \dots, w_{10}]^T$ that determines the activation of the node y_1 based on the input vector pattern, $[x_1, x_2, \dots, x_{10}]^T$.

3.2.2 Layer Size Reduction

Sparse Autoencoders (SAEs) produce specialized nodes using sparsity constraint as discussed in **Section 2.3.2**. Moreover, this property allows us to initialize SAEs with a hidden layer of size greater than that of the input layer. However, training an SAE with a large number of nodes could produce redundant nodes in the hidden layer. I developed the technique of correlation reduction to remove redundant nodes from the hidden layer. The technique of correlation reduction consists of three steps as shown in **Figure 3.2.** The first step involves computation of a pairwise correlation matrix between the weight vectors of

all the nodes in the hidden layer. The next step involves identifying redundant nodes by setting a threshold on the value of correlation coefficient. The nodes with correlation coefficient greater than the threshold are considered as redundant. The final step involves selecting the node with the maximum average activation as the representative node from every set of redundant nodes and dropping all the other nodes from the hidden layer. **Algorithm 3.1** summarizes the procedure for correlation reduction.

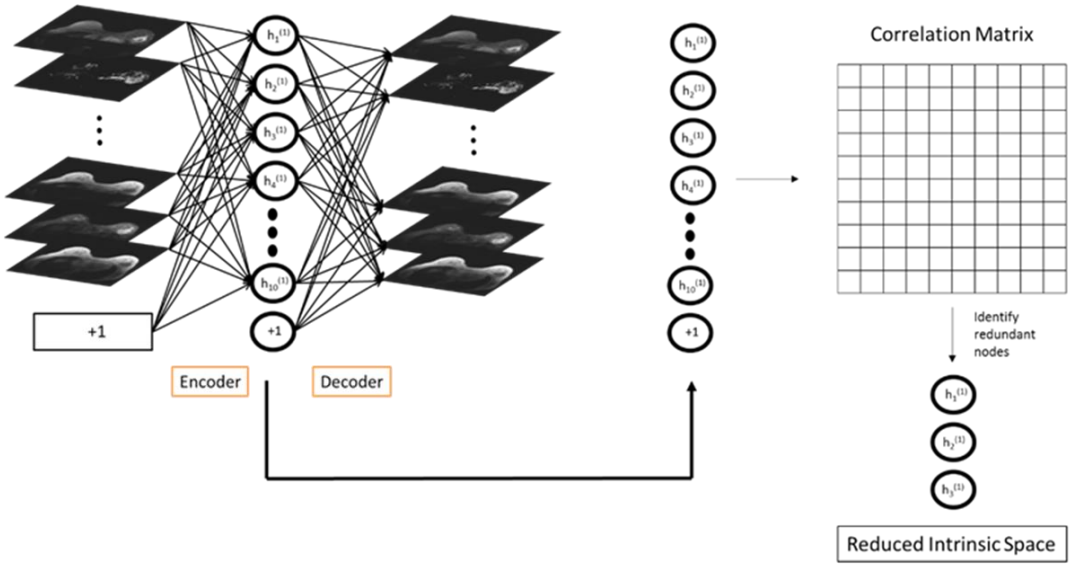


Figure 3.2 Flowchart demonstrating the correlation reduction algorithm. The correlation reduction algorithm identifies weight vectors in the hidden layer that are highly correlated and removes the redundant nodes to produce a smaller hidden layer.

Algorithm 3.1 Correlation Reduction algorithm

Inputs:
X: N multiparametric MRI datasets
t: correlation threshold for network reduction
Output: Trained SAE model with optimal number of nodes
1: Create a tissue signature database, S of tissue signatures from X
2: Train SAE on S
3: Extract weight vectors from the hidden layer of the trained SAE
4: Compute correlation coefficient matrix, CCM by computing pairwise correlation coefficient between all the pairs of weight vectors.
5: Identify redundant nodes on the CCM as the nodes with pairwise correlation coefficient $> t$
6: For every set of redundant nodes, updated SAE by keeping the node with maximum average activation on the training dataset.
7: return SAE

3.2.3 Reconstruction of inputs

The MPDL deep network uses saturated linear function as the activation function for each node. This allows us to approximate the input tissue signatures from the activation pattern of a single node, a hidden layer, or the final classification layer using linear reconstruction. The process of linear reconstruction does not provide an accurate estimation of the input tissue signature but generates an approximate reconstruction of the relationships between different input parameters that could be useful in determining how these relationships relate to the relationships learned by the human brain.

3.3 Experiments

3.3.1 Weight vector visualization for the first and second layers of the MPDL tissue signature model

I did a layer-wise decomposition of the MPDL SSAE network for visualization and analysis of the intrinsic representation learnt by the MPDL for different tissue types. **Figure 3.3a** illustrates the weight vectors learnt by the SSAE in the first layers. The node

corresponding to the weight vector illustrated in **Figure 3.3b** would activate when the input is a lesion or glandular tissue. This can be inferred from the assignment of positive weights to the high resolution DCE images. Similarly, all the different weight vectors in the layer 1 can be attributed to a single or a combination of closely related tissue types. Similarly, visualization of the layer 2 weights demonstrates how the nodes in layer 1 combine to generate more specialized activations as shown in **Figure 3.4.**

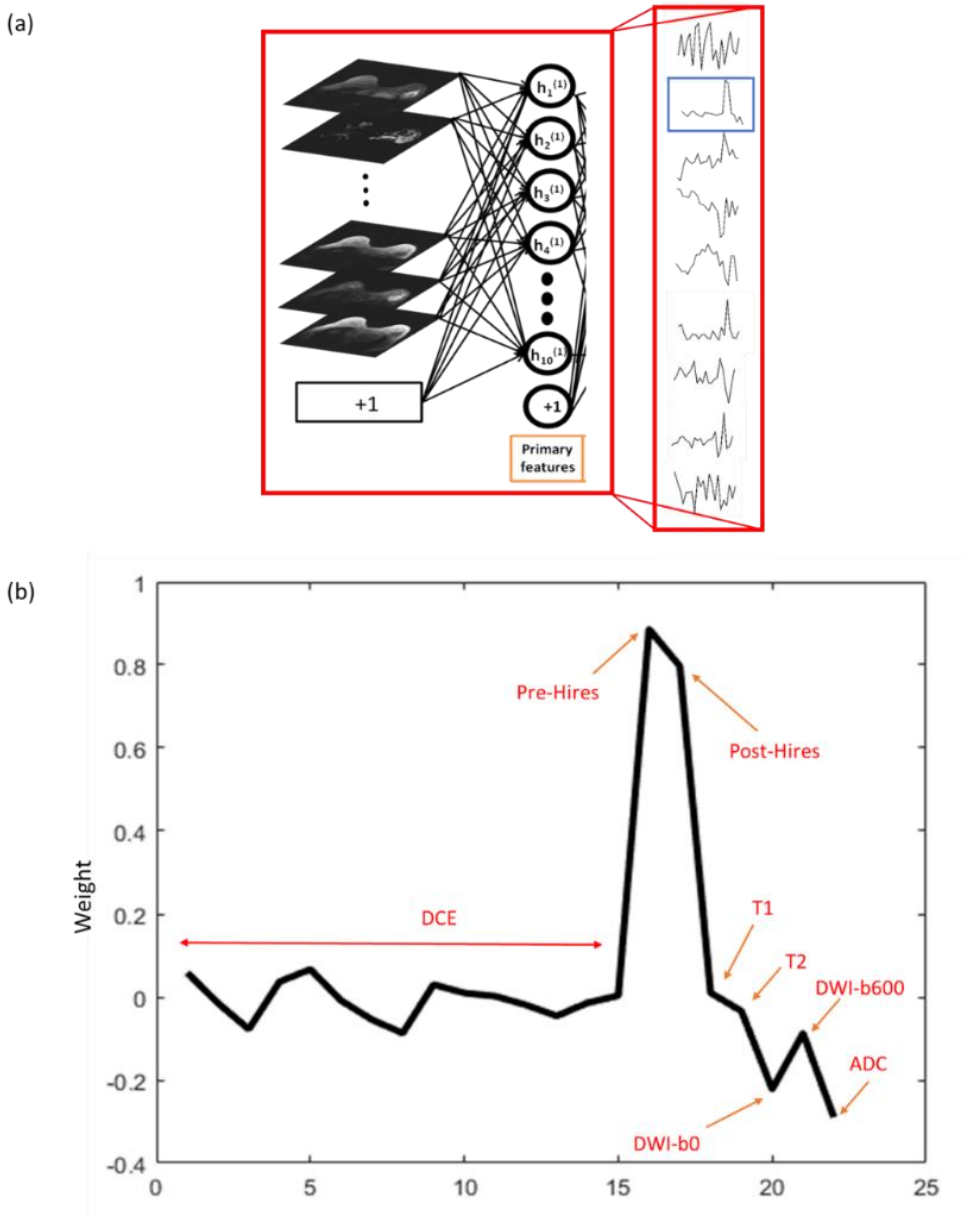
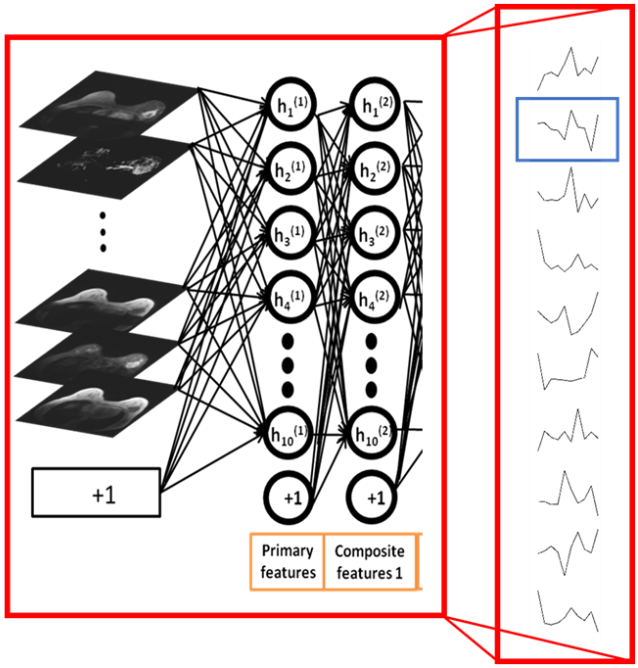


Figure 3.3 First Layer Decomposition of the MPDL neural network. (a) Expansion of the first layer nodes into their corresponding weight vectors (b) An example weight vector for the node selected in "a". The amount of weight assigned to each MRI parameters by the node is shown here. This particular node will get activated if the input tissue signature has high intensity values for high-resolution DCE images. This node is not a specialized node and will get activated for both glandular tissue and lesion tissue, but not fatty tissue.

(a)



(b)

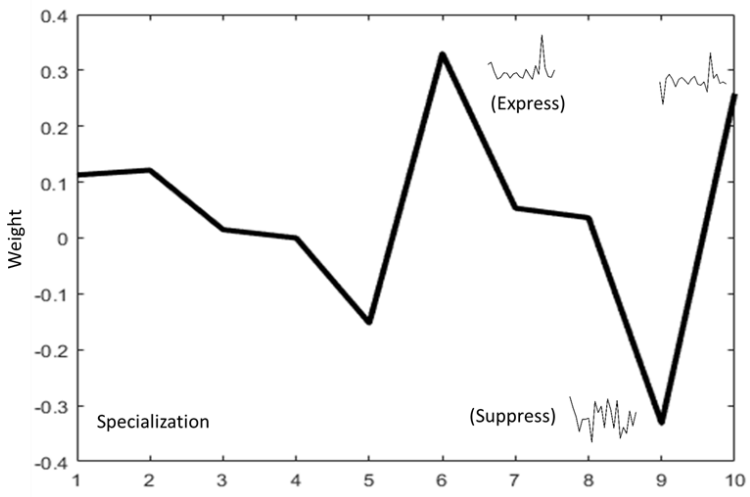


Figure 3.4 Second Layer Decomposition of the MPDL neural network. (a) Expansion of the second layer nodes into their corresponding weight vectors (b) An example weight vector for the node selected in "a". The amount of weight assigned to each of the previous layer nodes is shown here. This particular node produces a more specialized activation by selecting which activations to suppress and which to express.

3.3.2 Input Reconstruction from the deep and the classification layers of the MPDL tissue signature model

The weight vectors from deeper layers act as filters to produce a pure tissue signature representation that activates only for one particular tissue type as shown in **Figure 3.5.** **Figure 3.5** illustrates how the network learns intrinsic relationships between different MRI parameters and leverages that information to produce accurate tissue segmentations. Approximate linear reconstruction of the tissue signatures for three different tissue types (fat, glandular and lesion tissue) was generated from the softmax layer output of the SSAE for an example case as shown in **Figure 3.6.** The visualization in **Figure 3.6** gives an idea of how the SSAE network "sees" different tissue signatures.

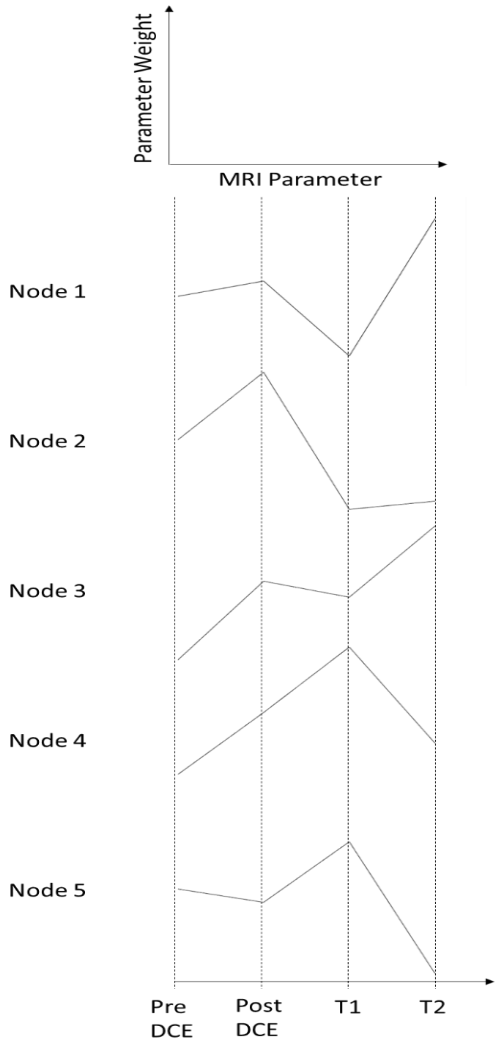


Figure 3.5 Illustration of the fifth layer encoding for a subset of MRI parameters: Pre- and Post- contrast DCE, T1, and T2. Here Y-axis represent the weights or importance assigned to each of the parameters by the fifth layer nodes with MRI parameters on the X-axis. We can see that the network learns the relationship between T1 and T2 (nodes 1 and 4) which is useful in distinguishing fat from glandular and lesion tissue. Similarly, the network learns the relationship between pre- and post- DCE (nodes 1-3) which is useful in distinguishing lesion from glandular tissue. The network learned these relationships without any input of prior knowledge, hand crafted features or domain expertise.

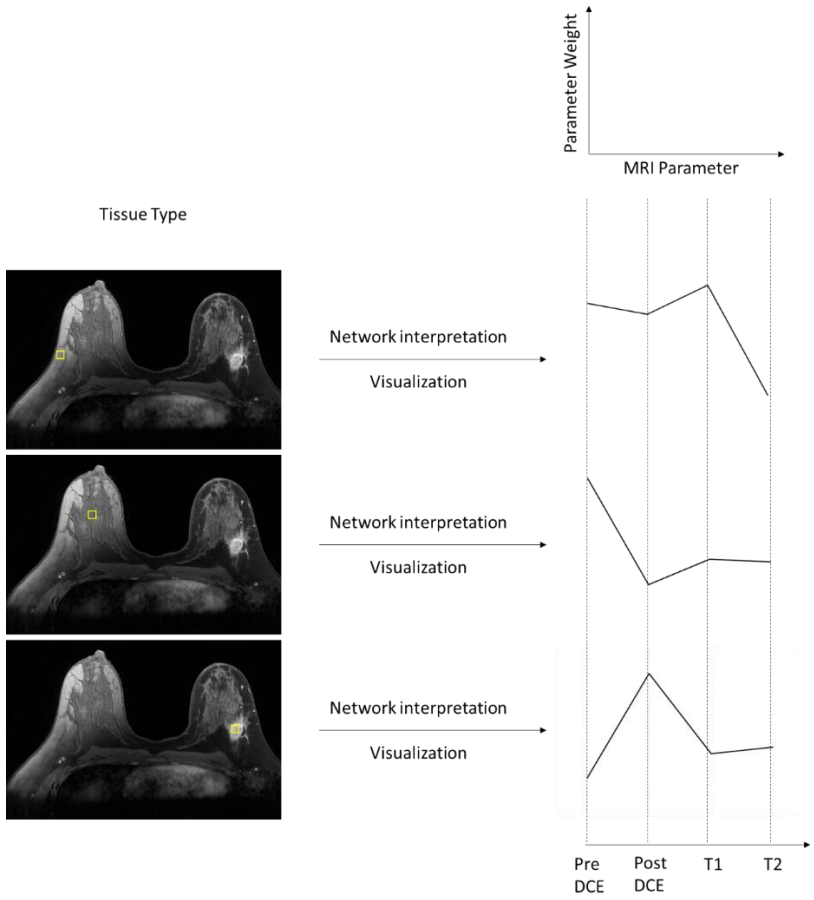


Figure 3.6 Visualization of MPDL reconstructed tissue signatures to understand how the network sees the input tissue signatures for fatty, glandular and lesion tissues for an example patient. The ROIs have been highlighted using a yellow box on each of the images.

3.3.3 Weight vector visualization from lesion tissue signatures

Kinetic curve analysis characterizes the wash-out pattern of gadolinium contrast agent into three categories. The three curve types are type I: persistently enhancing curve, type II: plateau, and type III: washout type. Type I is suggestive of benign tumor while type III is suggestive of malignancy and type II has an intermediate probability of malignancy. For comparison, I trained the unsupervised SAE algorithm with correlation reduction on the lesion

DCE-MRI tissue signatures obtained from benign and malignant patients to automatically extract kinetic curve patterns from DCE-MRI. **Figure 3.7a** demonstrates the different kinetic curve patterns discovered by the unsupervised SAE trained on the lesion tissue signatures of benign and malignant patients. The SAE was able to discern the DCE curve types from just the multiparametric imaging data without any prior knowledge of the underlying tissue type. The SAE curves unique to benign and malignant patients were obtained from back-projection of these curves on to tumors using the SAE decoder as illustrated in **Figures 3.7b and c**. The SAE encoding unique to benign tumors had significantly higher type I curves (66% for benign vs. 49% for malignant patients, $p = 0.01$) while the SAE encoding unique to malignant tumors had significantly higher type III curves (34% for benign vs. 51% for malignant patients, $p = 0.01$).

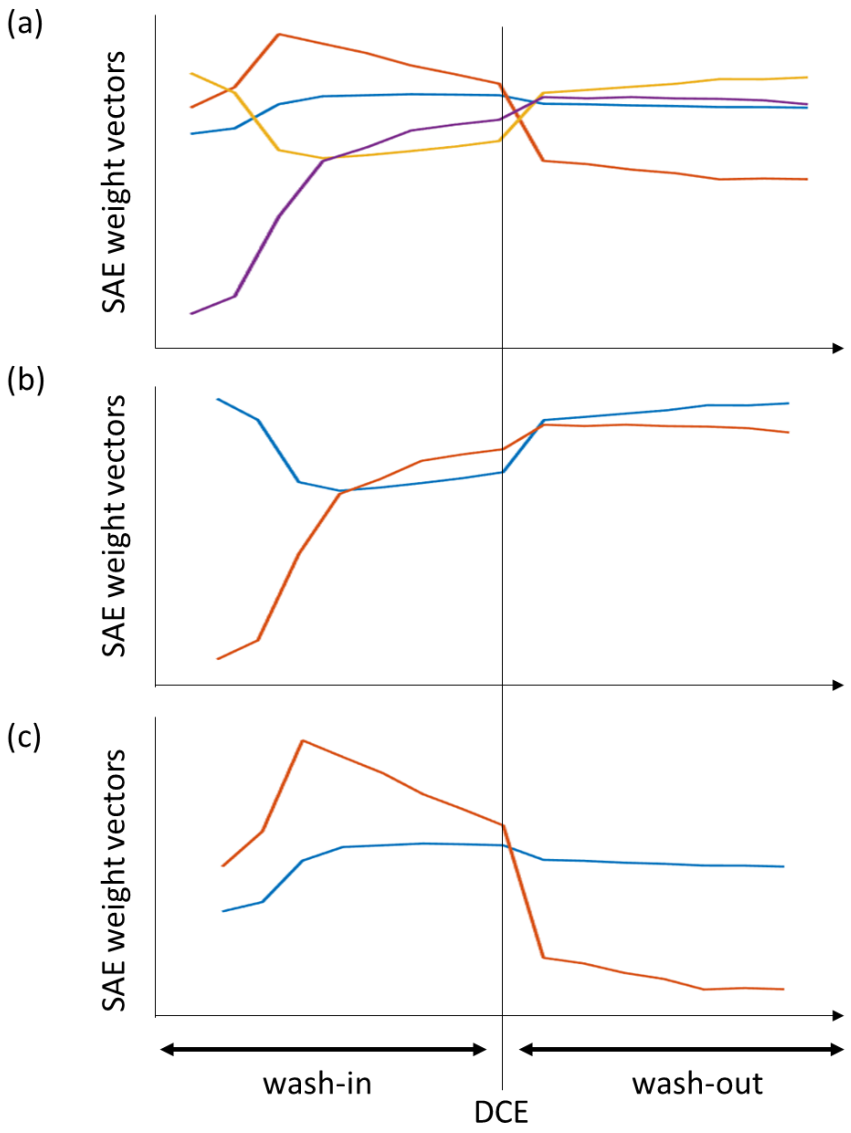


Figure 3.7(a) Illustration of DCE kinetic curve patterns discovered by the correlation reduction algorithm. The algorithm discovered four unique DCE kinetic curve patterns in breast tumors (b) SAE kinetic curve patterns that were unique to benign patients ($p = 0.01$) (c) SAE kinetic curve patterns that were unique to malignant patients ($p = 0.01$)

3.4 Discussion

I developed and tested multiple different approaches for visualization and analysis of the multiparametric deep learning (MPDL) tissue signature model developed in **Chapter 2**. The weight vector visualization and input reconstruction gave us valuable insights into how the MPDL model "looked" at the data and the intrinsic representation it formed within its network architecture.

The MPDL model learned relationships such as "Fat is bright on T1 and dark on T2" and "lesion is bright on post-contrast DCE and dark on pre-contrast DCE", which are very similar to radiologists. These results are very encouraging as the model was not trained with any prior knowledge on how these tissues appear on different radiological images. Furthermore, I was able to decode the intrinsic representations developed at each layer and how they relate to each other. The intrinsic representations got increasingly specific as the depth of the MPDL model increased. The first layer nodes could not distinguish between glandular tissue and lesion tissue whereas the final layer nodes had developed specific activation patterns that could distinguish between all types of tissue.

The unsupervised SAE with correlation reduction produced equally spectacular results on the tumor tissue signatures. Using only the database of tissue signatures, the SAE model was able to identify all the three types of DCE kinetic curves and their relationships with benign and malignant lesions. This demonstrates the potential of unsupervised artificial neural networks in identifying patterns from high-dimensional data with comparable performance to humans. Moreover, this report demonstrates that the SAE can detect heterogeneous zones within breast lesions. These heterogeneous regions can be used for further classification of breast tissue by quantitative ADC maps and/or PK-DCE

parameters. Finally, these results open up the possibility of deploying unsupervised neural networks across multiple different organs and pathologies for discovering patterns in high-dimensional radiological data.

In conclusion, "opening the MPDL black box" gave us valuable insights into how the MPDL organizes, integrates, and interprets the radiological imaging information and how we could use these insights to develop more general models for tissue segmentation and pattern discovery.

Chapter 4. Organ Invariant Tissue Segmentation: Preliminary Results

4.1 Introduction

In **Chapter 2**, the MPDL model for multiparametric radiological image segmentation was developed, validated, and tested for application to segmentation of different tissue types from multiparametric breast MRI. While the MPDL model resolves the problem of segmentation of different tissues from a multiparametric radiological imaging dataset, it is very narrow in that it is limited to segmentation of tissue from a specific organ. On the other hand, if we look at deep learning models developed for computer vision applications, they are more general. For example, object recognition algorithms based on deep learning are not limited to recognition of sub-groups of objects (e.g., vehicles, food, and furniture), but are generalized to detect all kinds of objects, irrespective of their type, scale, or origin [36]. The true potential of artificial intelligence in radiological applications can be achieved when the model is generalizable to multiple different tasks such as segmentation of all kinds of tissue types from any organ within the body.

Multiparametric Magnetic Resonance Imaging (mpMRI) is an ideal platform for development of a general AI because of its unique ability to form images influenced by different types of tissue parameters (i.e., proton density, T2, T1, diffusion, and perfusion weighted imaging (WI)). Each of these parameters interacts in a complex multidimensional space providing complementary information about the underlying tissue properties. The mpMRI complex interactions are based on the underlying MRI physics and remain

consistent across multiple applications. For example, the fluid is dark on T1WI and bright on T2WI irrespective of the application.

The consistency in complex interactions between different MRI parameters across different organs and pathologies provides an opportunity to model them into a single deep learning model. To that end, I utilized the intrinsic representations formed by the MPDL network architecture (**Chapter 2**) and evaluated their consistency across different organs for a generalized segmentation architecture. This is a feasibility study that investigates the stability of the MPDL tissue signature model and its capability in segmenting different tissue classes across different organs, in this case, breast and brain.

Section 4.2 describes the clinical brain data (brain tumor and stroke) acquired for this study and their corresponding acquisition protocols. The results from training the MPDL on breast data and testing the trained model on the brain data and vice versa have been detailed in **Section 4.3**. In the final section, I discuss the results from this study and the possibility of a universal deep learning framework for multi-organ, multi-tissue segmentation.

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

I designed the experiment for testing organ invariant tissue segmentation using MPDL, did the entire implementation, and data analysis.

4.2 Materials and Methods

4.2.1 Clinical Stroke Data

Five patients were part of this study (4 women and 1 man, age=70±19 years). Three of these patients with ischemic stroke underwent an acute stroke MRI protocol at different time-points divided into acute (0-12 hrs) and sub-acute (24-168 hrs) time-points. The other

two patients were also imaged at chronic timepoint (>168 hrs). The sample set consisted of 12 total studies (five acute, five sub-acute and two chronic time-points).

Multiparametric clinical stroke MRI imaging protocol

All patients were imaged on 1.5T clinical MRI system using a phased array head coil. The MRI parameters were: T1WI sagittal MPRAGE image ($TR/TE = 200/2.46ms$, $field\ of\ view\ (FOV) = 24\ cm \times 24\ cm$, $slice\ thickness\ (ST) = 5\ mm$), axial T2WI FLAIR ($TR/TE/TI = 9000/105/2500ms$, $FOV = 17.3\ cm \times 23\ cm$, $ST = 4\ mm$), axial DWI ($TR/TE = 9000/98ms$, $b - values = 1000\ and\ 0\ s/mm2$, $FOV = 23\ cm \times 23cm$, $matrix = 128 \times 128$, $ST = 4\ mm$) and echo planar T1WI perfusion ($TR/TE = 1350/30ms$, $FOV = 23\ cm \times 23cm$, $ST = 4\ mm$, total duration = 90 seconds). The contrast agent GdDTPA (Magnevist) was power injected at a dose of 0.1 mmol/kg and at a rate of 5 cc/sec.

4.2.2 Clinical Brain Tumor Data

Nine patients with grade IV glioblastoma (5 males and 4 females) were part of this study, with an average age of 59±13 years. MR images were obtained using a 3.0 Tesla Siemens Trio Tim system (Siemens Medical Solutions, Erlangen, Germany) with a 12-channel head matrix coil. Structural images included a 3D T1 MPRAGE sequence ($TR = 2300\ ms$, $TI = 900\ ms$, $TE = 3.5\ ms$, $flip\ angle\ (FA) = 9°$, $field\ of\ view\ (FOV) = 24\ cm \times 24\ cm$, $matrix = 256 \times 256$, $slice\ thickness\ (ST) = 1\ mm$) and a 2D T2-FLAIR axial sequence ($TR = 9310\ ms$, $TI = 2500\ ms$, $TE = 116\ ms$, $FA = 141°$, $FOV = 24\ cm \times 24\ cm$, $matrix = 320 \times 240$, $ST = 3\ mm$).

4.2.3 SSAE Evaluation and Statistical Analysis

The dice similarity metric between SSAE and EI segmentations was applied to evaluate the efficacy of SSAE segmentations on brain cancer and breast cancer mpMRI datasets. For stroke, I compared the SSAE segmentations with the EI segmentations at all three time-points (acute, sub-acute, and chronic) for both clinical and pre-clinical stroke mpMRI. The percentage difference in the size of lesion areas segmented from the SSAE and EI algorithms were computed. Similarly, the percentage differences in ADC values and TTP values of the areas segmented from the two algorithms were computed. Nonparametric Spearman rank correlations were performed between the SSAE maps and the segmented areas, quantitative ADC, and TTP values. The Bland-Altman technique was used to identify any systematic differences or biases between the EI and SSAE maps by plotting the difference versus results from EI algorithm [75]. Statistical significance was set at $p \leq 0.05$.

4.3 Experiments

4.3.1 Brain Tumor Segmentation

The grade IV brain tumors have tissue signatures that are very similar to breast tumors, in that, the dynamic contrast enhanced images demonstrate increased uptake in the tumor tissue. In addition, the fluid tissue signatures from glandular tissue in the breast should be adaptable to fluid tissue signature from cerebrospinal fluid (CSF). The MPDL trained on breast mpMRI (**Chapter 2**) was tested on brain mpMRI for segmentation of different tissue in the brain. The MRI parameters used were DCE-Pre, DCE-Post, and T2WI, which were acquired for both brain and breast mpMRI.

Furthermore, if the tissue signatures learned from breast are adaptable and transferrable to brain mpMRI, then the opposite should also be true. To test this hypothesis, I trained the MPDL network with same parameters on brain mpMRI and tested it on an example set of ten breast mpMRI cases (five benign and five malignant) for segmentation of different tissue in the breast.

Figure 4.1 demonstrates the results of MPDL trained on breast data for segmentation of brain tumor and vice versa for two example brain mpMRI and two example breast mpMRI datasets. The segmentations have been overlaid with a synthetic color scale ranging from blue to red, with each color representing a different tissue type as indicated in the color scale shown in **Figure 4.1**. The breast-trained MPDL segmented eight of nine brain mpMRI datasets with excellent dice similarity of 0.86 ± 0.09, but failed to segment the tumor on one dataset, shown in **Figure 4.2**.

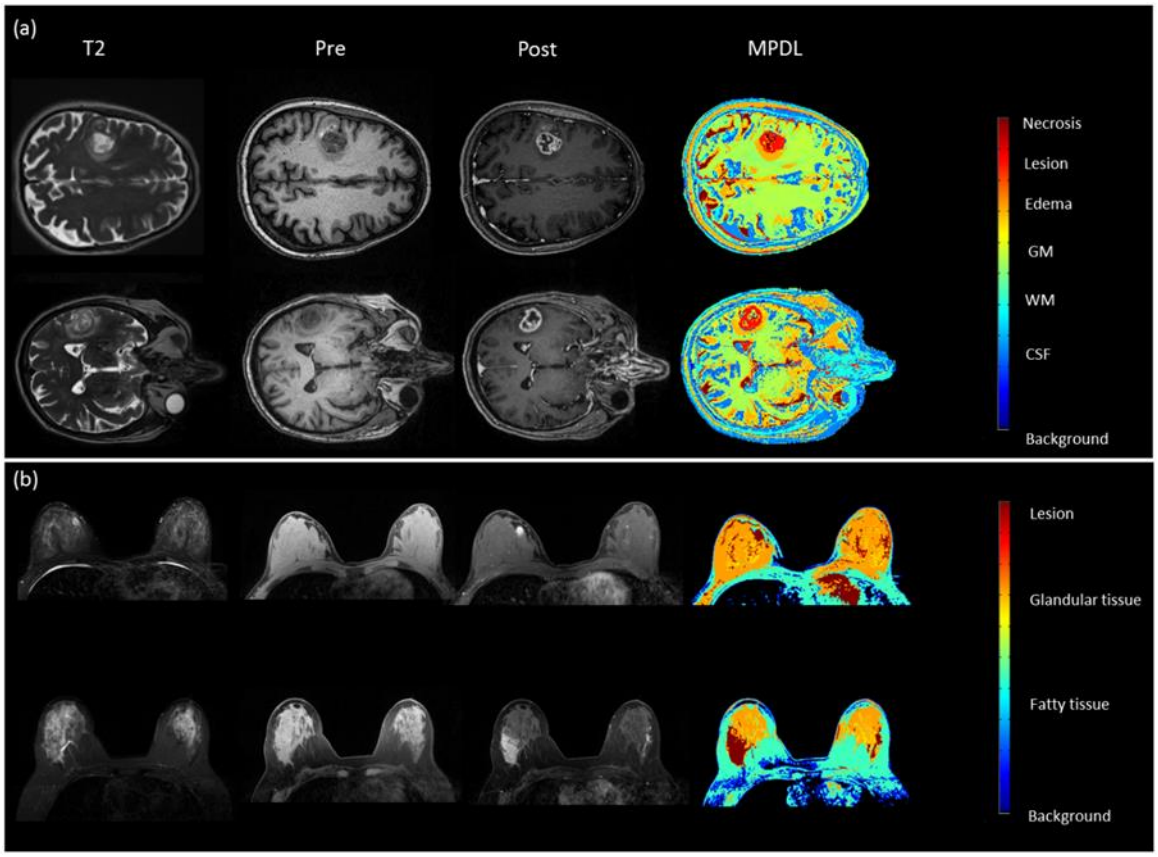


Figure 4.1 (a) Tissue segmentation of two example brain multiparametric MRI (mpMRI) datasets using the breast trained multiparametric deep learning (MPDL) model. (b) Tissue segmentation of two example breast mpMRI datasets using the brain trained MPDL model. The synthetic color scale to highlight different tissue types has been shown to the right of the two datasets.

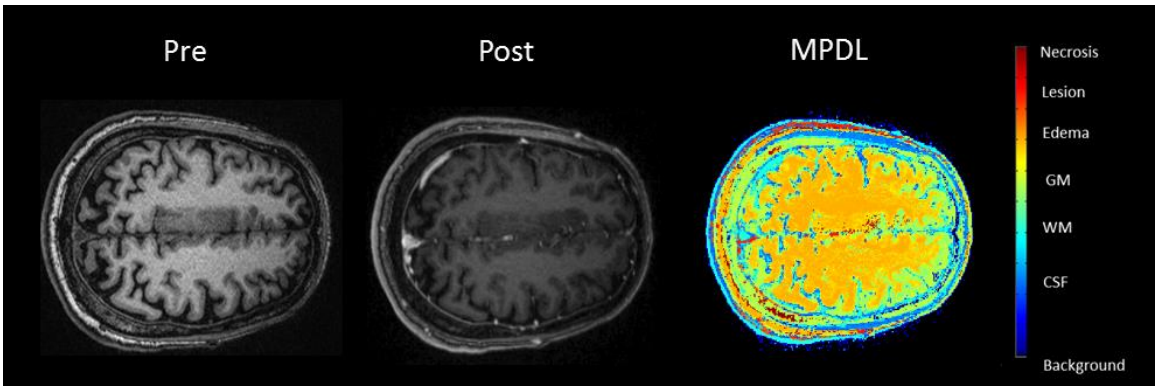


Figure 4.2 Tissue segmentation of the brain multiparametric MRI (mpMRI) dataset for which the breast-trained network failed to segment the tumor. There was no contrast enhancement in the tumor area due to which the MPDL failed to identify the tumor tissue correctly.

4.3.2 Brain Stroke Segmentation

For the clinical stroke mpMRI dataset, MRI parameters of perfusion-, diffusion-, T1-, and T2-weighted imaging (PWI, DWI, T1WI, and T2WI, respectively) were obtained from the stroke patients. There was a difference between the perfusion images obtained from breast and brain. Contrast wash-in and wash-out from the brain PWI is significantly faster than for breast DCE-MRI. As a result, the phase resolution of brain PWI was 50-60 images compared to the 15 images for breast DCE-MRI. The two dataset tissue signatures were matched by moving a sliding time window across the brain perfusion data to capture the appropriate time interval and then reslicing the imaging series within the sliding window to match the phase resolution of the breast DCE-MRI using wavelet decomposition [61]. The resliced PWI was combined with the remaining MR parameters to form the stroke MPDL tissue signature of the same size as that of breast tissue signatures. The MPDL

trained on breast mpMRI was then tested on the stroke MPDL tissue signatures obtained from all the patients.

The SSAE segmentation maps corresponding to the MPDL algorithm trained on breast mpMRI (**Chapter 2**) and tested on the independent clinical stroke data are shown in **Figure 4.3.** The segmentations were overlaid with a synthetic jet colormap such that red represented ischemic tissue and blue represented normal tissue.

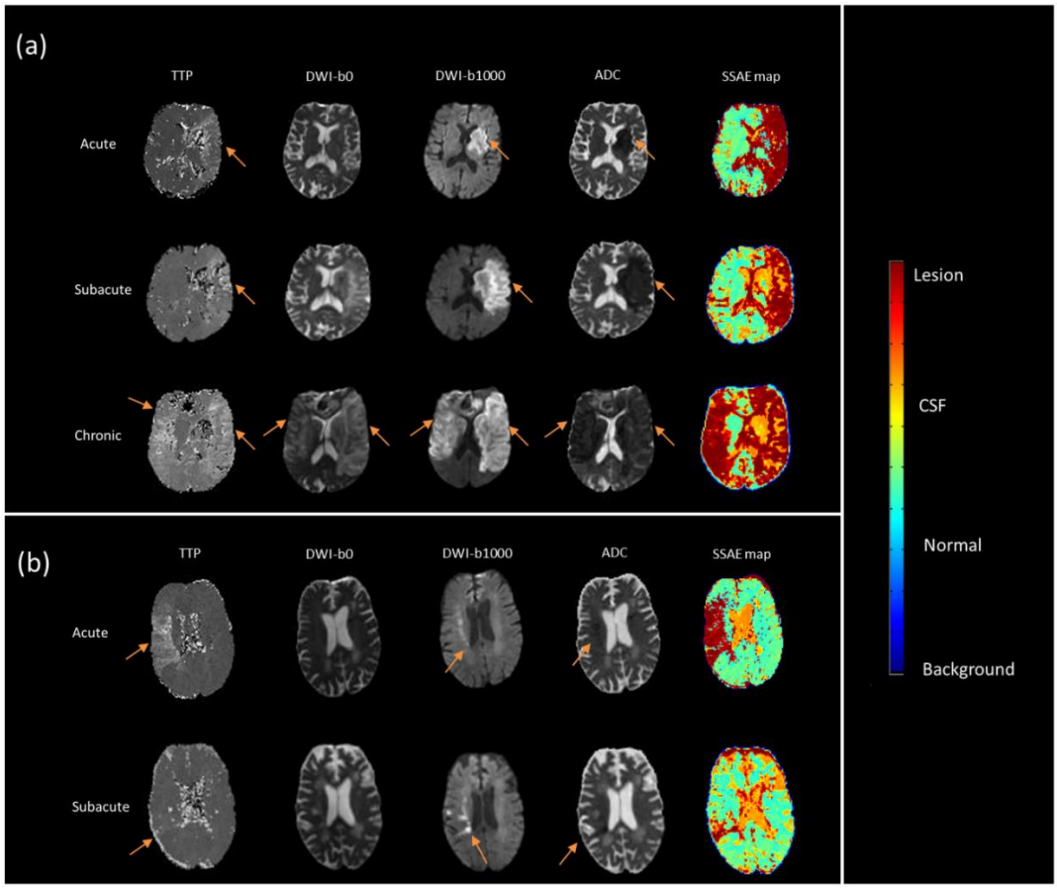


Figure 4.3. Illustration of the result of stacked sparse autoencoder (SSAE) network trained on axial breast mpMRI and applied to clinical stroke mpMRI in two representative stroke patients at different time-points after stroke. The color coding is shown to the right of the images.

Table 4.1 summarizes the areas segmented, ADC values, and TTP values of the clinical stroke lesion defined by the EI and SSAE maps for all the patients at different time-points. **Figure 4.4** illustrates the Bland Altman plots for the ischemic area along with quantitative ADC and TTP values characterized by the segmentations from the EI and SSAE maps. The Bland Altman plots show there was a good agreement between the ischemic areas segmented from the EI and SSAE. Furthermore, there was a high correlation between the quantitative metrics obtained from EI and SSAE segmentations ($areas$: $R = 0.99, 95\%CI = 0.99 - 1.0,\ slope = 1.08 \pm 0.016,\ p < 0.0001$; ADC: $R = 0.86,\ 95\%CI = 0.56 - 0.96,\ slope = 1.04 \pm 0.19,\ p = 0.0004$; TTP: $R = 0.99,\ 95\%CI = 0.98 - 1.0,\ slope = 0.99 \pm 0.04,\ p < 0.0001$). Finally, the percentage difference in the stroke areas, TTP and ADC values segmented from EI and MPDL were $6 \pm 2\%$, $3 \pm 3\%$, and $7 \pm 7\%$, respectively.

Table 4.1. Clinical stroke data. Ischemic areas and the quantitative values of apparent diffusion coefficient (ADC) and time-to-peak (TTP) segmented by the EI and SSAE algorithms.

Patient	**Timepoint**	**SSAE Area (mm^2)**	**EI Area (mm^2)**	**EI TTP (s)**	**SSAE TTP (s)**	**EI ADC (x 10^{-3} mm^2/sec)**	**SSAE ADC (x 10^{-3} mm^2/sec)**
Patient 1	Acute	19.44	19.14	10.92	11.01	659.14	626.94
	Sub-acute	37.35	36.12	12.28	12.31	522.87	547.06
	Chronic	87.21	81.15	13.44	13.79	600.70	606.96
Patient 2	Acute	16.98	18.39	12.73	12.57	776.58	798.50
	Sub-acute	4.17	4.41	10.78	10.85	733.32	792.61
Patient 3	Acute	5.16	5.58	8.73	8.18	796.83	1035.30
	Sub-acute	8.85	7.95	8.22	8.41	550.00	599.13
Patient 4	Acute	12.63	13.53	6.98	7.68	716.65	770.21
	Sub-acute	11.79	12.48	7.09	6.80	582.72	673.83
Patient 5	Acute	15.66	14.82	14.64	14.88	821.58	791.05
	Sub-acute	13.74	14.31	14.71	14.54	821.61	842.31
	Chronic	22.71	24.06	12.86	12.94	832.65	841.38

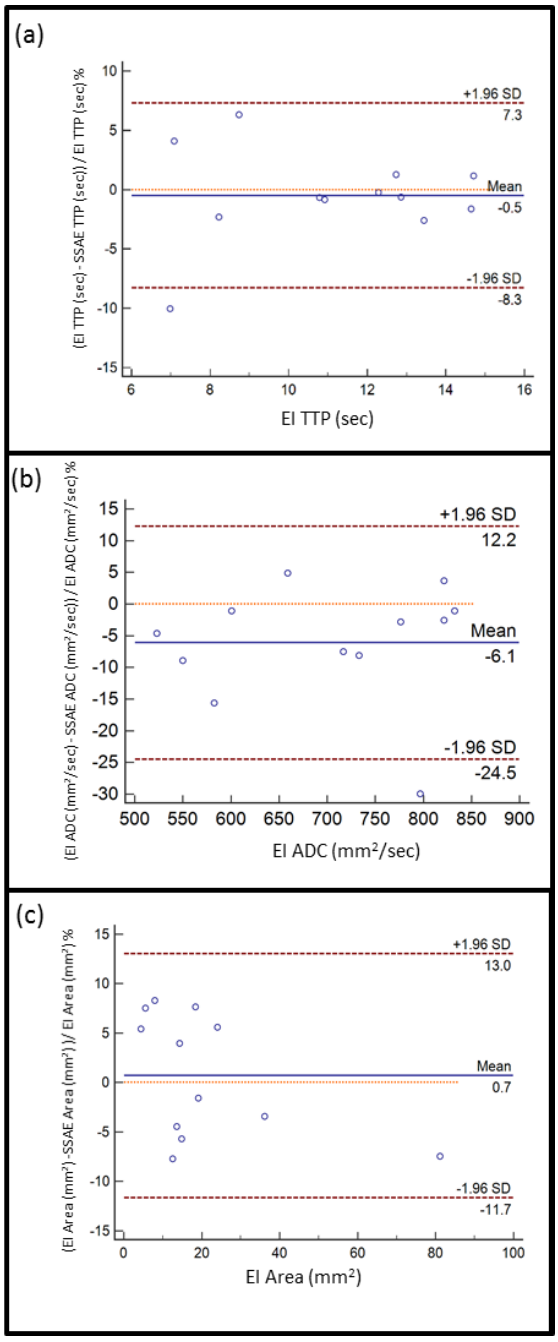


Figure 4.4. Bland Altman plots for comparison between the Eigenfilter (EI) segmented stroke lesions and stacked sparse autoencoder (SSAE) segmented stroke lesions. (a) Time to peak (TTP). (b) Apparent diffusion coefficient values. (c) Lesion segmented areas.

4.4 Discussion

The MPDL method based on an SSAE produced encouraging results for segmentation of different tissue types in breast tumor mpMRI, brain cancer mpMRI, and brain stroke mpMRI. The knowledge transfer between the MPDL trained on breast cancer and applied to brain cancer and stroke was excellent. Furthermore, the inverse knowledge transfer from brain to breast was also very encouraging. This feasibility study demonstrates the generalizability of the tissue signatures across different organs and the potential of a general AI that could be used to segment different tissue across multiple organs.

Deep learning has previously been successfully applied to segment and classify tissue types from brain, breast, prostate, and other organs [52]. In this study, I demonstrated the feasibility of deep learning in producing organ invariant tissue segmentations. The MPDL tissue signature model which was introduced in **Chapter 2** was able to segment different tissue types from pathologies (stroke) that the MPDL had never "seen" before. Moreover, the segmentations were irrespective of magnetic field strength (3T for breast data and 1.5T for clinical stroke data).

This study successfully demonstrated the capability of deep learning algorithms to learn the underlying MRI inter-parameter relationships. For example, the T1 and T2 of the breast and brain data were similar. The DWI and ADC mapping were the same, except for the number of b values. The MPDL method was able to adapt and apply the correct signatures. The biological composition of glandular tissue is fluid which appears bright on T2 and dark on T1. Similarly, for brain data, CSF presents the same MRI characteristic as it is also fluid. The MPDL was able to classify CSF as the same group as glandular tissue. Similarly, white matter (WM) in brain and fat in breast exhibit the same MRI

characteristics (bright on T1WI and dark on T2WI). As a result, both WM and fat result in the same class.

There were limitations to this study. This was a preliminary study to assess the possibility of organ invariant tissue segmentation using multiparametric MRI. The results from this study were encouraging but not comparable to the state-of-the-art results as demonstrated in my results from **Chapter 2** and other results reviewed in [52]. This is because the tissue signatures obtained from different organs have intrinsic properties that were not captured in our model. This could potentially be resolved by using all the tissue signatures from all the organs during the SSAE pre-training step. In the future, I plan to experiment with different parametrizations and inputs to the SSAE and optimize the organ invariant tissue segmentation workflow.

In conclusion, the MPDL method reveals the possibility of a universal deep learning model that can learn inter-parametric relationships corresponding to different tissue types across multiple organs based on the underlying MRI physics. In the future, I aim to develop a universal MPDL model which could automatically identify the underlying pathology and segment out the tissue accordingly for any organ or pathology.

Part II: Feature Extraction

Part I introduced the problem of image segmentation and my contributions to the segmentation of different tissues from multiparametric radiological imaging datasets. The MPDL algorithm developed in **Chapter 2** produced a segmentation map highlighting the different tissue types and their corresponding segmentation masks. The next step involves extraction of all the relevant information from the tissue of interest (TOI), for example, lesion or glandular tissue in breast. This section focuses on development of techniques for feature extraction from multiparametric radiological imaging datasets.

Feature extraction from multiparametric radiological images can be broadly categorized into two categories: inter-voxel analysis and inter-parametric analysis. I developed the methods of radiomic feature mapping and contribution scattergram for inter-voxel and inter-parametric analysis, respectively. The goal of radiomics is to capture the spatial appearance of the TOI (shape and texture) on each of the acquired images. On the other hand, the contribution scattergram characterizes how the images acquired using different imaging parameters relate to each other by modeling their relationships as a complex network model.

The majority of the feature extraction models developed in the literature focus on radiomics based approaches. As a result, I have devoted one complete chapter (**Chapter 5**) discussing the field of radiomics and the various applications of radiomics. My contributions to the field of radiomics have been detailed in **Chapters 6 and 7**. The final chapter (**Chapter 8**) of this section describes the technique of contribution scattergram and validates it on a multiparametric breast MRI dataset.

Chapter 5. A review on radiomics and its application to precision radiology

5.1 Introduction

Radiological imaging techniques are powerful noninvasive tools used for the detection, differentiation, and diagnosis of different tissue characteristics in patients. These imaging methods include X-Ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Nuclear Medicine (NM) Positron Emission Tomography (PET), and Ultrasound (US). Each of these modalities creates different tissue contrast based on whether the tissue is normal or abnormal. These different tissue contrast mechanisms are exploited by the radiologist to identify patterns to reach a diagnosis. However, each of the radiological images contains more information content not visible to the clinician's eye and this "hidden" information creates a "radiological texture" which can provide much more information about the tissue of interest than previously thought. Thus, radiomics was introduced as a "new" method to discover and translate the metrics obtained by using texture and other analysis techniques on radiological images. However, Radiomics is a new application using established techniques, specifically, entropy, mutual information, gray level co-occurrence matrix, and other features [24, 26, 27, 29]. The novel aspect of radiomics is the comparison of these measures to surrogate endpoints with the advent of increased computational power available today. The main idea is the information "hidden" within the radiological images can be extracted using texture feature analysis.

Intuitively, texture is used by everyone to define objects either visually or by touch. The terms commonly used to describe texture are roughness, smoothness, and coarseness

but for digital data, analytical methods were needed to better "describe" the information content in digital objects. Therefore, texture analysis, in part, is based on information theory developed in 1948 by Claude Shannon and extended by Haralick, Galloway, and others that incorporated different statistical measures to include gray level matrix operations on the input data [24, 26, 27, 29]. Historically, one of the first applications of texture analysis was in analyzing aerial photographs [76] then later applied to medical and other images [26, 77]. Currently, texture feature analysis is used in many different areas of research, especially, with increased computational power and digital storage capacity [78-81]. Moreover, given a region of interest, shape based features can also be extracted along with texture analysis [30].

Section 5.2 introduces the fundamental feature extraction methods followed by a review of historical and recent applications of radiomics techniques in detecting tumor characteristics in **Section 5.3**. This chapter concludes with an outlook on the potential future of radiomics in "personalized medicine" cancer therapies.

Publication from this work and author contributions

Publication

V. Parekh and M. A. Jacobs, "Radiomics: a new application from established techniques," Expert review of precision medicine and drug development, vol. 1, pp. 207-226, 2016.

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

5.2 Radiomics Feature Extraction

Radiomic feature extraction methods are based into primarily three categories consisting of statistical, filtering, and morphological features. In this chapter, I discuss the feature extraction methods from all the three categories; however, there are some other feature

extraction methods such as gray level size zone matrix based features and Minkowski functionals used in radiomics applications that have not been discussed here [17, 82].

5.2.1 Statistical texture features

Statistical features characterize the texture in an image using statistical measures. There are two levels of texture statistical methods, first and higher order as shown in **Figure 5.1**. These statistical radiomic methods have been applied to different imaging modalities used in several different diseases and are outlined below.

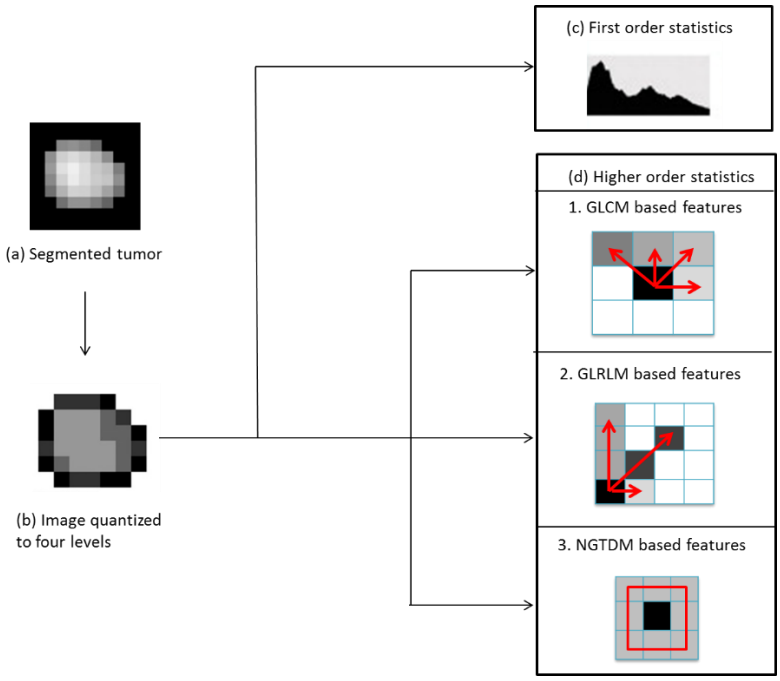


Figure 5.1 Illustration of statistical texture feature extraction. (a) Segmented tumor image (b) Segmented tumor image quantized to four intensity levels (c) First order statistical features corresponding to first order histogram (d) Higher order statistical features corresponding to 1. GLCM (gray level co-occurrence matrix), 2. GLRLM (gray level run length matrix) and 3. NGTDM (neighborhood gray tone difference matrix). Modified from [17].

a) First order texture statistics

First order texture statistics are based on the first order histogram that describes distribution of voxel intensities in an image. The normalized first order histogram (H) is computed by dividing the voxel intensities (I) in an image into B equally spaced bins and computing the proportion of voxels in each bin as defined by the equation.

$$H(i) = \frac{\#of\ pixels\ with\ gray\ levels\ in\ \{I \epsilon B_i\}}{\sum \#pixels\ in\ the\ image} \quad (1)$$

Entropy and uniformity are the two commonly used features computed using the histogram. Let H be the first order histogram with B bins, then the entropy [24] and the uniformity are given by the following equations:

$$Entropy = -K \sum_{i=1}^{B} H(i) \log H(i) \quad (2)$$

where K is a positive constant and is determined by the units of the application. However, since we are dealing with bits of information, $K = 1$ and the equation used in most digital applications is given as

$$Entropy = -\sum_{i=1}^{B} H(i) \log_2 H(i) \quad (3)$$

and

$$Uniformity = \sum_{i=1}^{B} H(i)^2 \quad (4)$$

Entropy measures the randomness in the gray level intensities of an image. The maximum value of entropy occurs when all the gray level intensities in an image occur with equal probability and is equal to $\log_2 B$. On the other hand, uniformity, as its name suggests, measures the uniformity of gray level intensities in an image. The maximum value of uniformity is one which occurs when all the pixels in an image have the same gray level intensity.

Both the first and second orders statistics can give different values based on the number of bins. The number of bins is a critical parameter defined by the user. If the number of bins is selected either too small or very large, then the histogram may not be able to correctly represent the underlying distribution. Several methods exist for determining the optimal number of bins without making any assumptions regarding the underlying data distribution [83, 84]. However, if different ROIs in a study have different number of bins in the first order histograms, then it is not possible to compare the results

directly. This is because number of bins, B, is an input parameter to both entropy and uniformity equations.

Another approach would be to use the same number of bins for all ROIs. Unfortunately, this leads to an issue of dependence of first order statistic features on the size of ROI. For example, if we consider two ROIs of sizes 128 and 1024 voxels and bin the gray level intensities within these two ROIs into 1024 bins (as we are using the same number of bins), then the entropy values of data distributions with maximum randomness (each voxel has a different intensity) would be equal to 7 and 10 for 128 voxel-sized and 1024 voxel-sized ROIs respectively. The issue of optimal binning is prominent for all statistical approaches as all the statistical approaches depend on the preprocessing step of image quantization. Thus, some form of feature normalization is required with respect to size in cases when the same number of bins is used for all the ROIs and with respect to number of bins when different number of bins is used for different ROIs. First order texture statistical features have been applied to a number of applications across different medical imaging modalities for characterization of biological tissue of interest [2-7, 17, 20, 30, 77, 85-121, 122]. **Table 5.1** summarizes the quantitative values of first order entropy for different organs and pathologies as reported in the literature. Depending on the organ being imaged and the imaging modality, the first order histograms may or may not have been the same across all the applications. However, a trend in the entropy values between pathological tissue and healthy tissue can be observed. The entropy values obtained from pathological tissue are relatively higher than the entropy values obtained from healthy tissues except for spinal cord and urinary bladder.

Table 5.1. Quantitative values of first order statistical entropy as reported in the literature. nMITR= normalized maximum intensity-time ratio, FA=fractional anisotropy and AD =Alzheimer's disease [1-9].

Reference	Tissue	Image/Object type	Pathology			
				Benign	Malignant	
Sinha et al. [1]	Breast	Normalized radial length of tumor		1.46	1.6	
Ertas et al. [2]	Breast	nMITR projection		5.4	6.77	
Suo et al. [9]	Urinary bladder	DWI, b=700		4.08	4.06	
	Urinary bladder	DWI, b=1500		3.97	3.78	
Kierans et al. [4]	Adnexal lesion	ADC		4.54	4.94	
				Low grade	High grade	
Ryu et al. [5]	Glioma	ADC		6.26	6.86	
				Volunteers	Healthy	
Brinkmann et al. [6]	Liver	MRS		7.6	4.7	
			Healthy adult	Elderly	CSM patients	
Cui et al. [7]	Spinal cord	FA map	6.07	6.01	5.32	
			Young	Middle aged	Non-demented elderly	Elderly (AD)
Chen et al. [3]	Brain	cortical surface structure	1.2	1.24	1.28	1.33

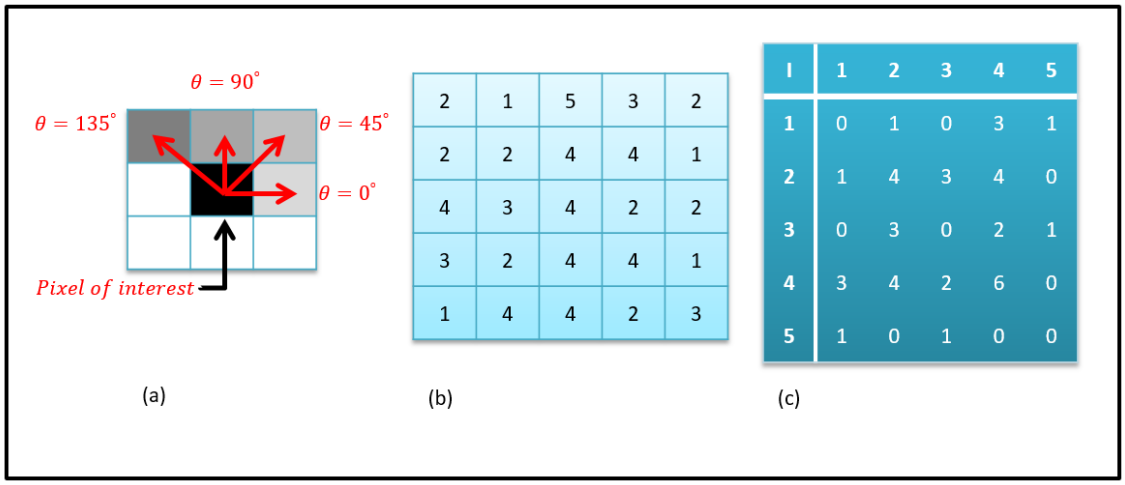
b) Higher order texture statistics

The features generated from first order statistics provide information about the distribution of voxel intensities in an image but they do not provide any information about the

inter-voxel relationships. The inter-voxel relationships in an image can be quantified using three different methods viz. gray level co-occurrence matrix, gray level run length matrix, and neighborhood gray tone difference matrix outlined below.

Gray level co-occurrence matrix

The spatial distribution of gray level intensities within an image can be extracted by using the gray level co-occurrence matrix (GLCM). The GLCM is constructed by considering the relationship between voxel pairs and the frequency of each intensity pairs within an image or a region of interest [26]. The relationship between voxel pairs is characterized by two user defined parameters, distance (d) and an angle (θ). **Figure 5.2a** illustrates the inter-voxel relationships characterized by the user defined parameter, θ. If the number of gray levels in the image is N_g, the number of possible voxel pairs would be $N_g \times N_g$. The frequency of each voxel pair in the image or a region of interest is computed and stored in the $N_g \times N_g$ GLCM matrix. The GLCM constructed is either symmetric [26] or asymmetric depending on the ordering of values in voxel pairs.



2	1	5	3	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

I	1	2	3	4	5
1	0	1	0	3	1
2	1	4	3	4	0
3	0	3	0	2	1
4	3	4	2	6	0
5	1	0	1	0	0

Figure 5.2 (a) Illustration of the inter-pixel relationships characterized by the user defined parameter, θ (b) An example 5 x 5 matrix with gray values ranging from 1 to 5. (c) The resultant symmetric gray level co-occurrence matrix (GLCM) obtained by multiplying the asymmetric GLCM with its transpose.

Mathematically GLCM of an image of size $N_x \times N_y$ with N_g gray levels is given by the following equation:

$$GLCM_d^{\theta}(i,j) = |\{((r,s),(t,v)): I(r,s) = i, I(t,v) = j\}| \; \forall \, i,j \in \{1,2,3,\dots,N_g\} \qquad (5)$$

where $(r,s),(t,v) \in Nx \times Ny$;

$$(t,v) = \begin{cases} r+d, s & if\ \theta = 0° \\ r+d, s+d & if\ \theta = 45° \\ r, s+d & if\ \theta = 90° \\ r-d, s+d & if\ \theta = 135° \end{cases}$$

I represents the image as a function I: $N_x \times N_y \rightarrow \{1,2,\dots,N_g\}$ and $|.|$ denotes the cardinality of a set.

The GLCM obtained from equation (5) is an asymmetric GLCM and can be converted to a symmetric GLCM by multiplying by the transpose ($Gsym = G \times G^T$).

The range of input parameters is $\theta\epsilon\{0°, 45°, 90°, 135°\}, d\,\epsilon\,\{1, 2, 3, \ldots, n\}$ for a two-dimensional image. For a three-dimensional image, the range for theta increases to 13 angles. An example 5×5 matrix with the corresponding symmetric GLCM for $\theta = 0°\ and\ d = 1$ is shown in **Figure 5.2**. Based on the different values of θ and d, a large number of GLCMs can be produced and for each GLCM, fourteen textural features were designed by Haralick et al. Subsequent research in texture analysis has led to construction of many more features, with most recent applications using twenty two GLCM based features [30].

Several studies have employed GLCM based features to determine benign form malignant lesions or treatment response and found no significant differences in changing the number of gray levels used for the features from the image [15].

There are twenty-two GLCM texture features [30]. The most commonly used features are given below:

$$ASM = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\left(G_{norm}(i,j)\right)^2 \tag{6}$$

$$Entropy = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} G_{norm}(i,j)\log_2(G_{norm}(i,j)) \tag{7}$$

$$Contrast = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}|i-j|^2 G_{norm}(i,j) \tag{8}$$

$$Correlation = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} ij(G_{norm}(i,j)-\mu_x(i)\mu_y(j)}{\sigma_x(i)\sigma_y(j)} \tag{9}$$

The notation used in the above equations is explained in **Table 5.2**.

Table 5.2. Notation used in the equations for computing texture features using gray level co-occurrence matrix.

G_{norm}	Normalized gray level co-occurrence matrix
G_x	Marginal probability matrix obtained by summing the rows of G_{norm}
G_y	Marginal probability matrix obtained by summing the columns of G_{norm}
μ_x	Mean of G_x
μ_y	Mean of G_y
σ_x	Standard deviation of G_x
σ_y	Standard deviation of G_y

The angular second moment (ASM) feature is the measure of uniformity in an image. A uniform image will have a few gray level transitions with the maximum uniformity of one representing only a single type of gray level transition through the image. Entropy feature, in contrast, increases as more and more different kinds of gray level transitions are present in an image. The maximum value of entropy is achieved when all the different gray level transitions in an image occur exactly the same number of times and is equal to $\log N_g^2$. Visually an image with higher entropy will appear more random or less homogeneous. Similarly, an image with higher contrast feature value will have higher frequency of large intensity differences between neighbors. An example high contrast image is a checkerboard. Finally, correlation measures the linear dependence between any two neighboring voxels throughout the image; i.e., an image with higher contrast will

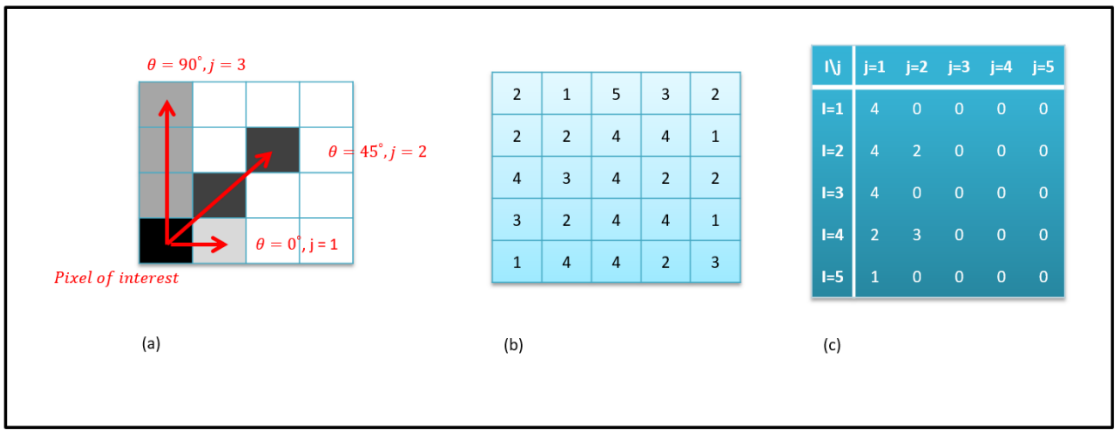
generally have lower correlation. Furthermore, texture features obtained from gray level co-occurrence matrix are not rotationally invariant. Rotationally invariant features are obtained by taking mean and range of texture values obtained in all directions (four for 2D images and thirteen for 3D volumes). Computation of rotationally invariant features for 3D is not straightforward because voxel spacing along the third dimension may not be same as voxel spacing in the other two dimensions. Thus, using the same value of d (GLCM input parameter) as voxel spacing in mm in all three dimensions is very important to achieve rotationally invariant texture features.

For many applications, GLCM based features are the most commonly used textural features applied to different medical imaging modalities for characterization of biological tissue of interest [1, 5, 13-17, 20, 30, 85-87, 91, 93, 95, 96, 99, 101, 103-105, 111-113, 116-121, 123-154]. These applications used symmetric GLCMs for texture analysis as defined by Haralick et al.; however, recent applications of GLCM based features have also used asymmetric GLCMs with some success [30]. Chen et al. compared the GLCMs generated from 2D slices vs. 3D volume on breast DCE-MRI and showed that 3D GLCM texture features performed significantly better than 2D GLCM texture features in classifying malignant from benign breast lesions [14].

Gray level run length matrix

The gray level run length matrix (GLRL) is defined as the number of contiguous voxels that have the same gray level value and it characterizes the gray level run lengths of different gray level intensities in any direction [27]. Elements (i, j) in the matrix represents the number of times, j, a gray level value, i, appears in the image. Based on the direction angle, θ (4 for two dimensions and 13 for three dimensions), different GLRL matrices can

be constructed. The inter-pixel relationship based on the user defined parameters of angle, θ and run length, j is illustrated in **Figure 5.3a.**



2	1	5	3	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

I\j	j=1	j=2	j=3	j=4	j=5
I=1	4	0	0	0	0
I=2	4	2	0	0	0
I=3	4	0	0	0	0
I=4	2	3	0	0	0
I=5	1	0	0	0	0

Figure 5.3 (a) Illustration of the inter-pixel relationships characterized by the user defined parameters, angle θ and run length j. (b) Example 5x5 matrix with values ranging from 1 to 5. (c) Resultant gray level run length matrix (GLRL) for run lengths of 1 to 5 and $\theta = 0°$.

Mathematically GLRL of an image of size $N_x \times N_y$ with N_g gray levels is given by the following equation:

$$GLRL_\theta(i,j) = |\{(m,n) : |\{(k,l) \in Nb(m,n,j,\theta) : I(k,l) = i\}| = j\}| \ \forall\, i,j \in \{1,2,3,\dots,N_g\} \qquad (10)$$

where $(m,n) \in Nx \times Ny$;

$$Nb(m,n,j,\theta) = \begin{cases} \{(m+1,n),(m+2,n),\dots,(m+j,n)\} & if\ \theta = 0° \\ \{(m+1,n+1),(m+2,n+2),\dots,(m+j,n+j)\} & if\ \theta = 45° \\ \{(m,n+1),(m,n+2),\dots,(m,n+j)\} & if\ \theta = 90° \\ \{(m-1,n+1),(m-2,n+2),\dots,(m-j,n+j)\} & if\ \theta = 135° \end{cases}$$

I represents the image as a function $: N_x \times N_y \rightarrow \{1,2,\ldots,N_g\}$ and $|.|$ denotes the cardinality of a set

Galloway et al. developed five features from the GLRL matrix viz. short runs emphasis (SRE), long runs emphasis (LRE), gray level nonuniformity (GLN), run length nonuniformity (RLN), and run percentage (RP) given by the following equations:

$$SRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{GLRL(i,j)}{j^2}}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}GLRL(i,j)} \quad (11)$$

$$LRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}j^2 GLRL(i,j)}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}GLRL(i,j)} \quad (12)$$

$$GLN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_r}GLRL(i,j)\right)^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}GLRL(i,j)} \quad (13)$$

$$RLN = \frac{\sum_{j=1}^{N_r}\left(\sum_{i=1}^{N_g}GLRL(i,j)\right)^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}GLRL(i,j)} \quad (14)$$

$$RP = \sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{GLRL(i,j)}{P} \quad (15)$$

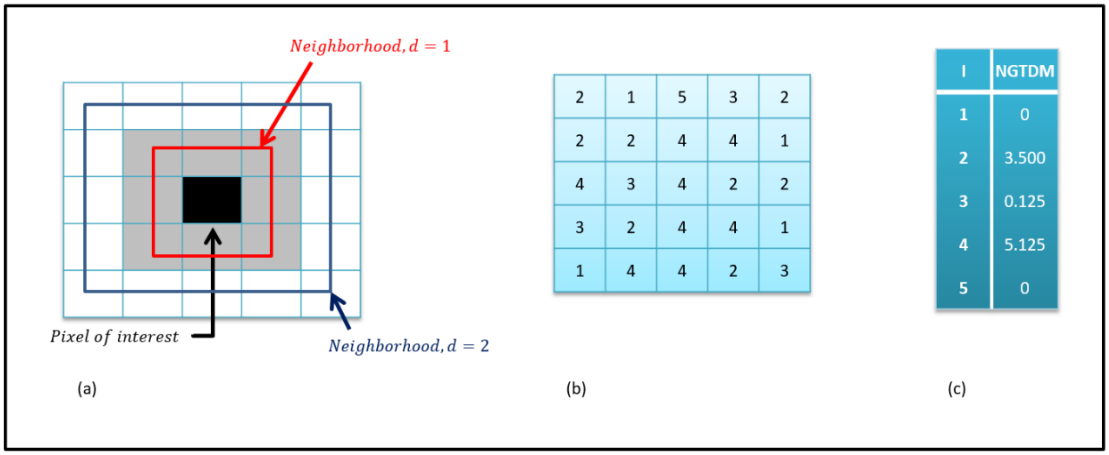
where N_r is the number of different run lengths measured to compute GLRL matrix and P is the total number of voxels in the image. A total of eleven features based on gray level run length matrices have been proposed in the literature [27, 155, 156]. Current applications in texture analysis use the complete set of eleven features derived from GLRL matrix [30].

An example 5x5 input matrix and the corresponding GLRL matrix for horizontal direction are shown in **Figures 5.3b** and **5.3c** respectively. The GLRL matrices are generally normalized (by the total number of entries in the matrix) in order to facilitate consistency between different scans of different patients, as well as different scans for same patient considered across a time interval. GLRL matrix based features have also been

applied to a number of applications across different medical imaging modalities for characterization of biological tissue of interest [17, 30, 85, 87, 93, 95, 96, 99, 101, 103, 105, 106, 113, 116-121, 125, 134].

Neighborhood gray tone difference matrix

The neighborhood gray tone difference matrix (NGTDM) is a texture analysis method based on the visual properties of an image [29]. Neighborhood gray tone difference matrix (NGTDM) is a one-dimensional matrix computed such that each gray level entry, defined as gt, in the NGTDM is the summation of the differences between all the pixels with gray level value, gt and the average gray level value of its neighborhood. The size of the neighborhood is defined by the user.



2	1	5	3	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

I	NGTDM
1	0
2	3.500
3	0.125
4	5.125
5	0

Figure 5.4 (a) Illustration of the neighborhood around the pixel of interest based on the user defined neighborhood parameter, d (b) Example 5x5 input matrix with values ranging from 1 to 5. (c) Neighborhood gray tone difference matrix for $d = 1$.

Mathematically, NGTDM of an image ($N_x \times N_y$) with N_g gray levels is given using the following set of equations:

$$ANGT(i,j) = \frac{1}{W-1}\left(\sum_{ik=-d}^{d}\sum_{jk=-d}^{d} I(i+ik, j+jk)\right), (ik, jk) \neq (0,0)$$

$$\forall\, i \in \{1,2,3,\dots,N_x\}\ and\ \ j \in \{1,2,3,\dots,N_y\} \tag{16}$$

$$NGTDM(gt) = \sum_{(i,j)\,|\,I(i,j)=gt} |gt - ANGT(i,j)| \quad \forall\, gt \in \{1,2,\dots,N_g\} \tag{17}$$

Here, $ANGT$ is the average neighborhood gray tone in the neighborhood of the pixel at position (i,j), $W = (2d+1)^2$ and $d\ =\ size\ of\ the\ neighborhood$ (e.g. 1 in case of 3 x 3 neighborhood). There are five features derived from the NGTDM and are given by the following equations:

$$Coarseness = \left[\epsilon + \sum_{i=1}^{N_g} P_i NGTDM(i)\right]^{-1} \tag{18}$$

$$Contrast = \left[\frac{1}{N_t(N_t-1)}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P_i P_j (i-j)^2\right]\left[\frac{1}{n^2}\sum_{i=1}^{N_g} NGTDM(i)\right] \tag{19}$$

$$Busyness = \frac{\left[\sum_{i=1}^{N_g} P_i NGTDM(i)\right]}{\left[\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} iP_i - jP_j\right]}, P_i \neq 0, P_j \neq 0 \tag{20}$$

$$Complexity = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\left\{\frac{|i-j|}{n^2(P_i+P_j)}\right\}\{P_i NGTDM(i) + P_j NGTDM(j)\}, P_i \neq 0, P_j \neq 0 \tag{21}$$

$$Texture\ strength = \frac{\left[\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(P_i+P_j)(i-j)^2\right]}{\left[\epsilon + \sum_{i=1}^{N_g} NGTDM(i)\right]} \tag{22}$$

where P_i is the probability of occurrence of a gray level value I, n is the number of entries in the NGTDM and N_t is the total number of gray levels present in the image. For example, **Figure 5.4** demonstrates the application to a 5x5 gray level input matrix.

The NGTDM features have been developed to correlate the quantitative values of texture features as closely as possible to the visual interpretation of texture by humans. For example, coarseness provides a quantitative measure of local uniformity while contrast provides quantitative information about the difference in the intensity levels of neighboring regions. Similarly, busyness tells us how "busy" (rapid intensity changes) are the

neighborhoods in a given image or a region of interest. Furthermore, NGTDM complexity, as the name suggests, quantifies the complexity of the spatial information present in an image. Finally, texture strength can be defined as characterizing the visual aesthetics of an image. Images with higher texture strength are generally more attractive to look at than images with low texture strength. The NGTDM features have been applied to some medical image applications [10, 17, 112].

5.2.2 Morphological texture features

Morphological features are used in many applications to define the shape of an object. The shape of the tumor can be quantified using fractal dimension [25, 157]. Fractal dimension is a measurement of irregularity in the shape of the tumor. Different methods proposed in the literature for computing the fractal dimension of a pattern are discussed in [158]. Other methods based on volume and surface area of the tumor have been summarized in [30]. However, different fractal sets having different textures may have the same fractal dimension values. Consequently, Mandelbrot introduced the concept of lacunarity for characterization of texture in an image [25]. Lacunarity quantifies how fractals fill space – larger the gaps, higher the lacunarity. Different methods proposed in the literature for computing the fractal dimension of a pattern are discussed in [158]. Li et al. [159] evaluated four methods for computing fractal dimension of parenchymal patterns obtained from mammography in assessment of breast cancer risk. The four methods evaluated included conventional box counting method, modified box counting method using linear discriminant analysis (LDA), global Minkowski method, and modified Minkowski method using LDA. The authors observed that the advanced methods using LDA resulted in better classification between low and high-risk patterns (AUC = 0.9 for modified box counting

method and AUC=0.93 for modified Minkowski method). Likewise, Guo et al. [12] compared five different methods for computing the fractal dimension as well as lacunarity analysis of breast regions obtained from mammography to classify breast masses from normal parenchyma. The fractal dimension methods tested in this study were reticular cell counting method [160], differential box counting method [161], blanket method [162], Fourier power spectrum method [163], and fractional Brownian motion (FBM) model [164]. The highest AUC of 0.84 was achieved for the FBM method. Moreover, when FBM method was combined with lacunarity, the AUC increased to 0.90. Morphological features based on Renyl fractal dimension, shape analysis, volume, and surface area of the tumor have also been implemented in the literature [30, 165, 166].

5.2.3 Filtering texture features

The basic idea of filtering approaches is that the original textured image is subjected to some form of linear or nonlinear transform, followed by statistical analysis of the transformed image to obtain texture information. A comparative study of different filtering approaches can be found in [167]. In this chapter, I discuss the filtering techniques commonly used for texture analysis of medical images.

a) Spatial filtering techniques

Spatial filtering techniques are based on neighborhood operations on the original textured input images. These neighborhood operations are based on filters or kernels of size $n \times n$, where n determines the size of the neighborhood considered by the kernel. Some examples commonly used filters for texture analysis include statistical filters like average filter, range filter, and entropy filter or edge filters like Prewitt filter, Sobel filter, Laplacian filter, and Laplacian of Gaussian (LoG) filter. The input image is convolved with the desired kernel

to produce filtered images highlighting specific texture information in the original texture image. The resultant filtered images are analyzed using first order statistics (mean, median, and standard deviation). Statistical and edge filtering techniques have been applied to medical images for texture analysis in a number of applications [20, 77, 168-176].

Apart from statistical and edge kernels, special kernels have also been designed for identifying different types of textures. For example, Laws designed three sets of one-dimensional convolution masks of different sizes corresponding to different types of textures such as level, edge, spot, wave, ripple, undulation, and oscillation [28, 177]. All the convolution masks were center weighted, symmetric or asymmetric and zero-sum except the level convolution masks. Using a vector product between the masks of the same sizes, different spatial domain filters or kernels of size 3x3, 5x5, and 7x7 were generated by Laws. Laws texture energy measures have been used in filtering input medical images for texture analysis in many applications [103, 178-189]. Another example of a specially designed kernel is the fractal dimension kernel designed by Al Kadi et al. [18]. Examples of different techniques for filtering in the spatial domain have been illustrated in **Figure 5.5.**

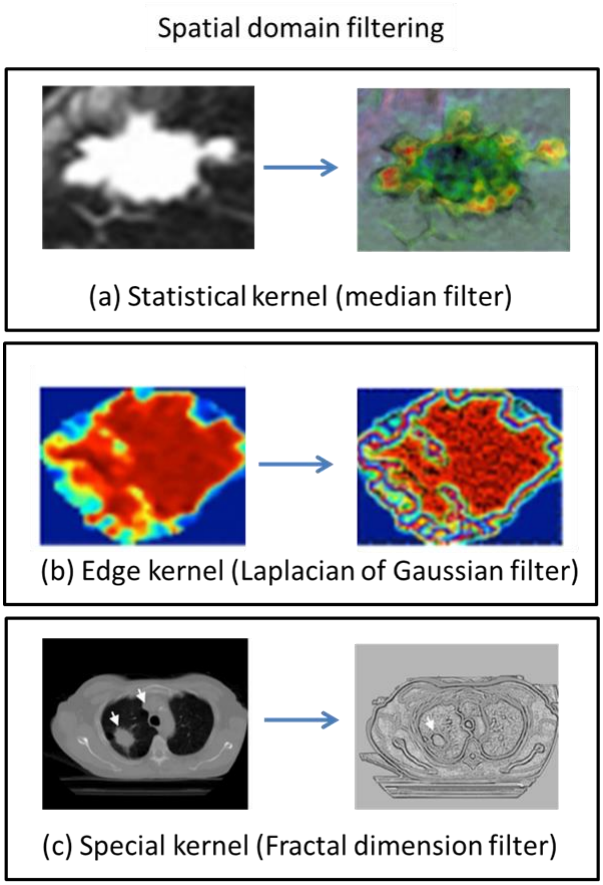


Figure 5.5 Illustration of different techniques used for spatial domain filtering (a) statistical kernel (e.g. median filter) (b) Edge kernel (e.g. Laplacian of Gaussian filter) (c) Special kernel (e.g. Fractal dimension filter). Modified from [18-20].

The value of n (size of the kernel neighborhood) is determined by the spatial resolution of the image being filtered. Higher the spatial resolution of the image being filtered, higher the value of neighborhood, n, used for filtering. However, further research is required to standardize the size of the neighborhood for different resolution images acquired using different imaging modalities.

b) Multi-resolution image scaling

The frequency of variations in the gray level values in a region of interest is dependent on the scale of the region of interest. The frequency content within an image can be analyzed

at different scales using wavelets [190-193]. Image texture can be analyzed at different scales by representing the image in a pyramid structure. Using discrete wavelet transform, four low resolution images can be obtained from the original image viz. I_{LL}, I_{LH}, I_{HL}, and I_{HH}. By repeatedly applying discrete wavelet transform on I_{LL} at each level, hierarchical pyramid structure for different resolutions can be created.

Texture analysis can be done by computing statistical texture features at each level or averaging the results across multiple resolutions. Wavelets have been extensively used in multiresolution texture analysis of medical images [20, 30, 102, 113, 117-121, 126, 129, 132, 142, 189, 194]. The use of multiresolution can decompose the data into different frequency components, thereby facilitating the study of each component with a resolution matched to its scale. **Figure 5.6** illustrates multiresolution scaling on an example breast MRI diffusion weighted image [21].

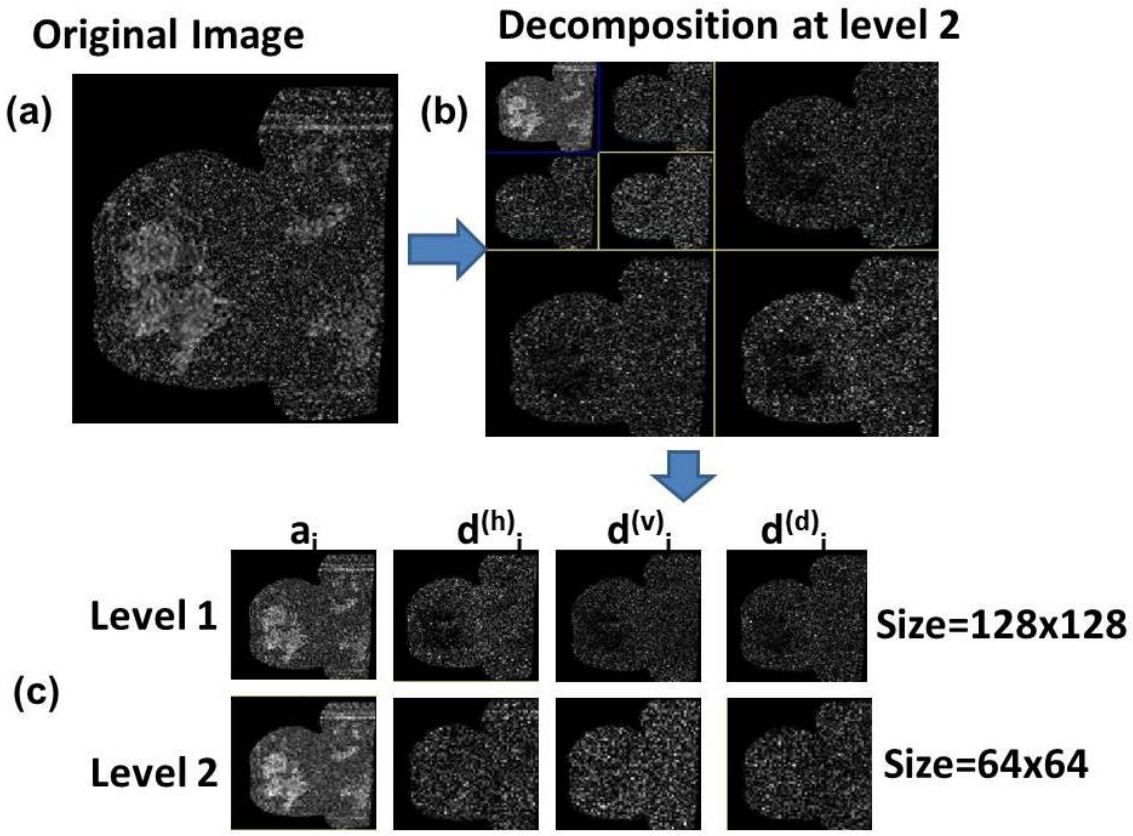


Figure 5.6. Multiresolution methods applied to a diffusion-weighted image ($b = 500$): a) the original size 256x256; b) compressed image (64x64) at different levels. c) For compression 2D biorthogonal spline wavelets were used. $d^{(h)}_j$, $d^{(v)}_j$ and $d^{(d)}_j$ respectively are detail components corresponding to vertical, horizontal, and diagonal. a_j, is the approximation (coarse) component at decomposition level. Modified from [21].

5.3 Radiomics applications

In early applications to medical imaging, the majority of methods used texture analysis, which, then evolved into the current application of radiomics. An extensive review on the applications of both texture and radiomics analysis in neurologic pathologies can be found in [195]. In this chapter, I review the application of these analyses to mostly body applications.

5.3.1 Lung

a) X-Ray

Texture analysis was first used in 1972 for automatic classification of pulmonary disease by Sutton and Hall [77]. They employed texture measures of symmetry, directional gradient, and Fourier domain energy on a dataset of 24 patients to classify between lung tissue using linear discriminant analysis. The authors observed that the texture features based on directional gradient gave the best classification accuracy. The training classification accuracy was achieved at 92% while the testing classification accuracy was achieved at 84%. This study was directed towards demonstrating the usefulness of texture in automatic screening of chest radiographs.

b) Computed Tomography

Al-Kadi and Watson implemented the differential box counting based fractal analysis method as well as lacunarity analysis on contrast enhanced (CE) CT images for differentiation of aggressive malignant lung tumors from the nonaggressive ones on a cohort of 15 patients [18]. The authors achieved an accuracy of 83.3% for distinguishing between these two groups using average fractal dimension. The authors also observed high correlation between the average fractal dimension and tumor uptake of 18FDG obtained using PET.

Ganeshan et al. [171] developed texture features using multiscale (fine to coarse) application of spatial domain filtering algorithm LoG (Laplacian of Gaussian) followed by extraction of statistical features of mean gray intensity (MGI) along with entropy (E) and uniformity (U) on unenhanced CT scans of 18 non-small cell lung cancer patients to correlate with tumor glucose metabolism and stage. The authors found significant

associations between coarse texture features and tumor standard uptake value (SUV) and fine texture features and tumor stage. In another study, Ganeshan et al. [174] used the texture feature of uniformity to associate with patient survival in a study of 54 patients. The authors observed that patients with coarse texture uniformity of less than 0.62 did not survive more than two and a half years. The same research group further extended their research to correlate texture features with histopathological markers of angiogenesis and hypoxia on a dataset of 14 patients with unenhanced as well as CE CT images [175]. The same texture analysis method was employed with different statistical features (standard deviation (SD) along with mean (MPP) and uniformity (UPP) of positive pixel distribution). The results indicated significant associations between texture features and average intensity of tumor staining with pimonidazole, tumor Glut-1 expression, and tumor CD34 expression. The relationships between image features and the histograms generated at different scales using spatial domain filtering algorithm, LoG have also been analyzed by the same group to provide clinical understanding of results [196].

Aerts et al. [30] used texture features based on first order statistics, GLCMs, GLRL matrices at multiple scales using wavelets on CT images of 1019 patients with non-small cell lung cancer or head-and-neck (H&N) cancer to associate radiomic features with primary tumor stage as well as patient survival. The radiomic features of first order statistics energy, shape compactness, and gray level nonuniformity (GLRL feature) obtained from unscaled data as well as wavelet HLH scaled data were found to be most significant. Good to moderate prediction concordance indices (CIs) of 0.65, 0.69, and 0.69 were achieved on one lung cancer validation dataset of 225 patients and two H&N datasets of 136 and 95 patients. The authors found significant associations between radiomic

features and gene expression patterns indicating the utility of radiomic features in characterizing underlying biological mechanisms. Strong correlations were observed between the radiomic features (GLRL gray level non-uniformity obtained from unscaled data as well as wavelet HLH scaled data) and cell cycling pathways, demonstrating higher proliferation for more heterogeneous tumors.

Recent studies by the same group achieved low AUCs of 0.66 (lung cancer) using random forest and CIs of 0.6 (lung cancer) and 0.68 (H&N cancer) using consensus clustering on the same set of radiomic features for predicting patient survival, which needs further investigation [120, 121]. Moreover, using consensus clustering, the authors also found either no or some association between consensus clusters and lung stage (AUC =0.61), lung histology (AUC=0.56), H&N stage (AUC=0.77), and H&N histology (AUC=0.58) based on the poor AUC. However, these are essentially low AUCs and more research is needed.

The same group used additional texture features based on first order statistics obtained after the application of LoG filter in addition to the other radiomic features used previously on a cohort of 182 patients to predict probability of distant metastasis in lung adenocarcinoma in addition to patient survival [20]. The authors found thirty-five radiomic features to be somewhat prognostic for distant metastasis (CI>0.6) and twelve features for patient survival (CI=0.55). The authors reported four features based on the LoG filter to be trending for potentially prognostic for distant metastasis and patient survival. A recent study by the same group tested two new features of entropy ratio and convexity for association with patient survival in lung adenocarcinoma using two independent patient

cohorts of 61 and 47 patients [122]. The authors found significant association between the two features and patient survival for only the first cohort.

In summary, statistical features obtained from LoG based spatial domain filtering of tumors at coarse level were found to be more predictive in different applications than statistical features obtained at fine level. A LoG filter highlights the edges within an ROI, suggesting that edges found by the filter using a wider Gaussian were more informative than edges produced using a narrow Gaussian. This may be because edges found using a narrow Gaussian are more susceptible to noise in the input image, Furthermore, LoG based features were also found to be more informative than statistical features based on first order statistics, GLCM and GLRL [20].

c) PET

Cook et al. [10] investigated the texture features obtained from NGTDM extracted from FDG PET images of 53 non-small cell lung cancer patients for association with chemoradiotherapy response and survival. The texture features were compared to the three survival groups (overall survival (OS), progression free survival (PFS), and local PFS (LPFS) obtained using RECIST (Response Evaluation Criteria in Solid Tumors) criteria. The authors found the texture features of coarseness, busyness, and contrast to significantly predict survival in the patient cohort as summarized in **Table 5.3.** The authors observed that responders had lower coarseness but higher contrast and busyness than non-responders. This suggests that the texture obtained for tumors corresponding to responders is less uniform with high frequency of intensity changes; i.e., they are more heterogeneous than tumors corresponding to non-responders.

Table 5.3 Summary of coarseness, contrast, busyness and complexity values corresponding to responders and non-responders of NSCLC obtained using Lung PET as reported in the literature [10].

Pathology	Coarseness	Contrast	Busyness	Complexity
Responders	0.01	0.11	0.76	1938
Non-responders	0.03	0.04	0.37	1926

d) PET-CT

Vaidya M et al. [111] used FDG-PET/CT dataset of 27 patients for characterization of radiotherapy tumor response in non-small cell lung cancer. Texture features of first order statistics and gray level co-occurrence matrix features were combined with the SUV/HU (Hounsfield units) measurements in this study. The model was built using logistic regression and resampling methods of cross validation and bootstrapping. The authors observed that the first order statistics features obtained from intensity volume histogram correlated more strongly with loco-regional control in contrast with the GLCM based features which correlated more strongly with local control.

5.3.2 Breast

a) Mammogram

Texture analysis was first used on mammograms in 1986 by Magnin et al. [123] to evaluate the risk for developing breast cancer. GLCM based texture features were used in this study. However, this study did not yield very good results with reproducibility barely reaching 80%. Wei et al. [126] successfully implemented GLCM based texture analysis at multiple scales using wavelets on 672 ROIs to classify between biopsy proven masses and normal parenchyma using linear discriminant analysis. They achieved an AUC of 0.89 on the

training set and an AUC of 0.86 on the test set. A study by Chan et al. [131] investigated the use of texture analysis on mammograms for associating the presence of clustered microcalcifications with malignant pathology. The authors implemented the GLCM texture features on a set of 86 mammograms followed by a backpropagation artificial neural network classifier to achieve an AUC of 0.88. Multiple studies have since investigated the use of texture analysis on mammograms for detection of masses [11, 12, 103, 132, 134, 136, 142, 165, 180, 183, 197-199] with an average AUC of 0.87 and a maximum AUC of 0.96 achieved with texture features obtained using first order statistics based features, GLCM based features, GLRL matrix based features and Law's texture energy measures. The quantitative values of the fractal dimension obtained from breast mammograms corresponding to normal, mass, and architectural distortion as reported in the literature are summarized in **Table 5.4.**

Table 5.4. Summary of fractal dimension values corresponding to different breast tissue classes obtained using mammogram as reported in the literature [11, 12].

Tissue class	Fractal dimension
Normal	2.64
Mass	2.39
Architectural distortion	2.52

b) Ultrasound

Garra et al. [86] performed texture analysis (first order statistics, GLCM, and fractal dimension) on breast ultrasound in a cohort of 80 patients. The authors were able to identify malignant lesions with a sensitivity of 100% and specificity of 80% (78% for fibroadenoma, 73% for cysts, and 91% for fibrocystic nodules). The best features identified

in this study were GLCM based contrast with $\theta = 45°$ and GLCM based correlation with $\theta = 0°$. Numerous studies have since used texture analysis to differentiate between benign and malignant breast lesions using ultrasound [138, 189, 194, 200].

c) MRI

The first application of texture analysis in breast MRI was in 1997 by Sinha et al. [1]. This study included 43 breast cases (23 benign and 20 malignant) and used eight texture features in combination with patient age and the DCE-MRI parameter of maximum enhancement to obtain sensitivity and specificity of 93% and 95%, respectively. In reality, it was radiomics, since they compared the results to an outcome or clinical variable. Moreover, Gibbs et al. [13] implemented GLCM based texture features to differentiate between benign and malignant breast tumors on post contrast MRI images from 79 patients. Texture features of variance, entropy, and sum entropy were found to be most significant using logistic regression analysis and the ROC analysis resulted in an excellent AUC of 0.92. Similarity, Ertas et al. [2] extracted first order statistics from normalized maximum intensity-time ratio (nMITR) projection generated using DCE-MRI to classify between benign and malignant breast tumors in 46 patients. The texture features of entropy, mean, standard deviation, and maximum were observed to be the most significant ($p<0.001$) with excellent AUCs between 0.86 and 0.97.

Nie et al. [143] used GLCM based features along with morphological features to distinguish between benign and malignant tumors using post contrast MRI images of 71 patients. The authors found the texture features of compactness, normalized radial length entropy, volume, and GLCM based features of entropy, sum average, and homogeneity to most significantly differentiate benign and malignant tumors and obtained an AUC of 0.82.

The same group compared the two feature selection and classification methods of logistic regression and artificial neural network for the task of classifying malignant breast tumors from benign breast tumors [146]. The authors found there was no significant difference in the results obtained from the two methods.

Instead of using post contrast MRI images, Karahaliou et al. [16] used parametric maps derived from DCE-MRI (e.g. signal enhancement ratio map) to extract GLCM based features. The authors found the GLCM ASM, GLCM entropy, and GLCM sum entropy obtained from signal enhancement ratio (SER) map as the most discriminative features producing an AUC of 0.92. Agner et al. [19] used kernels based on GLCM features along with other filters on the DCE MRI images in a dataset containing 41 patients for differentiating malignant from benign breast tumors. The authors created textural kinetic curves by computing the mean of the filtered image within the region of interest across the DCE sequence of images. The textural kinetic curves were classified using a probabilistic boosting tree and achieved sensitivity and specificity of 99% and 76%, respectively. In a recent study, Wang et al. [154] used morphological and GLCM texture features in combination with pharmacokinetic parameters obtained from DCE-MRI to classify between benign and malignant breast tumors. The authors found the texture feature of GLCM entropy, GLCM energy, and compactness along with pharmacokinetic parameters of rate constant (k_{ep}) and volume of plasma (v_p) to be the most discriminative with sensitivity and specificity of 91% and 92%, respectively. Cai et al. [152] used GLCM based texture features from DCE-MRI in combination with ADC, kinetic curve features, and morphological features to distinguish between benign and malignant breast tumors on a cohort of 234 patients. The authors achieved sensitivity and specificity of 85% and 89%,

respectively. They also tested the same set of features on a validation dataset consisting of 93 patients and achieved sensitivity and specificity of 69% and 91%, respectively [151].

Holi et al. [106] applied texture analysis (first order statistics, GLCM, GLRL) on T1-weighted pre contrast, post contrast and subtraction breast MRI datasets from twenty patients in order to associate texture features with histological types of invasive breast cancer (lobular vs. ductal). The authors identified the entropy based GLCM features to be the most effective features and were able to achieve a maximum accuracy of 100% using linear discriminant analysis (LDA) and nonlinear discriminant analysis (NDA) on the first subtraction and contrast images.

A recent study by Ahmed et al. [15] performed GLCM based features on breast MRI to predict chemotherapy response in 100 breast cancer patients. The authors found the texture features of contrast, variance, difference in variance, sum variance, sum entropy, sum average, cluster shade, and cluster prominence showed significant difference between responders and partial responders of chemotherapy when implemented on post contrast images. Parikh et al. [201] used multiscale LoG filter followed by extraction of first order statistical features from T2-weighted MRI of 36 patients to predict chemotherapy response. The authors found the texture features of entropy and uniformity showed significant different between responders and non-responders with an AUC of 0.84.

In summary, post contrast enhanced MRI was the most frequently used image for texture analysis of breast tumors. This is consistent with the clinical environment where radiologists also use post contrast enhanced MRI to discern textural features corresponding to breast tumors. The quantitative values of the texture features obtained using post contrast

enhanced breast MRI images for benign and malignant lesions reported in the literature are summarized in **Table 5.5**.

The quantitative texture values of entropy, energy, etc. obtained from various studies may be difficult to compare due to several reasons. For example, the use of different magnet strengths (1.5 or 3T), different preprocessing steps, binning methods, and gray level normalization. However, a trend in the quantitative values can still be observed, for example, entropy and contrast values tend to be increased for malignant compared to benign and/or normal tissue. These metrics may provide an insight into the heterogeneity of the tumor tissue. Based on the current literature, It may be inferred that malignant tumor tissue is more heterogeneous (high entropy, low uniformity, and high contrast) than benign tumor tissue and seen in pathological samples [202]. This correlation is the impetus for further radiomics research in developing noninvasive tools for clinical decision.

Table 5.5. Summary of quantitative values of four GLCM based features obtained using post contrast enhanced breast MRI images for benign and malignant lesions as reported in the literature [1, 13-16].

				Benign lesion				Malignant lesion			
Reference	Image analyzed	Pre-processing	GL	ASM	H	Con	Cor	ASM	H	Con	Cor
Sinha et al. [1]	Post contrast DCE-MRI	None	16	0.02	1.96	34.3	-	0.01	2.27	535.8	-
Ahmed et al. [15]	Post contrast DCE-MRI	H.Eq	16	-	-	-	-	0.01	7.13	8.5	0.8
Gibbs et al. [13, 14]	Post contrast DCE-MRI	H,Eq	32	0.004	5.85	25.5	0.84	0.004	5.94	23.2	0.86
Chen et al. [14]	Post contrast DCE-MRI	H.Eq	32	0.01	5.41	52.1	0.6	0.01	5.44	75.8	0.6
Chen et al. [14]	Post contrast DCE-MRI	H.Eq	32	0.01	5.12	67.5	0.5	0.01	5.26	69.2	0.55
Karahiliou et al. [16]	SER map	None	64	0.025	1.71	12.8	0.37	0.01	2.12	23.8	0.51
Parekh et al. [Chapter 6]	Post contrast DCE-MRI	None	256	0.002	6.44	351.28	0.69	0.001	7.24	294.53	0.79

H.Eq =histogram equalization
SER= signal enhancement ratio
GL = gray levels
ASM = angular second moment
H = GLCM Entropy
Con = Contrast
Cor = Correlation.

5.3.3 Liver

a) Computed Tomography

Mir et al. [125] implemented texture analysis based on GLCMs and GLRL on the liver CT images of 60 patients to classify the CT images as normal liver, clearly visible malignancy,

and invisible malignancy. The authors found the texture features of entropy (normal: 1.65±0.12, visible malignancy: 2.13±0.17, invisible malignancy: 1.64±0.08), local homogeneity (normal: 2.30±0.25, visible malignancy: 1.48±0.21, and invisible malignancy: 2.23±0.15), and gray level distribution (normal: 5.54±0.31, visible malignancy: 2.09±0.61, invisible malignancy: 4.72±0.54) to be the most significant. Chen et al. [135] used fractal analysis along with GLCM based features to classify the liver tumors as hemangioma or hepatoma in a dataset of 30 patients. Using modified probabilistic neural network classifier, the authors were able to achieve classification accuracy of 83%. In a study by Gletsos et al. [140], GLCM based features were used to classify the CT focal liver lesions of 147 patients into four classes (normal, hepatic cysts, hemangioma, and hepatocellular carcinomas). The authors used three sequentially placed feed forward neural networks and achieved excellent accuracies of 97%, 100%, and 82% in classifying the normal from abnormal, hepatic cysts from others, and hemangioma from hepatocellular carcinomas, respectively. Huang et al. [203] used autocovariance function on a dataset of 164 liver CT images to classify malignant (hepatocellular carcinomas and colorectal metastases) lesions from benign lesions and achieved moderate sensitivity and specificity of 75% and 88%, respectively.

b) MRI

Jirak et al. [91] extracted first order statistical and GLCM based texture features using T2-weighted MRI images obtained from 43 patients to classify between healthy and cirrhotic liver. The authors achieved classification error around 8%. Along with T2-weighted MRI images, Mayerhoefer et al. [105] also extracted texture features (first order statistics, GLCM, and GLRL matrix) from T1-weighted images to classify focal liver

lesions achieving error rates of 12-18% on T2-weighted images and 16-18% to T1-weighted images. Fujimoto et al. [110] used entropy of ADC to classify normal from abnormal fibrosis stage (AUC =0.94, $p<0.001$, entropy cutoff = 1.30). ADC maps were created using diffusion weighted MRI (b =0 and 1000 s/mm^2). O' Connor et al. [108] quantified tumor heterogeneity using first order statistics on voxel-wise K^{trans}, v_e, and v_p and fractal dimension analysis on DCE-MRI data obtained from 10 patients with 26 colorectal cancer liver metastases to predict shrinkage in tumor volume in response to bevacizumab and cytotoxic chemotherapy. The authors found the median v_e, tumor enhancing fraction (E_f), and microvascular uniformity obtained using fractal dimension to be the most significant features and the median classification error was achieved at 12%.

c) Ultrasound

Texture analysis was first implemented on liver ultrasound images in 1985 by Raeth et al. [85] to classify liver into normal, diffuse parenchymal, and malignant disease on a dataset of 71 patients with an accuracy of 96%. Wu et al. [124] used multiresolution fractal analysis to classify between normal liver, hepatoma, and cirrhosis on a dataset of 40 patients and achieved an accuracy of 90%. The authors observed that multiresolution fractal analysis features outperformed the GLCM based texture features as well as Laws texture energy measures. Sujana et al. [87] extracted first order statistical features along with GLCM and GLRL matrix based features to classify liver ultrasound images from 113 patients into normal, hemangioma, and malignant categories. Using artificial neural network, the authors were able to achieve classification accuracy of 100%. Horng et al. [139] developed a new texture analysis method called texture feature coding to classify liver ultrasound images into normal, hepatitis, and cirrhosis on 120 patients (30

training and 90 test images). The authors achieved an accuracy of 86.7% as compared to 75.7% obtained by the same group using GLCM, texture spectrum, and fractal dimension based features [127]. Yoshida et al. [94] used multiresolution analysis on a dataset of 44 patients and obtained an AUC of 0.92 in classifying malignant from benign lesions.

Numerous studies have since used texture analysis to classify focal hepatic lesions using ultrasound [181, 185, 204-208]. A comparative study of different texture analysis approaches implemented in the literature can be found in [209]. A recent implementation by Mitrea et al. [188] extracted GLCM based texture features from the Laws texture energy images obtained from filtering liver ultrasound images and achieved an accuracy of 90% in liver cancer detection.

5.3.4 Colorectal cancer

a) Computed Tomography

A study by Ganeshan et al. [168] investigated the use of texture analysis of liver on 28 patients with colorectal cancer. The authors used multiscale (fine to coarse) application of the spatial domain filtering algorithm LoG followed by extraction of first order statistical features. The authors observed that the relative scale texture parameter correlated inversely with the corresponding PET SUV metric ($r = -0.59,\ p = 0.007$) and hepatic phosphorylation index (HPFI) ($r = -0.59,\ p = 0.006$). There was a positive correlation with the total hepatic perfusion (THP) ($r = 0.51,\ p = 0.02$) and hepatic portal perfusion (HPP) ($r = 0.45, 0.05$) for patients without liver metastases. The same research group used the same set of texture features obtained from liver CT to predict survival in patients with colorectal cancer [170]. The texture feature of uniformity was obtained for the scale ratios of 1.5 (fine) to 2.5 (course) and 2 (medium) to 2.5(coarse) were observed to be significant

($p<0.005$) prognostic factors for surival. Goh et al. [210] extracted fractal dimension based features from colorectal perfusion CT images of 20 patients to differentiate between colon cancer and normal bowel. The authors found the fractal dimension and fractal abundance were significantly ($p \leq 0.001$) higher for colon cancer (1.71 ± 0.07 and 7.82 ± 0.62) than normal bowel (1.61 ± 0.07 and 6.89 ± 0.47). Cui et al. [109] used fractal dimension analysis on CT images obtained from 228 patients to classify the status of lymph nodes in rectal cancer as benign or malignant. The authors were able to classify the malignant nodes from benign nodes with an accuracy of 88%.

Ganeshan et al. [172] used their previously developed multiscale texture analysis algorithm on dynamic contrast enhanced CT of the liver obtained from 27 patients to classify between node negative and not positive non metastatic colorectal cancer. The authors observed significant difference in entropy and uniformity of the node negative and node positive patients. Using fine texture entropy of the images obtained between 26 and 30 seconds after contrast injection, the authors were able to achieve sensitivity and specificity of 100% and 71% respectively in identifying node positive patients (entropy cutoff ≤ 0.0807). The same group used texture analysis of the colorectal tumor obtained from contrast enhanced CT of 57 patients to define biomarkers for 5-year survival [176]. The authors used the texture features of fine scale entropy, uniformity, kurtosis, skewness, and standard deviation to define the five-year survival rate with the corresponding cutoffs at less than 7.89 for entropy, greater than 0.01 for uniformity, less than 2.48 for kurtosis, greater than 20.38 for skewness, and less than 61.83 for standard deviation.

5.3.5 Head and Neck

a) Computed Tomography

Leijenaar et al. [118] chose four radiomic features (Energy, compactness, GLRL non-uniformity and wavelet GLRL non-uniformity obtained by Aerts et al. [30] described previously to investigate their prognostic ability on 542 oropharyngeal squamous cell carcinoma (OPSCC) patients that underwent CT. The features were weighted in a Cox model [211] to develop and test a prognostic index for validation of the radiomic signatures. The authors observed that the Kaplan-Meiyer survival curves were significantly different ($p < 0.05$) between low and high radiomic signature predictions using a log-rank test, however, the C-index was 0.628 which is lower than reported in previous reports (C index=0.686 and 0.685) [30]

b) PET

El Naqa et al. [144] explored the GLCM texture and features on a cohort of nine head & neck patients and 14 cervical cancer patients undergoing chemoradiotherapy and imaged with PET. The most discriminative features from the GLCM included energy, contrast, local homogeneity, and entropy. The authors were able to achieve moderate to high AUCs of 0.76 (cervical cancer) and 1.0 (H&N) on the very small datasets indicating a potential usefulness of texture in PET imaging. For cervical cancer, the GLCM entropy of the CTV (clinical tumor volume) was much lower (3.6) than that of the tumor (5.3). A similar trend was observed for H&N cancer as well where entropy of the CTV was 3.8 and that of the tumor was 4.7.

5.3.6 Esophageal cancer

a) PET

Tixier et al. [17] used first order texture features from the GLCM, GLRL matrix and NGTDM on a PET dataset to classify between response in 41 esophageal cancer patients

treated with chemoradiation. The authors found the texture features of GLCM entropy (sensitivity=79%, specificity=91%) and GLCM angular second moment (sensitivity=88%, specificity=73%) along with size (sensitivity=76%, specificity=91%) and intensity variabilities (sensitivity=76%, specificity=91%) of uniform zones were the most discriminative between responders and non-responders compared to the SUV_{max} (sensitivity=46%, specificity=91%).

5.3.7 Adnexal lesion

a) MRI

Kierans et al. [4] extracted first order statistical features of entropy from ADC maps to classify between benign and malignant adnexal lesions in 37 patients. The average entropy ADC was significantly higher for malignant (4.94 ± 0.4) than benign (4.54 ± 0.44) lesions. However, the sensitivity (66.7%) and specificity (82.4%) of the entropy ADC was low in diagnostic performance.

5.3.8 Prostate lesion

a) MRI

Wibmer et al. [212] performed GLCM texture analysis on T2-weighted and diffusion weighted MRI of prostate obtained from 147 patients to associate the results with Gleason score as well as cancer detection. The authors used five GLCM features of inertia, entropy, energy, correlation, and homogeneity for analysis. According to the authors, the texture features of cluster shade and cluster prominence "over-emphasize" the other GLCM features such as energy, entropy, homogeneity, and contrast and hence were not used. There were a total of 186 cancerous lesions in the 147 patients with 143 lesions in the posterior zone (PZ) and 43 lesions in the transition zone (TZ). The authors observed a

similar trend in the GLCM texture features obtained from PZ tissue in both ADC map and T2WI as well as from TZ tissue in ADC map. For all these cases, the GLCM texture features of entropy and inertia were significantly higher for cancer while the remaining three features were significantly lower for cancer than non-cancerous tissue ($p < 0.0001 - 0.008$). Furthermore, the authors observed that only the GLCM texture features inertia ($p = 0.001$) and correlation ($p = 0.041$) obtained from T2 were significantly different between TZ cancerous and non-cancerous tissue. Moreover, when compared to Gleason score, GLCM energy was lower while GLCM entropy was higher for increased Gleason scores ($p < 0.05$ for all inter-score GLCM feature values).

5.4 Discussion

Radiomics and texture analysis have been widely used for diagnosis and prognosis in many different applications. In fact, some of the initial applications of texture analysis date back to 1970s. The most commonly implemented features across all the applications are based on GLCMs. The reason behind the popularity of GLCMs as the texture analysis method of choice is because the GLCM features analyze the inter-voxel relationships in both voxels that are in immediate neighborhood as well as voxels that are far apart. Moreover, GLCMs are easy to implement and the features are more intuitive to interpret. In spite of being the most popular method, the features identified as the most predictive GLCM features are not consistent across different applications. In fact, the most predictive GLCM features are not even consistent within the same organ, same modality, and the same diagnostic application. For example, in breast MRI, the texture features identified as the most important for diagnosis were not consistent across different implementations in the literature. For example, Sinha et al. [1] observed the GLCM features of difference entropy, contrast, sum

entropy, and inverse difference of moments to be the most predictive, Gibbs et al. [13] observed the GLCM features of variance, entropy and sum entropy as the most predictive and Nie et al. [143] observed the GLCM features of entropy, sum average, and homogeneity as the most predictive in classifying benign from malignant breast lesions. Moreover, all the studies demonstrated excellent results. The variations in the radiomics results may arise from the image preprocessing steps such as segmentation, image quantization (for statistical features), neighborhood size (for statistical and filtering features), or the modeling steps of feature selection and classification.

The development of a texture feature set that is consistent across different modalities, organs and pathologies may not be possible or practical. This is because different modalities highlight separate characteristics of the tissue of interest. Furthermore, different organs may have distinct inherent texture. However, it is very important to develop a consistent feature subspace for each application even though different applications may have different feature subspaces. Furthermore, consistency in feature subspaces is also required to understand and correlate the texture features with tissue biology. For example, if the entropy values of benign and malignant tumors are significantly different for one research and not for another, it is not possible to establish any correlation between tumor biology and texture values. In the following chapters (**Chapter 7 and 8**), I will discuss these shortcomings in further detail and develop methods to address these shortcomings.

Chapter 6. Radiomic feature mapping for visualization and characterization of texture in normal and pathological tissue

6.1 Introduction

In **Chapter 5**, I introduced the field of radiomics and its application to various organs for clinical decision support. The current radiomic methods extract information about the gray-scale patterns, inter-pixel relationships, and shape-based properties of the region of interest (ROI) [24-30]. However, these methods in radiomics have certain limitations as detailed below.

1. Radiomic features are dependent on the size of region of interest. For example, the range of values that the first order entropy feature can take varies between 0 and $\log_2 N$, where N is the number of voxels in the tissue. Therefore, larger homogeneous tumors could be misinterpreted as being heterogeneous when compared to a significantly smaller heterogeneous tumor.
2. Radiomic features are not standardized. The values a radiomic feature takes are dependent on size and preprocessing steps such as image quantization and image filtering.
3. Radiomic metrics for normal tissue has not yet been established. This is very important because a quantitative metric for normal tissue can be evaluated and used to standardize the metrics for abnormal tissue.

4. The understanding of what these features mean and how they correlate to tissue biology is limited.

This chapter presents the radiomic feature mapping (RFM) framework to overcome the limitations of current methods in radiomics. The RFM framework uses radiomic filters to transform radiological images into radiomic feature maps for visualization and analysis of textural information present in the images. The motivation behind the development of radiomic feature mapping is to empower the radiologists with the ability to "see" the hidden textural information present in the radiological images and correlate it with tissue biology.

I developed the RFM framework for radiomic analysis using breast multiparametric magnetic resonance imaging (mpMRI). The unique ability of mpMRI to better characterize tissue parameters using the different contrasts of each sequence provides us with an opportunity to develop standardized radiomic metrics and better correlate them to tissue biology. The mpMRI of breast involves acquisition of conventional T1- and T2-weighted images and advanced functional MRI parameters of dynamic contrast enhanced-MRI (DCE-MRI) and diffusion weighted imaging (DWI) with apparent diffusion coefficient (ADC) mapping. The DCE-MRI is a time series acquisition of T1-weighted scans that results in time intensity curves corresponding to the different tissue vascularity. Moreover, these differences can be evaluated using pharmacokinetic modeling (PK) of the temporal DCE. Similarly, radiomic analysis of the PK images could produce textural evolution curves and provide insight into the underlying vascular "texture" heterogeneity corresponding to different tissue types. Radiomic analysis applied to the ADC mapping obtained from DWI could help investigate the underlying cellular heterogeneity of the tissue of interest, where the DWI measures the intra- and extra-cellular water content.

In this chapter, I developed standardized temporal evolution curves from DCE-MRI RFMs for evaluation of vascular heterogeneity and standardized diffusion evolution curves from DWI RFMs for evaluation of cellular heterogeneity. Furthermore, I analyzed the diagnostic capabilities of RFM features for prediction of breast tumors as benign or malignant using a novel multiview feature embedding and classification model. The rest of the chapter is organized as follows: **Section 6.2** introduces the novel radiomic feature mapping (RFM) algorithm and the new radiomic metrics developed in this thesis, followed by the details of patient population, MRI parameters acquired and image analysis techniques in **Section 6.3**. The application of RFM on multiparametric MRI (mpMRI) datasets of breast for correlation with tissue biology and tumor diagnosis is discussed in **Section 6.4**. The final section of this chapter discusses the advantages and limitations of the different techniques developed in this chapter for the development of potential precision medicine biomarkers.

Publication from this work and author contributions

Publication

V. S. Parekh and M. A. Jacobs, "Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI," NPJ breast cancer, vol. 3, p. 43, 2017

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

I developed the algorithms for radiomics (RFM) and patient classification (IsoSVM), the mathematical formulae for standardized DCE (texture evolution curves) and DWI radiomics (texture evolution metric), and did the entire implementation. I performed all the data analysis described in this chapter.

6.2 Radiomic Feature Mapping algorithm

The radiomic feature mapping algorithm transforms any radiological image into a radiomic feature map (RFM) using a radiomic filter based on either first or higher order statistical features described in **Chapter 5**. A large number of RFMs can be generated from a radiological image by filtering it with different radiomic filters. The RFM algorithm is dependent on three parameters: $S \times S$, the size of radiomic filters; G, the number of gray levels used to quantize the input radiological image and F, the set of radiomic filters used. The size of the radiomic filters or the neighborhood scaling parameter, S is determined by the user depending on the spatial resolution of the input image. The value of G is determined by the user based on the range of intensities as well as the number of bits required to represent voxel intensity in the input image. The RFM algorithm is described in **Algorithm 6.1**.

Algorithm 6.1 Radiomics Feature Mapping algorithm

Inputs:
img: $N \times N$ sized radiological image
F: set of radiomic filters
S: filter size (3, 5, 7, 9, …)
G: number of gray levels to use for gray level quantization (2, 4, 8, 16, 32, 64, 128, …)
Output: $RFMS$: set of radiomic feature maps
Initialization: initialize each map in RFMS to N x N null matrix.
1: Quantize img to G gray levels
2: for each function, f in F
3: for each sliding window, W of size $S \times S$ in img
4: Evaluate W using statistical function, f to get radiomic metric for W
5: Set the center voxel position of W in $RFMS_f$ to the calculated radiomic metric
6: return $RFMS$

The means of the radiomic values were calculated from different regions of interest (ROI) in each RFM as features for classification and further analysis. Consequently, every RFM feature from every patient corresponds to the average value taken from sliding same

sized image window ($W \times W$) across the whole ROI ensuring there is no mathematical dependence between the computed RFM features and size of the ROI. The input parameters for this study were $G = 256$, $W = 9$ and $N = 30$ based on empirical analysis. Out of 30 RFMs, 7 were generated using first order statistics, 22 were generated using second order statistics, and one was generated using Laplacian of Gaussian (LoG) filter. The complete set of features used in this study has been enlisted in Appendix 6.A.

6.2.1 Standardized Radiomic Metrics

a) Textural evolution curves

The PK-DCE MRI RFMs provided information on the vascular heterogeneity within lesion tissue. The RFM textural evolution curves capture the time evolution of tissue heterogeneity as a function of contrast uptake using the time series derived from DCE images. The mean radiomic feature value across all the voxels within a region of interest in the RFMs was used to construct textural evolution curves for each tissue type. In order to compare the textural evolution curves across different ROIs, normalization of the radiomic feature values were applied.

Analysis of textural evolution curves from entropy feature maps was done using the time to peak and the textural wash out slope from the DCE RFMs. The time to peak is defined as the time it takes for the textural evolution curves to reach its maximum value. The textural wash out slope is as the textural wash out of the textural evolution curves within the lesion tissue. This was computed as the slope of the line connecting the peak texture enhancement in the first two minutes to the last time point including all the intermediate time points.

b) Textural evolution metric for DWI

The textural evolution metric (TEM) was developed to investigate the cellular of different tissue types. The TEM map for DWI is computed from RFMs generated from DWI obtained at b=0 and b>0 using the following equation:

$$TEM_{DWI} = \frac{RFM_{DWI_{b>0}}}{RFM_{DWI_{b0}}}$$

Similarly, a textural evolution metric (TEM) was defined for the high spatial resolution DCE-MRI dataset (HR-DCE-MRI) as shown here:

$$TEM_{HR-DCE-MRI} = \frac{Post_{HR-DCE-MRI}}{Pre_{HR-DCE-MRI}}$$

6.3 Materials and Methods

6.3.1 Clinical data

The radiomic feature mapping framework was tested on a multiparametric MRI dataset obtained from a retrospective cohort of 124 patients (mean age = 52, range = 24-80) to classify malignant and benign lesions. Out of the 124 patients, 98 had malignant lesions while 26 patients had benign lesions. Patients were selected for the study based on the potential for malignant breast lesions with a BIRADS score of 3 or greater from 2008-2010. This retrospective study was approved by the IRB at our facility and conforms to HIPAA requirements. The multiparametric MRI imaging protocol and the algorithms for registration, tissue signature generation, and image segmentation have been previously detailed in **Chapter 2**.

6.3.2 Dynamic MRI with Pharmacokinetic (PK) Contrast Enhancement

The vascularity of breast tissue was obtained using different semi-quantitative and quantitative metrics [213, 214]. The semi-quantitative metrics use the temporal evolution

of the time series curves from the DCE MR images and are scaled into three categories relating to the potential characterization of the tissue and other known metrics [53, 55, 58]. The PK-DCE quantitative metrics derived were volume transfer constant (K^{trans} (min^{-1})), the fractional volume of the extracellular extravascular space (EVF (V_e)), and the transfer rate constant (k_{ep} (min^{-1})) using commercial software DynaCad (InVivo, Gainesville, Florida) [58, 68]. For both benign and malignant patients, glandular and lesion tissue, the mean values and standard deviations of the transfer constant (K^{trans}) and extra-vascular volume fraction-EVF (V_e) were recorded.

6.3.3 Statistical analysis

I computed summary statistics (mean and standard error of the mean) for the radiomic metrics and functional metrics from mpMRI. An unpaired t-test (two tailed) was performed to compare the RFMs for the benign and malignant patient datasets. Statistical significance was set at $p \leq 0.05$. Univariate logistic regression analysis was used to find associations between the RFMs and the final diagnosis. Receiver operating characteristic (ROC) and total operating characteristic (TOC) [215] curve analysis was performed to assess the diagnostic performance of each RFM in characterizing benign versus malignant lesions.

6.4 Experiments

6.4.1 Experimental Summary

The radiomic feature maps were computed and analyzed for one hundred and twenty-four women with breast lesions that underwent mpMRI scan. The mean age of the patients was 52 years (range: 24-80 years). Ninety-eight women (79%) had malignant lesions and twenty-six women (21%) had benign lesions. The radiomic features were extracted using the radiomic feature mapping (RFM) method that creates whole breast texture images of

each feature. The overview of the radiomic feature mapping procedure for classification of a multiparametric radiological dataset as benign or malignant is illustrated in **Figure 6.1**.

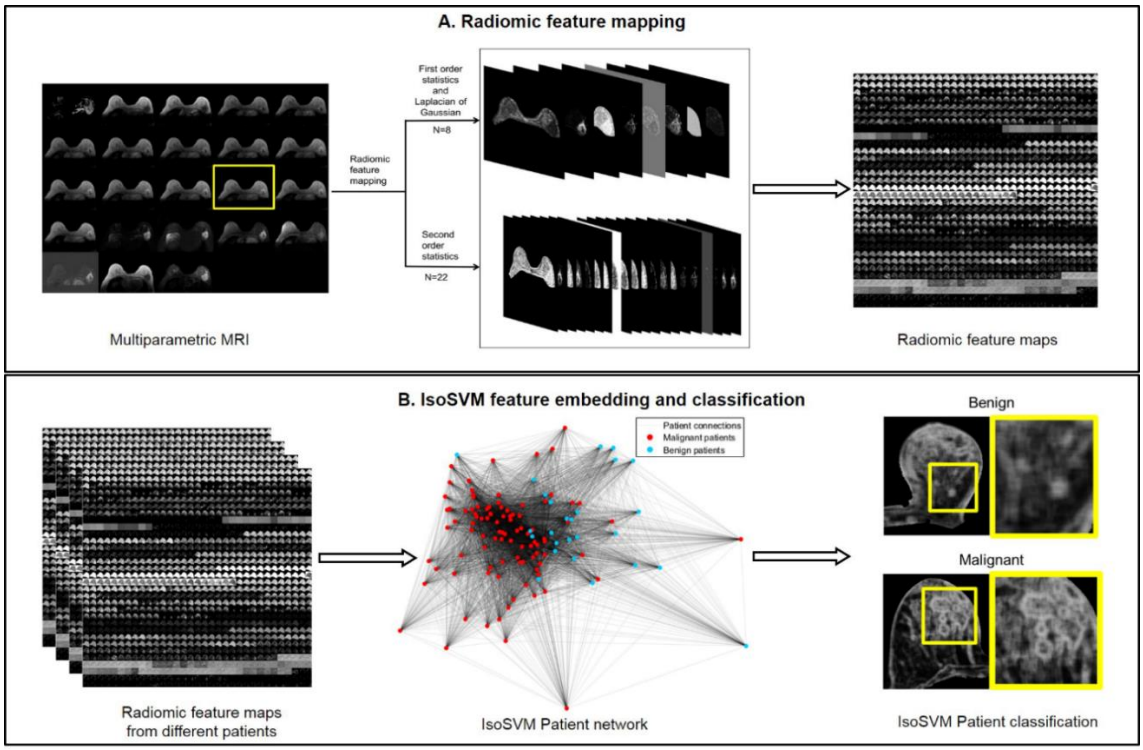


Figure 6.1 Concept of the radiomic feature mapping framework. **A.** The multiparametric radiological dataset ($N = 23$) is transformed into a high-dimensional radiomic feature space ($D = 690$) consisting of radiomic feature maps generated using Laplacian of Gaussian, texture statistical kernels ($n = 30$). **B.** The RFM space is first transformed to the patient network using the IsoSVM and then high-dimensional radiomic feature map space from each patient is classified as benign or malignant.

Figures 6.2 and **6.3** illustrate typical entropy feature maps corresponding to a benign and a malignant patient. A total of 690 RFMs were generated for the twenty-three image multiparametric MRI dataset of each patient. Regions of interest were defined on

each tissue type using the Eigen filter segmentation method and the MR radiomic feature maps were computed for different breast tissue. Finally, ROI size (area in cm^2), ADC value, and PK-DCE parameters for different regions of interest in each of these patients were obtained. There was no significant difference in the tumor size between benign and malignant patient groups (Benign size: $3.24 \pm 2.67\ cm^2$,Malignant size: $2.44 \pm 0.31\ cm^2$, $p = 0.77$).

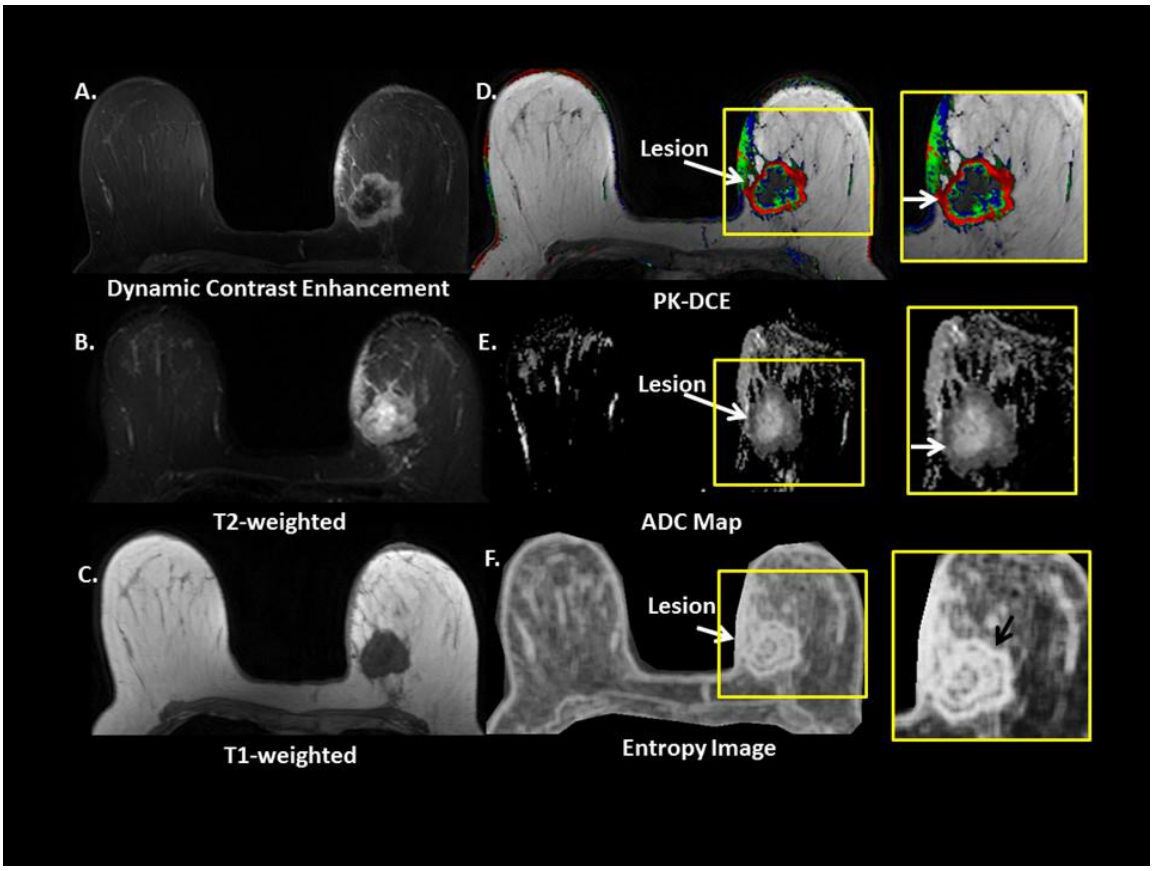


Figure 6.2 Typical multiparametric breast image of a malignant patient. **A.** Dynamic contract enhanced, **B.** T2-weighted, **C.** T1-weighted, **D.** Pharmacokinetic-DCE overlay of K^{trans} and EVF, where red indicates high K^{trans} and blue demonstrates low K^{trans} **E.** ADC maps, and **F.** whole breast entropy feature map.

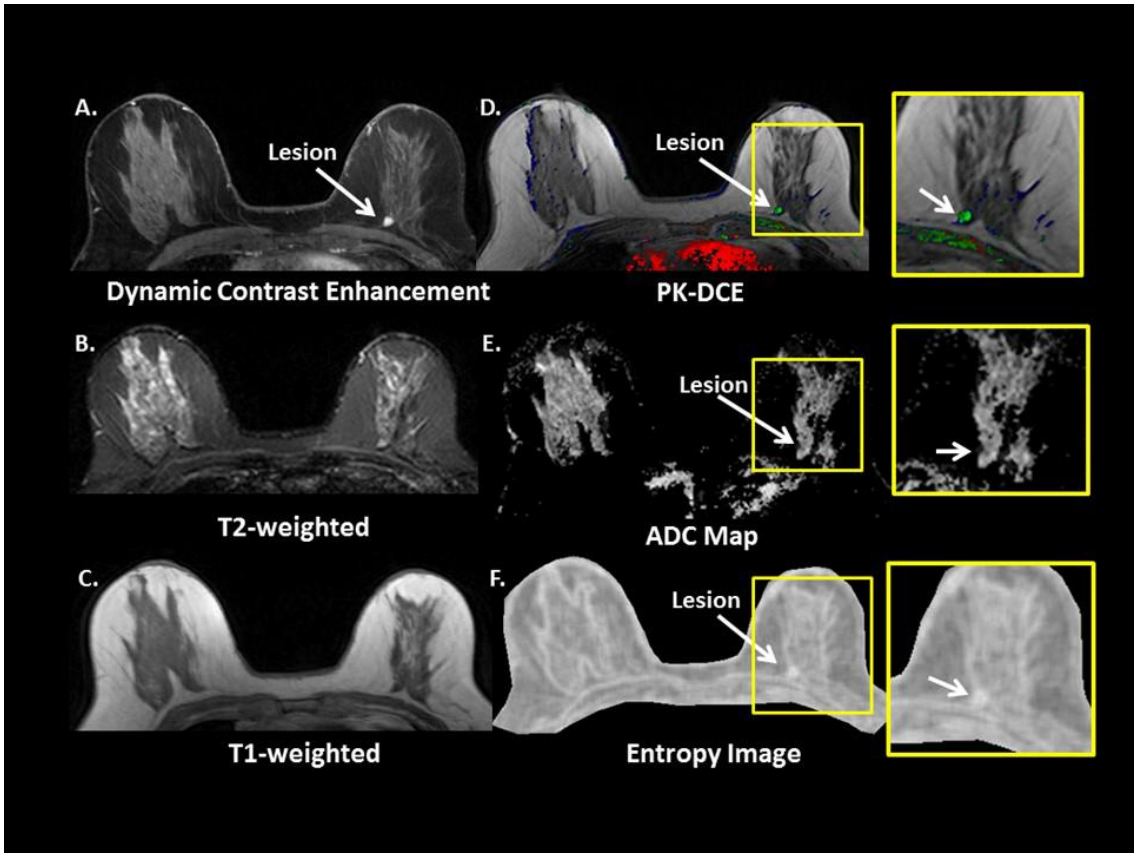


Figure 6.3 Typical multiparametric breast image of a benign patient. **A.** Dynamic contract enhanced, **B.** T2-weighted, **C.** T1-weighted, **D.** Pharmacokinetic-DCE overlay of K^{trans} and EVF, where red indicates high K^{trans} and blue demonstrates low K^{trans} **E.** ADC maps, and **F.** whole breast entropy feature map.

Table 6.1 summarizes the entropy values corresponding to the different regions of interest from the DCE-MRI and ADC map. Malignant lesions demonstrated significantly higher entropy on both post contrast DCE-MRI and ADC maps as compared to benign lesions (Benign DCE: 5.72 ± 0.12, Malignant DCE: 6.29 ± 0.06, $p = 0.0002$; Benign ADC: 5.65 ± 0.15, Malignant ADC: 6.20 ± 0.07, $p = 0.002$). There was no significant difference in the glandular tissue entropy values between the two groups (Benign DCE: 6.08 ± 0.10

Malignant DCE: 5.91 ± 0.05, $p = 0.16$; Benign ADC: 6.06 ± 0.32 Malignant ADC: $6.06 \pm 0.19, p = 1.00$).

Table 6.1 Summary of radiomic feature values and quantitative MpMRI metrics

Feature	Glandular tissue	Benign lesion	Malignant lesion	p value
Entropy (Post-Contrast DCE)	5.95±0.05	5.72±0.12	6.29±0.06	0.0002
Entropy (ADC map)	6.06±0.16	5.65±0.15	6.20±0.07	0.002
ADC Map (mm^2x10^{-3})	2.13±0.03	1.69±0.08	1.26±0.03	0.00001
K_{trans} (min^{-1})		0.27±0.21	0.69±0.45	0.001
EVF (V_e)		0.27±0.10	0.61±0.31	0.006
Tumor size (cm^2)		3.24±2.67	2.44±0.31	0.77

K_{trans} = volume transfer constant, EVF (V_e) = extracellular extravascular space

6.4.2 Correlation between tumor biology and radiomics using breast multiparametric MRI

a) Analysis of textural evolution curves

Figure 6.4 illustrates the textural evolution curves corresponding to the different radiomic features obtained from DCE MRI. **Figures 6.4(a)** and **6.4(b)** exhibit the textural evolution curves of the normalized mean values obtained from the entropy feature maps (top) corresponding to tumor and glandular tissue, respectively. The error bars in the textural evolution curves represent the standard error for the normalized mean of the entropy values. In **Figure 6.4(b),** there is no change in the texture for glandular tissue for both benign and malignant patients. However, the shapes of textural evolution curves were significantly different between benign and malignant lesions illustrating the difference in contrast uptake within benign and malignant lesions. The normalized entropy values during the wash-in phase were significantly ($p < 0.05$) higher for malignant than for benign lesions depicting a rapid textural enhancement for malignant lesions. Similarly, the normalized entropy values during the washout phase were significantly lower for malignant than for benign lesions depicting a rapid textural washout for malignant lesions. Moreover, similar trends were observed in the textural evolution curves obtained from range feature maps as illustrated in **Figure 6.4** (Bottom).

Preliminary analysis of textural evolution curves from entropy feature maps based on the time to peak and the textural wash out slope is shown below.

a. Time to peak: The average time to peak for benign lesions ($2.21 \pm 0.16\ mins$) was significantly longer ($p = 0.0003$) than for malignant lesions ($1.24 \pm 0.07\ mins$).
b. Textural wash out slope: The slope of the textural washout curves was also significantly different ($p = 0.001$) between benign (0.001 ± 0.001) and malignant lesions (-0.002 ± 0.0003).

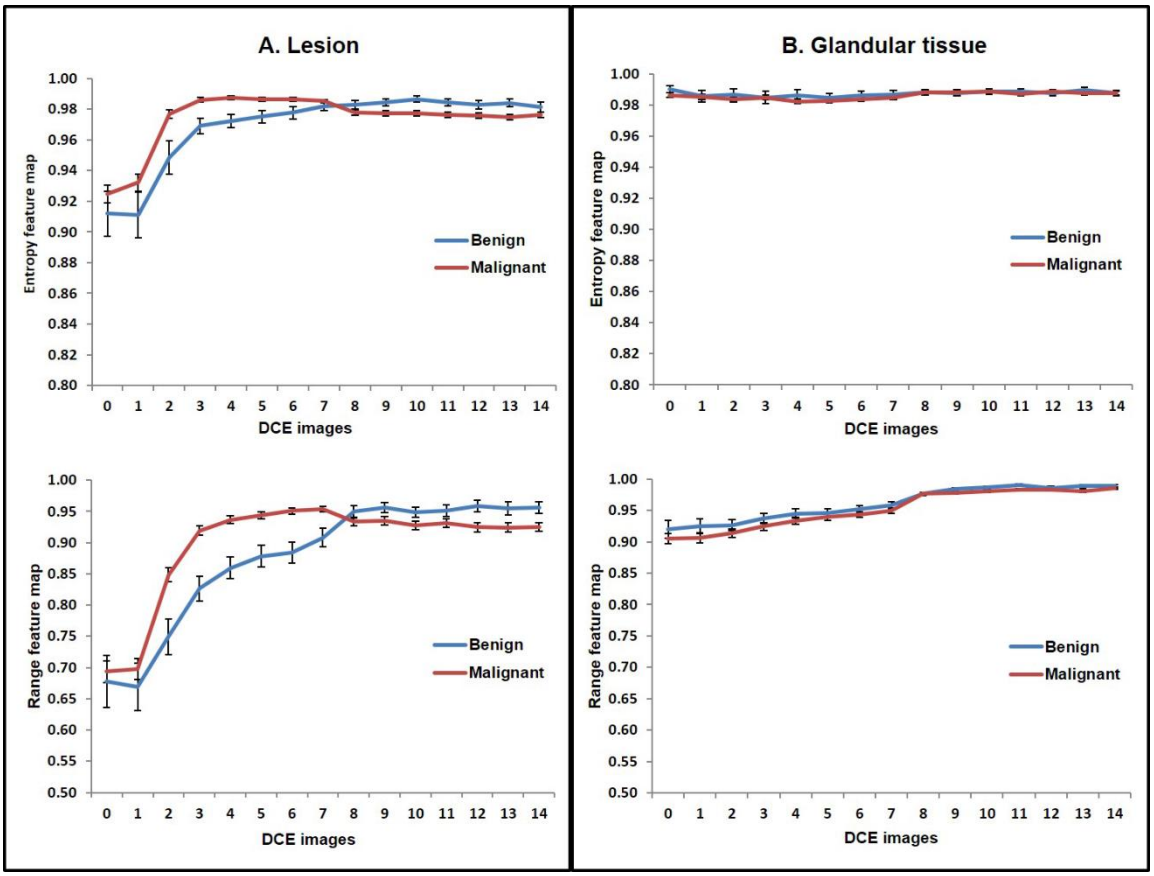


Figure 6.4. The DCE-MRI entropy evolution curves corresponding to the mean value of the entropy feature map and the range feature map. The range feature corresponds to the difference between the maximum and minimum intensity values of all the voxels within the sliding window. The error bars correspond to standard error. (Top) Normalized entropy and (bottom) range feature evolution curves. **(A)** Lesion graphs of benign (blue) and malignant (red). **(B)** Contralateral glandular tissue from benign (blue) and malignant patients (red). The shape of the radiomic feature evolution curves was significantly different between the benign and malignant lesions ($p < 0.05$). However, there was no significant difference between the contralateral glandular tissue from benign and malignant patients. Indicative of consistent radiomic features in normal tissue.

b) Analysis of textural evolution metric on DWI

I observed an increase in the first order energy of the lesion tissue from DWI-b0 to b600 with a texture evolution metric that was significantly higher ($p<0.001$) for malignant (3.09 ± 0.23) than for benign patients (1.84 ± 0.25). Similarly, the contrast in the lesion tissue also increased significantly ($p = 0.001$) for malignant (1.73 ± 0.14) than benign lesions (1.07 ± 0.13). The texture evolution metrics for five different radiomic feature maps portraying different textural characteristics have been summarized in **Table 6.2.**

Table 6.2. Summary of the texture evolution metric extracted from different RFMs DWI.

	Benign Lesion	**Malignant Lesion**	**p value**
Energy	1.84±0.25	3.09±0.23	<0.001
GLCM dissimilarity	0.98±0.06	1.26±0.05	0.001
GLCM contrast	1.07±0.13	1.73±0.14	0.001
GLCM homogeneity	1.10±0.05	0.99±0.02	0.08
First order entropy	0.93±0.02	0.97±0.01	0.12

GLCM = Gray Level Co-Occurrence Matrix

6.4.3 Breast Cancer Diagnosis

a) Multiview IsoSVM framework for feature embedding and classification

The radiomic feature maps were computed from a mpMRI dataset resulting in a high-dimensional feature space. Furthermore, the radiomic feature maps corresponding to different imaging sequences highlight different functional textural properties of the tissue of interest. Consequently, I developed a multiview feature embedding and classification framework termed IsoSVM by modifying and combining the Isomap and support vector machine (SVM) algorithms [216]. The overview of the IsoSVM framework is shown in **Figure 6.5.**

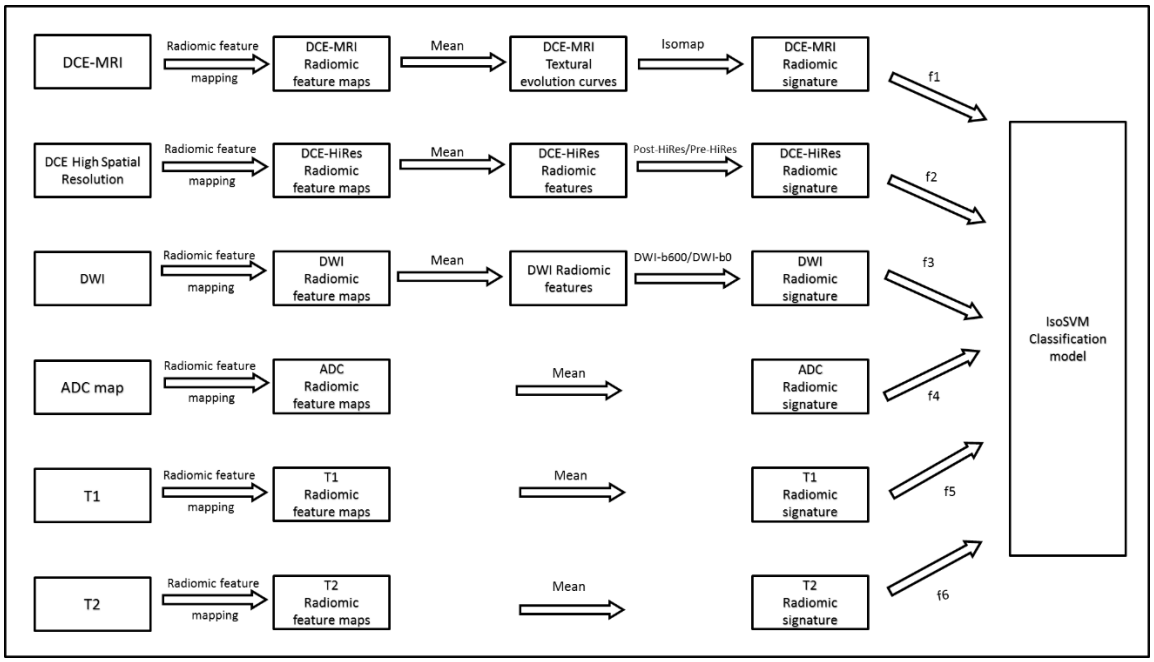


Figure 6.5 Illustration of the multiview feature embedding and classification framework. The six MRI datasets are first transformed into radiomic feature map (RFM) space using radiomic feature mapping. The RFMs for DCE-MRI are transformed into textural evolution curves, which are subsequently reduced to one-dimensional embedding using the Isomap algorithm. The vector of one-dimensional embedding corresponding to each RFM forms the 30-dimensional DCE-MRI radiomic signature. The RFMs for DCE High spatial resolution MRI and DWI are transformed into their respective radiomic signatures based on the textural evolution metric. The remaining datasets of ADC map, T1WI, and T2WI are directly transformed into radiomic signatures by calculating the mean of the RFMs. Finally, subsets of features ($f1, f2, \ldots, f6$) from each RFM signature form a unified RFM signature used to train the IsoSVM Classification model.

Computation of radiomic signatures

The high-dimensional mpMRI radiomic feature space was first analyzed to compute six different radiomics signatures as follows:

a. PK-DCE MRI radiomic signature: The textural evolution curves corresponding to the radiomic feature maps were transformed into a radiomic signature using the Isomap algorithm. For the PK-DCE RFM dataset, the 15-dimensional textural curves were transformed into a single-dimensional representation of the textural evolution curve characteristic. The correlation coefficient between the textural evolution curves of different patients was used as the distance metric to compute geodesics for the low-dimensional embedding.

b. DWI & DCE-MRI High spatial resolution radiomic signatures: The vector of the textural evolution metrics for the radiomic feature maps was used as the radiomic signature for both the datasets.

c. ADC map, T1WI, and T2WI radiomic signatures: Each one was a single image, making the vector of the mean of the radiomic feature maps, their radiomic signature.

Feature Selection

The set of radiomic signatures from the MRI datasets resulted in a 180-dimensional radiomic feature space. The 180-dimensional radiomic feature space was then transformed and modeled into an IsoSVM classification model as detailed below.

a. The feature set from each of the radiomic signatures was sorted from largest to smallest based on the area under the receiver operating characteristic curve obtained using univariate logistic regression.

b. A subset of top features $(f1, f2, \ldots, f6)$ from each of the radiomic signatures were selected to create a unified radiomic signature, $g = \cup f_i$

c. The unified radiomic signature, g was then transformed into a linearly separable, low-dimensional feature space, h using the Isomap algorithm. The feature transformation was executed using Isomap because the Isomap algorithm is not prone to overfitting because of its unsupervised nature and at the same time, accounts for the dependencies between different RFMs.

Classification

i. The support vector machine algorithm trains a classification model to classify between benign and malignant patients on the transformed feature space, h. Because SVM is a linear binary classification algorithm that attempts to create a hyperplane that best separates the different groups, the application of Isomap algorithm prior to SVM reduces the non-linearity in the data by transforming the feature space, g to h. The steps c and d combine to form the hybrid IsoSVM classification model. Mathematically, the hybrid IsoSVM classifier is represented using the following equation:

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i < \phi(x_i), \phi(x) > + b$$

where $\varphi()$ is the Isomap transformation function that maps the unified radiomic signature, g into a linearly separable space, h, N is the number of patients in the training set, α_i are the Lagrange multipliers, x_i represents the radiomic signatures of training set patients and x represents the radiomic signature of the test patient.

ii. For comparison, I tested five different SVM kernels on the unified radiomic signature, g including the hybrid IsoSVM kernel to classify the benign and malignant patients to determine the optimal kernel.

iii. Finally, due to class imbalance between the number of benign and malignant patients, the ratio between penalty for misclassification of different patient data sets was varied to identify the optimal penalty ratio and shown in the supplementary data [217].

The complete set of input parameters (input feature space: $f1, f2, \dots, f6$; Isomap neighborhood parameter, k; Isomap dimensionality, d; and misclassification penalty ratio) were estimated using leave-one-out and k fold cross validation ($k = 10$).

K-fold Cross validation

The set of benign and malignant patients were first separately divided into ten randomly sampled subsets due to an imbalance in the number of patients in each patient group. Next, the ten subsets from both categories were combined to form ten patient subsets. As a result, the ratio between the number of benign and malignant patients was maintained similar to the original patient cohort in the patient subsets. The ten-fold cross validation procedure was performed on these ten subsets. The complete procedure of generating ten subsets and performing ten-fold cross validation was repeated 100 times to avoid any bias that may occur due to specific partitioning of the data.

Receiver Operating Characteristic analysis

The value of predicted outcome using IsoSVM is given by $y = sign(w^T x + b)$, where w and b are the optimized values of the weights and intercept of the hyperplane, respectively. In order to plot the receiver operating characteristic (ROC) curve, I modified this function as follows:

$$y = \begin{cases} 1 & if\ w^T x + b \geq \eta \\ 0 & otherwise \end{cases}$$

Instead of $\eta = 0$, I varied the value of η to generate different values of sensitivity and specificity to plot on the ROC curve. The final values of sensitivity and specificity are still considered as the values corresponding to $\eta = 0$.

Total Operating Characteristic analysis

I used total operating characteristic (TOC) curves in addition to the ROC curves for assessing the diagnostic performance of the IsoSVM classifier [215]. The advantage of TOC over ROC curves is that the TOC curves allow us to visualize and compute true positive (TP), false positives (FP), true negatives (TN), and false negatives (FN) for each threshold directly from the curves.

The y-axis of a TOC plots TP and goes from zero to the percentage of malignant patients in the patient population. On the other hand, the x-axis of a TOC plots TP+FP and goes from zero to one (total population). Based on the number of benign and malignant patients in the patient population, a parallelogram is drawn to define the mathematically possible space for the TOC curve. At any point on the TOC curve, the values for TP, FN, TN, and FP can be inferred as follows. The TP corresponds to the value of the point's y-coordinate. The FP is calculated by subtracting TP from the value of the point's x-coordinate. The FN is calculated by subtracting TP from the maximum possible value of the y-axis. The TN is calculated by subtracting FP from the width of the parallelogram (percentage of benign patients in the patient population).

b) Results

Multiview feature embedding and classification

The multiview feature embedding and classification framework was set up as illustrated in **Figure 6.5 (Page 134)**. The optimal set of hyperparameters for the multiview classification framework, obtained using leave-one-out-cross-validation based grid search, were $f1 = 18, f2 = 0, f3 = 6, f4 = 8, f5 = 0, f6 = 0$, neighborhood parameter $k = 45$, dimensionality $d = 10$, and $misclassification\ penalty\ ratio = 2.5{:}\ 1$. The parameter space for each of the input parameters was set as follows:

a. The subset of features, f_i selected from each MRI dataset were iteratively selected based on the area under the ROC curve computed using univariate logistic regression.
b. The neighborhood parameter was varied from 5 through 120 in steps of 5.
c. The dimensionality of the transformed feature space was varied between one and ten.
d. The misclassification penalty ratio between benign and malignant classes was selected from the set $\{2{:}\ 1, 2.5{:}\ 1, 3{:}\ 1, 3.5{:}\ 1, 4{:}\ 1\}$.

The multiview feature embedding and classification model trained using leave-one-out cross validation resulted in sensitivity and specificity of 93% and 85% respectively with an AUC of 0.91 in classifying benign from malignant lesions. The ROC curves for the IsoSVM classification model and other kernels are shown in **Figure 6.6.** The TOC curve for the IsoSVM classification model is shown in **Figure 6.7**. The search space for the misclassification penalty parameter for the SVM kernels was increased to all the ratios in the set $\{2{:}\ 1, 2.5{:}\ 1, 3{:}\ 1, 3.51, 4{:}\ 1, 4.5{:}\ 1, 5{:}\ 1, 5.5{:}\ 1, 6{:}\ 1\}$ The resultant sensitivity, specificity, and AUC from all the SVM classifiers are shown in **Table 6.3**. The multiview feature transformation and classification framework was further tested using ten-fold cross

validation performed across 100 trials. The optimal set of hyperparameters obtained with ten-fold cross validation concurred with the previously obtained optimal set of hyperparameters using leave one out cross validation. The average sensitivity and specificity achieved from ten-fold cross validation experiment were 91% and 82% respectively with an AUC of 0.87. The result from ten-fold cross validation ascertains the stability of the unified RFM signature as well as the IsoSVM classifier. For comparison, the classification of benign from malignant using tumor size alone produced an AUC of 0.77 which was significantly lower than the AUC of 0.91 obtained from the unified RFM signature using the IsoSVM classifier.

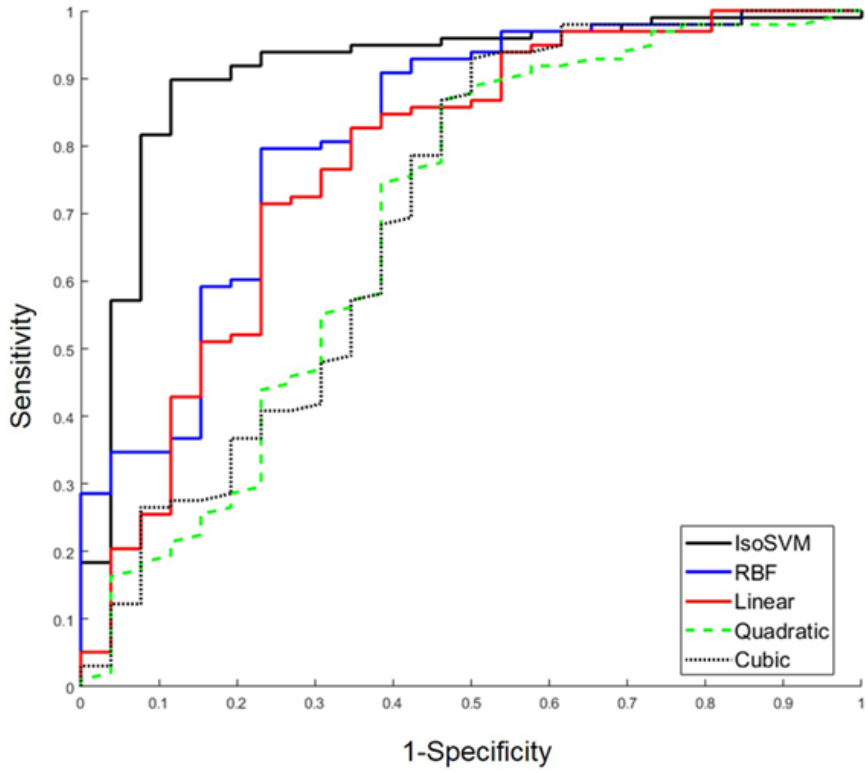


Figure 6.6 The receiver operating characteristic curves corresponding to the IsoSVM classification (black), radial basis function (RBF) kernel SVM (blue), linear kernel SVM (red), quadratic kernel SVM (dashed green), and the cubic kernel SVM kernel (dotted black) evaluated using leave one out cross validation. The area under the ROC were obtained at 0.91, 0.82, 0.78, 0.65, and 0.71 for IsoSVM, RBF, linear, quadratic, and cubic kernel SVMs, respectively.

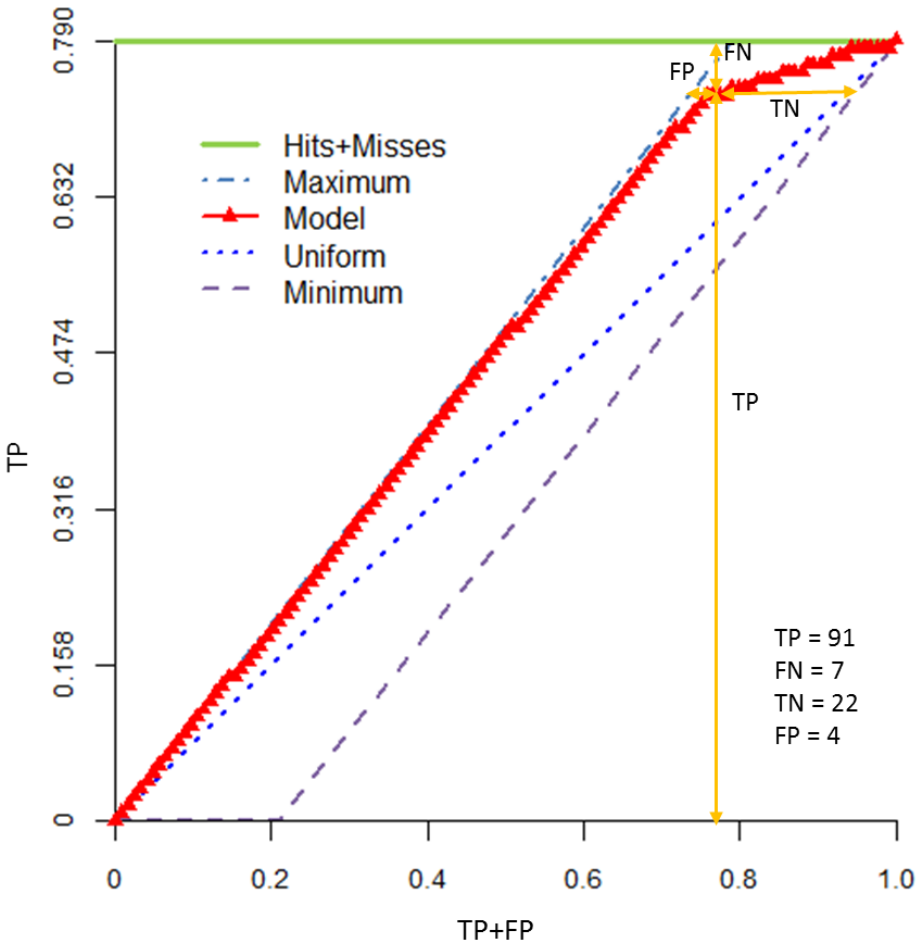


Figure 6.7 The total operating characteristic (TOC) curve for the IsoSVM classifier. The TOC curve allows us to compute the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) at every point on the curve. The TP, FN, TN, and FP for the optimal IsoSVM hyperplane were 91, 7, 22, and 4, respectively.

Table 6.3. Summary of sensitivity, specificity, and AUC for the IsoSVM classifier and various SVM kernels.

Classifier	Input parameters	Sensitivity	Specificity	AUC
IsoSVM	$k = 45;\ d = 10;\ PR = 2.5:1$	0.93	0.85	0.91
Radial basis function SVM	$sigma = 19;\ PR = 5:1$	0.80	0.77	0.82
Linear SVM	$PR = 3.5:1$	0.85	0.62	0.78
Quadratic SVM	$PR = 5.5:1$	0.85	0.54	0.65
Cubic SVM	$PR = 5:1$	0.93	0.50	0.71

AUC = Area under the Curve
IsoSVM=Hybrid Isomap and Support Vector Machine
PR= The misclassification penalty ratio between benign and malignant classes

6.5 Discussion

I have demonstrated that the radiomic feature maps for visualization and evaluation of radiological texture in radiological images produced excellent features that were correlated to breast tissue biology and compared with quantitative metrics of radiological parameters. Malignant lesions demonstrated increased entropy compared to benign lesions for both ADC maps and DCE MRI. In contrast, glandular tissue entropy was similar across all subjects. Furthermore, the radiomic feature maps (RFMs) demonstrated excellent sensitivity and specificity in classifying benign from malignant lesions. Moreover, this is the first study to relate the quantitative metrics of ADC and PK-DCE to radiomics values. In addition, I also demonstrated that multiparametric radiomic features distinguished responders from non-responders with excellent sensitivity and specificity.

Radiomic features such as entropy have been shown to classify between benign and malignant tumors in addition to predicting patient survival and treatment response in previous studies as reviewed in **Chapter 5**. However, my work explored the whole image visualization and interpretation of these quantitative radiomic values employing RFMs. Indeed, the entropy feature maps exhibit higher entropy and intra-tumor heterogeneity for malignant tumors compared to benign tumors. The RFMs would provide the radiologists with a tool for visual interpretation of the radiomic feature values. Furthermore, radiomic feature maps provide a visualization of intra-tumor heterogeneity as opposed to a single quantitative value provided by quantitative radiomic analysis. In addition, radiomic feature maps produce voxel-wise radiomic values improving the quantitative measure. In contrast, single quantitative value corresponding to the whole tumor region may not define the entire tumor. This study investigated the relationship between RFMs and underlying tissue biology derived from the radiological images. Preliminary analysis of RFMs corresponding to DCE-MRI suggest that time evolution of RFMs is indicative of heterogeneity in the vasculature of tissue of interest. The textural evolution curves obtained from mean value of the radiomic feature maps had significantly different curve characteristics for benign and malignant tumors. Furthermore, the curve characteristic for the glandular tissue corresponding to benign and malignant patients demonstrated no shape difference indicating there is no textural evolution with contrast uptake within glandular tissue. The radiomic features provided new metrics for comparison of the different tissue types. Moreover, the vascular parameters of K_{trans} and EVF have been shown to be different between benign and malignant tumors. In concordance, the radiomic values also demonstrated significant differences. In the previous studies, the ADC value and PK-DCE

for a given region of interest have been established as excellent biomarkers in classification between benign and malignant breast tumors [57, 59, 218-223]. Here, I established radiomic entropy (and others) of the ADC map and DCE-MRI within the tumor ROI as a biomarker for correlation with cellular and vascular heterogeneity. The ADC and DCE entropy values were significantly different between benign and malignant tumors. Furthermore, the entropy ADC feature map provides more insight into the cellular distribution within the tumor, whereas, the DCE radiomic metric provides information about the vascularity texture of the tissue. Additionally, a metric for quantification of tissue heterogeneity evolution with increasing b value was developed and analyzed. A subset of the texture evolution metrics for DWI were significantly different between benign and malignant lesions indicating a potential biomarker in the texture evolution metric. In addition, the glandular ADC and radiomics values were similar across all subjects. These findings lay the groundwork in radiomic metrics to describe normal vs abnormal tissue, which is needed for increased use in clinical applications.

The training efficacy of most machine learning algorithms depend on the balance between the number of instances corresponding to each class. Typically, benign breast tumors are more frequently observed in clinical setting as compared to malignant tumors. However, in research setting, MRI for malignant breast tumors are more frequently obtained than for benign breast tumors producing a class imbalance that may result in performance bias of the trained classifier towards one class. Class imbalance is a frequent occurrence in health care machine learning applications. My work approached the problem of class imbalance by assigning different misclassification penalty to each class type. My

results indicate that setting an appropriate misclassification penalty significantly improves the classification accuracy.

This work has certain limitations. First, the radiomic feature maps were created and classified on a retrospective data and no separate validation data was used. Second, this study evaluated radiomic feature maps corresponding to only first and second order statistical features. Other statistical radiomic methods such gray level run length matrix features, Neighborhood gray tone difference matrix feature have not been evaluated in this study. Third, RFMs provide voxel wise heterogeneity information of the whole tissue of interest. However, the feature used in the texture evolution curves and classification model was the mean derived from the RFMs. In addition, the radiomic features were extracted corresponding to each radiological image in the mpMRI datasets, which could lead to a large number of features and potential overfitting. I discuss this issue in **Chapter 7** and develop a method for multiparametric radiomic feature extraction.

In summary, RFMs present a new powerful tool for analysis of textural information present within radiological images and may provide a new perspective into the biological information radiomics is capable of providing and the potential it holds in future diagnostic applications.

APPENDIX

6.A List of first and second order statistical features extracted using radiomic feature mapping

Table 2.A.1 List of all the features used for extraction of radiomic feature maps. The mathematical formulation of these features can be found in [30].

S. No.	First Order Statistics Features
1	Entropy
2	Energy
3	Kurtosis
4	Range
5	Skewness
6	Uniformity
7	Variance
	Grey Level Co-occurrence Matrix Features
1	Autocorrelation
2	Cluster Prominence
3	Cluster Shade
4	Cluster Tendency
5	Contrast
6	Correlation
7	Difference Entropy
8	Dissimilarity
9	Energy
10	Entropy
11	Homogeneity 1
12	Homogeneity 2
13	Inverse Difference Moment Normalized
14	Inverse Difference Normalized
15	Informational Measure of Correlation 1
16	Informational Measure of Correlation 2
17	Inverse Variance
18	Maximum Probability
19	Sum Average
20	Sum Entropy
21	Sum Variance
22	Variance
1	Laplacian of Gaussian

Chapter 7. Multiparametric Radiomics

7.1 Introduction

I introduced the radiomic feature mapping for radiomic analysis of mpMRI with improved visualization and interpretability of feature metrics in **Chapter 6**. However, the RFM framework is still based on extraction of textural patterns from a single image or volume and is not capable of extraction of textural patterns from multimodal or multiparametric datasets consisting of multiple image sequences. For example, multiparametric magnetic resonance imaging (mpMRI) produces soft tissue contrasts of the tissue, where each imaging sequence (e.g., proton density, T2, T1, diffusion, and perfusion weighted imaging (WI)) provides a specific representation of the tissue based on the underlying physics and gray level. Integrating the imaging information from different imaging modalities and parameters provides a more holistic view of the underlying tissue characteristics. Correspondingly, texture analysis on a high-dimensional multi-sequence data would provide information about the "true texture" of the tissue rather than from a specific point of view. To that end, I developed a multiparametric imaging radiomics (MPRAD) framework for extracting radiomics information from multiparametric and multimodal imaging data.

In a multiparametric setting, tissue signatures (TS) encode the characteristics of a tissue instead of a grayscale value. Moreover, different imaging parameters that form the tissue signature interact with each other in a complex high-dimensional space forming a complex interaction network. Probing the complex interaction network could provide information that was not possible to extract using conventional radiomic features. The

MPRAD framework analyzes both the spatial distribution of TS in addition to the complex interaction network within a region of interest (ROI) to compute multiparametric imaging radiomic features. In this chapter, I present the theory of the MPRAD framework, develop radiomic features for multi-sequence images and implement the MPRAD framework in two different applications. First, I applied the MPRAD framework to multiparametric breast MRI for classification of benign from malignant lesions and compare the obtained results to single image radiomics. Next, the MPRAD framework was applied to multiparametric brain MRI for classification of diffusion perfusion mismatch in stroke patients [224-228].

Publication from this work and author contributions

Publication

V. S. Parekh and M. A. Jacobs, "MPRAD: A Multiparametric Imaging Radiomics Framework," Submitted, Nature Communications, 2018

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

I developed the concept of multiparametric radiomics, all the four methods discussed in this chapter, and their corresponding mathematical formulae. I did the entire implementation for extraction and visualization of multiparametric radiomics from multiparametric radiological images. I performed all the data analysis described in this chapter.

7.2 Multiparametric Radiomic Framework

7.2.1 Tissue Signature

The concept of tissue signature was previously established in **Chapter 2** for the development of the multiparametric deep learning (MPDL) tissue signature model. Here, I only provide a brief overview of tissue signatures as a refresher. In multiparameter imaging settings, a tissue signature represents a composite feature representation of a tissue type based on the physical modeling of different imaging parameters acquired. For N different imaging parameters, a tissue signature at a voxel position, p is defined as a vector of gray level intensity values at that voxel position across all the images in the data sequence and given by the following equation

$$S_p = \left[I_p^{(1)}, I_p^{(2)}, I_p^{(3)}, \ \ldots, I_p^{(N)}\right]^T$$

7.2.1 Tissue Signature Probability Matrix Features

The tissue signature probability matrix (TSPM) characterizes the spatial distribution of tissue signatures in an ROI. The mathematical formulation of TSPM is as follows:

Suppose that the intensity values representing each voxel are quantized to G levels, then the total number of possible tissue signatures in a dataset consisting of N images will be equal to G^N. We define a function $f: T \rightarrow M$, where T is the set of all tissue signatures in the dataset and M is a N-dimensional matrix with edges of length G where each tissue signature is represented as a cell. The function f populates each cell of the matrix M with the frequency of occurrence of the corresponding tissue signature in the set T. The matrix M is called the tissue signature probability matrix (TSPM).

The information content of the N-dimensional multiparametric imaging dataset $(X_1, X_2, \dots X_N)$ can be analyzed by computing the joint entropy, uniformity and mutual information of the resultant TSPM [229]. These are defined below.

1. The TSPM entropy, H is given by the following equation

$$H(X_1, X_2, \dots X_N) = -\sum_{i_1=1}^{N_g}\sum_{i_2=1}^{N_g}\dots\sum_{i_N=1}^{N_g} TSPM(i_1, i_2, \dots, i_N)\log_2 TSPM(i_1, i_2, \dots, i_N)$$

2. The TSPM uniformity, U is given by the following equation

$$U(X_1, X_2, \dots X_N) = \sum_{i_1=1}^{N_g}\sum_{i_2=1}^{N_g}\dots\sum_{i_N=1}^{N_g} TSPM(i_1, i_2, \dots, i_N)^2$$

3. The TSPM mutual information, I is given by

$$I(X_1; X_2; \dots; X_N) = \big(H(X_1) + H(X_2) + \cdots + H(X_N)\big) - \cdots + \cdots (-1)^{N-1} H(X_1, X_2, \dots, X_N)$$

By choosing different possible subsets $Y \subseteq \{X_1, X_2, \dots, X_N\}$, different values of $H(Y)$, $U(Y)$ and $I(Y)$ can be obtained producing a large number of multiparametric imaging radiomic features.

7.2.3 Tissue Signature Co-occurrence Matrix Features

The tissue signature co-occurrence matrix (TSCM) characterizes the spatial relationship between tissue signatures within an ROI. The TSCM is defined similar to the gray level co-occurrence matrix (GLCM) using two input parameters, distance (d) and angle (θ) between the two tissue signature locations. Mathematically the GLCM between any two tissue signatures, S_i and S_j is given by the following equation

$$GLCM_d^{\theta}(S_i, S_j, m, n) = \big|\{r : S_i(r) = m, S_j(r) = n\}\big| \; \forall\, m, n \in \{1,2,3, \dots, G\}$$

where $r \in N$ $(number\ of\ imaging\ sequences)$ and $|\,.\,|$ denotes the cardinality of a set.

Given a distance, d and angle, (θ), the co-occurrence matrix for all such possible pairs of tissue signatures is given as follows:

$$TSCM_d^\theta(m,n) = \Sigma_{i,j} GLCM_d^\theta(S_i, S_j, m, n)$$

$$\forall\ i,j\ satisfied\ by\ d\ and\ \theta$$

Here, $TSCM_d^\theta$ is the tissue signature co-occurrence matrix. The TSCM can then be analyzed to extract twenty-two different TSCM features using the equations developed by Haralick et al. [26].

7.2.4 Tissue Signature Complex-Interaction Network Analysis Features

The tissue signature complex interaction network (TSCIN) analysis characterizes the complex interactions that define the inter-parametric relationships between different imaging parameters using statistical analysis. The TSCIN features are extracted by transforming a high-dimensional multiparametric radiological imaging data into a radiomic feature map using first or higher order statistical analysis of the tissue signature vectors, S_p at each voxel position.

The TSCIN feature maps are transformed into a single radiomic quantitative value corresponding to a region of interest using a summary statistical metric such as mean, median or standard deviation.

a) First order TSCIN features

The first order TSCIN features are straightforward and calculated directly from the tissue signature. For example, the TSCIN entropy at a voxel position, p is given by the following equation:

$$Entropy_{TSCIN} = entropy\ (S_p)$$

Similarly, all the other first order TSCIN features follow.

b) Second order TSCIN features

The second order TSCIN features characterize the inter-parameter relationships within the tissue signature by computing a TSCIN relationship matrix (TSRM). Mathematically, TSRM for a N-dimensional tissue signature at voxel position, p with N imaging sequences quantized to G gray levels is given by the following equation:

$$TSRM_d^p(i,j) = \left|\left\{k: I_p^{(k)} = i, I_p^{(k+d)} = j\right\}\right| \ \forall\, i,j \in \{1,2,3,\dots,G\}, k \in \{1,2,\dots,N\}$$

Here, d represents the distance between the two imaging parameters, $I^{(k)}$ and $I^{(k+d)}$.

The TSRM is dependent on the relative location of different imaging parameters within the tissue signature. As a consequence, TSRM requires the input imaging series to have an intrinsic relationship between different imaging sequences, for example, pharmacokinetic dynamic contrast enhanced (PK-DCE) imaging and diffusion weighted imaging (DWI) sequences. The structure of the TSRM is similar to a $G \times G$ gray level co-occurrence matrix, thereby allowing us to utilize all the twenty-two equations established to extract relevant features from such matrices [26].

All the four classes of features developed in this manuscript have been illustrated in **Figure 7.1** on an example multiparametric breast MRI dataset.

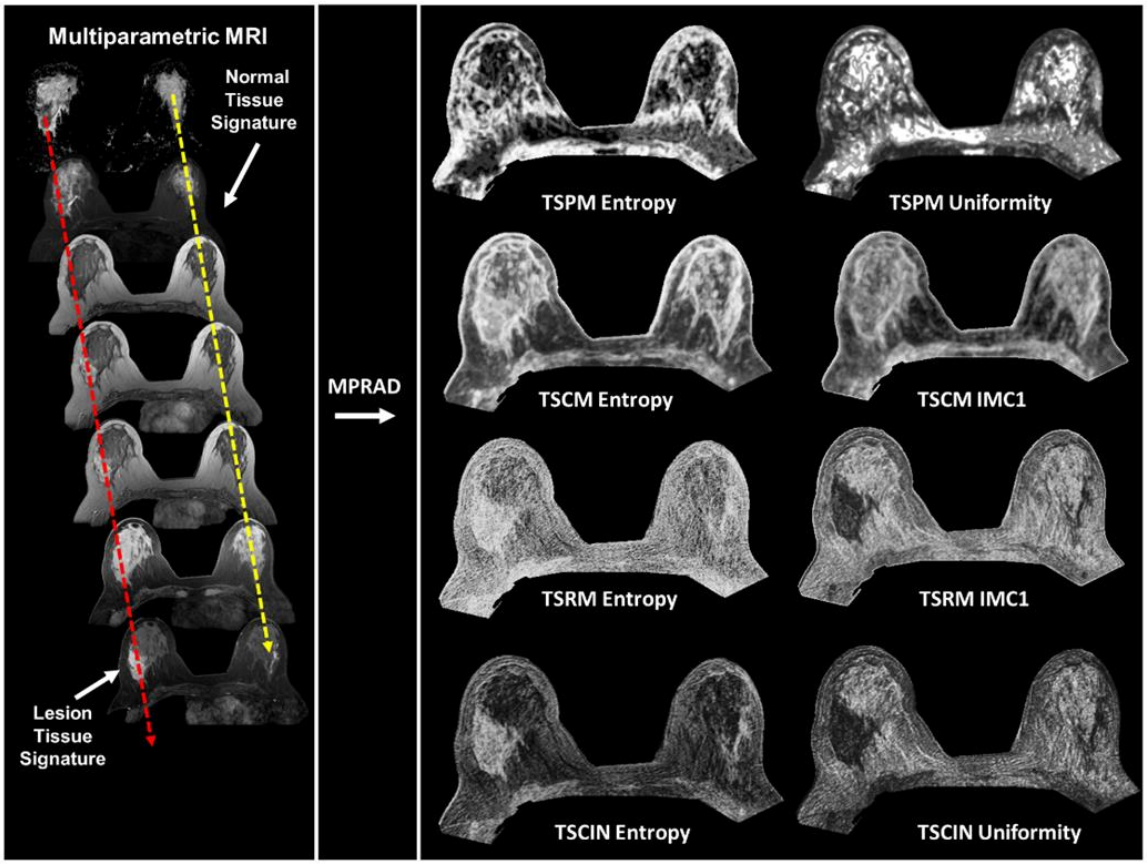


Figure 7.1 Illustration of the four different types of multiparametric imaging radiomic features based on the first and second order statistical analysis. The tissue signature probability matrix (TSPM) and tissue signature co-occurrence matrix (TSCM) features are based on the spatial relationship between tissue signatures while the tissue signature complex interaction network (TSCIN) and tissue signature relationship matrix (TSRM) features evaluate the inter-parameter complex interactions. A typical tissue signature is demonstrated using a yellow arrow that runs through all the images in the multiparametric imaging dataset and tumor in the images is indicated using an orange arrow.

7.3 Experiments

7.3.1 Digital Phantom

a) Experimental Setup

The multiparametric imaging radiomic feature extraction methods developed in this manuscript were first tested on a digital phantom shown in **Figure 7.2A and 7.2B.** The texture phantom images (**Figure 7.2A1 and A2**) consists of a composite mixture of several raw texture images of grass, sand, wool, water, and others derived by Brodatz [230]. Using these composite images ground truth texture images can be determined to demonstrate the effectiveness of any radiomic method [28, 177]. The ground truth images are shown in the top row in **Figure 7.2A1 and A2**. I stacked a series of these texture phantom images to demonstrate the effectiveness of the MPRAD tissue signature model to accurate segment each of the different textures. I applied single image radiomics to each image and MPRAD to stacked images to compare the results from the two methods. The radiomic parameters of neighborhood and gray level quantization were set to 15x15 and 256 gray levels, respectively.

a) Results

The MPRAD features on each USC composite texture image are shown **in Figure 7.2.** The reference texture ground truth images provide a method to evaluate the radiomic features from known objects with high texture. From the two composite images, the MPRAD feature of TSCM mutual information (IMC1) was able to produce a 100 percent match with the reference images, confirming the method on an independent data set. Moreover, the multiparametric radiomic features were able to capture the differences in the shape and intensity distribution of both single parameter radiomic images with excellent detail of the underlying structure. The higher order entropy (GLCM) values for the highlighted

subregion in **Figure 7.2** were 9.9 and 9.7 for the single radiomic images corresponding to the square and mosaic, respectively while the MPRAD TSCM entropy was 10.23.

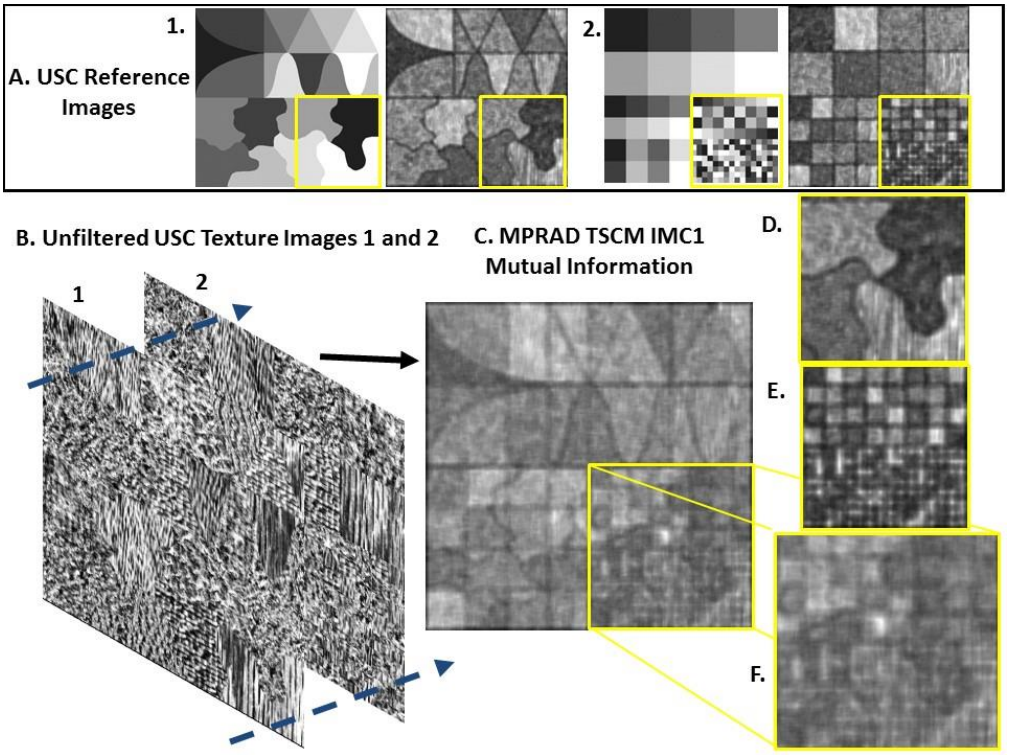


Figure 7.2 A. USC reference texture ground truth images. 1. Reference image made out of a composite of several different shapes and textures and (right) single radiomic image. 2. The composite Reference image and (right) single radiomic image **B.** Multiparametric USC composite images. **C.** mpRadiomics image of USC images. **D and E.** Enlarged radiomic images from reference images 1 and 2. **F.** Enlarged mpRadiomic image from the combination of the images. The multiparametric radiomic features were able to capture the differences in both shape and intensity distribution of both single parameter radiomic images with excellent detail of the underlying structure.

7.3.2 Multiparametric Breast MRI Dataset

a) Experimental Setup

Multiparametric breast MRI dataset consisted of a cohort of 138 patients (Age = 52±11) to classify between malignant and benign lesions. Of the 138 patients, there were 97 patients with biopsy proven malignancy whereas 41 patients had benign lesions. All the studies were performed in accordance to the institutional guidelines for clinical research under IRB approved protocol. MRI scans were performed on a 3T magnet, using a dedicated phased array breast coil with the patient lying prone with the breast in a holder to reduce motion. The methods for MRI Image Processing, registration, segmentation, and classification have all been detailed in **Chapter 2**. For radiomic analysis, the radiological images were quantized to 128 gray levels and the neighborhood was set to 5x5.

b) Results

Figure 7.3 illustrates both the single and MPRAD feature maps from a patient with a malignant lesion in the upper outer quadrant of the right breast with a benign appearing cyst superior and more medial to the lesion (curved yellow arrow). The cyst is uniformity bright on T2 and the ADC map consistent with known tissue characteristics associated with cysts. Similarly, the cyst is dark on T1 and no contrast enhancement on the DCE image indicating no vascularity. Moreover, the lesion appears to heterogenous on all the MRI images with decreased ADC and increased DCE characteristics. The single radiomic images do show some texture features, however, there is a striking difference in the textural representation shown by the MPRAD radiomics. In particular, the cyst is shown with decreased entropy in the MPRAD compared to single radiomic images. The lower entropy in the cyst is consistent with the fact, that the homogenous object has less disorder and

hence lower entropy. This is clearly evident when looking at the lesion which has much higher entropy values.

Figure 7.4 illustrates both the single and MPRAD feature maps from a benign patient. There was a clear difference between the textural representation of the lesion and glandular tissue. Furthermore, the tissue characterization of lesion and glandular tissue was consistent for both the benign and the malignant patient. **Table 7.1** summarizes the quantitative values from single parameter and TSPM entropy for individual and MPRAD features on benign and malignant patient cohorts demonstrating improved tissue characterization using MPRAD. The MPRAD TSPM entropy computed using all the MRI parameters was significantly different between benign and malignant lesions (Benign: 7.06 ± 0.27, Malignant: 8.93 ± 0.17, $p < 0.00001$). Furthermore, the univariate AUC of TSPM entropy was 0.82, 9% higher than the maximum AUC (0.75 for post contrast DCE) obtained from univariate analysis of first order entropy computed from different imaging parameters. More importantly, there were no significant differences between the single and multiparametric radiomic features values in the contralateral glandular tissue from benign and malignant patients as shown in **Table 7.2.**

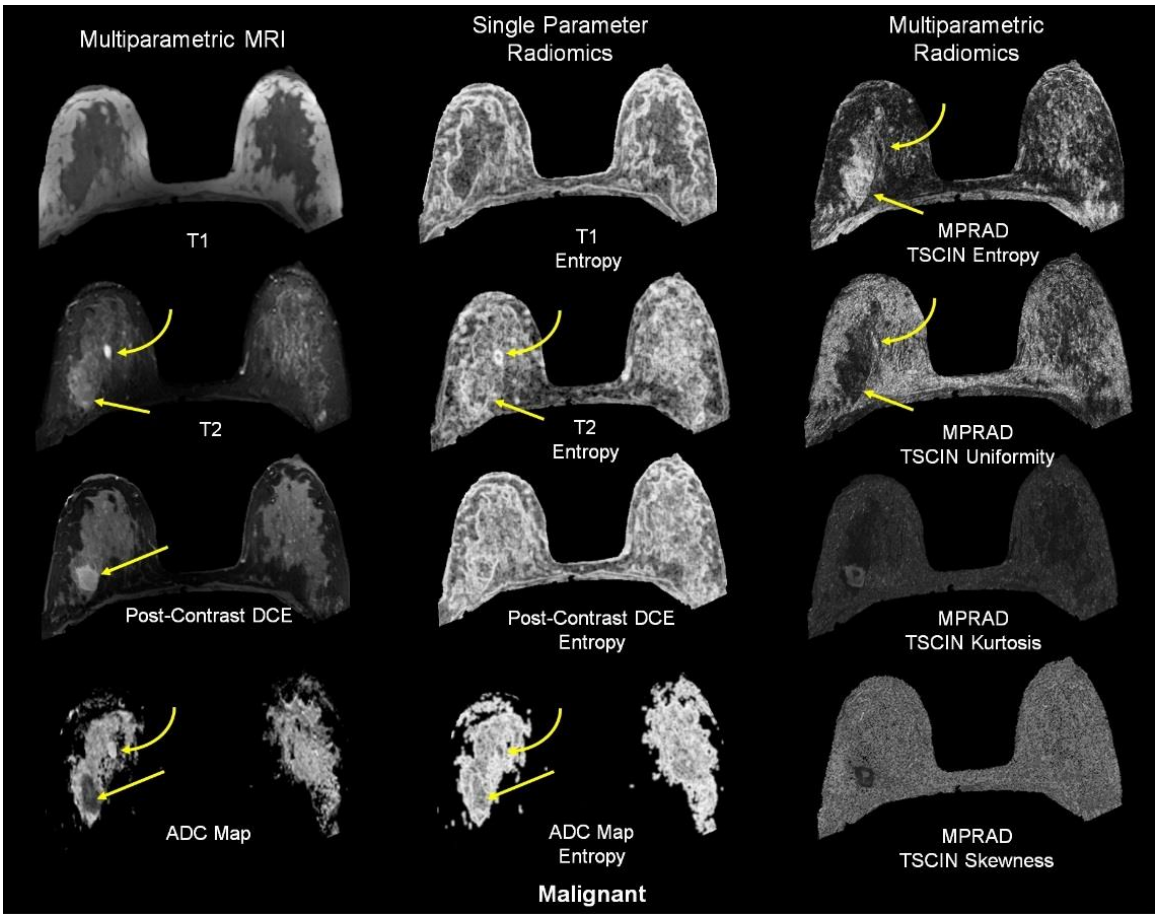


Figure 7.3. The radiomic feature maps (RFM) obtained from single and multiparametric radiomic analysis in a malignant patient. The straight yellow arrow highlights the lesion. The curved arrow demonstrates a benign cyst. **A.** Multiparametric MRI parameters **B.** Single radiomic gray level co-occurrence matrix (GLCM) entropy features maps of each MRI parameter. **C.** The MPRAD RFMs tissue signature co-occurrence matrix (TSCM) and tissue signature complex interaction network (TSCIN) radiomic features. Note, the improved tissue delineation between the different tissue types using MPRAD.

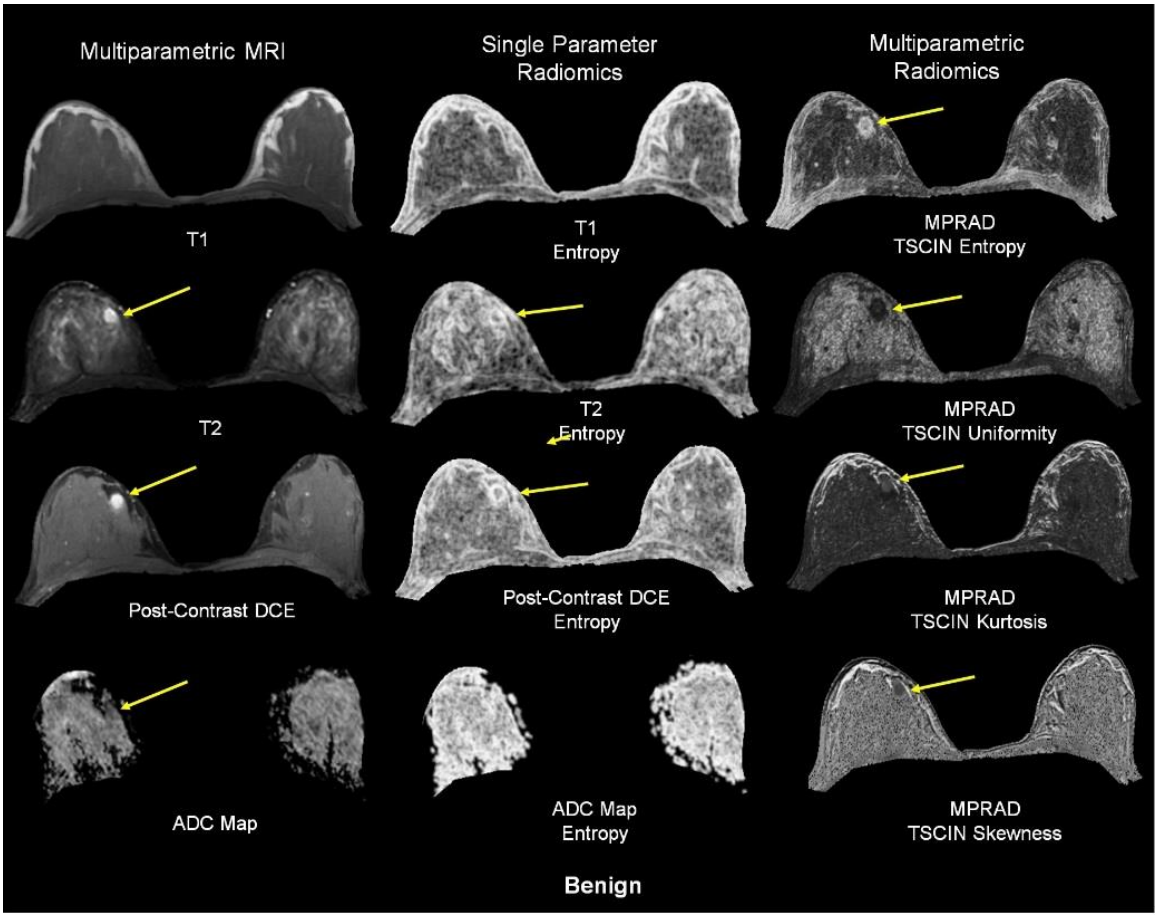


Figure 7.4. The radiomic feature maps (RFM) obtained from single and multiparametric radiomic analysis in a benign patient. The yellow arrow highlights the lesion. A. Multiparametric MRI parameters B. Single radiomic gray level co-occurrence matrix (GLCM) entropy features maps of each MRI parameter. C. The MPRAD RFMs tissue signature co-occurrence matrix (TSCM) and tissue signature complex interaction network (TSCIN) radiomic features.

Table 7.1. Summary of single and multiparametric entropy values corresponding to benign and malignant breast tumors

Single Parameter Entropy	Benign Tumor	Malignant Tumor	p Value	AUC
Entropy T1	4.14±0.11	4.66±0.06	0.00008	0.72
Entropy T2	4.98±0.12	5.42±0.06	0.002	0.68
Entropy b0	4.44±0.17	5.06±0.09	0.002	0.67
Entropy b600	3.00±0.20	3.77±0.09	0.0009	0.67
Entropy ADC	4.90±0.12	5.40±0.06	0.0004	0.70
Entropy Post-Contrast DCE (High Spatial Resolution)	5.00±0.10	5.54±0.05	0.00001	0.75
Entropy PK-DCE Pre	4.32±0.12	4.65±0.05	0.02	0.62
Entropy PK-DCE Post (wash-in)	4.89±0.08	5.30±0.05	0.00006	0.72
Entropy PK-DCE Post (wash-out)	4.90±0.09	5.24±0.04	0.00007	0.69
Multiparametric Entropy				
TSPM Entropy (all Parameters)	7.06±0.27	8.93±0.17	P<0.00001	0.82
TSPM Entropy (PK-DCE)	7.06±0.27	8.92±0.17	P<0.00001	0.82
TSPM Entropy (High Spatial Resolution DCE)	6.74±0.19	8.28±0.12	P<0.00001	0.82
TSPM Entropy (DWI)	6.66±0.22	8.20±0.15	P<0.00001	0.78

DWI: Diffusion Weighted Imaging
ADC: Apparent Diffusion Coefficient
PK: Pharmacokinetic
DCE: Dynamic Contrast Enhancement
FOS: First Order Statistics
TSPM: Tissue Signature Probability Matrix

Table 7.2. Summary of single and multiparametric entropy values corresponding to contralateral glandular tissue in patients with benign and malignant breast tumors

Single Parameter Entropy	Glandular Tissue Benign Patients	Glandular Tissue Malignant Patients	p Value
Entropy T1	5.29±0.11	5.12±0.06	0.20
Entropy T2	5.37±0.10	5.32±0.06	0.68
Entropy b0	5.19±0.24	4.89±0.10	0.27
Entropy b600	3.46±0.24	3.13±0.10	0.20
Entropy ADC	5.27±0.28	5.39±0.16	0.71
Entropy Post-Contrast DCE (High Spatial Resolution)	5.13±0.10	5.00±0.06	0.26
Entropy PK-DCE Pre	5.24±0.12	5.12±0.05	0.38
Entropy PK-DCE Post (wash-in)	5.28±0.11	5.18±0.05	0.40
Entropy PK-DCE Post (wash-out)	5.30±0.10	5.24±0.05	0.60
Multi-sequence entropy			
TSPM Entropy (all Parameters)	10.93±0.34	10.64±0.17	0.46
TSPM Entropy (PK-DCE)	10.92±0.34	10.64±0.17	0.47
TSPM Entropy (High Spatial Resolution DCE)	9.17±0.17	9.04±0.10	0.51
TSPM Entropy (DWI)	9.31±0.35	9.06±0.18	0.54

DWI: Diffusion Weighted Imaging
ADC: Apparent Diffusion Coefficient
PK: Pharmacokinetic
DCE: Dynamic Contrast Enhancement
FOS: First Order Statistics
TSPM: Tissue Signature Probability Matrix

Table 7.3 Top multiparametric radiomic features for classification of malignant from benign breast tumors.

MIRAD Radiomic feature	Benign Tumor	Malignant Tumor	p Value	AUC
TSPM entropy (all parameters)	7.06±0.27	8.93±0.17	p<0.00001	0.82
TSPM entropy (DCE)	7.06±0.27	8.92±0.17	p<0.00001	0.82
TSPM entropy (HiRes)	6.74±0.19	8.28±0.12	p<0.00001	0.82
TSPM entropy (DWI)	6.66±0.22	8.20±0.15	p<0.00001	0.78
TSCIN DWI max maximum	0.44±0.02	0.34±0.01	p<0.00001	0.77
TSCIN DWI standard deviation	0.18±0.01	0.12±0.00	p<0.00001	0.79
TSCIN DWI range	0.34±0.02	0.24±0.01	p<0.00001	0.79
TSCIN DWI median absolute deviation	0.13±0.01	0.09±0.00	p<0.00001	0.78
TSCIN DCE kurtosis	2.63±0.14	3.37±0.08	0.00004	0.76
TSCIN DCE skewness	-0.69±0.07	-1.06±0.04	0.00001	0.75

DWI: Diffusion Weighted Imaging
PK: Pharmacokinetic
DCE: Dynamic Contrast Enhancement
FOS: First Order Statistics
TSPM: Tissue Signature Probability Matrix

The top MPRAD features for differentiating benign patients from malignant patients have been encapsulated in **Table 7.3**. Using IsoSVM (**Chapter 6**) with leave-one-out cross validation, these features produced a sensitivity and specificity of 82.5% and 80.5% respectively with an AUC of 0.87. The optimal IsoSVM parameters were $k = 20$, $d = 1$ with an imbalance ratio of 3:1 (benign:malignant). The ROC curves demonstrating the predictive power of single and multiparametric radiomic features as well as the IsoSVM

model have been displayed in **Figure 7.5. Figure 7.6** illustrates the TOC curve for the IsoSVM model for improved visualization.

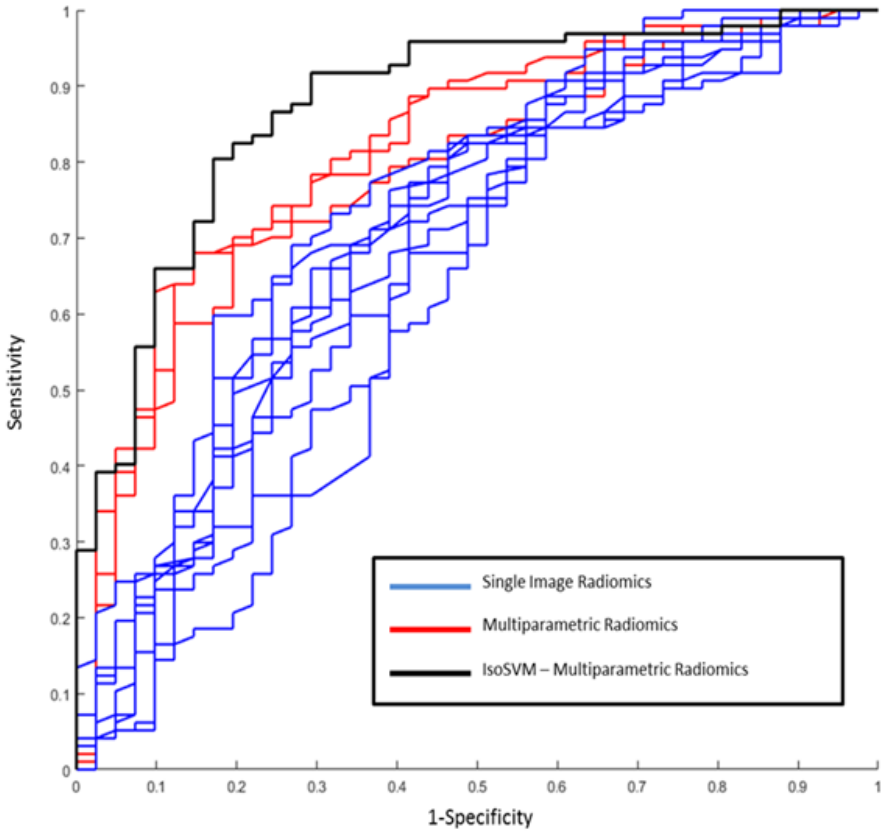


Figure 7.5 Comparison between the predictive accuracy of the single parameter based radiomics features and multiparametric radiomic features using receiver operating characteristic (ROC) curve analysis. The multiparametric radiomic feature ROC curves (displayed in red) produced area under the ROC curve (AUC) values that were 9%-28% greater than the AUCs obtained for single parameter radiomics (ROC curves displayed in blue). The ROC curve obtained from applying IsoSVM (**Chapter 6**) for classification of benign from malignant patients is displayed in black. The area under the ROC curve (AUC) for IsoSVM was obtained at 0.87.

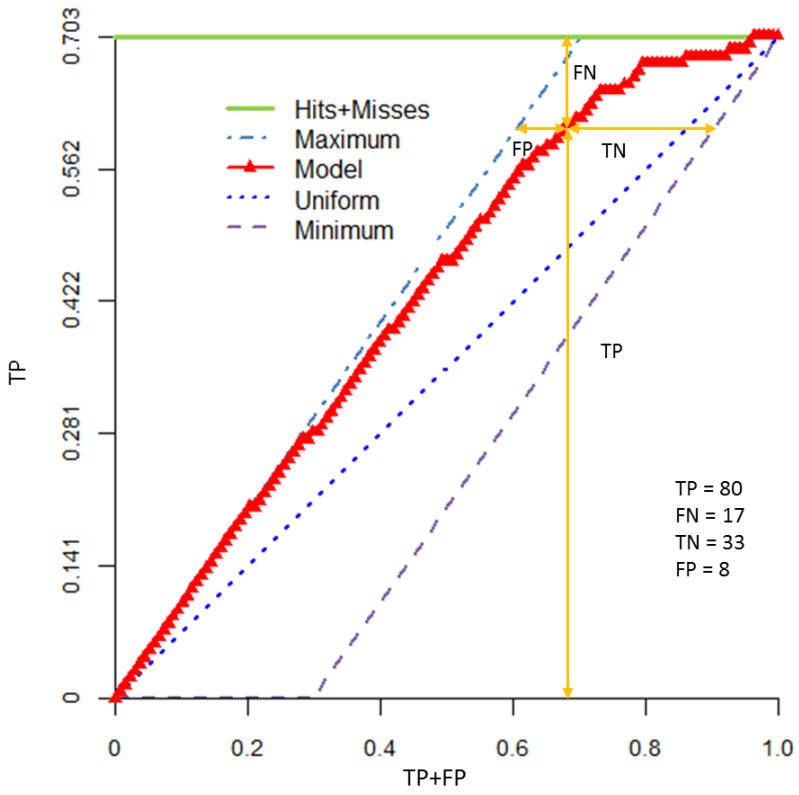


Figure 7.6 The total operating characteristic (TOC) curve for the IsoSVM classifier applied to MPRAD features. The true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP) for the optimal IsoSVM hyperplane were 80, 17, 33, and 8, respectively.

7.3.4 Multiparametric Brain Stroke MRI Dataset

a) Experimental Setup

Our stroke data consisted of ten patients (n=10, 5 women and 5 men, age=64±19 years) that were imaged at the acute time point (<12h) on a 1.5T clinical MRI system using a phased array head coil. The sample set consisted of 10 total studies. The mean age of patients was 64±19 ranging between 36-87. The MRI parameters were: T1WI sagittal MPRAGE image (TR/TE $200/2.46ms$, $field\ of\ view\ (FOV) = 24\ cm \times 24\ cm$,

$slice\ thickness\ (ST) = 5\ mm$), axial T2WI FLAIR ($TR/TE/TI = 9000/105/2500ms$, $FOV = 17.3cm\ x\ 23cm$, $ST = 4\ mm$), axial DWI ($TR/TE = 9000/98ms$, $b - values = 1000\ and\ 0\ s/mm^2$, $FOV = 23cm\ x\ 23cm$, $matrix = 128x128$, $ST = 4\ mm$) and echo planar T1WI perfusion ($TR/TE = 1350/30ms$, $FOV = 23\ cm\ x\ 23cm$, $ST = 4\ mm$, $total\ duration = 90\ seconds$). The contrast agent GdDTPA (Magnevist) was power injected at a dose of 0.1 mmol/kg and at a rate of 5 cc/sec.

The infarcted tissue was segmented from the diffusion DWI and ADC map using the Eigen filter algorithm [69, 231]. The input to Eigen filter algorithm was the selection of white matter, gray matter, cerebrospinal fluid, potential infarcted tissue, and the tissue at risk using pixels from each tissue type and respective sequence. The Eigen filter then used Gram-Schmidt orthonormalization to segment complete areas for each tissue type. Similarly, the tissue at risk was segmented from the PWI using the Eigen filter algorithm. The areas for the two tissue types were calculated by counting the number of pixels within the Eigen filter followed by multiplication with the pixel resolution in mm^2. Furthermore, the ROIs defined by the Eigen filter were overlaid on the ADC map and TTP map to obtain quantitative measurements.

b) Results

Four MPRAD features were extracted from ten patients with stroke imaged at acute timepoint (<12h). The mean age of patients was 64±19 ranging between 36-87. **Figures 7.7 and 7.8** illustrates both the single and multi-parametric radiomic feature maps from the DWI and PWI MRI. As shown in **Figure 7.7**, there is a striking difference in the textural representation shown by single and multi-parametric radiomics from the DWI and ADC maps. The majority of the single radiomic second order features (GLCM) did not

show any significant textural difference between infarcted tissue and tissue at risk on the ADC map. Whereas the same second order multiparametric radiomic features (TSPM) were significantly different for multiparametric complete DWI dataset. These results have been tabulated in **Table 7.4**.

Similarly, multiparametric radiomic values for the TTP and perfusion dataset (all parameters, dimensionality > 50) demonstrated excellent results for the MPRAD as shown in **Figure 7.8**. For example, The MPRAD TSPM Entropy exhibited a significant difference between infarcted tissue and tissue-at-risk: (6.6 ± 0.5 vs 8.4 ± 0.3, $p = 0.01$). These results are summarized in **Table 7.5** detailing the comparison between the single and multiparametric radiomics on PWI. **Table 7.6** summarizes the results from multiparametric radiomics applied to all the parameters.

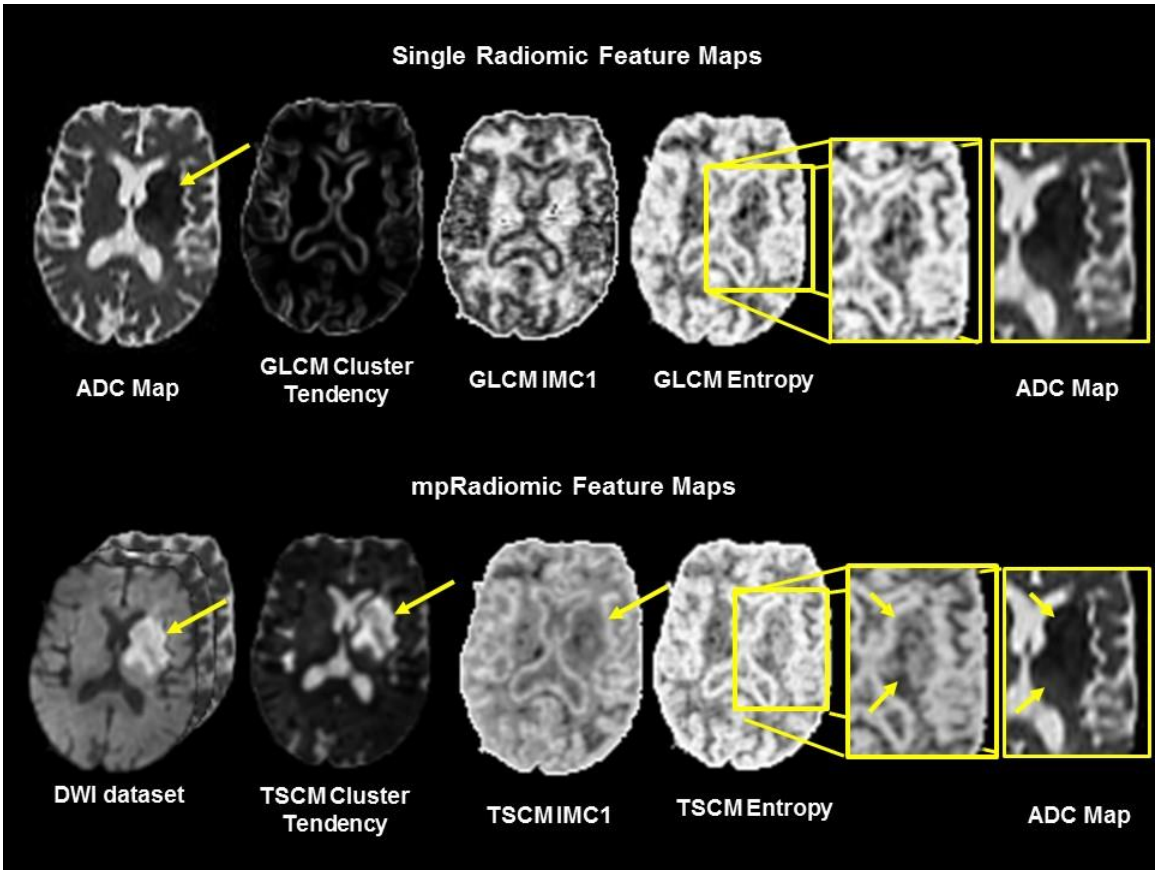


Figure 7.7. Illustration of radiomic feature maps (RFM) obtained from single and multiparametric radiomic analysis of an acute stroke patient with mpMRI Diffusion weighted imaging and ADC mapping. **Top Row.** ADC map with the yellow arrow showing the densely ischemic tissue. The RFMs in the illustrate different gray level co-occurrence matrix (GLCM) single radiomic features maps for the ADC map. The delineation of the infarcted tissue is hard to discern. **Bottom Row.** MPRAD of the DWI data set with yellow arrows showing the infarcted tissue. The MPRAD demonstrates excellent delineation of the infarcted tissue. The enlarged area shows the heterogeneity of the lesion.

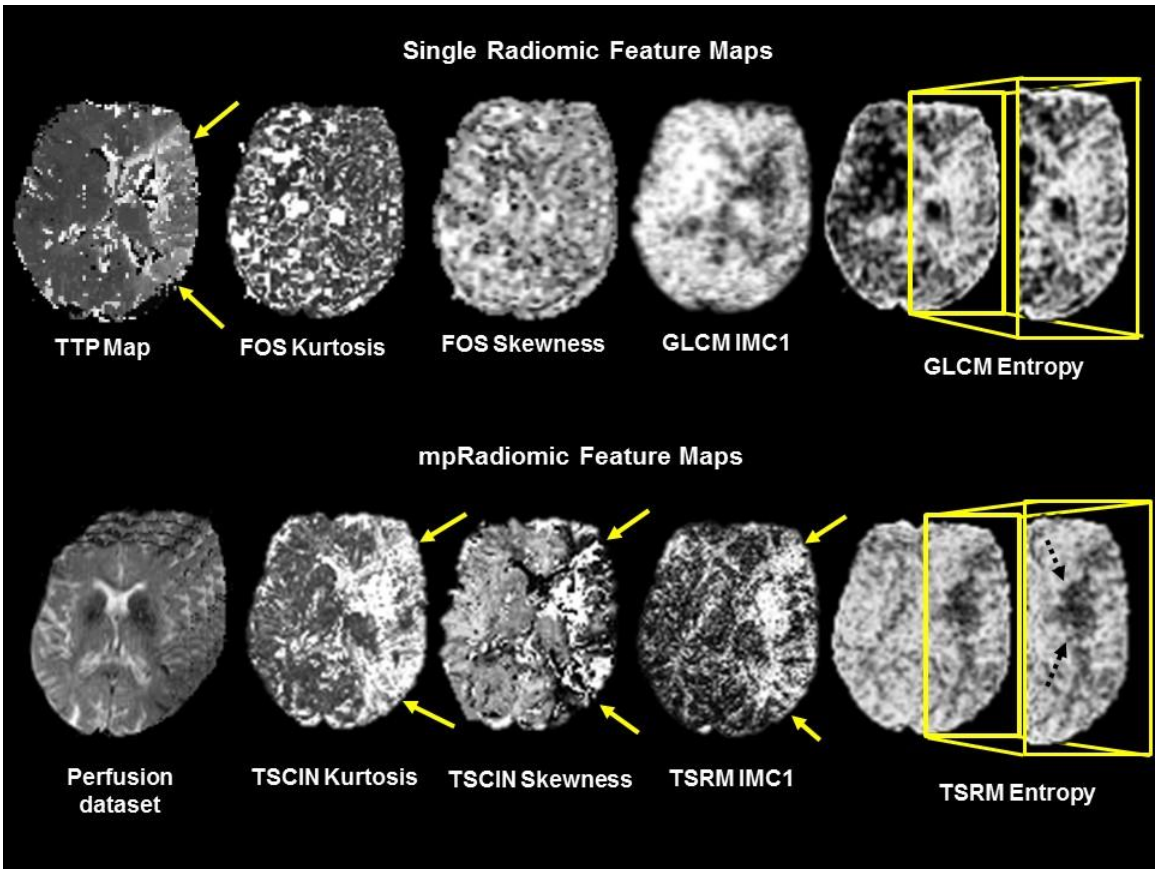


Figure 7.8. Illustration of radiomic feature maps (RFM) obtained from single and multiparametric radiomic analysis of an acute stroke patient with mpMRI perfusion weighted imaging. **Top Row.** Time to Peak (TTP) map from the perfusion MRI with the yellow arrow showing potential "tissue at risk". The first order (FOS) RFMs illustrate the different gray level single radiomic and co-occurrence matrix (GLCM) maps from the TTP. **Bottom Row.** MPRAD images from perfusion MRI illustrates the power mpRadiomics and the striking difference in the "tissue at risk" delineation in both the tissue signature matrix (TSCIN) and tissue signature relationship matrix (TSRM) radiomic features. The black dotted arrows show the infarcted tissue in the caudate putamen and internal capsule.

Table 7.4 Summary of multiparametric radiomic features for diffusion weighted imaging in stroke infarcted and tissue at risk

	Radiomic Feature	Infarcted Tissue	Tissue at Risk	White matter	Gray matter	p value (Infarcted vs Tissue at Risk)
Single Parameter Radiomics (ADC map)	Mean ADC value	0.66±0.04	0.88±0.05	0.91±0.04	1.14±0.06	0.003
	GLCM Autocorrelation	42.36±5.09	71.61±7.66	74.59±7.30	111.20±11.17	0.01
	GLCM Cluster Tendency	2.64±0.69	7.29±3.81	1.65±0.40	14.15±2.27	0.26
	GLCM Contrast	0.83±0.12	1.20±0.24	0.61±0.11	5.02±0.90	0.19
	GLCM Homogeneity1	0.74±0.02	0.72±0.01	0.79±0.02	0.54±0.02	0.43
	GLCM Variance	42.76±5.08	72.20±7.69	74.87±7.30	113.69±11.10	0.01
	GLCM Entropy	2.10±0.20	2.59±0.20	1.73±0.15	3.24±0.13	0.10
	GLCM IMC1	-0.18±0.03	-0.19±0.03	-0.18±0.02	-0.28±0.03	0.75
Multiparametric Imaging Radiomics (all DWI + ADC)	TSCM Autocorrelation	199.66±23.35	133.96±17.33	110.95±19.87	137.16±18.54	0.04
	TSCM Cluster Tendency	206.16±29.37	90.61±18.50	73.95±33.10	66.26±15.38	0.01
	TSCM Contrast	4.20±0.94	1.92±0.48	0.58±0.08	5.50±0.18	0.05
	TSCM Homogeneity1	0.64±0.02	0.70±0.02	0.80±0.01	0.58±0.02	0.05
	TSCM Variance	201.73±23.40	134.90±17.41	111.22±19.89	139.88±19.08	0.04
	TSCM Entropy	3.94±0.17	3.73±0.19	2.70±0.13	4.21±0.15	0.43
	TSCM IMC1	-0.43±0.02	-0.44±0.02	-0.54±0.02	-0.31±0.01	0.74

DWI: Diffusion Weighted Imaging
ADC: Apparent Diffusion Coefficient
GLCM: Gray-level co-occurrence matrix
TSCM: Tissue Signature Co-occurrence Matrix

Table 7.5 Summary of multiparametric radiomic features for perfusion weighted imaging in stroke infarcted and tissue at risk

	Radiomic Feature	Infarcted Tissue	Tissue at Risk	White matter (contralateral)	Gray matter (contralateral)	p value (Infarcted vs Tissue at Risk)
Single Parameter Radiomics (TTP map)	Mean TTP value	11.19±1.90 (sec)	10.81±1.34 (sec)	8.23±1.06 (sec)	7.61±1.09 (sec)	0.88
	FOS Entropy	2.96±0.72	2.83±0.33	1.22±0.36	2.12±0.35	0.88
	FOS Uniformity	0.26±0.16	0.21±0.05	0.54±0.12	0.36±0.09	0.77
	FOS Kurtosis	0.24±0.08	0.40±0.03	3.26±0.34	10.72±2.52	0.12
	FOS Skewness	1.64±0.57	1.11±0.46	0.14±0.33	0.22±1.16	0.49
	GLCM Correlation	0.18±0.03	0.31±0.04	0.20±0.09	0.21±0.06	0.04
	GLCM IMC1	-0.28±0.04	-0.17±0.03	-0.19±0.10	-0.26±0.05	0.05
Multiparametric Imaging Radiomics (all PWI series)	TSPM Entropy	6.62±0.53	8.41±0.33	6.16±0.41	6.46±0.15	0.01
	TSPM Uniformity	0.016±0.005	0.004±0.001	0.019±0.005	0.012±0.002	0.05
	TSCIN Kurtosis	4.70±0.58	4.63±0.54	5.93±0.79	5.08±0.50	0.94
	TSCIN Skewness	-0.47±0.25	-0.59±0.17	-1.11±0.28	-0.84±0.24	0.71
	TSCM Correlation	0.76±0.06	0.85±0.04	0.86±0.05	0.78±0.03	0.24
	TSCM IMC1	-0.23±0.03	-0.30±0.03	-0.37±0.05	-0.22±0.03	0.11

TTP: time-to-peak
PWI: Perfusion Weighted Imaging
FOS: First Order Statistics
GLCM: Gray-level co-occurrence matrix
TSCM: Tissue Signature Co-occurrence Matrix
IMC1: Informational Measure of Correlation 1

Table 7.6 Summary of multiparametric radiomic features for the complete dataset consisting of diffusion and perfusion weighted imaging in stroke infarcted and tissue at risk

	Radiomic Feature	Infarcted Tissue	Tissue at Risk	White matter (contralateral)	Gray matter (contralateral)	p value (Infarcted vs Tissue at Risk)
Multiparametric Imaging Radiomics (PWI+ DWI)	TSPM Entropy	6.65±0.52	8.41±0.33	6.16±0.41	6.47±0.15	0.01
	TSPM Uniformity	0.015±0.004	0.004±0.001	0.019±0.005	0.012±0.002	0.02
	TSCM Correlation	0.88±0.03	0.92±0.01	0.94±0.01	0.81±0.03	0.26
	TSCM IMC1	-0.33±0.03	-0.27±0.02	-0.42±0.04	-0.23±0.02	0.07

PWI: Perfusion Weighted Imaging
TSPM: Tissue Signature Probability Matrix
TSCM: Tissue Signature Co-occurrence Matrix
IMC1: Informational Measure of Correlation

7.4 Discussion

I have developed and validated a novel multiparametric imaging radiomics (MPRAD) framework that integrates all the data to define different tissue characteristics. MPRAD outperformed single radiomic features in both synthetic and clinical datasets. The MPRAD features captured the underlying tissue texture based on tissue signatures rather than individual imaging parameter intensities. In addition, MPRAD produces full texture images for visualization of normal and lesion heterogeneity, thereby providing radiologists with a tool for visualization and quantization of the true underlying tissue heterogeneity.

In multiparametric imaging settings, single radiomic features from each individual image can result in large numbers of texture features creating a high-dimensional dataset across all images for analysis. These single radiomic features may not reflect the true underlying tissue contrast, heterogeneity, or homogeneity and only provide a limited information corresponding to the physical modeling of each imaging parameter. The

MPRAD framework extracts radiomic features that consider the complete multiparametric dataset, hence producing more meaningful features and textural visualization of the underlying tissue. In addition, the MPRAD framework allows us to analyze the complex interactions between different imaging parameters, consequently opening up a completely new source of information that did not exist with conventional radiomic features.

In breast, consistent with other reports, malignant lesions had increased entropy compared to benign lesions (**Chapter 5**). Importantly, there were no differences in the normal glandular tissue between each group (**Chapter 6**). The MPRAD radiomic features delineated different tissue types better than the single radiomic features, for example, in cysts, normal, and peri-tumoral regions. Finally, the MPRAD demonstrated excellent sensitivity and specificity with increased AUC compared to single radiomic features comparable with those achieved by radiologists.

In stroke patients, predicating stroke outcome is very challenging and the evaluation and treatment are determined by the ability to identify the ischemic penumbra, oligemic, and potential salvageable tissue. Ischemic penumbra refers to brain tissue which is at risk for infarction [224, 225, 232, 233]. Damage to this salvable tissue can be potentially prevented or reversed using thrombolytic therapy [227]. The brain tissue corresponding to ischemic penumbra and oligemia can be identified using advanced MRI parameters of diffusion-weighted imaging (DWI) and perfusion-weighted imaging (PWI) [228, 233, 234]. The MPRAD was able to accurately separate the diffusion-perfusion mismatch at the acute time point. The diffusion-perfusion mismatch maybe indicative of the extent of the stroke and if therapeutic intervention needed and MPRAD will provide new quantitation and visualization tools for use in stroke patients.

The integration of advanced MPRAD features with the ADC map and perfusion metrics could provide important information about the spatial distribution and characteristics of the tissue. For example, the ADC map and entropy values for infarcted tissue were decreased. The perfusion values were prolonged and where the entropy values were higher. This is consistent with the known biology of ischemic tissue, where the tissue is dead or dying and will exhibit a more uniform pattern. In contrast, tissue at risk is highly variable with different tissue characteristics and has increased ADC and perfusion values. The MPRAD radiomic entropy values were increased, reflecting this tissue heterogeneity. Thus, by combining these new radiomic features could be very helpful in clinical decision to give treatment or withhold it

In general, multiparametric imaging for applications such as brain, breast, and prostate MRI produces a large number of images (>50) corresponding to each slice location, thereby producing a high-dimensional image space. Extracting radiomic features from each image in such datasets may not provide complete information about the tissue. In addition, given the high-dimensionality of the dataset, there is an increase in both the computational and space complexity of the radiomic process making radiomics analysis impractical for such cases. The MPRAD framework resolves this issue by extracting radiomic features that not only analyze the progression of tissue texture with time but also evaluate the overall tissue texture in large data sets.

There are, however, some technical limitations to the use the MPRAD in practice. First, there is a need for graphical processor units (GPU) with large memory and user-friendly software for processing. These items may not be widely available. More specific to the present study, any assessment of the clinical value of MPRAD framework

will require additional prospective studies. These studies would have subsequent follow-up and pathological correlation using MPRAD in breast. In stoke, these types of studies would provide us with new MPRAD data to predict final infraction volume or identify markers of hemorrhagic transformation over time.

In conclusion, I have demonstrated that MPRAD framework shows a great potential in analysis of textural information in multimodal and multiparametric imaging settings. With increasing use of multiparametric imaging in the clinical setting, MPRAD provides an ideal framework for future clinical decision support systems.

Chapter 8. Contribution Scattergram: A complex network model of inter-parametric relationships for high-dimensional multiparametric radiological imaging data

8.1 Introduction

Multiparametric data acquisition techniques such as multiparametric magnetic resonance imaging (MRI) or multispectral imaging produce a diversity of different parameters with each parameter quantifying a certain aspect of the structural and functional property of the underlying data type. Different parameters acquired interact with each other based on complex high-level relationships which could provide important insight into the data being acquired. I am interested in modeling the relationships between these imaging parameters and understand how they interact with each other.

In the previous chapter (**Chapter 7**), I discussed the technique of multiparametric radiomics (MPRAD), which characterized the inter-parametric relationships between different imaging parameters using statistical methods. In addition, quantitative imaging methods such as pharmacokinetic analysis for DCE-MRI and ADC mapping for DWI also model multiparametric relationships using mathematical models of underlying tissue biology. However, all of these methods model the inter-parametric relationships based on a specific model. As a result, they are not capable of uncovering the complete underlying network structure of multiparametric MRI. In addition, the previously developed techniques are based on voxel-wise analysis and do not model the tissue of interest as a

whole. To that end, we developed the technique of contribution scattergram (CSg) to uncover the underlying multidimensional complex network for multiparametric imaging data [235, 236].

The contribution scattergram is constructed by transforming the high-dimensional multiparametric space into a complex network model such that the vertices on CSg represent the imaging parameters and the edges, relationship between different parameters. The CSg is visualized in a lower dimensional space using manifold learning algorithms. Manifold learning algorithms transform high-dimensional multiparametric imaging datasets into a lower dimensional representation of its intrinsic dimensionality. Moreover, the intrinsic dimensionality of the CSg is obtained by calculating residual variance between the high- and low- dimensional geodesic networks.

The contribution scattergram is analyzed using graph theoretic techniques to determine the contribution of each parameter and quantify the complex interactions between different parameters. Graph theoretic analysis of complex networks has received extensive attention in the recent years following the seminal papers by Watts and Strogatz on small world phenomenon and by Barabasi and Albert on scale free property of complex networks. Graph theoretic techniques have been very effective in extracting the intrinsic properties of complex networks in many different applications; namely, genetic networks, brain networks, social media networks, and traffic networks [237-244].

Multiparametric breast MRI involves acquisition of anatomical parameters of T1-weighted imaging (T1WI) and T2-weighted imaging (T2WI) and advanced functional MRI parameters of dynamic contrast enhanced-MRI (DCE-MRI) and diffusion weighted imaging (DWI). These MRI parameters provide complimentary noninvasive information

about the underlying tissue biology. Specifically, DWI provides information about the cellularity of the underlying tissue through computation of apparent diffusion coefficient (ADC) map while DCE-MRI provides information about the vascularity of the tissue. I used the contribution scattergram to analyze the complex network formed by the anatomic and functional MRI parameters, quantify the inter-parameter relationships as well as their contribution to breast cancer diagnosis. I implemented the IsoSVM algorithm as the breast tissue classifier (discussed in **Chapter 6**). IsoSVM is a nonlinear classifier that implements an Isomap kernel for the support vector machine (SVM) classifier [216].

The background on manifold learning and graph theoretic techniques used for contribution scattergram analysis have been described in **Section 8.2**. **Section 8.3** details the contribution scattergram algorithm. The breast cancer clinical data, multiparametric MRI acquisition protocol, and image analysis techniques used in this chapter have been explained in **Section 8.4** followed by the results from application of contribution scattergram on this data. This chapter concludes with a discussion on the potential of contribution scattergram in precision medicine applications.

Publication from this work and author contributions

Publication

V. Parekh, A. Akhbardeh, M. Jacobs, "Contribution Scattergram: A complex network model based on graph theory and manifold learning for creating a unique signature for high-dimensional multiparametric data," Ready for Submission, 2018

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

Dr. Jacobs and Dr. Akhbardeh developed the concept of contribution scattergram for visualization of relationships between different imaging parameters (**Section 8.2.1**). I developed the complex network model for analyzing inter-parametric relationships and standardized network metrics for quantitative evaluation of the tissue of interest as a whole and for correlation to tissue biology.

I solely implemented all the algorithms for complex network analysis. The algorithms for manifold learning were jointly implemented by me (Isomap and Laplacian Eigenmap) and Dr. Akhbardeh (Diffusion map and Locally Linear Embedding). The code for visualization of multiparametric MRI in a two-dimensional space was also jointly developed by Dr. Akhbardeh and me, while the code for visualization in a three-dimensional space was solely developed by me. I solely performed all the data analysis described in the chapter.

8.2 Background

Dr. Jacobs and Dr. Akhbardeh developed the concept of contribution scattergram using manifold learning for visualization of relationships between different imaging parameters ***(Section 8.2.1)***

8.2.1 Manifold learning

Manifold learning algorithms transform high-dimensional datasets into a lower dimensional representation of its intrinsic dimensionality, d. A d-dimensional manifold is defined as a Hausdorff, second countable topological space, X for which every point $x \in X$ has an open neighborhood, N_x which is homeomorphic to a d-dimensional Euclidean space R^d [245, 246]. Let us consider a high-dimensional MRI dataset, $X = \{x_1, x_2, \dots, x_N\} \subset R^D$ where, x_i represents an MRI image or volume (e.g. T1, T2, DCE, and DWI), N represents

the number of MRI images, and D represents the number of voxels in each MRI image. The goal of manifold learning is to transform X into a d-dimensional representation, $Y = \{y_1, y_2, \dots y_N\} \subset R^d$ corresponding to the intrinsic dimensionality of X. Here, Y represents the contribution scattergram and y_i represents each MRI image in the d-dimensional contribution scattergram.

Many different algorithms such as Isometric feature mapping (Isomap) [43], Locally Linear Embedding (LLE) [42], Diffusion maps [41] and Laplacian Eigenmaps [247] have been proposed in the literature for manifold learning and are divided into global and local methods [248].

The algorithms for manifold learning were jointly implemented by me (Isomap and Laplacian Eigenmap) and Dr. Akhbardeh (Diffusion map and Locally Linear Embedding).

a) Isomap

Isomap is a metric extension of classical multidimensional scaling (MDS), where geodesic distances represent the interpoint distances on the weighted graph instead of Euclidean distances [43, 249]. If the high-dimensional data lies on or near a curved smooth manifold, Euclidean distance does not consider the distribution of the neighboring data points and might consider two data points as close, whereas their distance over the manifold is much larger than the typical interpoint distance. Isomap overcomes this problem by preserving pairwise geodesic (or curvilinear) distances between data points using a neighborhood graph.

Geodesic distance (GD) is defined as the distance between two points measured over the manifold. GD is estimated using a shortest path algorithm such as Dijkstra's or Floyd's algorithm [250, 251]. These algorithms construct a neighborhood graph, G and

calculate shortest path distances (GD) from every point to every other point on the graph based on the neighborhood connectivity. Once the GDs for the complete data is calculated, MDS is applied such that it attempts to preserve the interpoint geodesic distance of the higher dimension as the interpoint Euclidean distance in the lower dimensional space, Y. Topological instability is one major weakness of the Isomap algorithm [252]. The Isomap algorithm is highly sensitive to the selection of the neighborhood parameter. The solution to this is proposed which is a parameterless implementation of the Isomap algorithm. Another important weakness of Isomap is that it can fail if the manifold is non-convex.

b) Diffusion maps

Diffusion maps are based on diffusion distances as opposed to geodesic distances used in Isomap [41]. The diffusion distances are based on defining a Markov random walk on a Laplacian graph. Diffusion maps find the subspace that best preserves the interpoint diffusion distances. Diffusion maps are more robust to short circuits than Isomap owing to the fact that the diffusion distance between any two points on the graph is computed using several paths between the two points instead of just the shortest path (geodesic distance). The procedure for nonlinear dimensionality using diffusion maps is as follows:

In the first step, a similarity matrix, K is computed for every pair of points (i, j) on the input dataset using the gaussian kernel:

$$K_{ij} = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad 1 \leq i, j \leq N$$

Here, N equals the number of multidimensional points and σ is the kernel scale. In the next step, the matrix, K is normalized such that its rows add to 1. The resultant matrix, p is given as:

$$p_{ij} = \frac{K_{ij}}{\sum_{n=1}^{N} K_{in}}$$

The elements of matrix, p characterize the forward transition probability of a single step taken from point i to point j. For characterizing the transition probability of a t-step random walk from i to j, powers of p are computed as follows:

$$p_{ij}^{t} = \frac{K_{ij}}{\sum_{n=1}^{N} K_{in}}$$

Increasing powers of p allows us to explore the data at different scales (local and global). Consequently, the diffusion distance between any two points i and j is defined as follows:

$$D_t(x_i, x_j)^2 = \sum_{r=1}^{N} \frac{\left(p_{ir}^{(t)} - p_{rj}^{(t)}\right)^2}{\psi(x_r)}$$

$$\psi(x_m) = \frac{\sum_{j=1}^{N} p_{jm}}{\sum_{j=1}^{N}\sum_{k=1}^{N} p_{jk}}$$

The goal of diffusion mapping is to preserve the diffusion distances when going from higher dimensional space to a lower dimensional space. The lower dimensional representation, Y can be obtained by computing the top d nontrivial eigenvectors of the matrix, p [41]

$$Y: x \rightarrow \{\lambda_2 V_2, \ldots, \lambda_d V_d\}$$

The DfM graph is fully connected, therefore eigen vector v_1 of the largest eigenvalue ($\lambda_1 = 1$) is discarded, and the remaining eigenvectors are normalized by their corresponding eigenvalues.

c) Locally Linear Embedding

Locally Linear Embedding (LLE) is a local method for nonlinear dimensionality reduction that aims to preserve local properties of the data [42, 253]. LLE breaks the data space into a number of small connected overlapping regions which can be assumed to be approximately locally linear. For every point in the data, k nearest neighbors are chosen to represent the local neighborhood of that point. LLE attempts to model the geometric property of the local neighborhood around every point by attempting to represent that point as a weighted summation of its k nearest neighbors. The reconstruction weights for every neighborhood are computed by minimizing the following error function:

$$Err(W) = \sum_{i=1}^{L} \| x_i - \sum_{j \in N(i)} w_{ij} x_j \|^2$$

subject to two constraints, $\sum_{j \in N(i)} w_{ij} = 1$ and $w_{ij} = 0$ when j $\notin$ N(i). Here, X is the input data, L is the number of points and $N(i)$ represents the neighborhood of i. The optimal weights can be computed by solving a set of constrained least squares problems [253].

The idea behind LLE is to preserve these reconstruction weights when going from higher dimensional space to lower dimension. Therefore, to find the reduced (d) dimensional data representation Y, the following cost function is minimized for each point y_i:

$$\phi(Y) = \sum_{i=1}^{L} \| y_i - \sum_{j \in N(i)} w_{ij} y_j \|^2$$

The lower dimensional embedding, Y is obtained by computing the eigenvectors corresponding to the smallest d nonzero eigenvalues of the following E, $E =$

$(I - W)^T . (I - W)$. Here, I is the identity matrix and W is the weight matrix ($L \times k$). The major problem with LLE is that it does not work when the manifold has holes.

d) Laplacian Eigenmaps

Laplacian Eigenmaps, just like LLE, aim to find a low-dimensional data representation by preserving local properties of the manifold [247]. Laplacian Eigenmaps aim to preserve the pairwise distance between nearest neighbors. This is done by attaching weights to the cost function based on the nearness between the data points. Just like diffusion maps, Laplacian Eigenmaps use the Gaussian Kernel to estimate weights (K) of the edges in the graph. The cost function in computing low-dimensional representations y_i is given by the following equation.

$$E(Y) = \sum_{ij} (y_i - y_j)^2 K_{ij}$$

The eigenvalues and eigenvectors for the generalized eigenvector problem $Ly = \lambda Dy$ are computed to determine the low-dimensional embedding, Y. Here, D is the diagonal weight matrix formed from row sums of K, $D_{jj} = \sum_i K_{ij}$ and $L = D - K$ is the Laplacian matrix. The low-dimensional embedding, Y is obtained by computing the eigenvectors corresponding to the d smallest non-zero eigenvalues.

8.2.2 Complex Network Analysis

Complex network analysis is the study of complex, irregular and dynamic networks that are evolving in time. Complex network analysis has received considerable attention since the seminal paper by Erdos and Renyi on random graphs. In the past decade, the papers by Watts and Strogatz on small world phenomenon and by Barabasi and Albert on scale free property have sparked a great amount of interest in the field of dynamically evolving complex social network analysis. Complex network analysis techniques have shown

incredible success in the analysis of social networks, world wide web, brain networks, and traffic networks [237-244]. The complex network can be probed using a number of different metrics that highlight important characteristics of the underlying network.

a) Graph summary statistics

The average path length and diameter are the basic statistical metrics computed for any complex network. A path is defined as the set of edges connecting any two nodes and the sum of weights of these edges represent the path length. Average path length, as the name suggests, is the average of the path lengths across all pairs of nodes. Diameter, on the other hand, is the maximum value among all the path lengths.

b) Clustering coefficient

The clustering coefficient defines the connectedness of the neighborhood of an imaging parameter. Clustering coefficient for an imaging parameter ranges from zero to one with zero representing completely disconnected neighborhood and one representing completely connected neighborhood. Mathematically, clustering coefficient of an imaging parameter, I is given by the following equation:

$$CC(I) = \frac{2e_I}{k_I(k_I - 1)}$$

Here e_i is the number of connected edges neighborhood of I, k_I is the number of nodes in the neighborhood of I making $\frac{k_I(k_I-1)}{2}$ the maximum possible number of edges in the neighborhood of I.

c) Degree distribution

The degree distribution is the most fundamental metric calculated for any complex network. The analysis of degree distribution enables us to identify whether the network is a scale-free network or not. The degree of an imaging parameter is defined as the total

number of imaging parameters it is directly connected to. Furthermore, degree distribution, $P(k)$ is determined using the following equation

$$P(k) = \frac{1}{N}\sum_{I=1}^{N} \mathbf{1}\{\deg(I) = k\} \ \forall k \in \{1,2,\dots,N-1\}$$

The probability that an imaging parameter can interact with k other imaging parameters can be determined from the degree distribution. If the degree distribution varies as a power law $P(k) \propto k^{-\gamma}$, the complex network can be considered scale free. One important feature of scale free networks is the presence of few highly connected important nodes that influence the network properties. Furthermore, the probability of any incoming node connecting to these important nodes is significantly higher than connecting to other nodes. These important nodes may correspond to the key imaging parameters that guide the diagnosis or prognosis of pathology.

d) Centrality measures

Centrality measures are calculated to determine the importance of each informatics parameter in the complex informatics network. The concept of centrality dates back to 1950 [254], however, the most widely used measures of centrality are the ones developed by Freeman in his seminal papers [255, 256]. In this chapter, I calculated three centrality measures, each centrality measure highlighting a different importance property as elaborated below.

Betweenness centrality

The betweenness centrality of an MRI image (BC) is a measure of its centrality in the network [255]. Betweenness centrality quantifies the amount of information that flows through a particular MRI image. Betweenness centrality for an MRI image is defined as

the number of shortest paths that pass through that MRI image. Betweenness centrality of an MRI image, I in the contribution scattergram is given by the following equation:

$$B(I) = \sum_{s \neq I \neq t} \frac{N_{st}(I)}{N_{st}}$$

Here, N_{st} is the total number of shortest paths between the MRI images s and t and $N_{st}(I)$ is the total number of shortest paths between s and t that pass through I.

Closeness centrality

Closeness centrality, as defined for each MRI image is the inverse of sum of all geodesic distances from that MRI image to all the other MRI images [256]. Closeness centrality of an MRI image can be interpreted as the amount of time it would take for information to spread to other MRI images from that MRI image. Mathematically, closeness centrality of an MRI image, I is given by the following equation:

$$C(I) = \frac{1}{N-1}\left(\frac{1}{\sum_{\substack{s=1 \\ s \neq I}}^{N} G(s,I)}\right)$$

Here, N is the total number of MRI images in the contribution scattergram and $G(s,I)$ is the geodesic distance between the MRI images s and I. The value of closeness centrality, C as computed using the above formula may lead to the value 0 when the contribution scattergram has disconnected components. For example, if there is no path from image, s to image, t, then the geodesic distance, $G(s,t) = \infty$ and as a result their closeness centrality values will be 0.

To resolve the problem of zero closeness centrality, the measure of closeness centrality can be modified in two ways.

$$C(I) = \frac{1}{N-1}\left(\frac{1}{\sum^{N}_{\substack{s=1 \\ s \neq I \\ G(s,I) \neq \infty}} G(s,I)}\right)$$

$$H(I) = \frac{1}{N-1}\sum_{\substack{S=1 \\ S \neq I}}^{N} \frac{1}{G(S,I)}$$

Here, $H(I)$ is also called the harmonic closeness centrality. I implemented the modified closeness centrality for contribution scattergram.

Degree centrality

Degree centrality is defined as the total number of MRI images each image is connected to in the contribution scattergram. For weighted networks, it is defined as the summation of edge weights of all the MRI images, each image is connected to in the contribution scattergram.

8.3 Contribution Scattergram algorithm

We have developed a multidimensional complex network model called the contribution scattergram using manifold learning. A typical contribution scattergram is shown in **Figure 8.1**. The development and analysis of contribution scattergram is primarily composed of two steps: Establishing the network topology using manifold learning and evaluating contribution scattergrams using graph theoretic techniques.

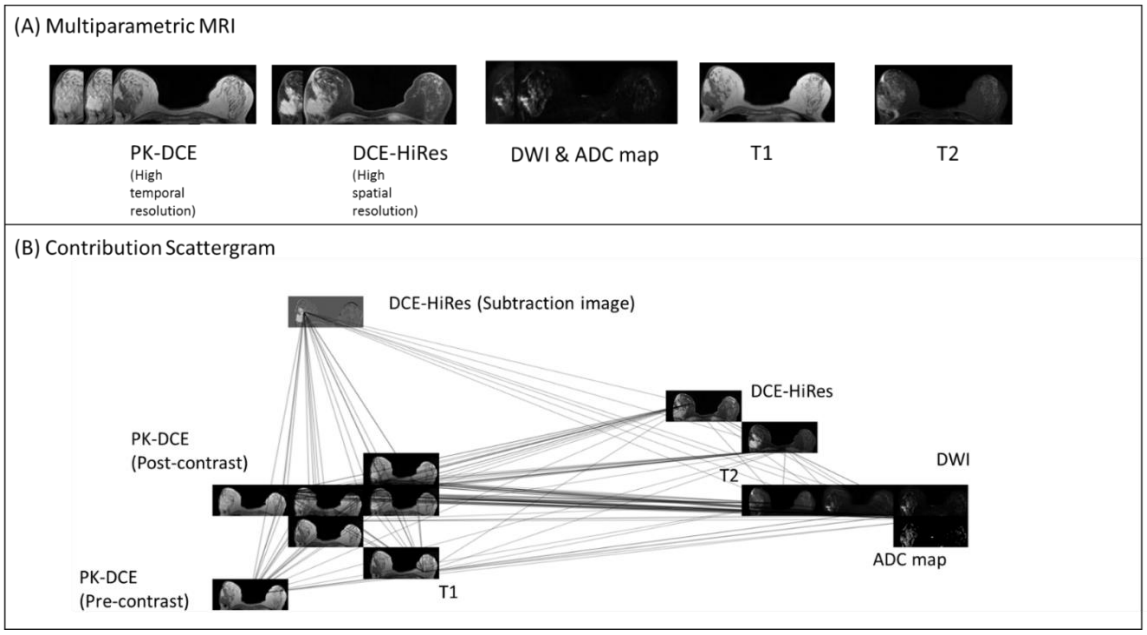


Figure 8.1. (a) An example multiparametric MRI dataset, X in a D-dimensional space. Here D corresponds to the number of pixels in each MRI image. (b) The manifold learning algorithm learns the manifold representation of the D-dimensional multiparametric MRI dataset and represents the learned manifold as a graph called the contribution scattergram. The intrinsic dimensionality of the manifold is two in this example.

I implemented the Isomap algorithm for manifold learning of the multiparametric MRI dataset as it is a global method and provides a theoretical guarantee for correctness. Furthermore, the graph theoretic measures of average path length, centrality metrics and clustering coefficients are based on the geodesic distances and neighborhood parameters as calculated in the Isomap procedure, making Isomap the appropriate choice for manifold learning. The contribution scattergram is characterized by geodesic distances (GD), where geodesic distances represent the interpoint distances over the manifold. The GD can be estimated using a shortest path algorithm such as Dijkstra's or Floyd's algorithm [250, 251].

These algorithms construct a neighborhood graph, G and calculate shortest path distances (GD) from every point to every other point on the graph based on the neighborhood connectivity. There are two main computationally expensive components of the Isomap algorithm – calculation of shortest path distances and classical multidimensional scaling. Given a multiparametric MRI dataset of N images with D voxels each, the computational complexity of the Isomap algorithm is $O(N^3)$ when using Floyd's or Dijkstra's algorithm

The graph theoretic analysis of contribution scattergram is based on the geodesic distances computed by the shortest path algorithms. The shortest paths along with the final geodesic distances between different MRI images on the contribution scattergram are evaluated using graph theory to determine the graph metrics described in the previous section. The procedure for contribution scattergram is represented as a pseudocode in **Algorithm 8.1.**

Algorithm 8.1 Contribution Scattergram algorithm

Inputs:
X: multiparametric imaging dataset, $X = \{x_1, x_2, \dots, x_N\} \in R^D$. Here x_i are the imaging parameters and D is the number of voxels in each image.
k: Isomap neighborhood
Output: Graph metrics and contribution scattergram
1: Transform X to a complex network model $G = (V, E)$:
2: // v_i represent the imaging parameters
3: // E_{ij}: Euclidean distance between v_i and v_j if they are connected, otherwise infinity
4: Compute pairwise geodesic distance matrix, GD using k
5: Run multidimensional scaling on GD to compute multiple low-dimensional embeddings, Y
6: Compute residual variance, R for each embedding in Y
7: Calculate the intrinsic dimensionality, d using the elbow method
8: Generate the contribution scattergram using the correct low-dimensional embedding from Y
9: Compute graph metrics using the complex network, G and geodesic distances, GD
10: return contribution scattergram and graph metrics

8.3.1 Estimating the intrinsic dimensionality

The intrinsic dimensionality of the contribution scattergram is computed by residual variance analysis [43]. The residual variance for the Isomap algorithm is given by $R = 1 - C(G_X, G_Y)$ where C is the correlation coefficient between the vectors G_X and G_Y, G_X represents the vector of all the pairwise geodesic distances in the original high-dimensional space (D), G_Y represents the vector of all the pairwise geodesic distances in the lower dimensional space (d).

Using this equation, the residual variance of the contribution scattergram at $d = 1$ through $d = D$ is plotted as shown in **Figure 8.2**. The intrinsic dimensionality of the contribution scattergram is estimated by identifying the elbow in the resulting curve. The elbow is defined as the value of d at which there is no significant decrease in the curve with added dimensions [43]. It can be seen from **Figure 8.2** that elbow occurs at $d = 2$.

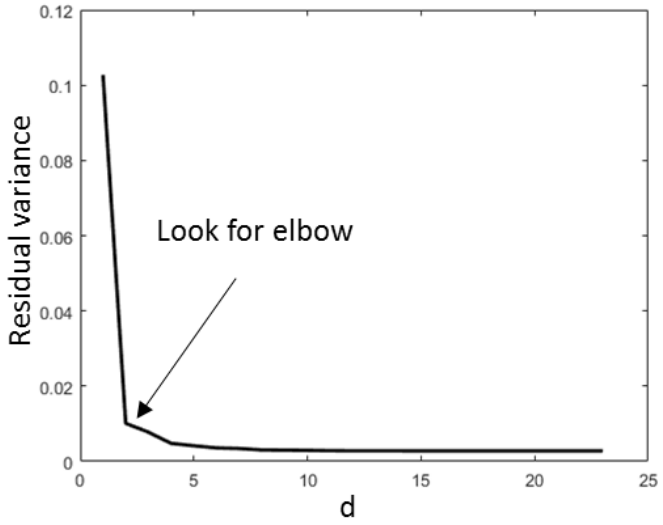


Figure 8.2 Example plot representing the residual variance analysis procedure for estimating the intrinsic dimensionality of the multiparametric MRI manifold X. It can be seen that elbow for the plot occurs at $d = 2$.

8.4 Materials and Methods

8.4.1 Clinical data

I computed, visualized and analyzed contribution scattergrams corresponding to the multiparametric MRI data obtained from one hundred and thirty-nine women (mean age = 52±11, age range = 24-80). Of the one hundred and thirty-nine, ninety-six patients had biopsy proven breast cancer, thirty-nine patients had benign lesions and the remaining four patients had no lesion. The methods for MRI Image Processing, registration, segmentation, and classification have all been detailed in **Chapter 2**.

8.4.2 Statistical analysis

Student ttest was performed between the graph theoretic metrics corresponding to different patient groups (benign and malignant) to determine which MRI parameters were significantly different for the two patient groups. Logistic regression was used to determine associations between the graph theoretic metrics and the final clinical diagnosis. The ROC curve analysis was performed to assess the diagnostic accuracy of each graph theoretic metric. The P values ≤ 0.05 were considered significant.

8.4.3 Multiparametric MRI network analysis using contribution scattergram

I constructed contribution scattergrams for the multiparametric MRI data corresponding to the lesion and glandular tissue of all the patients. Prior to construction of contribution scattergram, the intensity values for all the MRI parameters were normalized in the range of zero to one. The intensity normalization step allows us to compare the inter-image distances with one another and prevents any potential bias in the contribution scattergram representation that may occur due to difference in the image intensity ranges. However, the inter-image distances in the contribution scattergrams would still need to be scaled based on the tumor

size, D to enable the comparison of contribution scattergrams between different patients. To that effect, the pairwise inter-image distances between different MRI images were normalized with the $\sqrt{D}$, effectively transforming the inter-image distances in the range of zero to one.

I applied the manifold learning algorithm, Isomap to transform each patient's multiparametric MRI into a low-dimensional manifold representation of its intrinsic dimensionality. The graph theoretic metrics of degree centrality and betweenness centrality were used to determine the optimal range of the neighborhood parameter, k. The neighborhood connectivity, k is a very critical parameter as it determines the structure of the contribution scattergram and subsequently affects the graph theoretic metrics. If the value of k is very small, the resultant contribution scattergram will be a disconnected graph containing different clusters and we would not be able to characterize the flow between these clusters. On the other hand, if the value of k is very large, the Isomap algorithm would produce spurious short circuit connections resulting in an incorrect contribution scattergram embedding. I computed the optimal range of the neighborhood parameter, k as the values of k which produced statistically significantly different graph theoretic metrics for benign and malignant patients.

Wilcoxon's rank sum test was employed to determine if the graph theoretic measures of degree and betweenness centrality were significantly different between benign and malignant patients. **Figures 8.3 (a) and (b)** illustrate the identified optimal range of k values for the average betweenness centrality and the average degree centrality, respectively. The value of the neighborhood parameter, k was chosen as 8 from the optimal range {6,7,8,10} as it was one of the median optimal values. The value of d computed using the residual

variance procedure was obtained at $d = 2$ corresponding to the neighborhood parameter, $k = 8$.

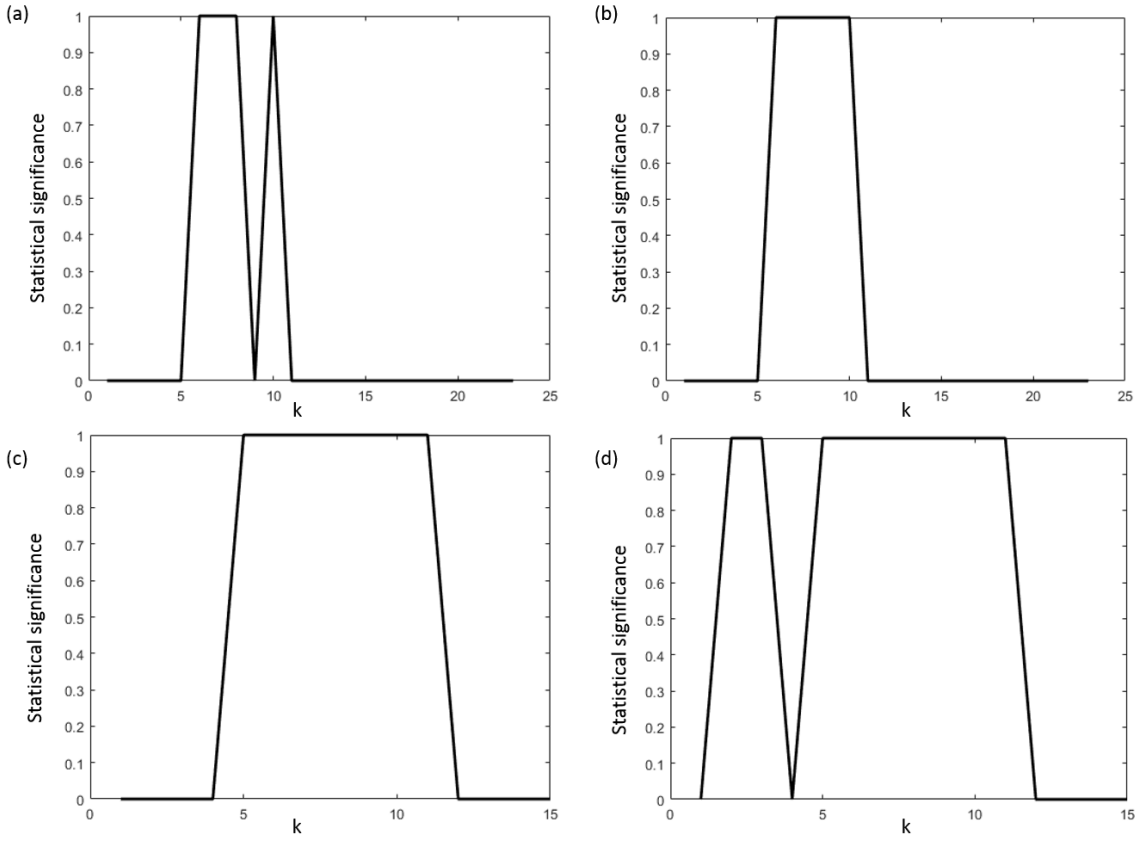


Figure 8.3 Plot of statistical significance vs. k for different graph centrality measures. The statistical significance value was set at one if the Wilcoxon rank sum test between benign and malignant sets produced a p value ≤ 0.05 and zero otherwise. (a) The plot for average betweenness centrality for the CS consisting of all the MR parameters. The notch corresponds to $p = 0.18$ (b) The plot for average degree centrality for the CS consisting of all the MR parameters (c) The plot for average betweenness centrality for the DCE-MRI CS (d) The plot for average degree centrality for the DCE-MRI CS. The notch here corresponds to $p = 0.06$.

I further computed the inter-image geodesic distances on the contribution scattergram to characterize the complex interactions between different MRI images while the contributions of each MRI image were characterized using the graph centrality measures along with clustering coefficient.

Additionally, I analyzed two subgraphs corresponding to the functional MRI parameters of DCE-MRI and DWI-ADC map. The DCE-MRI CSg included the fifteen high temporal resolution DCE images and the DWI-ADC CSg included the two DWI images and the ADC map. **Figures 8.3(c) and (d)** illustrate the identified optimal range of k values for the average betweenness centrality and degree centrality, respectively for the DCE-MRI subgraph. The optimal range of k values was determined as $K \in \{5,6,7,8,9,10,11\}$. The neighborhood parameter, k was again chosen as 8 as it was the median of K. The graph theoretic measures of betweenness centrality, degree centrality, closeness centrality and clustering coefficients were computed for the DCE-MRI CSg to identify the contribution of the DCE images at different time points after contrast injection as well as how the wash-in and wash-out phases clustered within the subgraph. The DWI-ADC subgraph consisted of only three images, leaving only one possible value of k; i.e., $k = 1$. Furthermore, closeness centrality alone was used to identify the contribution of the DWI and ADC map in the DWI-ADC CSg as the DWI-ADC CSg just had three nodes and one path.

8.4.4 Patient classification

I implemented patient classification using the hybrid IsoSVM feature transformation and classification algorithm based on the Isomap [43] and the Support Vector Machine (SVM) algorithms [216] as described in **Chapter 6**.

The imbalance in the number of benign and malignant patients may result in the SVM being trained in favor of the dominant group, which in this case was the malignant patient group. The penalty for misclassification of benign patients was set higher than the penalty for misclassification of malignant patients while training the SVM to avoid training bias. Finally, the optimal values of Isomap neighborhood parameter and the misclassification penalty ratio (benign : malignant) were estimated using leave-one-out and k-fold cross validation ($k = 10$).

8.4.5 K-fold Cross validation

The benign and the malignant patients were individually separated into ten randomly sampled subsets as the number of patients in the two groups were highly unbalanced. The ten patient subsets were then formed by combining the ten subsets from the two groups. Consequently, the ratio between the number of benign and malignant patients was maintained in each of the ten subsets used for ten-fold cross validation. Finally, the procedure for ten-fold cross validation was repeated a hundred times to overcome any bias occurring due to specific data partitions.

8.5 Results

8.5.1 Experimental Summary

The contribution scattergrams were computed and analyzed for one hundred and thirty-nine women that underwent multiparametric MRI scan. The mean age of the patients was 52 years (range: 24-80 years). Ninety- six women (69%) had malignant lesions, thirty-nine women (28%) had benign lesions and four patients had no lesion (3%). The average contribution scattergrams for the benign, malignant and normal patient groups are illustrated in **Figure 8.4**

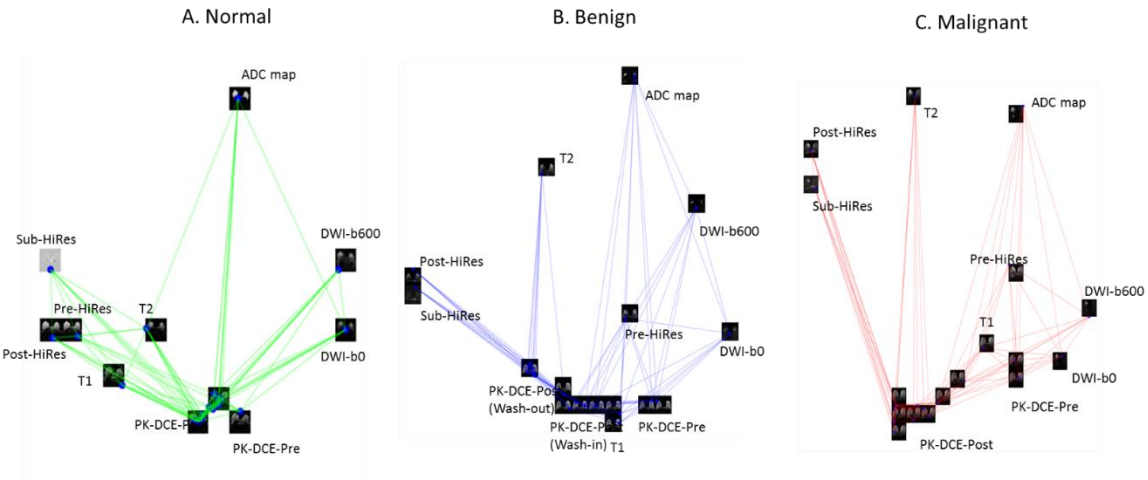


Figure 8.4 Visualization of the average contribution scattergrams obtained for (a) Normal (b) Benign and (c) Malignant patients. There are significant structural differences in the three contribution scattergrams. For example, the degree distribution across different MRI parameters is significantly different between normal, benign and malignant patients ($p<0.05$). Moreover, the sub graphs for DWI and DCE-MRI are also significantly different ($p<0.05$) as elaborated in the chapter.

The contribution scattergrams corresponding to benign and malignant lesions demonstrated significantly different average degree centrality (DC) (Benign DC: 10.55±0.08; Malignant DC: 10.82±0.05; $p=0.004$) while the glandular tissue contribution scattergrams exhibited no significant difference in the average degree centrality between the normal, benign and malignant patients (Normal DC: 11.17±0.17; Benign DC: 11.17±0.09; Malignant DC: 11.13±0.05; $p=0.71$). Similar trend was observed in average betweenness centrality metric as shown in **Table 8.1**.

Furthermore, the average path length (APL) and the diameter of the lesion contribution scattergrams were significantly different from the glandular tissue contribution scattergrams indicating a different structural contribution scattergram representation for normal and abnormal tissue types (Lesion APL: 0.21±0.003, Glandular APL: 0.14±0.003, $p<0.0001$; Lesion diameter: 0.63±0.01, Glandular diameter: 0.48±0.01, $p<0.0001$). The difference in the graph theoretic metrics between benign lesions, malignant lesions, and the glandular tissue have been encapsulated in **Table 8.1**.

Table 8.1. Graph summary statistics corresponding to the benign, malignant and glandular contribution scattergrams.

Graph metric	Benign	Malignant	Glandular	p value (Benign vs. Malignant)	p value (Lesion vs. Glandular)
Average path length	0.20±0.007	0.21±0.004	0.14±0.003	0.44	<0.00001
Diameter	0.64±0.02	0.63±0.01	0.48±0.01	0.71	<0.00001
Average degree centrality	10.55±0.08	10.82±0.05	11.14±0.04	0.004	<0.00001
Average betweenness centrality	7.98±0.18	7.46±0.10	6.62±0.07	0.02	<0.00001
Average degree weight	1.19±0.04	1.24±0.02	0.96±0.02	0.2	<0.00001

8.5.2 Correlation between breast tissue biology and contribution scattergram

a) DCE-MRI CSg analysis

The DCE-MRI CSg was analyzed using the neighborhood parameter, $k = 8$. The average path length (APL) between the pre-contrast DCE and the post-contrast wash-in phase DCE images was significantly (p=0.0001) higher for malignant (0.18±0.01) than for benign (0.13±0.01) patients. Furthermore, the percentage difference between the APL-wash-in (APL: DCE-pre to wash-in DCE images) and the APL-wash-out (DCE-pre to the wash-out DCE images) was significantly higher (p=0.001) for benign (96.69±12.6) than malignant patients (43.60±4.00), as also evident in **Figure 8.4**. The results from different graph theoretic measures have been summarized in **Table 8.2** and elaborated in the following subsections.

Betweenness Centrality

I first computed betweenness centrality (BC) of the subgraph to analyze the network flow within the DCE-MRI CSg. As illustrated in **Figure 8.5(a),** the betweenness centrality across the DCE images followed a similar trend for both benign and malignant patients with post contrast DCE image six and seven having a significantly higher value of betweenness centrality than the other DCE images (DCE post 6-7 BC: 5.19±0.39; DCE pre, post 1-5,8-14 BC: 1.46±0.05; p<0.0001); i.e., maximum number of paths flow through these images. The post contrast DCE images six and seven are the final images obtained during the contrast wash-in phase. However, the malignant patients had significantly (p<0.05) lower number of paths (BC: 3.36±0.58) passing through the post contrast DCE image seven than benign patients (Betweenness centrality: 9.077±1.41). On the other hand,

the number of paths flowing through the last post contrast images (DCE-14) were significantly higher for malignant (BC: 1.55±0.35) than benign patients (BC: 0.26±0.16). This result indicates that the paths linking the images formed during the wash-out phase are different for benign and malignant patients. This is consistent with the known behavior of kinetic curves following different wash-out trajectories for benign and malignant lesions.

In summary, the average betweenness centrality of the malignant and benign lesions was significantly different (Benign BC: 2.09±0.03; Malignant BC: 1.91±0.02; p=0.00001).

Degree centrality

The degree centrality (DC) exhibited a similar trend to betweenness centrality as shown in **Figure 8.5(b)** indicating the post contrast DCE images six and seven have a significantly higher degree or number of connections than the rest of the DCE images (DCE 6-7 DC: 13.19±0.11; DCE pre,1-5,8-14 DC: 9.62±0.04; $p<0.0001$). There were significant differences ($p<0.05$) in the degree distributions of both the wash-in and wash-out DCE images as presented in **Figure 8.5(b)**. In addition, the average degree centrality was significantly different between benign and malignant lesions (Benign DC: 9.83±0.07; Malignant DC: 10.21±0.03; p=0.00001).

Closeness centrality

Closeness centrality curves corresponding to benign and malignant lesions are displayed in **Figure 8.5(c).** The closeness centralities for the pre and first post contrast DCE-MRI were much higher for benign lesions than malignant lesions. Closeness centrality reveals the nearness of a DCE-MRI image to the rest of the network, which implies that the

pre-contrast and the first post-contrast images are significantly closer to the rest of the network for benign than malignant patients (**Pre** Benign CCt: 7.05±0.59 Malignant CCt: 5.18±0.24, p=0.02; **First Post** Benign CCt: 7.41±0.66 Malignant CCt: 5.46±0.29, p=0.03).

Clustering coefficient

The clustering coefficient (CC) metric indicated that the images on either side of post contrast DCE images six and seven formed separate clusters exhibiting the wash-in and wash-out DCE clusters as illustrated in **Figure 8.5(d)**. Consequently, the clustering coefficient of the DCE images six and seven is significantly lower than the rest of the DCE images (DCE post 6-7 CC: 0.71±0.01; DCE pre, post 1-5,8-14 CC: 0.88±0.004; $p<0.0001$). Furthermore, the clustering coefficients of the wash-in and wash-out DCE images were also significantly different between benign and malignant lesions as illustrated in **Figure 8.5(d).**

To summarize, the post contrast DCE images obtained at the end of the wash-in phase were identified as the most important using all the graph theoretic measures. Furthermore, the graph theoretic measures also highlighted the differences in the wash-in and wash-out characteristics of the benign and malignant lesions concurring with the known biological knowledge of the DCE-MRI characteristics.

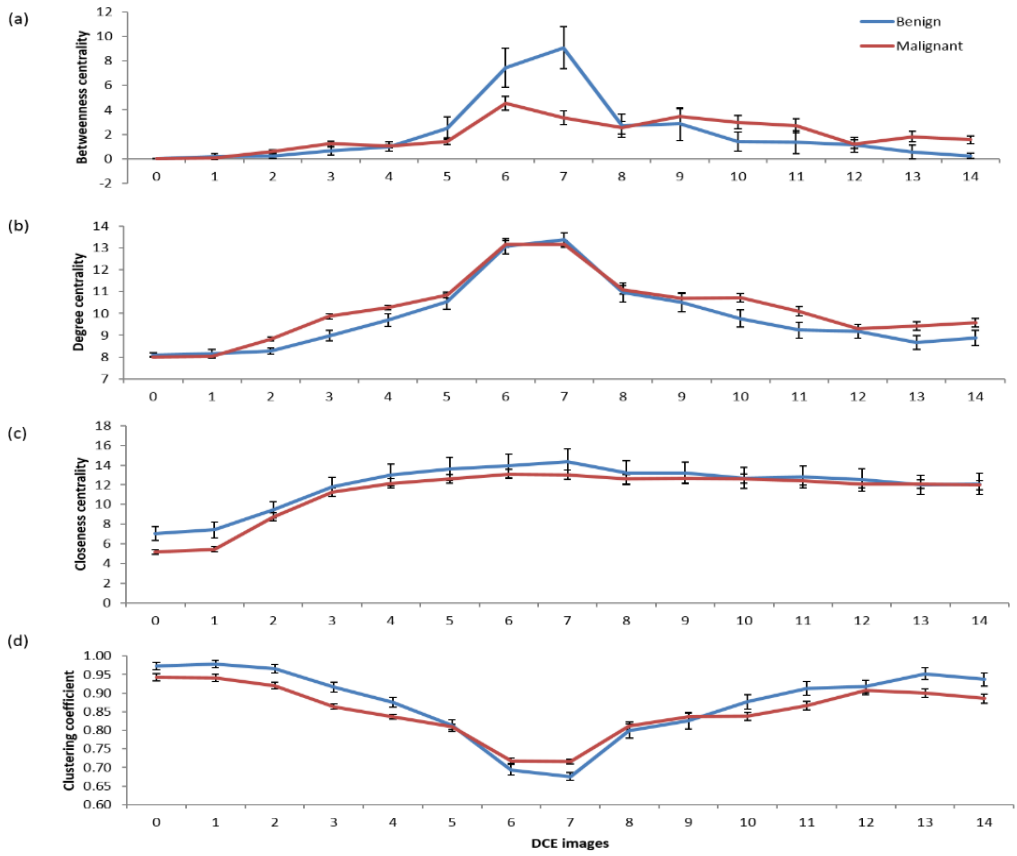


Figure 8.5 The results from different centrality measures and the clustering coefficient have been illustrated here. The DCE images 6 and 7 were identified as the most central images. Furthermore, there were significant differences in the centrality values of the wash-in and wash-out images between benign patients and malignant patients. (a) The plot of betweenness centrality values for all the high temporal resolution DCE images. (b) The plot of degree centrality values for all the high temporal resolution DCE images. (c) The plot of closeness centrality values for all the high temporal resolution DCE images. (d) The plot of clustering coefficients for all the high temporal resolution DCE images. The DCE images formed two clusters on either side of DCE images 6 and 7 depicting the wash-in and wash-out clusters.

Table 8.2. (a) Betweenness centrality metrics corresponding to the benign and malignant DCE contribution scattergrams.

DCE images	Betweenness centrality		
	Benign	Malignant	P value
0	0.00±0.00	0.02±0.01	0.158
1	0.18±0.18	0.09±0.04	0.644
2	0.23±0.16	0.61±0.14	0.081
3	0.64±0.27	1.24±0.20	0.082
4	1.00±0.31	1.04±0.15	0.907
5	2.51±0.74	1.43±0.28	0.179
6	7.44±1.29	4.52±0.58	0.047
7	9.08±1.41	3.36±0.58	0.001
8	2.72±0.76	2.52±0.51	0.832
9	2.85±1.10	3.48±0.60	0.621
10	1.44±0.64	3.00±0.53	0.065
11	1.36±0.75	2.71±0.57	0.159
12	1.15±0.50	1.21±0.33	0.929
13	0.56±0.46	1.81±0.43	0.052
14	0.26±0.16	1.55±0.35	0.001

Table 8.2. (b) Degree centrality metrics corresponding to the benign and malignant DCE contribution scattergrams.

DCE images	Degree Centrality		
	Benign	Malignant	p value
0	8.10±0.08	8.00±0.00	0.2100
1	8.15±0.15	8.02±0.01	0.3947
2	8.28±0.11	8.83±0.10	0.0003
3	8.97±0.19	9.86±0.11	0.0002
4	9.69±0.23	10.26±0.11	0.0289
5	10.54±0.28	10.84±0.14	0.3455
6	13.08±0.29	13.18±0.15	0.7650
7	13.38±0.24	13.18±0.15	0.4734
8	10.95±0.36	11.08±0.20	0.7461
9	10.51±0.36	10.70±0.22	0.6630
10	9.77±0.32	10.72±0.21	0.0155
11	9.23±0.30	10.09±0.21	0.0215
12	9.18±0.27	9.31±0.18	0.6823
13	8.67±0.26	9.42±0.20	0.0244
14	8.87±0.28	9.58±0.19	0.0430

Table 8.2. (c) Closeness centrality metrics corresponding to the benign and malignant DCE contribution scattergrams.

DCE images	Closeness centrality		
	Benign	Malignant	p value
0	7.05±0.59	5.18±0.24	0.02
1	7.41±0.66	5.46±0.29	0.03
2	9.46±0.70	8.74±0.38	0.45
3	11.80±0.80	11.28±0.44	0.64
4	12.97±0.90	12.16±0.45	0.51
5	13.59±0.95	12.60±0.45	0.44
6	13.91±0.99	13.08±0.47	0.53
7	14.36±1.04	13.01±0.46	0.34
8	13.22±0.99	12.59±0.47	0.64
9	13.19±0.88	12.64±0.48	0.65
10	12.68±0.87	12.62±0.47	0.96
11	12.80±0.90	12.41±0.45	0.75
12	12.51±0.94	12.09±0.43	0.74
13	12.00±0.81	12.06±0.43	0.96
14	12.09±0.88	11.97±0.48	0.92

Table 8.2. (d) Clustering coefficient metrics corresponding to the benign and malignant DCE contribution scattergrams.

DCE images	Clustering coefficient		
	Benign	Malignant	p value
0	0.97±0.01	0.94±0.01	0.028
1	0.98±0.01	0.94±0.01	0.010
2	0.97±0.01	0.92±0.01	0.002
3	0.92±0.01	0.86±0.01	0.001
4	0.88±0.01	0.84±0.01	0.011
5	0.81±0.02	0.81±0.01	0.837
6	0.69±0.01	0.72±0.01	0.123
7	0.68±0.01	0.72±0.01	0.002
8	0.80±0.02	0.81±0.01	0.551
9	0.83±0.02	0.84±0.01	0.637
10	0.88±0.02	0.84±0.01	0.093
11	0.91±0.02	0.87±0.01	0.043
12	0.92±0.02	0.91±0.01	0.570
13	0.95±0.02	0.90±0.01	0.011
14	0.94±0.02	0.89±0.01	0.023

b) DWI-ADC CSg

The DWI-ADC CSg consisted of DWI-b0, DWI-b600, and the corresponding ADC map. The closeness centrality (CCt) of DWI-b0 was significantly higher than the DWI-b600 and ADC map (DWI-b0 CCt: 7.14±0.37; DWI-b600, ADC map CCt: 5.30±0.24; $p<0.0001$) across all the patients as displayed in **Figure 8.6**. Furthermore, the closeness centralities of both the DWI b values were significantly higher for malignant than benign lesions (**DWI-b0**: Benign CCt: 5.10±0.46 Malignant CCt: 7.69±0.43, $p=0.0001$; **DWI-b600** Benign CCt: 4.24±0.38 Malignant CCt: 5.63±0.27, $p=0.005$). Closeness centrality is an indication of the speed with which information travels from a given node to the rest of the network. Here, the meaning of closeness centrality can be translated to water diffusion in the region of interest. Moreover, the APL for the benign DWI-ADC CSg (0.26±0.08) was significantly higher (0.0007) than the APL for the malignant CSg (0.20±0.08) which again agrees with the biological knowledge that water molecules diffuse a larger distance for benign than malignant lesions. The results are summarized in **Table 8.3.**

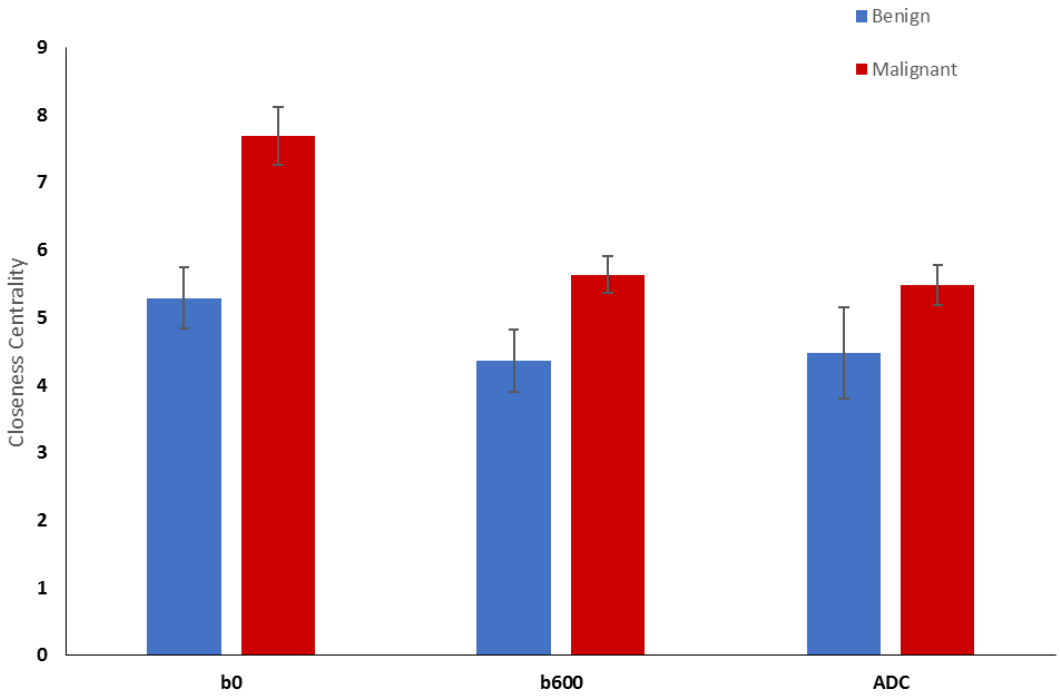


Figure 8.6 Plot of closeness centrality values for all the DWI images and ADC map. The closeness centrality values were significantly higher ($p < 0.05$) for benign patients than for malignant patients across all the images.

Table 8.3. Graph centrality metrics and summary statistics corresponding to the benign and malignant DWI-ADC contribution scattergrams.

	Closeness centrality			Average path length	Diameter
	b0	b600	ADC		
Benign	5.29±0.37	4.36±0.38	4.48±0.55	0.26±0.02	0.39±0.02
Malignant	7.69±0.43	5.63±0.27	5.48±0.30	0.20±0.01	0.29±0.01
p value	0.0001	0.009	0.115	0.0001	0.0001

8.5.3 Breast Cancer Diagnosis

The IsoSVM model was trained using ten graph theoretic metrics to establish the usefulness of CSg representations in classifying malignant from benign breast tumors. The ten features were based on APLs from the DCE-MRI and DWI-ADC CSg in addition to betweenness and degree centralities from DCE-MRI CSg and closeness centrality from DWI-ADC CSg. The IsoSVM model trained using leave-one-out cross validation resulted in sensitivity and specificity of 89% and 82% respectively with an AUC of 0.87 in classifying benign from malignant lesions. The optimal IsoSVM parameters obtained were d=9, k=85, and misclassification penalty ratio = 2.25:1 (benign: malignant) corresponding to the maximum performance of the IsoSVM algorithm. Moreover, the ten-fold cross validation performed across 100 trials yielded an AUC of 0.86. **Table 8.4** shows a comparison of the performance of the IsoSVM with other linear and nonlinear kernels used with the SVM algorithm using leave one out cross validation. The ROC curves from all the SVM kernels

and TOC curve for the IsoSVM kernel are illustrated in **Figure 8.7 and Figure 8.8**, respectively.

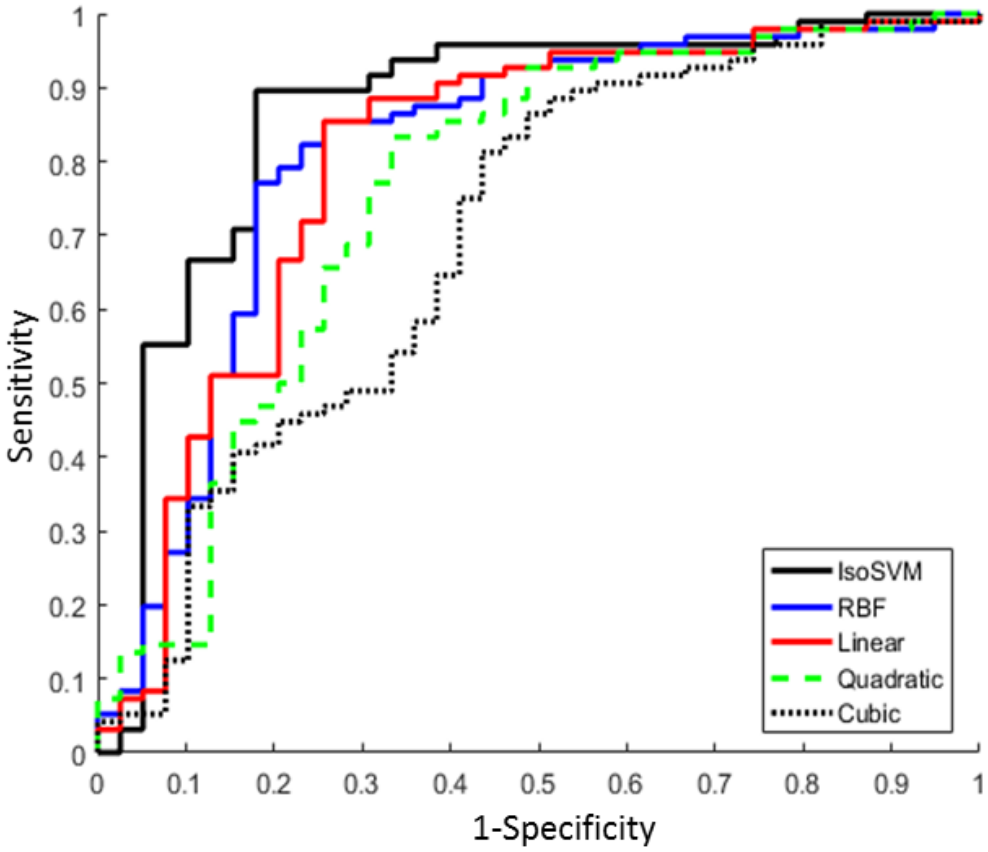


Figure 8.7 The receiver operating characteristic curves corresponding to the hybrid IsoSVM kernel classification (black), radial basis function (RBF) kernel SVM (blue), linear kernel SVM (red), quadratic kernel SVM (dashed green), and the cubic kernel SVM kernel (dotted black) evaluated using leave one out cross validation. The area under the ROC were obtained at 0.87, 0.81, 0.80, 0.76, and 0.70 for IsoSVM, RBF, linear, quadratic, and cubic kernel SVMs, respectively.

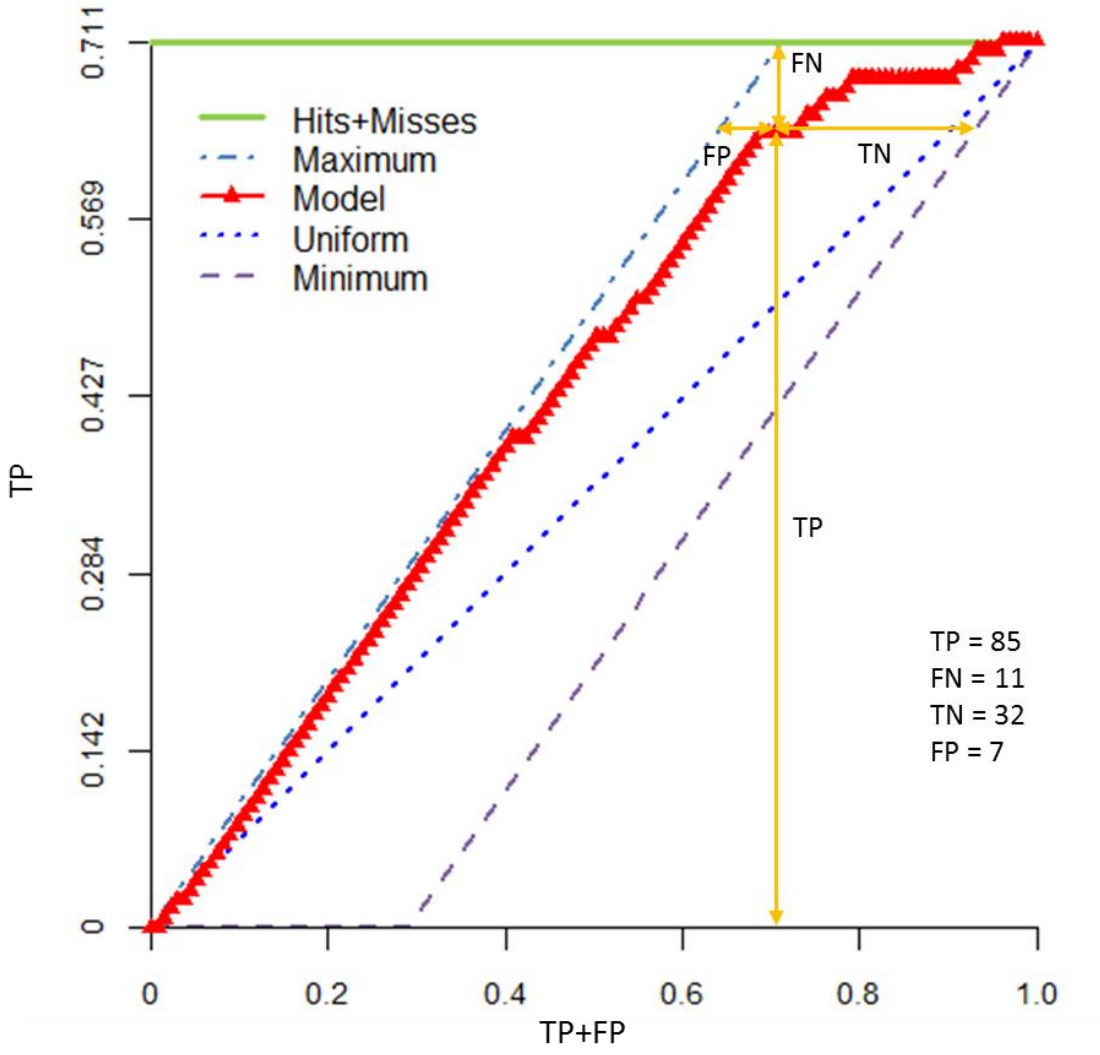


Figure 8.8 The total operating characteristic (TOC) curve for the IsoSVM classifier. The true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP) for the optimal IsoSVM hyperplane were 85, 11, 32, and 7, respectively.

Table 8.4. Summary of sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve (AUC) for the hybrid IsoSVM kernel and various linear and nonlinear SVM kernels. The optimal input parameters obtained for each classifier are also mentioned. The Isomap neighborhood parameter and the dimensionality of the transformed low-dimensional embedding are indicated by k and d, respectively. Finally, the sigma used for the radial basis kernel is denoted as sigma.

Kernel	Misclassification penalty (Benign:Malignant)	Classification parameters	Sensitivity	Specificity	AUC
RBF	3:1	sigma = 3.2	82	77	0.81
Linear	2:1		88	69	0.80
Quadratic	1:1		83	62	0.76
Cubic	3.5:1		81	54	0.70
IsoSVM	2.25:1	k=85, d=9	89	82	0.87

8.6 Discussion

We have demonstrated that the novel manifold-based representation of contribution scattergram for visualization and analysis of multiparametric MRI produced excellent network representations that were correlated to breast tissue biology and patient diagnosis. The graph summary statistics of average path length and diameter distinctly distinguished lesion tissue (benign or malignant) from normal tissue (glandular tissue). Furthermore, the complex network analysis of contribution scattergrams demonstrated excellent sensitivity and specificity in classifying benign from malignant lesions as compared to the state-of-the-art methods.

The major motivation behind the development of contribution scattergrams was its potential application to the field of precision medicine where the goal is to understand individual properties of each patient and treat them accordingly. The contribution scattergram visualizes and identifies the organizational principles of the multiparametric MRI image space corresponding to different patients and pathologies, bringing a new perspective to radiological diagnosis. Furthermore, the organizational principles of multiparametric MRI image space can be established at the level of individual patients, providing an opportunity for clinicians to identify graphical properties specific to each contribution scattergram and use these properties to produce a personalized diagnosis and treatment options for patients.

A large number of studies spanning different fields of applications such as gene co-expression networks, traffic networks, brain networks, and social networks have demonstrated the efficacy of complex network analysis in finding the most contributing parameters in the network known as the network hubs [237-244]. Here, I extended the

concept of complex network analysis to multiparametric MRI networks and identified the contribution of different MRI parameters to the network representation. The most important images in the DCE network were identified as the final images during the wash-in phase of the DCE-MRI which concur with the most important image in the DCE-MRI as recognized by the radiologists. Moreover, the complex network analysis of each individual patient networks allowed us to identify how the contributions of different MRI images varied from patient to patient.

The DCE-CSg can be considered as a multidimensional integrated kinetic curve corresponding to a hypervoxel comprising all the voxels in a region of interest (lesion or glandular tissue). The complex network analysis of the DCE-CSg provides us with an opportunity to study the complete tumor as a whole and understand the complex interactions between the tumor representations at different time points post contrast injection. In addition, kinetic curve analysis can also be done on the DCE-CSg representation to produce the integrated metrics of wash-in rate, wash-out rate, kinetic curve type corresponding to the tumor sized hypervoxel.

The DWI-ADC CSg quantified the relationships between the different diffusion weighted images of b0 and b600. The average path length and closeness centrality metrics extracted from the DWI-ADC CSg were significantly different between benign and malignant patients suggesting their utility as potential integrated DWI biomarkers for the complete tumor.

I used the geodesic distance to model the network connections in the contribution scattergram. However, other similarity and dissimilarity metrics such as normalized mutual information, correlation coefficient, and angular distance can also be used to model

network connections in the contribution scattergram. Furthermore, the Isomap algorithm was used to visualize the contribution scattergram in the lower dimensional space. There are many different dimensionality reduction algorithms such as diffusion map, LLE, and force atlas that can also be used to visualize the contribution scattergram [41, 253, 257]. The effects of visualization and analysis of contribution scattergrams using different proximity measures and different manifold learning algorithms would be an interesting direction for future work.

In conclusion, I proposed a novel complex network model approach for learning intrinsic manifold representations in multiparametric imaging data called the contribution scattergram. Furthermore, a graph theoretic framework was implemented for the extraction of intrinsic properties of individual contribution scattergrams for clinical diagnosis. The IsoSVM model trained on the extracted graph theoretic metric achieved excellent sensitivity and specificity as compared to the state of the art even with huge difference in the number of benign and malignant patients. Moreover, the correlation between the biological knowledge about the breast tissue and contribution scattergram was established in this chapter. In the future, I plan to extend the contribution scattergram to multimodal imaging settings. The value of the neighborhood parameter, k along with the types of subgraphs being analyzed would have to be adjusted depending on the organ being imaged, imaging modality and the diagnostic application. In conclusion, the contribution scattergram provides a powerful tool for visualization and complex interaction analysis of multiparametric radiological images with broad future applicability.

Part III: Feature Modeling

Chapter 9: IRIS: Integration of radiological, clinical and histopathological parameters into a personalized decision support framework.

9.1 Introduction

The previous chapters have laid the foundation for extracting useful information from high-dimensional multiparametric imaging data. The complete set of features that extracted from radiological images include biological features such as cellularity and vascularity obtained from functional imaging, radiomic features such as texture and shape of the tissue of interest and finally, the contribution scattergram features that highlight the complex interactions between different imaging parameters. Furthermore, these features can be combined with clinical and histopathological parameters that are acquired during the clinical work up.

The advancement in computational sciences in recent years, more specifically in areas of high-performance computing and memory management has given rise to an era of "big data" in medicine. The high dimensionality of the dataset obtained by integrating different information sources as well as the complexity within the data presents significant challenges for statistical data analysis. Conventional machine learning algorithms such as support vector machines and discriminant analysis concatenate all the different parameters from disparate sources into a single vector for analysis which can lead to overfitting and does not provide any biological meaning since each information source has an intrinsic

property. Furthermore, the features acquired from different information sources could be entangled in a highly correlated complex network, making it desirable to understand the complex interactions between the informatics parameters as well as identifying the importance of each of the parameters for the task at hand. The challenge here is to combine both mpMR images with clinicopathologic parameters to stratify patients and provide personalized clinical decision support.

To answer this challenge, I have developed a new machine-learning informatic method termed Integrated Radiomic Information System (IRIS) that integrates clinical parameters, radiological parameters, morphological parameters, and patient information (age, sex) to build a clinical decision support model. Herein, I will call this complete set of parameters as informatics parameters. The technique of nonlinear dimensionality reduction (NLDR) makes up the backbone of IRIS. By definition, dimensionality reduction (DR) is a mathematical mapping that transforms the high-dimensional data into a meaningful representation of its intrinsic dimensionality (lower dimensional representation). The intrinsic dimensionality refers to a lower number of variables that can represent the underlying true structure of the data.

The purpose of this chapter is to integrate all the features developed in this thesis with the quantitative features manually extracted from imaging and clinical workup into a multiview clinical decision support system for predicting breast cancer recurrence and compare with the OncotypeDX 21-gene array assay. Oncotype DX is based on the mRNA expression by RT-PCR for estrogen receptor (ER) positive disease without the human growth factor receptor 2 (HER2-nu) overexpression and is tested with biopsy or surgical samples [258]. Oncotype DX has been validated in prospective-retrospective studies as a

prognostic tool in ER-positive patients treated with tamoxifen. The Oncotype DX score stratifies the patients into three risk groups: low, medium, and high-risk. The Oncotype DX score has been shown to identify patients most likely to benefit from the addition of adjuvant chemotherapy to endocrine therapy [258, 259]. For example, the side effects of undergoing chemotherapy outweigh its benefits for patients with low risk.

The major disadvantage of the OncotpyeDX test is its cost, which is $4000 per test. However, if we can accurately predict the Oncotype DX score with high confidence using the routinely acquired clinical, radiological, and pathological features, the clinician could opt to forego ordering the OnctoypeDX text. The Oncotype DX test will only need to be ordered in cases when the model predicts the outcome with a low confidence score. To that end, I developed the IRIS technique and applied it to the prediction of breast cancer recurrence risk score using all the information available from routinely acquired data.

The remainder of the chapter is organized as follows: **Section 9.2** establishes the theory for clinical decision support heatmap. A brief overview of image acquisition, analysis different feature extraction methods used in this study constitute **Section 9.3**, followed by application of IRIS for prediction of breast cancer recurrence risk in **Section 9.4**. This chapter concludes with a discussion on the advantages and disadvantages of IRIS for application to precision medicine.

Publication from this work and author contributions

Publication

*M. A. Jacobs, **V. Parekh**, C. Umbricht, K. J. Macura, R. EI-Khouli, S. Harvey, I. Kamel, A. C. Wolff, "Novel advanced machine learning informatics modeling using clinical and radiological imaging metrics for characterizing breast tumor characteristics with the OncotypeDX gene array," Ready for Submission, 2018*

Author contributions

I wrote the complete chapter. Dr. Jacobs reviewed it and helped with the editing.

I developed the Informatics Radiomics Integration System (IRIS) and did the entire implementation. The data obtained from radiological imaging (ADC map and PK-DCE) and histological phenotyping (Ki-67) was analyzed by Dr. Jacobs (**Sections 9.3.1 (a)** and **9.3.2 (a-c)**). I analyzed the features extracted from radiomics and contribution scattergram. I integrated all the data for visualization and patient classification.

The clinical data was acquired by Dr. Umbricht, Dr. Macura, Dr. EI-Khouli, Dr. Harvey, Dr. Kamel, Dr. Wolff, and Dr. Jacobs. This included patient recruitment, scanning, and radiological reporting (**Section 9.3.1)**.

9.2 Information Radiomics Integration system (IRIS)

The goal of the IRIS clinical decision support system is to transform the patient space, represented as $X = \{x_1, x_2, \dots, x_n\} \subset R^D$ where, x_i represents the i^{th} patient, n represents the number of patients and D represents the number of informatics parameters, into an integrated heatmap as shown in **Figure 9.1** using the following procedure.

In the first step, the D-dimensional patient space is distributed into N subspaces where each subspace corresponds to a data view e.g. clinical, histopathological, radiomics, and contribution scattergram. The second step involves transformation of each subspace, $S_i = \{x_1, x_2, \dots, x_n\} \in R^{d_i}\ \forall i \in \{1,2, \dots, N\}$ into a risk score, $Y_i = \{x_1, x_2, \dots, x_n\} \in R^1\ \forall i \in \{1,2, \dots, N\}$ using a nonlinear dimensionality reduction algorithm such as Isomap, Diffusion map and locally linear embedding) [41-43, 247, 248]. The risk score embedding from each subspace is concatenated in the third step to produce the IRIS decision support

heatmap representation as shown in **Figure 9.1**. The columns of IRIS correspond to different patients while the rows correspond to different feature subspaces. Each cell is color coded to represent the risk predicted by the corresponding feature subspace for that patient. The color coding is based on the jet color scale (blue–red) where in blue represent low risk and red represents high risk. In the fourth step, support vector machine (SVM) [216] classifier is trained on the N-dimensional transformed space, Y to produce patient prognosis.

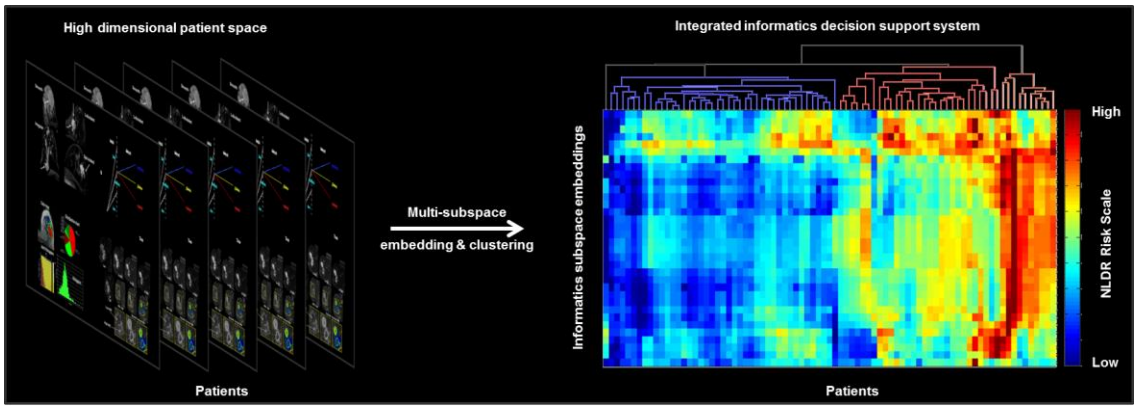


Figure 9.1 Illustration of the IRIS clinical decision support system. The high-dimensional patient space consisting of different patients and their corresponding informatics parameter subspaces (left) are transformed into an integrated informatics decision support system, IRIS (right). The IRIS is represented using a heatmap where the color-scale (blue–red) indicates risk as identified by each embedding (low, intermediate, and high-risk clusters).

9.3 Materials and methods

*The clinical data was acquired by Dr. Umbricht, Dr. Macura, Dr. EI-Khouli, Dr. Harvey, Dr. Kamel, Dr. Wolff, and Dr. Jacobs (**Section 9.3.1**). The data obtained from radiological imaging (ADC map and PK-DCE) and histological phenotyping (Ki-67) was analyzed by Dr. Jacobs (**Sections 9.3.1 (a)** and **9.3.2 (a-c)**).*

9.3.1 Clinical Subjects

All studies were performed in accordance with the institutional guidelines for clinical research under a protocol approved by the Institutional Review Board (IRB) and all HIPAA agreements were followed, no consent was signed. Seventy patients were analyzed as part of a retrospective breast research study at our institution.

a) Histological Phenotyping

Breast lesions were categorized by histological phenotyping based upon hormonal markers: Estrogen and progesterone receptors (ER and PR), HER2-Nu by FISH, and Ki-67 proliferation index (%). Histopathological data was obtained from the clinical reports. Moreover, the patients were matched to the clinical criteria for the Oncotype DX test (ER positive).

b) Multiparametric Breast Imaging

Patients were scanned on a 3T MRI system (3T Achieva, Philips Medical Systems, Best, The Netherlands) using a bilateral, dedicated four-channel, phased array breast coil (InVivo, Orlando, FL) with the patient in the prone position. Multiparametric breast imaging for all the three risk groups as defined by OncotypeDX have been illustrated in **Figure 9.2.** The multiparametric MRI imaging protocol and the algorithms for registration, tissue signature generation, and image segmentation have been previously detailed in **Chapter 2**.

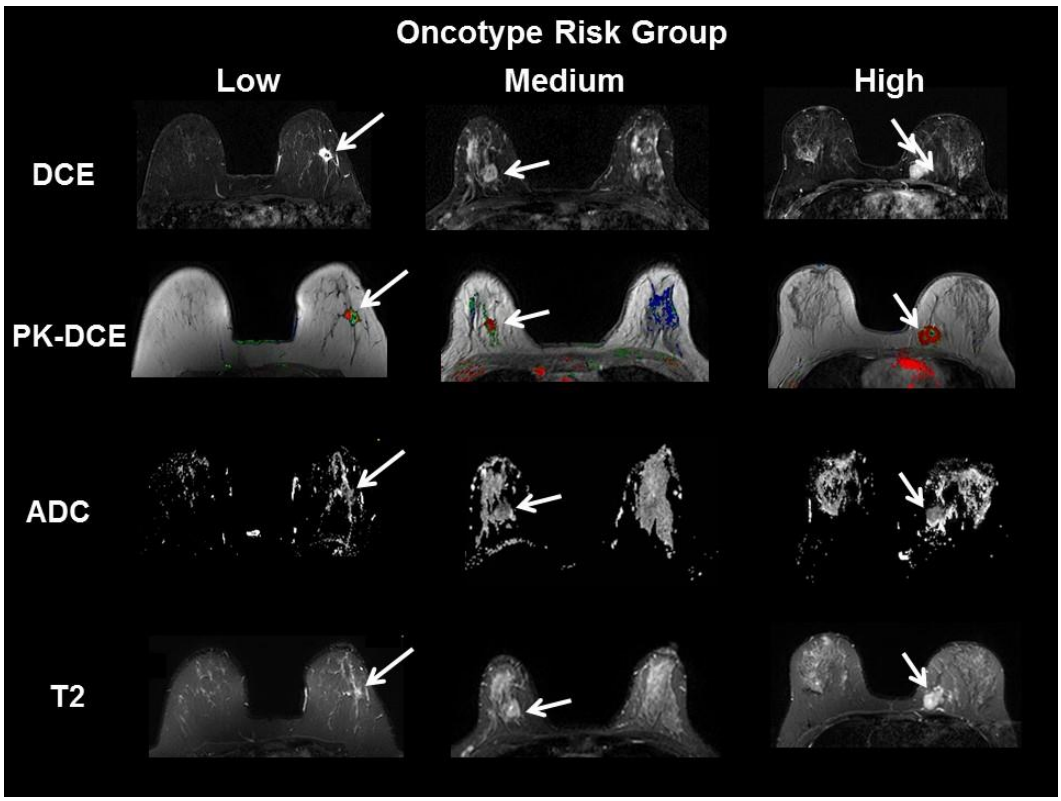


Figure 9.2 Demonstration of multiparametric MRI imaging of each risk group defined by the Oncotype DX. Left) typical imaging of the low-risk group. Middle) typical imaging of the medium-risk group, and Right) typical imaging of the high-risk group. Note, the PK-DCE all demonstrate malignant phenotype, however, by integrating all the data using IRIS, I was able to separate each group.

9.3.2 MRI data analysis

a) Clinical breast lesion classification methods

Breast lesions were identified by the radiologist on the breast MRI who was blinded to pathological results defined by the BIRADS lexicon [260]. Breast density was defined as markedly or heterogeneously dense, moderately glandular, or fatty. Lesions were classified

into mass or non-mass like enhancement (NMLE). Morphologic assessment was defined for masses, margins (1=smooth/focal, 2=lobulated/regional, 3=irregular, or 4=spiculated/ductal/segmental) and enhancement patterns (1=homogenous, 2=heterogeneous/stippled/reticular, or 3=rim/clumped). For NMLE, distribution (1=diffuse/focal, 2=regional, 3=ductal, 4=reticular, and 5=segmental) and enhancement pattern (homogenous, heterogeneous/stippled/reticular, or clumped) were recorded. The lesion morphology was defined into seven classes (1=focal NMLE, 2=regional NMLE, 3=ductal or segmental NMLE, 4=smooth mass, 5=lobulated mass, 6=irregular mass, or 7=spiculated mass).

b) Pharmacokinetic Contrast Enhancement Metrics

Pharmacokinetic kinetic DCE MRI provides quantitative metrics of the volume transfer constant (K_{trans} (min^{-1})) which characterize uptake of the contrast agent, the leakage within the extracellular extravascular space (v_e (%)), and the transfer rate constant (k_{ep} (min^{-1})). Post-processing of the DCE exam was performed by a combined Brix and Tofts model [213, 214, 261-263] using DynaCad (InVivo, FL) software from the identified breast lesions.

c) ADC Mapping

Regions of Interest (ROI) were drawn on normal appearing glandular tissue and breast lesions defined by DCE MRI. Means and standard deviations were calculated for both tissue types. Ratios of lesion ADC to glandular tissue ADC (L/GT) were calculated from the following equation on lesion and glandular tissue [59].

$$\text{ADC Ratio} = \frac{\text{ADC value of Lesion}}{\text{ADC value of glandular tissue}}$$

d) Radiomics

I extracted 51 radiomic features each from T1, T2, DWI and DCE imaging sequences, divided into five categories - First Order Statistics (14 features), Gray Level Co-occurrence Matrix (GLCM, 18 features), Gray Level Run Length Matrix (GLRLM, 11 features), Neighborhood Gray Tone Difference Matrix (NGTDM, 5 features), fractal dimension features (2 features) and convexity. In broad terms, first order features describe the statistics of voxel intensity distributions within the ROI while second order features such as GLCM, GLRLM, and NGTDM seek to quantify textural features. GLCM and GLRLM were evaluated in four directions (0°, 45°, 90°,135°) with values averaged to achieve rotational invariance, and the GLCM parameter, d, was set to one voxel in all directions. I used 64 equally sized bins for First Order Statistics, and gray levels were quantized into 64 levels for GLCM, GLRLM and NGTDM calculations. Both binning and gray levels were systematically varied between 8 to 512 (specifically 8, 16, 32, 64, 128, 256, 512), and 64 was empirically chosen in each case.

e) Contribution Scattergram Analysis

I constructed contribution scattergrams for the multiparametric MRI data corresponding to the lesion. The neighborhood parameter, k was set to 7 as empirically chosen in **Chapter 8**. Furthermore, the sub-graphs of PK-DCE and DWI-ADC were evaluated for sub-graph analysis with neighborhood parameter, k set at 7 and 1 for PK-DCE and DWI-ADC sub-graphs, respectively.

9.4 Results

9.4.1 Clinical Demographics

The IRIS model was tested on a breast cancer patient cohort of 70 patients who had both multiparametric MRI imaging and Oncotype assay. Of the 70 patients, there were 19 (27%) patients with low risk (0-17), 42 (60%) patients with intermediate risk (18-31), and 9 (13%) patients with high risk (>31). There was no age difference (51-56y/o) between the patients.

9.4.2 Radiological Findings

The high-risk group had the largest tumor size ($7.6 \pm 5.8 cm^2$), followed by the low-risk group tumor size ($5.8 \pm 9.0 cm^2$) and the intermediate-risk group ($4.6 \pm 5.4 cm^2$). For advanced MRI parameters, there were clear differences in each parameter and Oncotype risk groups. The PK-DCE parameters (K_{trans}, EVF). The K_{trans} values for the high- and intermediate-risk groups were higher ($0.28\ and\ 0.41\ (1/min)$) compared to the low-risk group ($0.16(1/min)$). Similar results were noted for the other PK-DCE parameters. The maximum contrast enhancement was largest for the high-risk group ($523 \pm 145 s$), compared to the intermediate-risk ($434 \pm 138 s$), and low-risk groups ($489 \pm 139 s$). Similarly, the ADC map values from the high- and intermediate-risk patients in the lesion tissue were significantly lower ($p < 0.05$) than those for the low-risk patients ($1.11\ vs\ 1.49 x 10^{-3} mm^2/s$) as summarized in **Table 9.1**. However, the ADC map values in glandular tissue remained constant across all groups ($2.14 \pm 0.03 x 10^{-3}\ mm^2/s$). The bar graphs are shown in **Figure 9.3.**

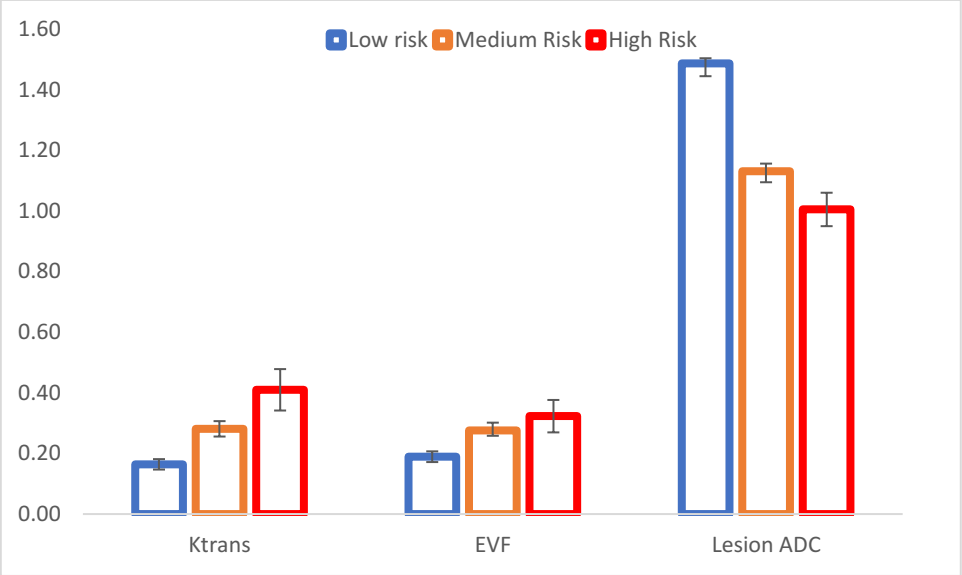


Figure 9.3. Bar graphs of quantitative multiparametric MRI parameters from the IRIS model demonstrating significant differences between low, medium and high-risk patient groups.

9.4.3 Contribution Scattergram

Figure 9.4 illustrates the average contribution scattergram for low-risk group and medium/high-risk group. The contribution scattergrams were further analyzed using sub graph analysis of DWI-ADC CSg and DCE-MRI CSg.

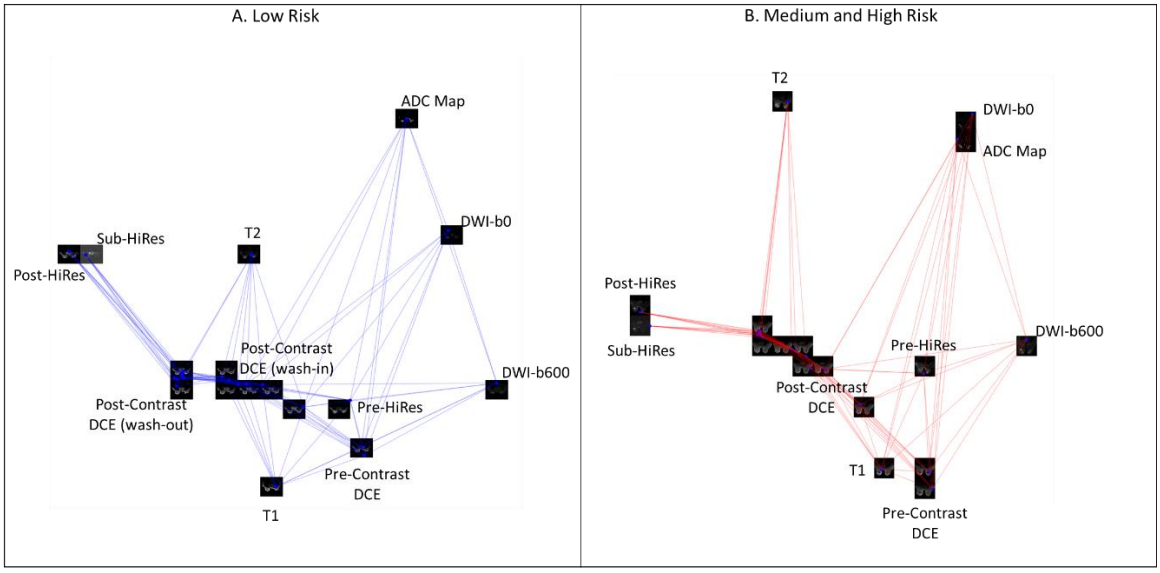


Figure 9.4 Visualization of the average contribution scattergrams obtained for (a) Low-Risk (b) Medium and High-Risk Patients. There are significant structural differences in the two contribution scattergrams ($p < 0.05$) as elaborated in this chapter.

a) DWI-ADC CSg

The average closeness centrality of the DWI-ADC subgraph for low-risk patients (5.15 ± 0.30) was significantly ($p = 0.01$) lower than that of medium and high-risk patients (6.31 ± 0.31). The closeness centrality for all the three images is shown in **Figure 9.5.** Similarly, the average path length between the high diffusion b value (DWI b600) and ADC map was significantly ($p < 0.05$) higher for low-risk patients (0.31 ± 0.02) than for medium and high-risk patients (0.31 ± 0.02). These results indicate that water diffuses for a larger distance in low-risk patients than for high-risk patients.

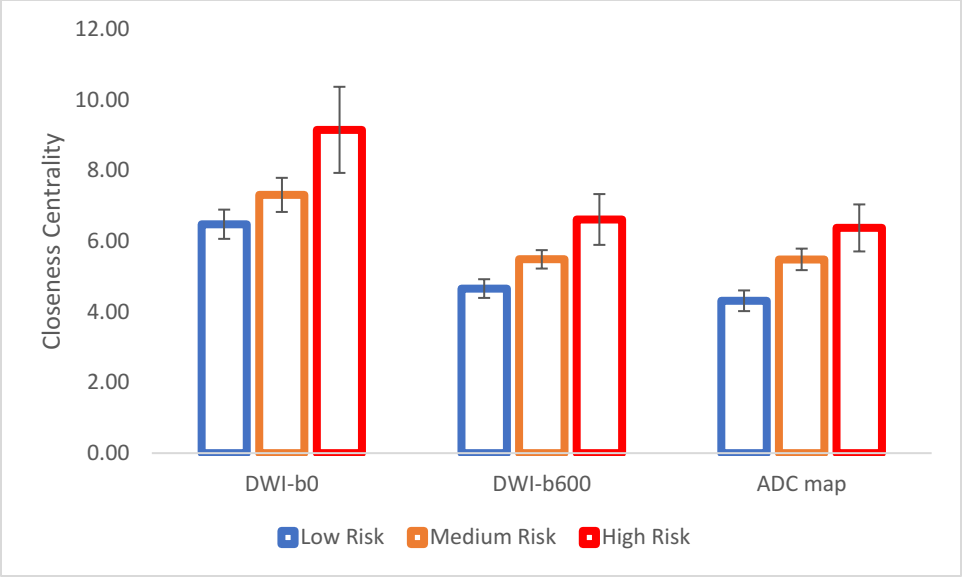


Figure 9.5 Closeness centrality values for all the DWI images and the ADC map. The closeness centrality values were found to be significantly higher ($p < 0.05$) for medium and high-risk patients as compared to low-risk patients.

b) DCE-MRI CSg

The sub-graph analysis of DCE-MRI revealed significant differences in the wash-out characteristics of patients belonging to different groups. The average path length for the wash-out sub-graph of DCE-MRI was significantly higher for medium and high-risk patients (0.03 ± 0.002) than that of low-risk patients (0.02 ± 0.002). Similarly, degree centrality revealed significant differences between the low-risk and medium/high-risk patients during the washout phase as shown in **Figure 9.6**

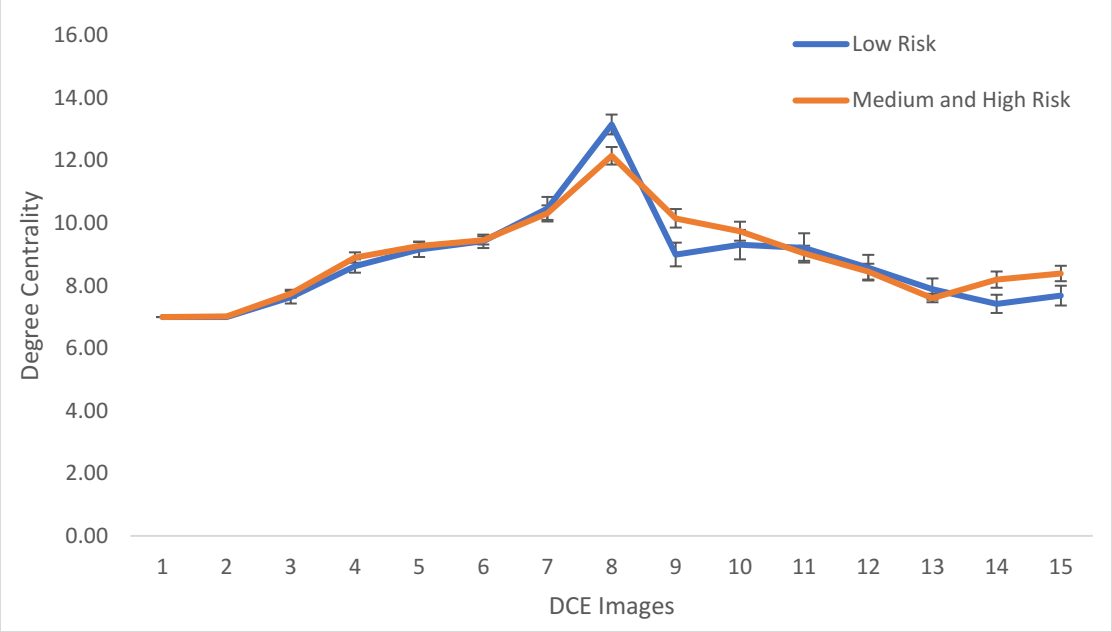


Figure 9.6 Degree centrality values across the complete DCE temporal evolution. The degree centrality of the images in the wash-out phase was significantly different between low-risk and medium/high-risk patients.

9.4.4 Radiomics

The ADC map and DCE-MRI produced the top radiomic features for distinguishing between low, medium, and high-risk patients, while the radiomic features derived from T1 and T2 weighted imaging failed to significantly differentiate between the three groups. The top DCE features obtained were gray-level run length matrix features of short run emphasis (SRE), short run low gray level emphasis (SRLGE) and short run high gray level emphasis (SRHGE) as illustrated in **Figure 9.7.** For ADC radiomics, top features were obtained to be first order statistical features of root mean square and skewness in addition to gray level co-occurrence matrix (GLCM) features of autocorrelation, sum average and variance as illustrated in **Figure 9.8** and tabulated in **Table 9.1.**

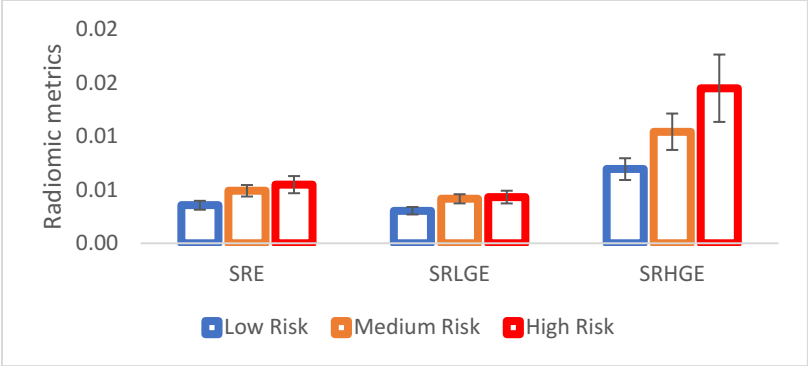


Figure 9.7 Illustration of the top DCE radiomic features for distinguishing between low, medium, and high-risk patient groups. Here SRE refers to short run emphasis, SRLGE refers to short run low gray level emphasis, and SRHGE refers to short run high gray level emphasis.

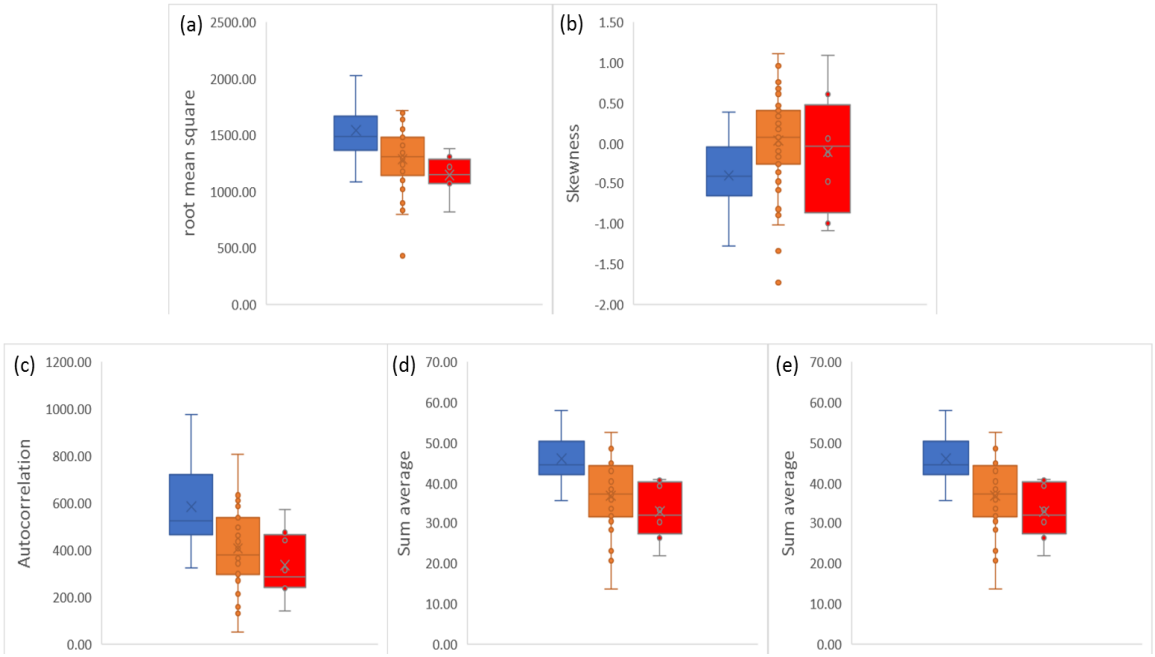


Figure 9.8 Illustration of the top ADC radiomic features for distinguishing between low, medium, and high-risk patient groups.

Table 9.1 Top features from each feature group for classification of low, medium and high-risk breast cancer recurrence patients.

Features		Low Risk	Medium Risk	High Risk	p value
ADC Radiomics	Root mean square	1539.45±57.71	1260.75±36.89	1145.79±57.67	0.0004
	Skewness	-0.40±0.10	0.01±0.09	-0.11±0.25	0.006
	Autocorrelation	585.12±42.80	393.95±22.38	335.37±48.11	0.001
	Sum average	45.93±1.58	36.10±1.21	32.91±2.31	p<0.0001
	Variance	593.79±43.22	401.60±22.48	341.44±48.80	0.001
Contribution Scattergram	APL Washout DCE	0.02±0.00	0.03±0.00	0.03±0.01	0.009
	Average degree (6-7) DCE	11.82±0.13	11.24±0.17	10.61±0.48	0.010
	APL b600-ADC	0.31±0.02	0.23±0.01	0.21±0.02	0.001
	Closeness centrality ADC	4.31±0.29	5.64±0.28	6.38±0.67	0.003
	Average closeness centrality (b600-ADC)	4.49±0.24	5.66±0.25	6.50±0.66	0.002
PK-DCE Features	K_{trans}	0.16±0.02	0.30±0.03	0.41±0.07	p<0.0001
	EVF	0.19±0.02	0.29±0.02	0.32±0.05	0.001
Quantitative DWI	Lesion ADC	1.49±0.04	1.11±0.03	1.01±0.06	p<0.0001
	Glandular tissue ADC	2.17±0.06	2.13±0.04	2.14±0.04	0.622
	ADC Ratio	0.69±0.02	0.53±0.02	0.47±0.03	p<0.0001
	Ki-67	6.83±1.04	20.04±2.13	30.00±5.21	p<0.0001

ADC: Apparent Diffusion Coefficient
DCE: Dynamic Contrast Enhancement
APL: Average Path length
PK: Pharmacokinetics
K_{trans}: volume transfer constant
EVF: Extracellular extravascular fraction
ADC_L: Mean Apparent Diffusion Coefficient of the lesion
ADC_G: Mean Apparent Diffusion Coefficient of the glandular tissue

9.4.3 IRIS Model

The IRIS heatmap demonstrating the risk profile of each patient is shown in **Figure 9.9**. The top five feature embeddings included Radiomic features, Contribution Scattergram features, PK-DCE metrics, ADC metrics, and Ki-67. The IRIS heatmap visualizes the individual contribution of each informatic parameter subspace using a color scale (low-risk: blue to high-risk: red). As shown in **Figure 9.9**, the intermediate and high-risk patients demonstrated significantly higher risk predictions as compared to low-risk patients. The SVM applied to this five-dimensional embedding produced a sensitivity and specificity of 89.47% and 84.31% respectively with an AUC of 0.93 using leave-one-out cross validation. The ROC and TOC curves for the analysis are shown in **Figure 9.10 and Figure 9.11**, respectively.

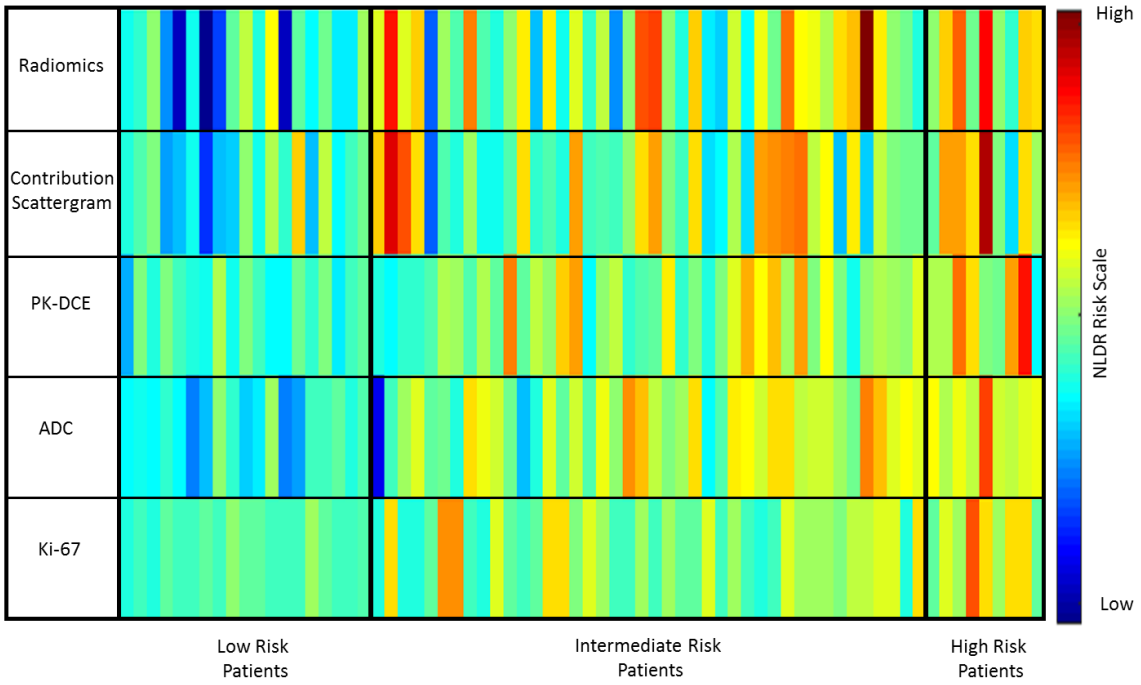


Figure 9.9 The IRIS clinical decision support system for breast cancer prognosis. The IRIS heatmap here comprises of five feature subspaces – radiomics, contribution scattergram, PK-DCE metrics, ADC metrics, and Ki-67. The IRIS is represented using a heatmap where the color-scale (blue–red) indicates risk as identified by each embedding (low, intermediate, and high-risk clusters).

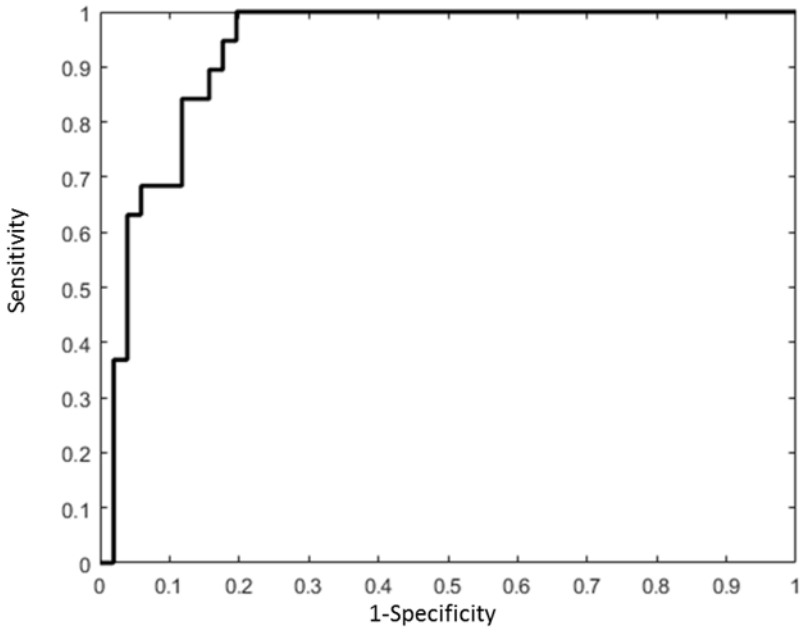


Figure 9.10 The receiver operating characteristic (ROC) curve of the five-dimensional multiview feature embedding. The sensitivity was 89.47% and specificity was 84.31% with an AUC of 0.93

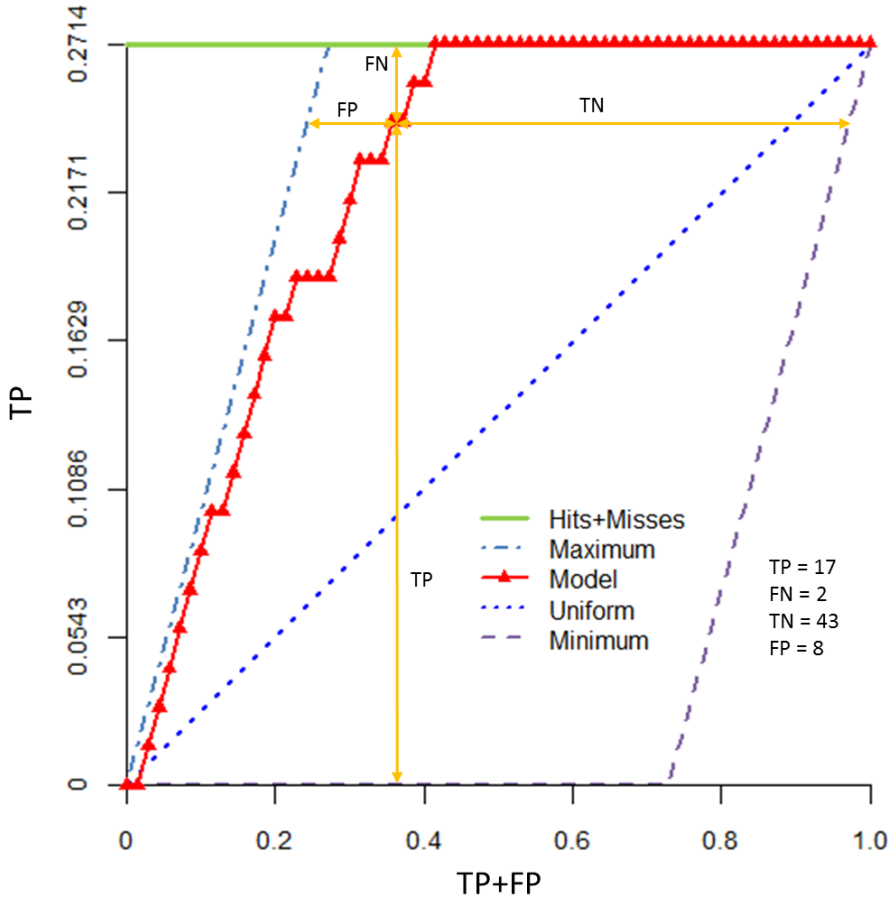


Figure 9.11 The total operating characteristic (TOC) curve of the five-dimensional multiview feature embedding. The true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP) for the optimal IsoSVM hyperplane were 17, 2, 43, and 8, respectively.

9.5 Discussion

I have introduced and demonstrated an advanced multiview feature embedding and visualization model to visualize and analyze the relationships and interactions between mpMRI parameters, clinical, and histological variables compared with OncotypeDX assay. The IRIS model was able to group patients into the three different groups from integration

of the data. More importantly, I defined important radiological and clinical variables for tumor recurrence and visualized the interaction between each variable with the OncotypeDX, which could lead to a non-invasive test for recurrence. These radiological variables provide a unique window into the pathophysiology of these tumors and concrete measures of comprehensive tumor characteristics. For example, the parameters of radiomics, contribution scattergram, ADC map values, PK-DCE metrics, and Ki-67, which reflect the cellularity, vascularity and proliferation index of the tumor in addition to their intrinsic textural heterogeneity and relationships between these parameters. This integrated model of radiological, clinical, and histological data will be useful in determining personalized characteristics of patients. The ability to combine these quantitative measures would be an important step in ensuring that "the right patient receives the right treatment".

I developed and implemented an integrated informatics decision support system (IRIS) for the purpose of diagnosis or prognosis. Furthermore, the IRIS heatmap provides a visualization of relationship between different patients along with quantifiable embedding metrics. Using the IRIS heatmap, we would be able to identify a patient or a group of patients with the most similar embedding metrics to a new patient with an unknown risk of recurrence. Understanding these complex relationships between different embeddings can provide an insight on how these embeddings are related at biological level for predicting recurrence of breast cancer.

In conclusion, these initial studies provide insight into the molecular underpinning of the surrogate radiologic features and provide the foundation to relate these changes to the OncotypeDX score, and eventually, apply these methods to the assessment of treatment response for improved personalized medicine.

Chapter 10: Thesis Summary and future work

10.1 Thesis Summary

This thesis presents the integrated graph theoretic, radiomics, and deep learning (I-GRAD) framework, which is an end-to-end system for modeling the structural and biological composition of different pathologies and the complex interactions between different features and integrating them into a personalized clinical decision support system. There were three major challenges in developing an end-to-end personalized clinical decision support framework.

1. Develop an algorithm for automatic segmentation and classification of different tissue types in the radiological imaging dataset.

   My first goal was to develop a supervised approach for segmentation of different tissue from radiological imaging datasets. However, my training dataset consisting of labeled voxels corresponding to different tissue types across all the patients was sparse and under-sampled due to high cost and time required to manually create segmentation masks. I overcame this issue by implementing stacked sparse autoencoders (SSAE). The SSAE is a deep neural network that can be pre-trained in an unsupervised fashion by using all the input data (labeled and unlabeled) to learn an intrinsic representation of the input and encode it within its architecture. The pre-trained SSAE can then be fine-tuned using the labeled data to accurately segment different tissue types. The multiparametric deep learning (MPDL) model based on SSAE accurately segmented lesion, glandular, and fatty tissue from 192 breast cancer patients. The average dice similarity between eigen-filter segmented lesions and MPDL

segmented lesions was 0.87 ± 0.05 for malignant patients and 0.84 ± 0.07 for benign patients. Of the 192 patients, 142 patients were imaged at 3T and in axial plane while the remaining 50 were imaged in sagittal plane at 1.5T. The MPDL was optimized on 3T data using two-fold cross validation and was tested on the 1.5T data. My results demonstrated the robustness of the MPDL model to imaging orientation, MRI parameters acquired, and magnetic field strength.

The MPDL model resolved the problem of tissue segmentation from multiparametric breast MRI using a sparsely labeled dataset. Next, I analyzed how the MPDL model encoded the multiparametric breast MRI datasets. This is important because this would help us predict when the network will fail. Moreover, we can compare how the intrinsic representations learnt by MPDL relate to the intrinsic representations learnt by humans. For this purpose, I visualized the weights connecting different layers and reconstructed inputs from different layer encodings. The encodings learnt by the MPDL model were very similar to how humans encode different relationships in their brain, e.g. Tumors are brighter on post-contrast DCE as compared to pre-contrast DCE, Fat is bright on T1WI and dark on T2WI. Furthermore, the SAE correlation reduction algorithm was able to extract DCE kinetic curve patterns from breast dataset concurrent with the established literature on DCE-MRI.

The inter-parametric relationships encoded by the MPDL are consistent irrespective of the underlying application (organ or pathology). As a result, the intrinsic representations learnt for one application could be used for another application. As my final contribution to segmentation of multiparametric

radiological imaging datasets, I evaluated the robustness of the intrinsic relationships encoded by MPDL when applied to segmentation of datasets acquired for different organs with different pathology. This study produced encouraging results on test datasets of brain tumor mpMRI and brain stroke mpMRI opening up the possibility of a universal deep learning framework for segmentation of different tissue types across multiple organs.

2. Extract the texture, shape and inter-parameter complex interactions from the tissue of interest segmented from the radiological images.

   I developed the techniques of radiomic feature mapping (RFM) and contribution scattergram (CSg) to model the spatial appearance and inter-parameter relationships corresponding to a tissue of interest (TOI) segmented from multiparametric radiological images. The radiomic feature mapping (RFM) algorithm was discussed in the sixth chapter. The current methods in radiomics are not standardized. In fact, radiomic features extracted for the same application (e.g. breast diagnosis) have been found to be different across different groups, as discussed in **Chapter 5**. To that end, I developed standardized features for characterizing the cellular and vascular heterogeneity of the TOI. The vascular heterogeneity was computed by analyzing the time evolution of texture with contrast enhancement in DCE-MRI. The radiomic features were normalized to negate the effect of size on the textural evolution curves. Similarly, I developed a textural evolution metric (TEM) for characterizing the evolution of texture with increasing b values. In addition, the RFM features were correlated with tissue biology quantitative metrics of ADC

and PK-DCE. The RFM features modeled the underlying cellular and vascular heterogeneity in the breast cancer dataset of one hundred and twenty-four patients for classification between benign and malignant tumors. The RFM features produced excellent classification accuracy with sensitivity and specificity of 93% and 85% respectively with an AUC of 0.91.

The RFM features were based on texture analysis of single radiological images, which resulted in hundreds of texture features creating a huge dataset across all images for analysis. These radiomic features do not reflect the true underlying tissue contrast, heterogeneity, or homogeneity and only provide limited information corresponding to the physical modeling of each imaging parameter. As a result, I extended the RFM method to develop multiparametric radiomics (MPRAD) to characterize the texture of the complete multiparametric imaging datasets. The MPRAD framework modeled the "true" texture of the underlying tissue by combining the multiview information obtained from different imaging parameters. In addition, the MPRAD framework also extracted features based on the complex interactions between different imaging parameters, consequently opening up a completely new source of information that did not exist with conventional radiomic features.

My final contribution to the field of feature extraction from multiparametric radiological images was the technique of contribution scattergram (CSg). The CSg modeled the inter-parametric relationships between different imaging parameters as a complex network model. The CSg was distinct from previously developed methods for extraction of inter-parametric relationships, in that, the

CSg modeled the complete network structure governing the inter-parametric relationships as opposed to modeling a specific biological characteristic (e.g. pharmaco-kinetic modeling for DCE).

The contribution scattergram algorithm is based on manifold learning and complex network analysis for visualization and evaluation of the graphical network formed by the inter-parameter interactions, respectively. I developed CSg metrics based on graph theoretic analysis centrality measures to capture the vascularity (PK-DCE) and cellularity (DWI) of the tumor as a whole in contrast to single voxel measurements evaluated using conventional methods. The CSg was applied to model the inter-parameter relationships in a breast cancer patient cohort of one hundred and thirty-nine patients. The CSg classified the benign lesions from malignant lesions with excellent sensitivity and specificity of 89% and 82% respectively with an AUC of 0.87.

3. Model the extracted information into an intrinsic representation that would be useful for identifying the most similar patients as well as classifying the patients based on predefined classes.

   My final contribution to this thesis entails development of the multiview feature integration model, IRIS for end-to-end clinical decision support. The IRIS was tested on a breast mpMRI dataset for prediction of breast cancer recurrence risk. The breast cancer recurrence risk is traditionally measured using Oncotype DX score, which costs around $4000. The goal of this work was to develop and implement IRIS on features extracted from data that is routinely collected and accurately predict breast cancer recurrence risk. Specifically, I wanted to model

the radiomics, contribution scattergram, clinical, and pathological features (extracted routinely in clinical practice) into an end-to-end clinical decision support system.

The IRIS transforms each subspace formed by a feature group such as radiomics and contributions scattergram into a risk score that can be evaluated for final decision support. The IRIS system visualized each embedding on a heatmap where the risk scores are color-coded from blue (low-risk) to red (high-risk). Using the IRIS heatmap, most similar patients as well as an overall risk score were obtained. The IRIS system demonstrated excellent accuracy in prediction of breast cancer recurrence risk in a cohort of 70 patients with sensitivity and specificity of 89.47% and 90.20% respectively with an AUC of 0.93.

Apart from a high accuracy in predicting breast cancer recurrence risk score, IRIS could potentially be used to predict "true" risk score for patients with intermediate risk score. The intermediate-risk patients are the patients for which we do not know the true risk group (low vs. high). Using IRIS, we can identify similar patients from history as well evaluate risk score corresponding to individual biomarkers to compute the true risk score for the patient in question.

In summary, the I-GRAD framework presented in this thesis can be used to model any pathology using radiological imaging data or in combination with clinical and histopathological data and produce a personalized patient diagnosis, prognosis, or treatment plan.

10.2 Future Work

1. **Universal Organ Invariant Tissue Segmentation**

   In this thesis, the multiparametric deep learning tissue signature model was developed to automatically segment different breast tissue types. In addition, detailed inspection of each layer of the MPDL trained for breast tissue segmentation revealed that it was able to learn the basic features of imaging physics such as "fluid in bright on T2 and dark on T1". These features are universally present for the same set of imaging parameters for any organ, which was reasserted by my work in **Chapter 4**. In the future, my aim is to develop a universal organ invariant tissue signature segmentation database that can potentially segment any tissue type from any MRI dataset.

2. **Unsupervised tissue segmentation**

   I plan to develop a completely unsupervised counterpart of the multiparametric deep learning (MPDL) tissue signature model which can be used to segment any multiparametric imaging dataset without prior training dataset. This is especially useful in radiological applications as it is difficult to have a training dataset where all kinds of possible tissue types can be pre-labeled.

3. **Radiomic and Contribution Scattergram synthesis using deep learning**

   Both the textural features and the inter-parameter relationships can be captured by a single 3D convolutional neural network (CNN) model, provided we gather enough training data. The 3D CNN can capture the 2D texture information in the initial layers of the neural network in addition to capturing the inter-parameter relationships across the third dimension. However, such a neural network model

would require a huge training dataset owing to the large number of trainable parameters. This is one direction that I plan to extend my work into as it would enable to develop a single model to extract all the information present in a radiological dataset.

4. **Precision Radiology Applications**

   I plan to implement and validate the I-GRAD framework for precision radiology across multiple different applications enlisted below:

   - Breast cancer treatment response assessment to neoadjuvant therapy
   - Distinguishing true progression from radiation necrosis after stereotactic radiotherapy for brain metastasis
   - Distinguishing true progression from pseudo progression in WHO Grade IV glioblastoma patients
   - Characterization of prostate lesions with histopathology

BIBLIOGRAPHY

[1] S. Sinha, F. A. Lucas-Quesada, N. D. Debruhl, J. Sayre, D. Farria, D. P. Gorczyca*, et al.*, "Multifeature analysis of Gd-enhanced MR images of breast lesions," *Journal of Magnetic Resonance Imaging,* vol. 7, pp. 1016-1026, 1997.

[2] G. Ertas, H. O. Gulcur, and M. Tunaci, "Improved lesion detection in MR mammography: three-dimensional segmentation, moving voxel sampling, and normalized maximum intensity-time ratio entropy," *Acad Radiol,* vol. 14, pp. 151-61, Feb 2007.

[3] M. Q. Cao, S. T. Suo, X. B. Zhang, Y. C. Zhong, Z. G. Zhuang, J. J. Cheng*, et al.*, "Entropy of T2-weighted imaging combined with apparent diffusion coefficient in prediction of uterine leiomyoma volume response after uterine artery embolization," *Acad Radiol,* vol. 21, pp. 437-44, Apr 2014.

[4] A. S. Kierans, G. L. Bennett, T. C. Mussi, J. S. Babb, H. Rusinek, J. Melamed*, et al.*, "Characterization of malignancy of adnexal lesions using ADC entropy: comparison with mean ADC and qualitative DWI assessment," *Journal of Magnetic Resonance Imaging,* vol. 37, pp. 164-171, 2013.

[5] Y. J. Ryu, S. H. Choi, S. J. Park, T. J. Yun, J. H. Kim, and C. H. Sohn, "Glioma: application of whole-tumor texture analysis of diffusion-weighted imaging for the evaluation of tumor heterogeneity," *PLoS One,* vol. 9, p. e108335, 2014.

[6] G. Brinkmann, U. H. Melchert, G. Lalk, L. Emde, J. Link, C. Muhle*, et al.*, "The total entropy for evaluating 31P-magnetic resonance spectra of the liver in healthy volunteers and patients with metastases," *Invest Radiol,* vol. 32, pp. 100-4, Feb 1997.

[7] J. L. Cui, C. Y. Wen, Y. Hu, T. H. Li, and K. D. Luk, "Entropy-based analysis for diffusion anisotropy mapping of healthy and myelopathic spinal cord," *Neuroimage,* vol. 54, pp. 2125-31, Feb 1 2011.

[8] Y. Chen and T. D. Pham, "Sample entropy and regularity dimension in complexity analysis of cortical surface structure in early Alzheimer's disease and aging," *J Neurosci Methods,* vol. 215, pp. 210-7, May 15 2013.

[9] S. T. Suo, X. X. Chen, Y. Fan, L. M. Wu, Q. Y. Yao, M. Q. Cao*, et al.*, "Histogram analysis of apparent diffusion coefficient at 3.0 T in urinary bladder lesions: correlation with pathologic findings," *Acad Radiol,* vol. 21, pp. 1027-34, Aug 2014.

[10] G. J. Cook, C. Yip, M. Siddique, V. Goh, S. Chicklore, A. Roy*, et al.*, "Are Pretreatment 18F-FDG PET Tumor Textural Features in Non–Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy?," *Journal of Nuclear Medicine,* vol. 54, pp. 19-26, 2013.

[11] G. D. Tourassi, D. M. Delong, and C. E. Floyd Jr, "A study on the computerized fractal analysis of architectural distortion in screening mammograms," *Physics in medicine and biology,* vol. 51, pp. 1299-1312, 2006.

[12] Q. Guo, J. Shao, and V. F. Ruiz, "Characterization and classification of tumor lesions using computerized fractal-based texture analysis and support vector machines in digital mammograms," *International journal of computer assisted radiology and surgery,* vol. 4, pp. 11-25, 2009.

[13] P. Gibbs and L. W. Turnbull, "Textural analysis of contrast-enhanced MR images of the breast," *Magnetic Resonance in Medicine,* vol. 50, pp. 92-98, 2003.

[14] W. Chen, M. L. Giger, H. Li, U. Bick, and G. M. Newstead, "Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images," *Magnetic Resonance in Medicine,* vol. 58, pp. 562-571, 2007.

[15] A. Ahmed, P. Gibbs, M. Pickles, and L. Turnbull, "Texture analysis in assessment and prediction of chemotherapy response in breast cancer," *Journal of Magnetic Resonance Imaging,* vol. 38, pp. 89-101, 2013.

[16] A. Karahaliou, K. Vassiou, N. S. Arikidis, S. Skiadopoulos, T. Kanavou, and L. Costaridou, "Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis," *The British Journal of Radiology,* vol. 83, pp. 296-309, 2010.

[17] F. Tixier, C. C. Le Rest, M. Hatt, N. Albarghach, O. Pradier, J.-P. Metges*, et al.*, "Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer," *Journal of Nuclear Medicine,* vol. 52, pp. 369-378, 2011.

[18] O. S. Al-Kadi and D. Watson, "Texture analysis of aggressive and nonaggressive lung tumor CE CT images," *IEEE Transactions on Biomedical Engineering,* vol. 55, pp. 1822-1830, 2008.

[19] S. C. Agner, S. Soman, E. Libfeld, M. McDonald, K. Thomas, S. Englander*, et al.*, "Textural kinetics: a novel dynamic contrast-enhanced (DCE)-MRI feature for breast lesion classification," *Journal of Digital Imaging,* vol. 24, pp. 446-463, 2011.

[20] T. P. Coroller, P. Grossmann, Y. Hou, E. R. Velazquez, R. T. Leijenaar, G. Hermann*, et al.*, "CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma," *Radiotherapy and Oncology,* vol. 114, pp. 345-350, 2015.

[21] A. Akhbardeh and M. A. Jacobs, "Comparative analysis of nonlinear dimensionality reduction techniques for breast MRI segmentation," *Medical physics,* vol. 39, pp. 2275-2289, 2012.

[22] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. Aerts, "Artificial intelligence in radiology," *Nature Reviews Cancer,* p. 1, 2018.

[23] G. Downing, "Biomarkers Definitions Working Group. Biomarkers and Surrogate Endpoints," *Clinical Pharmacology & Therapeutics,* vol. 69, pp. 89-95, 2001.

[24] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal.,* vol. 27, pp. 379-423, July 1948.

[25] B. B. Mandelbrot, *The fractal geometry of nature* vol. 173: Macmillan, 1983.

[26] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics,* pp. 610-621, 1973.

[27] M. M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics and Image Processing,* vol. 4, pp. 172-179, 1975/06/01 1975.

[28] K. I. Laws, "Rapid texture identification," *Proc. SPIE 0238, Image Processing for Missile Guidance,* pp. 376-381, 1980.

[29] M. Amadasun and R. King, "Textural features corresponding to textural properties," *IEEE Transactions on Systems, Man and Cybernetics,* vol. 19, pp. 1264-1274, 1989.

[30] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho*, et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature communications,* vol. 5, p. 4006, 2014.

[31] S. Sun, "A survey of multi-view machine learning," *Neural Computing and Applications,* vol. 23, pp. 2031-2038, 2013.

[32] A. D. Barker, C. C. Sigman, G. J. Kelloff, N. M. Hylton, D. A. Berry, and L. J. Esserman, "I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy," *Clin Pharmacol Ther,* vol. 86, pp. 97-100, Jul 2009.

[33] L. J. Esserman, D. A. Berry, A. DeMichele, L. Carey, S. E. Davis, M. Buxton*, et al.*, "Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY 1 TRIAL--CALGB 150007/150012, ACRIN 6657," *J Clin Oncol,* vol. 30, pp. 3242-9, Sep 10 2012.

[34] N. M. Hylton, J. D. Blume, W. K. Bernreuter, E. D. Pisano, M. A. Rosen, E. A. Morris*, et al.*, "Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy--results from ACRIN 6657/I-SPY TRIAL," *Radiology,* vol. 263, pp. 663-72, Jun 2012.

[35] V. S. Parekh, K. J. Macura, S. Harvey, I. Kamel, R. EI-Khouli, D. A. Bluemke*, et al.*, "Multiparametric Deep Learning Tissue Signatures for a Radiological Biomarker of Breast Cancer: Preliminary Results," *arXiv preprint arXiv:1802.08200,* 2018.

[36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[37] V. Parekh and M. A. Jacobs, "Radiomics: a new application from established techniques," *Expert review of precision medicine and drug development,* vol. 1, pp. 207-226, 2016.

[38] V. S. Parekh and M. A. Jacobs, "Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI," *NPJ breast cancer,* vol. 3, p. 43, 2017.

[39] V. Parekh and M. Jacobs, "MPRAD: A Multiparametric Imaging Radiomics Framework," *Submitted, Nature Communications,* 2018.

[40] J. M. Lee, *Riemannian manifolds : an introduction to curvature*. New York: Springer, 1997.

[41] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner*, et al.*, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 102, pp. 7426-7431, May 24 2005.

[42] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science,* vol. 290, pp. 2323-2326, December 22, 2000 2000.

[43] J. Tenenbaum, V. Silva, and J. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science,* vol. 290, pp. 2319-2323, 2000.

[44] V. Parekh, A. Akhbardeh, and M. Jacobs, "Contribution Scattergram: A complex network model based on graph theory and manifold learning for creating a unique signature for high dimensional multiparametric data," *Ready for submission,* 2018.

[45] P. Athamanolap, V. Parekh, S. I. Fraley, V. Agarwal, D. J. Shin, M. A. Jacobs*, et al.*, "Trainable high resolution melt curve machine learning classifier for large-scale reliable genotyping of sequence variants," *PloS one,* vol. 9, p. e109094, 2014.

[46] L. Peng, V. Parekh, P. Huang, D. D. Lin, K. Sheikh, B. Baker*, et al.*, "Distinguishing True Progression from Radionecrosis after Stereotactic

Radiotherapy for Brain Metastases with Machine Learning and Radiomics," *International Journal of Radiation Oncology• Biology• Physics,* 2018.

[47] M. A. Jacobs, V. Parekh, C. Umbricht, K. J. Macura, R. EI-Khouli, S. Harvey*, et al.*, "Novel advanced machine learning informatics modeling using clinical and radiological imaging metrics for characterizing breast tumor characteristics with the OncotypeDX gene array," *Ready for submission,* 2018.

[48] V. S. Parekh, J. R. Jacobs, and M. A. Jacobs, "Unsupervised nonlinear dimensionality reduction machine learning methods applied to multiparametric MRI in cerebral ischemia: preliminary results," in *International Society for Optics and Photonics, Medical Imaging*, 2014, p. 90342O.

[49] V. S. Parekh and M. A. Jacobs, "A multidimensional data visualization and clustering method: Consensus similarity mapping," in *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 420-423.

[50] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition," *Medical Physics,* vol. 20, pp. 1033-1048, 1993.

[51] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annual Review of Biomedical Engineering,* vol. 2, pp. 315-337, 2003.

[52] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian*, et al.*, "A survey on deep learning in medical image analysis," *arXiv preprint arXiv:1702.05747,* 2017.

[53] C. K. Kuhl, P. Mielcareck, S. Klaschik, C. Leutner, E. Wardelmann, J. Gieseke*, et al.*, "Dynamic breast MR imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions?," *Radiology,* vol. 211, pp. 101-110, 1999.

[54] D. A. Bluemke, C. A. Gatsonis, M. H. Chen, G. A. DeAngelis, N. DeBruhl, S. Harms*, et al.*, "Magnetic Resonance Imaging of the Breast Prior to Biopsy," *JAMA,* vol. 292, pp. 2735-2742, December 8, 2004 2004.

[55] M. D. Schnall, J. Blume, D. A. Bluemke, G. A. DeAngelis, N. DeBruhl, S. Harms*, et al.*, "Diagnostic architectural and dynamic features at breast MR imaging: multicenter study," *Radiology.,* vol. 238, pp. 42-53., 2006.

[56] K. J. Macura, R. Ouwerkerk, M. A. Jacobs, and D. A. Bluemke, "Patterns of enhancement on breast MR images: interpretation and imaging pitfalls," *Radiographics,* vol. 26, pp. 1719-34; quiz 1719, Nov-Dec 2006.

[57] S. C. Partridge, W. B. DeMartini, B. F. Kurland, P. R. Eby, S. W. White, and C. D. Lehman, "Quantitative diffusion-weighted imaging as an adjunct to conventional breast MRI for improved positive predictive value," *AJR Am J Roentgenol,* vol. 193, pp. 1716-22, Dec 2009.

[58] R. H. El Khouli, K. J. Macura, P. B. Barker, M. A. Jacobs, and D. A. Bluemke, "Dynamic Contrast Enhanced Magnetic Resonance Imaging of the Breast: Effect of Temporal Resolution. ," *J Magn Reson Imaging,* vol. in press, 2009.

[59] R. H. Ei Khouli, M. A. Jacobs, S. D. Mezban, P. Huang, I. R. Kamel, K. J. Macura*, et al.*, "Diffusion-weighted imaging improves the diagnostic accuracy of conventional 3.0-T breast MR imaging," *Radiology,* vol. 256, pp. 64-73, Jul 2010.

[60] S. H. Park, W. K. Moon, N. Cho, J. M. Chang, S. A. Im, I. A. Park*, et al.*, "Comparison of diffusion-weighted MR imaging and FDG PET/CT to predict pathological complete response to neoadjuvant chemotherapy in patients with breast cancer," *Eur Radiol,* vol. 22, pp. 18-25, Jan 2012.

[61] A. Akhbardeh and M. A. Jacobs, "Methods and systems for registration of radiological images," *U.S. Patent No. 9008462,* April 14, 2015 2015.

[62] L. Van Der Maaten, E. Postma, and J. Van den Herik, "Dimensionality reduction: a comparative review," *J Mach Learn Res,* vol. 10, pp. 66-71, 2009.

[63] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science,* vol. 313, pp. 504-507, 2006.

[64] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature,* vol. 521, pp. 436-444, 2015.

[65] A. Y. Ng and M. I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," in *Advances in neural information processing systems*, 2002, pp. 841-848.

[66] T. L. Paine, P. Khorrami, W. Han, and T. S. Huang, "An analysis of unsupervised pre-training in light of recent advances," *arXiv preprint arXiv:1412.6597,* 2014.

[67] J. P. Windham, M. A. Abd-Allah, D. A. Reimann, J. W. Froelich, and A. M. Haggar, "Eigenimage filtering in MR imaging," *Journal of computer assisted tomography,* vol. 12, pp. 1-9, 1988.

[68] R. H. El Khouli, K. J. Macura, M. A. Jacobs, T. Khalil, I. Kamel, A. Dwyer*, et al.*, "Dynamic contrast-enhanced MRI of the breast: quantitative method for kinetic curve type assessment," *AJR Am J Roentgenol,* vol. 193, pp. W295-300, Oct 2009.

[69] M. A. Jacobs, R. A. Knight, J. P. Windham, Z. G. Zhang, H. Soltanian-Zadeh, A. V. Goussev*, et al.*, "Identification of cerebral ischemic lesions in rat using eigenimage filtered magnetic resonance imaging," *Brain Research,* vol. 837, pp. 83-94, 1999.

[70] H. Soltanian-Zadeh, R. Saigal, J. P. Windham, A. E. Yagle, and D. O. Hearshen, "Optimization of MRI protocols and pulse sequence parameters for eigenimage filtering," *IEEE Trans Med Imag,* vol. 13, pp. 161-175, 1994.

[71] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology,* vol. 26, pp. 297-302, 1945.

[72] J. W. Park, M. C. Liu, D. Yee, C. Yau, L. J. van 't Veer, W. F. Symmans*, et al.*, "Adaptive Randomization of Neratinib in Early Breast Cancer," *N Engl J Med,* vol. 375, pp. 11-22, Jul 07 2016.

[73] F. Rosenblatt, *Principles of Neurodynamics: Principles of Neurodynamics*. Washington D: Spartan Books, , 1961.

[74] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics,* vol. 36, pp. 193-202, April 01 1980.

[75] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *lancet,* vol. 1, pp. 307-310, 1986.

[76] H. Kaizer, "A quantification of textures on aerial photographs," *Boston University Research Lab, Boston, Massachusetts, Tech. Note,* vol. 121, 1955.

[77] R. N. Sutton and E. L. Hall, "Texture measures for automatic classification of pulmonary disease," *IEEE Transactions on Computers,* pp. 667-676, 1972.

[78] J. M. Bardeen, B. Carter, and S. W. Hawking, "The Four laws of black hole mechanics. ," *Commun.Math.Phys.,* vol. 31, pp. 161-170, 1973.

[79] J. D. Bekenstein, "Black Holes and Entropy," *Physical Review D,* vol. 7, pp. 2333-2346, 04/15/ 1973.

[80] R. Sonntag, C. Borgnakke, and G. Van Wylen, *Fundamentals of thermodynamics*. the University of Michigan: Wiley, 1998.

[81] J. Eisert, M. Cramer, and M. B. Plenio, " Area laws for the entanglement entropy," *Reviews of Modern Physics,* vol. 82, pp. 277-306, 02/04/ 2010.

[82] T. J. Larkin, H. C. Canuto, M. I. Kettunen, T. C. Booth, D. E. Hu, A. S. Krishnan*, et al.*, "Analysis of image heterogeneity using 2D Minkowski functionals detects tumor responses to treatment," *Magnetic Resonance in Medicine,* vol. 71, pp. 402-410, 2014.

[83] K. H. Knuth, "Optimal data-based binning for histograms," *arXiv preprint physics/0605197,* 2006.

[84] H. Shimazaki and S. Shinomoto, "A method for selecting the bin size of a time histogram," *Neural computation,* vol. 19, pp. 1503-1527, 2007.

[85] U. Raeth, D. Schlaps, B. Limberg, I. Zuna, A. Lorenz, G. Van Kaick*, et al.*, "Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease," *Journal of Clinical Ultrasound,* vol. 13, pp. 87-99, Feb 1985.

[86] B. S. Garra, B. H. Krasner, S. C. Horii, S. Ascher, S. K. Mun, and R. K. Zeman, "Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis," *Ultrasonic Imaging,* vol. 15, pp. 267-285, 1993.

[87] H. Sujana, S. Swarnamani, and S. Suresh, "Application of artificial neural networks for the classification of liver lesions by image texture parameters," *Ultrasound in medicine & biology,* vol. 22, pp. 1177-1181, 1996.

[88] N. A. Mayr, W. T. Yuh, J. C. Arnholt, J. C. Ehrhardt, J. I. Sorosky, V. A. Magnotta*, et al.*, "Pixel analysis of MR perfusion imaging in predicting radiation therapy outcome in cervical cancer," *Journal of Magnetic Resonance Imaging,* vol. 12, pp. 1027-1033, 2000.

[89] A. Bernasconi, S. B. Antel, D. L. Collins, N. Bernasconi, A. Olivier, F. Dubeau*, et al.*, "Texture analysis and morphological processing of magnetic resonance imaging assist detection of focal cortical dysplasia in extra-temporal partial epilepsy," *Annals of neurology,* vol. 49, pp. 770-775, 2001.

[90] C. Hayes, A. R. Padhani, and M. O. Leach, "Assessing changes in tumour vascular function using dynamic contrast-enhanced magnetic resonance imaging," *Nmr in Biomedicine,* vol. 15, pp. 154-163, Apr 2002.

[91] D. Jirak, M. Dezortová, P. Taimr, and M. Hájek, "Texture analysis of human liver," *Journal of Magnetic Resonance Imaging,* vol. 15, pp. 68-74, 2002.

[92] A. Bernasconi, "Advanced MRI analysis methods for detection of focal cortical dysplasia," *Epileptic Disord,* vol. 5 Suppl 2, pp. S81-4, Sep 2003.

[93] L. Bonilha, E. Kobayashi, G. Castellano, G. Coelho, E. Tinois, F. Cendes*, et al.*, "Texture analysis of hippocampal sclerosis," *Epilepsia,* vol. 44, pp. 1546-1550, 2003.

[94] H. Yoshida, D. D. Casalino, B. Keserci, A. Coskun, O. Ozturk, and A. Savranlar, "Wavelet-packet-based texture analysis for differentiation between benign and

malignant liver tumours in ultrasound images," *Physics in Medicine and Biology,* vol. 48, pp. 3735-3753, Nov 21 2003.

[95] F. Chabat, G.-Z. Yang, and D. M. Hansell, "Obstructive Lung Diseases: Texture Classification for Differentiation at CT 1," *Radiology,* vol. 228, pp. 871-877, 2003.

[96] S. Herlidou-Meme, J. Constans, B. Carsin, D. Olivie, P. Eliat, L. Nadal-Desbarats*, et al.*, "MRI texture analysis on texture test objects, normal brain and intracranial tumors," *Magnetic resonance imaging,* vol. 21, pp. 989-993, 2003.

[97] Y. C. Chang, C. S. Huang, Y. J. Liu, J. H. Chen, Y. S. Lu, and W. Y. Tseng, "Angiogenic response of locally advanced breast cancer to neoadjuvant chemotherapy evaluated with parametric histogram from dynamic contrast-enhanced MRI," *Phys Med Biol,* vol. 49, pp. 3593-3602, Aug 21 2004.

[98] Q. G. de Lussanet, W. H. Backes, A. W. Griffioen, A. R. Padhani, C. I. Baeten, A. van Baardwijk*, et al.*, "Dynamic contrast-enhanced magnetic resonance imaging of radiation therapy-induced microcirculation changes in rectal cancer," *Int J Radiat Oncol Biol Phys,* vol. 63, pp. 1309-15, Dec 1 2005.

[99] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. Hoffman, "MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies," *IEEE Transactions on Medical Imaging,* vol. 25, pp. 464-475, 2006.

[100] A. Jackson, J. P. O'Connor, G. J. Parker, and G. C. Jayson, "Imaging tumor vascular heterogeneity and angiogenesis using dynamic contrast-enhanced magnetic resonance imaging," *Clinical Cancer Research,* vol. 13, pp. 3449-3459, 2007.

[101] J. J. Caban, J. Yao, N. A. Avila, J. R. Fontana, and V. C. Manganiello, "Texture-based computer-aided diagnosis system for lung fibrosis," *Proc. SPIE 6514, Medical Imaging: Computer-Aided Diagnosis,* p. 651439, 2007.

[102] A. Depeursinge, D. Sage, A. Hidki, A. Platon, P.-A. Poletti, M. Unser*, et al.*, "Lung tissue classification using wavelet frames," *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. ,* pp. 6259-6262, 2007.

[103] A. Karahaliou, S. Skiadopoulos, I. Boniatis, P. Sakellaropoulos, E. Likaki, G. Panayiotakis*, et al.*, "Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis," *Br J Radiol,* vol. 80, pp. 648-656, Aug 2007.

[104] D. Kontos, P. R. Bakic, A.-K. Carton, A. B. Troxel, E. F. Conant, and A. D. Maidment, "Parenchymal texture analysis in digital breast tomosynthesis for breast cancer risk estimation: a preliminary study," *Academic radiology,* vol. 16, pp. 283-298, 2009.

[105] M. E. Mayerhoefer, W. Schima, S. Trattnig, K. Pinker, V. Berger-Kulemann, and A. Ba-Ssalamah, "Texture-based classification of focal liver lesions on MRI at 3.0 Tesla: A feasibility study in cysts and hemangiomas," *Journal of Magnetic Resonance Imaging,* vol. 32, pp. 352-359, 2010.

[106] K. Holli, A.-L. Lääperi, L. Harrison, T. Luukkaala, T. Toivonen, P. Ryymin*, et al.*, "Characterization of breast cancer types by texture analysis of magnetic resonance images," *Academic radiology,* vol. 17, pp. 135-141, 2010.

[107] T. K. Chuah, C. L. Poh, and K. Sheah, "Quantitative texture analysis of MRI images for detection of cartilage-related bone marrow edema," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2011, pp. 5112-5, 2011.

[108] J. O'Connor, C. Rose, A. Jackson, Y. Watson, S. Cheung, F. Maders*, et al.*, "DCE-MRI biomarkers of tumour heterogeneity predict CRC liver metastasis shrinkage following bevacizumab and FOLFOX-6," *British journal of cancer,* vol. 105, pp. 139-145, 2011.

[109] C. Cui, H. Cai, L. Liu, L. Li, H. Tian, and L. Li, "Quantitative analysis and prediction of regional lymph node status in rectal cancer based on computed tomography imaging," *European radiology,* vol. 21, pp. 2318-2325, 2011.

[110] K. Fujimoto, T. Tonan, S. Azuma, M. Kage, O. Nakashima, T. Johkoh*, et al.*, "Evaluation of the mean and entropy of apparent diffusion coefficient values in chronic hepatitis C: correlation with pathologic fibrosis stage and inflammatory activity grade," *Radiology,* vol. 258, pp. 739-748, 2011.

[111] M. Vaidya, K. M. Creach, J. Frye, F. Dehdashti, J. D. Bradley, and I. El Naqa, "Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer," *Radiotherapy and Oncology,* vol. 102, pp. 239-245, 2012.

[112] S. Chicklore, V. Goh, M. Siddique, A. Roy, P. K. Marsden, and G. J. Cook, "Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis," *European journal of nuclear medicine and molecular imaging,* vol. 40, pp. 133-140, 2013.

[113] R. T. Leijenaar, S. Carvalho, E. R. Velazquez, W. J. Van Elmpt, C. Parmar, O. S. Hoekstra*, et al.*, "Stability of FDG-PET Radiomics features: An integrated analysis

of test-retest and inter-observer variability," *Acta Oncologica,* vol. 52, pp. 1391-1397, 2013.

[114] P. Foroutan, J. M. Kreahling, D. L. Morse, O. Grove, M. C. Lloyd, D. Reed*, et al.*, "Diffusion MRI and novel texture analysis in osteosarcoma xenotransplants predicts response to anti-checkpoint therapy," *PLoS One,* vol. 8, p. e82875, 2013.

[115] J. R. Sato, D. Y. Takahashi, M. Q. Hoexter, K. B. Massirer, and A. Fujita, "Measuring network's entropy in ADHD: a new approach to investigate neuropsychiatric disorders," *Neuroimage,* vol. 77, pp. 44-51, Aug 15 2013.

[116] S. Suoranta, K. Holli-Helenius, P. Koskenkorva, E. Niskanen, M. Kononen, M. Aikia*, et al.*, "3D texture analysis reveals imperceptible MRI textural alterations in the thalamus and putamen in progressive myoclonic epilepsy type 1, EPM1," *PLoS One,* vol. 8, p. e69905, 2013.

[117] C. Parmar, E. R. Velazquez, R. Leijenaar, M. Jermoumi, S. Carvalho, R. H. Mak*, et al.*, "Robust radiomics feature quantification using semiautomatic volumetric segmentation," *PloS one,* vol. 9, p. e102107, 2014.

[118] R. T. Leijenaar, S. Carvalho, F. J. Hoebers, H. J. Aerts, W. J. van Elmpt, S. H. Huang*, et al.*, "External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma," *Acta Oncologica,* pp. 1-7, 2015.

[119] K. M. Panth, R. T. Leijenaar, S. Carvalho, N. G. Lieuwes, A. Yaromina, L. Dubois*, et al.*, "Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells," *Radiotherapy and Oncology,* vol. 116, pp. 462-466, 2015.

[120] C. Parmar, P. Grossmann, J. Bussink, P. Lambin, and H. J. Aerts, "Machine Learning methods for Quantitative Radiomic Biomarkers," *Scientific reports,* vol. 5, p. 13087, 2015.

[121] C. Parmar, R. T. Leijenaar, P. Grossmann, E. R. Velazquez, J. Bussink, D. Rietveld, *et al.*, "Radiomic feature clusters and Prognostic Signatures specific for Lung and Head & Neck cancer," *Scientific reports,* vol. 5, p. 11044, 2015.

[122] O. Grove, A. E. Berglund, M. B. Schabath, H. J. Aerts, A. Dekker, H. Wang, *et al.*, "Quantitative Computed Tomographic Descriptors Associate Tumor Shape Complexity and Intratumor Heterogeneity with Prognosis in Lung Adenocarcinoma," *PloS one,* vol. 10, p. e0118261, 2015.

[123] I. E. Magnin, F. Cluzeau, C. L. Odet, and A. Bremond, "Mammographic texture analysis: an evaluation of risk for developing breast cancer," *Optical Engineering,* vol. 25, pp. 156780-156780, 1986.

[124] C.-M. Wu, Y.-C. Chen, and K.-S. Hsieh, "Texture features for classification of ultrasonic liver images," *IEEE Transactions on Medical Imaging,* vol. 11, pp. 141-152, 1992.

[125] A. Mir, M. Hanmandlu, and S. Tandon, "Texture analysis of CT images," *IEEE Engineering in Medicine and Biology Magazine,* vol. 14, pp. 781-786, 1995.

[126] D. Wei, H. P. Chan, M. A. Helvie, B. Sahiner, N. Petrick, D. D. Adler, *et al.*, "Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis," *Medical Physics,* vol. 22, pp. 1501-1513, 1995.

[127] Y. Sun, M.-H. Horng, X. Lin, and J.-Y. Wang, "Ultrasonic image analysis for liver diagnosis," *IEEE Engineering in Medicine and Biology Magazine,* vol. 15, pp. 93-101, Nov-Dec 1996.

[128] R. Lucht, G. Brix, and W. Lorenz, "Texture analysis of differently reconstructed PET images," *Physics in medicine and biology,* vol. 41, pp. 2207-2219, 1996.

[129] N. Petrick, H. P. Chan, D. Wei, B. Sahiner, M. A. Helvie, and D. D. Adler, "Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification," *Medical Physics,* vol. 23, pp. 1685-1696, 1996.

[130] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M. Helvie, D. D. Adler*, et al.*, "Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images," *IEEE Transactions on Medical Imaging,* vol. 15, pp. 598-610, 1996.

[131] H.-P. Chan, B. Sahiner, N. Petrick, M. A. Helvie, K. L. Lam, D. D. Adler*, et al.*, "Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network," *Physics in Medicine and Biology,* vol. 42, pp. 549-567, 1997.

[132] D. Wei, H.-P. Chan, N. Petrick, B. Sahiner, M. A. Helvie, D. D. Adler*, et al.*, "False-positive reduction technique for detection of masses on digital mammograms: Global and local multiresolution texture analysis," *Medical Physics,* vol. 24, pp. 903-914, 1997.

[133] P. Freeborough and N. C. Fox, "MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease," *IEEE Transactions on Medical Imaging,* vol. 17, pp. 475-478, 1998.

[134] B. Sahiner, H.-P. Chan, N. Petrick, M. A. Helvie, and M. M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis," *Medical Physics,* vol. 25, pp. 516-526, 1998.

[135] E. Chen, P.-C. Chung, C.-L. Chen, H.-M. Tsai, and C.-I. Chang, "An automatic diagnostic system for CT liver image classification," *Biomedical Engineering, IEEE Transactions on,* vol. 45, pp. 783-794, 1998.

[136] N. R. Mudigonda, R. M. Rangayyan, and J. L. Desautels, "Gradient and texture analysis for the classification of mammographic masses," *IEEE Transactions on Medical Imaging,* vol. 19, pp. 1032-1043, 2000.

[137] V. Kovalev, F. Kruggel, H.-J. Gertz, and D. Y. Von Cramon, "Three-dimensional texture analysis of MRI brain datasets," *IEEE Transactions on Medical Imaging,* vol. 20, pp. 424-433, 2001.

[138] R. Sivaramakrishna, K. A. Powell, M. L. Lieber, W. A. Chilcote, and R. Shekhar, "Texture analysis of lesions in breast ultrasound images," *Computerized medical imaging and graphics,* vol. 26, pp. 303-307, 2002.

[139] M. H. Horng, Y. N. Sun, and X. Z. Lin, "Texture feature coding method for classification of liver sonography," *Comput Med Imaging Graph,* vol. 26, pp. 33-42, Jan-Feb 2002.

[140] M. Gletsos, S. G. Mougiakakou, G. K. Matsopoulos, K. S. Nikita, A. S. Nikita, and D. Kelekis, "A computer-aided diagnostic system to characterize CT focal liver lesions: design and optimization of a neural network classifier," *Information Technology in Biomedicine, IEEE Transactions on,* vol. 7, pp. 153-162, 2003.

[141] D. Mahmoud-Ghoneim, G. Toussaint, J.-M. Constans, and D. Jacques, "Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas," *Magnetic resonance imaging,* vol. 21, pp. 983-987, 2003.

[142] R. R. Pereira Jr, P. M. A. Marques, M. O. Honda, S. K. Kinoshita, R. Engelmann, C. Muramatsu*, et al.*, "Usefulness of texture analysis for computerized classification of breast lesions on mammograms," *Journal of digital imaging,* vol. 20, pp. 248-255, 2007.

[143] K. Nie, J.-H. Chen, J. Y. Hon, Y. Chu, O. Nalcioglu, and M.-Y. Su, "Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI," *Academic radiology,* vol. 15, pp. 1513-1525, 2008.

[144] I. El Naqa, P. Grigsby, A. Apte, E. Kidd, E. Donnelly, D. Khullar*, et al.*, "Exploring feature-based approaches in PET images for predicting cancer treatment outcomes," *Pattern recognition,* vol. 42, pp. 1162-1171, 2009.

[145] A. Kassner, F. Liu, R. E. Thornhill, G. Tomlinson, and D. J. Mikulis, "Prediction of hemorrhagic transformation in acute ischemic stroke using texture analysis of postcontrast T1-weighted MR images," *Journal of Magnetic Resonance Imaging,* vol. 30, pp. 933-941, 2009.

[146] C. E. McLaren, W. P. Chen, K. Nie, and M. Y. Su, "Prediction of malignant breast lesions from MRI features: a comparison of artificial neural network and logistic regression techniques," *Acad Radiol,* vol. 16, pp. 842-51, Jul 2009.

[147] M. E. Mayerhoefer, G. H. Welsch, G. Riegler, T. C. Mamisch, A. Materka, M. Weber*, et al.*, "Feasibility of texture analysis for the assessment of biochemical changes in meniscal tissue on T1 maps calculated from delayed gadolinium-enhanced magnetic resonance imaging of cartilage data: comparison with conventional relaxation time measurements," *Invest Radiol,* vol. 45, pp. 543-7, Sep 2010.

[148] P. D. Korfiatis, A. N. Karahaliou, A. D. Kazantzi, C. Kalogeropoulou, and L. Costaridou, "Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector CT," *IEEE Transactions on Information Technology in Biomedicine,* vol. 14, pp. 675-680, 2010.

[149] M. E. Mayerhoefer, D. Stelzeneder, W. Bachbauer, G. H. Welsch, T. C. Mamisch, P. Szczypinski*, et al.*, "Quantitative analysis of lumbar intervertebral disc abnormalities at 3.0 Tesla: value of T(2) texture features and geometric parameters," *NMR Biomed,* vol. 25, pp. 866-72, Jun 2012.

[150] R. A. Gatenby, O. Grove, and R. J. Gillies, "Quantitative imaging in cancer evolution and ecology," *Radiology,* vol. 269, pp. 8-14, 2013.

[151] H. Cai, Y. Peng, C. Ou, M. Chen, and L. Li, "Diagnosis of breast masses from dynamic contrast-enhanced and diffusion-weighted MR: a machine learning approach," *PLoS One,* vol. 9, p. e87387, 2014.

[152] H. Cai, L. Liu, Y. Peng, Y. Wu, and L. Li, "Diagnostic assessment by dynamic contrast-enhanced and diffusion-weighted magnetic resonance in differentiation of breast lesions under different imaging protocols," *BMC Cancer,* vol. 14, p. 366, 2014.

[153] J. N. Stember, F. M. Deng, S. S. Taneja, and A. B. Rosenkrantz, "Pilot study of a novel tool for input-free automated identification of transition zone prostate tumors using T2- and diffusion-weighted signal and textural features," *J Magn Reson Imaging,* vol. 40, pp. 301-5, Aug 2014.

[154] T. C. Wang, Y. H. Huang, C. S. Huang, J. H. Chen, G. Y. Huang, Y. C. Chang*, et al.*, "Computer-aided diagnosis of breast DCE-MRI using pharmacokinetic model and 3-D morphology analysis," *Magn Reson Imaging,* vol. 32, pp. 197-205, Apr 2014.

[155] A. Chu, C. M. Sehgal, and J. F. Greenleaf, "Use of gray value distribution of run lengths for texture analysis," *Pattern Recognition Letters,* vol. 11, pp. 415-419, 1990.

[156] B. V. Dasarathy and E. B. Holder, "Image characterizations based on joint gray level—run length distributions," *Pattern Recognition Letters,* vol. 12, pp. 497-502, 1991.

[157] B. B. Mandelbrot, "How long is the coast of Britain," *Science,* vol. 156, pp. 636-638, 1967.

[158] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: a review," *Medical image analysis,* vol. 13, pp. 634-649, 2009.

[159] H. Li, M. L. Giger, O. I. Olopade, and L. Lan, "Fractal analysis of mammographic parenchymal patterns in breast cancer risk assessment," *Academic radiology,* vol. 14, pp. 513-521, 2007.

[160] J. Gagnepain and C. Roques-Carmes, "Fractal approach to two-dimensional and three-dimensional surface roughness," *Wear,* vol. 109, pp. 119-126, 1986.

[161] N. Sarkar and B. Chaudhuri, "An efficient differential box-counting approach to compute fractal dimension of image," *IEEE Transactions on Systems, Man and Cybernetics,* vol. 24, pp. 115-120, 1994.

[162] S. Peleg, J. Naor, R. Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pp. 518-523, 1984.

[163] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pp. 661-674, 1984.

[164] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM review,* vol. 10, pp. 422-437, 1968.

[165] R. M. Rangayyan, N. R. Mudigonda, and J. L. Desautels, "Boundary modelling and shape analysis methods for classification of mammographic masses," *Medical and Biological Engineering and Computing,* vol. 38, pp. 487-496, 2000.

[166] C. J. Rose, S. J. Mills, J. P. O'Connor, G. A. Buonaccorsi, C. Roberts, Y. Watson, *et al.*, "Quantifying spatial heterogeneity in dynamic contrast-enhanced MRI parameter maps," *Magnetic Resonance in Medicine,* vol. 62, pp. 488-499, 2009.

[167] T. Randen and J. H. Husoy, "Filtering for texture classification: A comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 21, pp. 291-310, 1999.

[168] B. Ganeshan, K. A. Miles, R. Young, and C. Chatwin, "Hepatic entropy and uniformity: additional parameters that can potentially increase the effectiveness of contrast enhancement during abdominal CT," *Clinical radiology,* vol. 62, pp. 761-768, 2007.

[169] B. Ganeshan, K. A. Miles, R. C. Young, and C. R. Chatwin, "In search of biologic correlates for liver texture on portal-phase CT," *Academic radiology,* vol. 14, pp. 1058-1068, 2007.

[170] K. A. Miles, B. Ganeshan, M. R. Griffiths, R. C. Young, and C. R. Chatwin, "Colorectal Cancer: Texture Analysis of Portal Phase Hepatic CT Images as a Potential Marker of Survival 1," *Radiology,* vol. 250, pp. 444-452, 2009.

[171] B. Ganeshan, S. Abaleke, R. C. Young, C. R. Chatwin, and K. A. Miles, "Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage," *Cancer Imaging,* vol. 10, pp. 137-143, 2010.

[172] B. Ganeshan, K. Burnand, R. Young, C. Chatwin, and K. Miles, "Dynamic contrast-enhanced texture analysis of the liver: initial assessment in colorectal cancer," *Investigative radiology,* vol. 46, pp. 160-168, 2011.

[173] C. Wachinger and N. Navab, "Entropy and Laplacian images: Structural representations for multi-modal registration," *Medical Image Analysis,* vol. 16, pp. 1-17, 2012.

[174] B. Ganeshan, E. Panayiotou, K. Burnand, S. Dizdarevic, and K. Miles, "Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival," *European radiology,* vol. 22, pp. 796-802, 2012.

[175] B. Ganeshan, V. Goh, H. C. Mandeville, Q. S. Ng, P. J. Hoskin, and K. A. Miles, "Non–small cell lung cancer: histopathologic correlates for texture parameters at CT," *Radiology,* vol. 266, pp. 326-336, 2013.

[176] F. Ng, B. Ganeshan, R. Kozarski, K. A. Miles, and V. Goh, "Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival," *Radiology,* vol. 266, pp. 177-184, 2013.

[177] K. I. Laws, "Textured image segmentation," Ph.D. Ph.D. Dissertation, Image Processing Institute, University of Southern California, Los Angeles, California, 1980.

[178] G. Cox, F. Hoare, and G. de Jager, "Experiments in lung cancer nodule detection using texture analysis and neural network classifiers," *Third South African Workshop on Pattern Recognition,* vol. 31, pp. 136-142, 1992.

[179] P. Miller and S. Astley, "Classification of breast tissue by texture analysis," *Image and Vision Computing,* vol. 10, pp. 277-282, 1992.

[180] Y. Chu, L. Li, D. B. Goldgof, Y. Qui, and R. A. Clark, "Classification of masses on mammograms using support vector machine," *Proc. SPIE 5032, Medical Imaging: Image Processing,* pp. 940-948, 2003.

[181] S. Poonguzhali and G. Ravindran, "Automatic classification of focal lesions in ultrasound liver images using combined texture features," *Information Technology Journal,* vol. 7, pp. 205-209, 2008.

[182] J. Awad, A. Krasinski, G. Parraga, and A. Fenster, "Texture analysis of carotid artery atherosclerosis from three-dimensional ultrasound images," *Medical Physics,* vol. 37, pp. 1382-1391, 2010.

[183] J. Dheeba and S. Tamil Selvi, "Classification of malignant and benign microcalcification using SVM classifier," *International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT),* pp. 686-690, 2011.

[184] C. Barata, J. S. Marques, and T. Mendonça, "Bag-of-features classification model for the diagnose of melanoma in dermoscopy images using color and texture descriptors," *Image Analysis and Recognition,* pp. 547-555, 2013.

[185] J. Virmani, V. Kumar, N. Kalra, and N. Khandelwa, "Pca-SVm based caD System for Focal liver lesions using B-mode ultrasound Images," *Defence Science Journal,* vol. 63, pp. 478-486, 2013.

[186] L. C. Pereyra, R. M. Rangayyan, M. Ponciano-Silva, and P. M. Azevedo-Marques, "Fractal analysis for computer-aided diagnosis of diffuse pulmonary diseases in HRCT images," *IEEE International Symposium on Medical Measurements and Applications (MeMeA)* pp. 1-6, 2014.

[187] S. Dilger, A. Judisch, J. Uthoff, E. Hammond, J. Newell, and J. Sieren, "Improved pulmonary nodule classification utilizing lung parenchyma texture features," *Proc. SPIE 9414, Medical Imaging: Computer-Aided Diagnosis,* p. 94142T, 2015.

[188] D. Mitrea, S. Nedevschi, and M. Abrudean, "Classification of the liver tumors Using Co-Occurrence Matrices of textural Microstructures," *Journal of Communication and Computer,* vol. 12, pp. 6-12, 2015.

[189] B. K. Singh, K. Verma, and A. Thoke, "Adaptive Gradient Descent Backpropagation for Classification of Breast Tumors in Ultrasound Imaging," *Procedia Computer Science,* vol. 46, pp. 1601-1609, 2015.

[190] A. Haar, "Zur theorie der orthogonalen funktionensysteme," *Mathematische Annalen,* vol. 69, pp. 331-371, 1910.

[191] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on pure and applied mathematics,* vol. 41, pp. 909-996, 1988.

[192] I. Daubechies, *Ten lectures on wavelets* vol. 61: Society for Industrial and Applied Mathematics, 1992.

[193] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 11, pp. 674-693, 1989.

[194] D.-R. Chen, R.-F. Chang, W.-J. Kuo, M.-C. Chen, and Y.-L. Huang, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound in medicine & biology,* vol. 28, pp. 1301-1310, 2002.

[195] A. Kassner and R. Thornhill, "Texture analysis: a review of neurologic MR imaging applications," *American Journal of Neuroradiology,* vol. 31, pp. 809-816, 2010.

[196] K. A. Miles, B. Ganeshan, and M. P. Hayball, "CT texture analysis using the filtration-histogram method: what do the measurements mean?," *Cancer Imaging,* vol. 13, pp. 400-6, 2013.

[197] R. H. de Melo, E. d. A. Vieira, and A. Conci, "Characterizing the lacunarity of objects and image sets and its use as a technique for the analysis of textural patterns," *Advanced concepts for intelligent vision systems,* pp. 208-219, 2006.

[198] R. Nandi, A. K. Nandi, R. M. Rangayyan, and D. Scutt, "Classification of breast masses in mammograms using genetic programming and feature selection," *Medical and Biological Engineering and Computing,* vol. 44, pp. 683-694, 2006.

[199] R. M. Rangayyan and T. M. Nguyen, "Fractal analysis of contours of breast masses in mammograms," *Journal of Digital Imaging,* vol. 20, pp. 223-237, 2007.

[200] D.-R. Chen, R.-F. Chang, Y.-L. Huang, Y.-H. Chou, C.-M. Tiu, and P.-P. Tsai, "Texture analysis of breast tumors on sonograms," *Seminars in Ultrasound, CT and MRI,* vol. 21, pp. 308-316, 2000.

[201] J. Parikh, M. Selmi, G. Charles-Edwards, J. Glendenning, B. Ganeshan, H. Verma*, et al.*, "Changes in primary breast cancer heterogeneity may augment midtreatment MR imaging assessment of response to neoadjuvant chemotherapy," *Radiology,* vol. 272, pp. 100-12, Jul 2014.

[202] M. Gerlinger , A. J. Rowan , S. Horswell , J. Larkin , D. Endesfelder , E. Gronroos *, et al.*, "Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing," *New England Journal of Medicine,* vol. 366, pp. 883-892, 2012.

[203] Y.-L. Huang, J.-H. Chen, and W.-C. Shen, "Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images," *Academic radiology,* vol. 13, pp. 713-720, 2006.

[204] S. Poonguzhali and G. Ravindran, "Performance evaluation of feature extraction methods for classifying abnormalities in ultrasound liver images using neural network," *IEEE 28th Annual International Conference of the Engineering in Medicine and Biology Society, 2006.,* pp. 4791-4794, 2006.

[205] D. Mittal, V. Kumar, S. C. Saxena, N. Khandelwal, and N. Kalra, "Neural network based focal liver lesion diagnosis using ultrasound images," *Computerized Medical Imaging and Graphics,* vol. 35, pp. 315-323, 2011.

[206] J. Virmani, V. Kumar, N. Kalra, and N. Khadelwal, "A rapid approach for prediction of liver cirrhosis based on first order statistics," *2011 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT),* pp. 212-215, 2011.

[207] J. H. Jeon, J. Y. Choi, S. Lee, and Y. M. Ro, "Multiple ROI selection based focal liver lesion classification in ultrasound images," *Expert Systems with Applications,* vol. 40, pp. 450-457, 2013.

[208] G.-m. Xian, "An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM," *Expert Systems with Applications,* vol. 37, pp. 6737-6741, 2010.

[209] J. Virmani, V. Kumar, N. Kalra, and N. Khandelwal, "A comparative study of computer-aided classification systems for focal hepatic lesions from B-mode ultrasound," *Journal of medical engineering & technology,* vol. 37, pp. 292-306, 2013.

[210] V. Goh, B. Sanghera, D. M. Wellsted, J. Sundin, and S. Halligan, "Assessment of the spatial pattern of colorectal tumour perfusion estimated at perfusion CT using

two-dimensional fractal analysis," *European radiology,* vol. 19, pp. 1358-1365, 2009.

[211] D. R. Cox, "Regression Models and Life-Tables," *Journal of the Royal Statistical Society Series B-Statistical Methodology,* vol. 34, pp. 187-220, 1972.

[212] A. Wibmer, H. Hricak, T. Gondo, K. Matsumoto, H. Veeraraghavan, D. Fehr*, et al.*, "Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores," *European radiology,* vol. 25, pp. 2840-2850, 2015.

[213] P. S. Tofts, "Modeling tracer kinetics in dynamic Gd-DTPA MR imaging," *J Magn Reson Imaging,* vol. 7, pp. 91-101, Jan-Feb 1997.

[214] P. S. Tofts, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp*, et al.*, "Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols," *J Magn Reson Imaging,* vol. 10, pp. 223-32, Sep 1999.

[215] R. G. Pontius Jr and K. Si, "The total operating characteristic to measure diagnostic ability for multiple thresholds," *International Journal of Geographical Information Science,* vol. 28, pp. 570-583, 2014.

[216] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning,* vol. 20, pp. 273-297, 1995.

[217] C. Elkan, "The foundations of cost-sensitive learning," *International joint conference on artificial intelligence,* vol. 17, pp. 973-978, 2001.

[218] H. Degani, V. Gusis, D. Weinstein, S. Fields, and S. Strano, "Mapping pathophysiological features of breast tumors by MRI at high spatial resolution," *Nat Med,* vol. 3, pp. 780-782, 1997.

[219] 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,* vol. 210, pp. 233-40, 1999.

[220] Y. Guo, Y. Q. Cai, Z. L. Cai, Y. G. Gao, N. Y. An, L. Ma*, et al.*, "Differentiation of clinically benign and malignant breast lesions using diffusion-weighted imaging," *J Magn Reson Imaging,* vol. 16, pp. 172-8., 2002.

[221] R. Woodhams, K. Matsunaga, K. Iwabuchi, S. Kan, H. Hata, M. Kuranami*, et al.*, "Diffusion-weighted imaging of malignant breast tumors: the usefulness of apparent diffusion coefficient (ADC) value and ADC map for the detection of malignant breast tumors and evaluation of cancer extension," *J Comput Assist Tomogr,* vol. 29, pp. 644-9, Sep-Oct 2005.

[222] M. J. Park, E. S. Cha, B. J. Kang, Y. K. Ihn, and J. H. Baik, "The role of diffusion-weighted imaging and the apparent diffusion coefficient (ADC) values for breast tumors," *Korean J Radiol,* vol. 8, pp. 390-6, Sep-Oct 2007.

[223] N. M. deSouza, S. F. Riches, N. J. Vanas, V. A. Morgan, S. A. Ashley, C. Fisher*, et al.*, "Diffusion-weighted magnetic resonance imaging: a potential non-invasive marker of tumour aggressiveness in localized prostate cancer," *Clin Radiol,* vol. 63, pp. 774-82, Jul 2008.

[224] T. Neumann-Haefelin, H.-J. Wittsack, F. Wenserski, M. Siebler, R. J. Seitz, U. Mödder*, et al.*, "Diffusion-and perfusion-weighted MRI The DWI/PWI mismatch region in acute stroke," *Stroke,* vol. 30, pp. 1591-1597, 1999.

[225] P. W. Schaefer, G. J. Hunter, J. He, L. M. Hamberg, A. G. Sorensen, L. H. Schwamm*, et al.*, "Predicting cerebral ischemic infarct volume with diffusion and perfusion MR imaging," *AJNR Am J Neuroradiol,* vol. 23, pp. 1785-94, Nov-Dec 2002.

[226] A. E. Hillis, R. J. Wityk, N. J. Beauchamp, J. A. Ulatowski, M. A. Jacobs, and P. B. Barker, "Perfusion-weighted MRI as a marker of response to treatment in acute and subacute stroke," *Neuroradiology,* vol. 12, p. 12, 2003.

[227] H. Ma, P. Wright, L. Allport, T. G. Phan, L. Churilov, J. Ly*, et al.*, "Salvage of the PWI/DWI mismatch up to 48 h from stroke onset leads to favorable clinical outcome," *Int J Stroke,* vol. 10, pp. 565-70, Jun 2015.

[228] W. T. C. Yuh, M. D. Alexander, T. Ueda, M. Maeda, T. Taoka, K. Yamada*, et al.*, "Revisiting Current Golden Rules in Managing Acute Ischemic Stroke: Evaluation of New Strategies to Further Improve Treatment Selection and Outcome," *American Journal of Roentgenology,* vol. 208, pp. 32-41, 2017/01/01 2016.

[229] A. D. Wyner, "A definition of conditional mutual information for arbitrary ensembles," *Information and Control,* vol. 38, pp. 51-59, 1978/07/01/ 1978.

[230] P. Brodatz, *Textures: a photographic album for artists and designers*: Dover Pubns, 1966.

[231] M. A. Jacobs, P. Mitsias, H. Soltanian-Zadeh, S. Santhakumar, A. Ghanei, R. Hammond*, et al.*, "Multiparametric MRI Tissue Characterization in Clinical Stroke With Correlation to Clinical Outcome:Part 2," *Stroke,* vol. 32, pp. 950-957, 2001.

[232] K. Butcher, M. Parsons, L. Allport, S. B. Lee, P. A. Barber, B. Tress*, et al.*, "Rapid assessment of perfusion-diffusion mismatch," *Stroke,* vol. 39, pp. 75-81, Jan 2008.

[233] M. Motta, A. Ramadan, A. E. Hillis, R. F. Gottesman, and R. Leigh, "Diffusion-perfusion mismatch: an opportunity for improvement in cortical function," *Front Neurol,* vol. 5, p. 280, 2014.

[234] O. Maier, C. Schröder, N. D. Forkert, T. Martinetz, and H. Handels, "Classifiers for ischemic stroke lesion segmentation: a comparison study," *PloS one,* vol. 10, p. e0145118, 2015.

[235] M. A. Jacobs and A. Akhbardeh, "Advanced treatment response prediction using clinical parameters and advanced unsupervised machine learning: the contribution scattergram," *United States patent application US20160171695 A1,* 2016.

[236] V. Parekh and M. Jacobs, "Multidimensional Imaging Radiomics-Geodesics: A Novel Manifold Learning Based Automatic Feature Extraction Method for Diagnostic Prediction in Multiparametric Imaging," *Medical Physics,* vol. 43, pp. 3373-3374, 2016.

[237] W. W. Powell, K. W. Koput, and L. Smith-Doerr, "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology," *Administrative science quarterly,* pp. 116-145, 1996.

[238] M. E. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences,* vol. 98, pp. 404-409, 2001.

[239] M. E. Newman, "The structure and function of complex networks," *SIAM review,* vol. 45, pp. 167-256, 2003.

[240] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature,* vol. 435, pp. 814-818, 2005.

[241] B. Zhang and S. Horvath, "A general framework for weighted gene co-expression network analysis," *Statistical applications in genetics and molecular biology,* vol. 4, p. 1128, 2005.

[242] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics reports,* vol. 424, pp. 175-308, 2006.

[243] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience,* vol. 10, pp. 186-198, 2009.

[244] J. Scott, "Social network analysis," 2012.

[245] J. M. Lee, *Riemannian manifolds: an introduction to curvature* vol. 176: Springer Science & Business Media, 2006.

[246] J. Lee, *Introduction to topological manifolds* vol. 940: Springer Science & Business Media, 2010.

[247] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Advances in neural information processing systems*, 2002, pp. 585-591.

[248] V. D. Silva and J. B. Tenenbaum, "Global versus local methods in nonlinear dimensionality reduction," in *Advances in neural information processing systems*, 2002, pp. 705-712.

[249] M. Bernstein, V. De Silva, J. C. Langford, and J. B. Tenenbaum, "Graph approximations to geodesics on embedded manifolds," Technical report, Department of Psychology, Stanford University2000.

[250] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik,* vol. 1, pp. 269-271, 1959.

[251] R. W. Floyd, "Algorithm 97: shortest path," *Communications of the ACM,* vol. 5, p. 345, 1962.

[252] M. Balasubramanian and E. L. Schwartz, "The isomap algorithm and topological stability," *Science,* vol. 295, pp. 7-7, 2002.

[253] L. K. Saul and S. T. Roweis, "Think globally, fit locally: unsupervised learning of low dimensional manifolds," *The Journal of Machine Learning Research,* vol. 4, pp. 119-155, 2003.

[254] A. Bavelas, "Communication patterns in task-oriented groups," *Journal of the acoustical society of America,* 1950.

[255] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry,* pp. 35-41, 1977.

[256] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks,* vol. 1, pp. 215-239, 1978.

[257] T. M. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Software: Practice and experience,* vol. 21, pp. 1129-1164, 1991.

[258] S. Paik, S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin*, et al.*, "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer," *New England Journal of Medicine,* vol. 351, pp. 2817-2826, 2004.

[259] S. Paik, G. Tang, S. Shak, C. Kim, J. Baker, W. Kim*, et al.*, "Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor–positive breast cancer," *Journal of clinical oncology,* vol. 24, pp. 3726-3734, 2006.

[260] D. M. Ikeda, N. M. Hylton, K. Kinkel, M. G. Hochman, C. K. Kuhl, W. A. Kaiser*, et al.*, "Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging studies," *Journal of Magnetic Resonance Imaging,* vol. 13, pp. 889-895, 2001.

[261] G. Brix, W. Semmler, R. Port, L. R. Schad, G. Layer, and W. J. Lorenz, "Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging," *Journal of computer assisted tomography,* vol. 15, pp. 621-628, 1991.

[262] A. Radjenovic, B. Dall, J. Ridgway, and M. Smith, "Measurement of pharmacokinetic parameters in histologically graded invasive breast tumours using dynamic contrast-enhanced MRI," *Measurement,* vol. 81, 2008.

[263] P. S. Tofts, B. Berkowitz, and M. D. Schnall, "Quantitative Analysis of Dynamic Gd-DTPA Enhancement in Breast Tumors Using a Permeability Model," *Magnetic Resonance in Medicine,* vol. 33, pp. 564-568, 1995.

CODE BASE

I have wrapped all the algorithms developed in this thesis into a graphical user interface (GUI) using MATLAB. The GUI serves as a user-friendly interface for physicians, radiologists and medical students who are interested in running complex data analysis (e.g. radiomics, contribution scattergram, deep learning) on their data, but do not have an expertise in programming. However, the computational radiology GUI still requires users to have a thorough understanding of the various algorithms for generation of meaningful results. The computational radiology software can be found on https://github.com/vishwaparekh/Computational_Radiology.git. Drop me an email at vishwaparekh@jhu.edu if you would like to get access to the source code or contribute to the project.

Vita

Vishwa is a PhD candidate in Computer Science at the Johns Hopkins University. He is primarily advised by Dr. Michael Jacobs and co-advised by Dr. Russell Taylor and Dr. Jerry Prince. Vishwa received a B.E. in Computer Science from BITS, Pilani in 2011 and a M.S.E in Computer Science from JHU in 2013.

Vishwa's research interest lies in developing techniques that enable us to "see" patterns in high dimensional imaging data that are not visually perceivable to naked eye. During his Ph.D., Vishwa published 5 journal papers, 2 conference papers, 8 abstracts and filed 4 patents. His research in manifold and deep learning was covered in AuntMinnie.com for "on the road to RSNA" for years 2015 and 2017. In addition, his work on manifold learning in prostate imaging was selected for power pitch presentation (top 2%) at The International Society for Magnetic Resonance in Medicine in 2017.